

Same Forecasts, Different Portfolios: Factor Normalization and Regularization

Mingyang Liu

Department of Finance, Imperial Business School, Imperial College London

yang.liu19@imperial.ac.uk

October 2026

Abstract

Portfolio Sharpe ratios from statistical factor models can reflect allocation choices even when return forecasts coincide. Adding a ridge adjustment to factor covariances penalizes factor exposures in units set by normalization. We derive the transformation that preserves portfolio positions across equivalent model representations. Two separately estimated IPCA normalizations deliver nearly identical forecasts but gross Sharpe ratios of 0.69 and 1.07 under the same numerical penalty; transforming the penalty consistently aligns their portfolios. Using the full fitted mean, including characteristic-associated intercepts, produces a portfolio with a gross Sharpe ratio of 2.51 and little sensitivity to the penalty.

Keywords: conditional factor models; firm characteristics; portfolio choice; covariance regularization; factor model intercepts.

JEL classification: G11, G12, C38, C58.

I. Introduction

A factor model can forecast the same stock returns and yet lead an investor to hold a different portfolio. The difference need not come from forecasting. A mean–variance portfolio depends on both the model’s expected returns and the risk matrix used to turn those forecasts into positions. The latter is often regularized, and the economic meaning of a regularization choice can depend on how a statistical factor model is represented. We study when this dependence changes portfolio holdings, realized Sharpe ratios, and comparisons across factor models. The central question is whether a reported difference in portfolio performance reflects different return forecasts or a difference in the portfolio construction step.

The distinction is particularly sharp for latent factors. Their loadings and factor realizations can be rescaled without changing a single fitted stock return. Suppose one factor’s realizations are doubled while its stock loadings are halved. Its contribution to fitted returns and to the stock return covariance is unchanged. But adding the same positive constant to the variance of that factor in both representations does not make the same adjustment to stock risk. With B_t denoting the stocks’ fitted factor loadings and w a vector of portfolio positions, a ridge adjustment λI to the factor covariance adds

$$\frac{\lambda}{2} \|B_t' w\|^2$$

to the quadratic risk term of the mean–variance objective. Under the rescaling, this penalty on any given portfolio’s aggregate factor exposures is quartered. The fitted model is unchanged; the portfolio objective is not.

This observation yields a simple criterion for comparing portfolios across equivalent representations. A latent-factor normalization fixes the coordinates in which factor exposures are reported, but it does not fix an economically meaningful penalty in those coordinates. Proposition 1 shows that if factor loadings are re-expressed as $B_t J^{-1}$, an adjustment R to their factor covariance must be carried over as $J R J'$ to preserve the stock-level risk adjustment. Applying the same λI in both

representations instead preserves the portfolio only when J is orthogonal. One can therefore specify the adjustment in a reference set of factor units and transform it consistently, or specify a penalty directly on stock positions. Neither procedure identifies a preferred penalty strength; both prevent a change of factor normalization from inadvertently changing the investor’s objective.

We measure the consequence in rolling portfolios of U.S. stocks constructed from 132 firm characteristics. We estimate instrumented principal components analysis (IPCA; Kelly, Pruitt, and Su, 2019) with three latent factors and five observed Fama–French factors, using 120-month estimation windows. Two independently estimated specifications differ only in how the latent loadings are normalized. The Euclidean normalization makes their characteristic-loading coefficients orthonormal, while the Gram normalization uses the cross-sectional second moments of the characteristics to normalize the stock exposures generated by those coefficients (Liu, 2026a). Their forecasts agree closely: the median relative difference across monthly stock-forecast vectors is 4.8×10^{-6} . Nevertheless, at a factor covariance penalty of $\lambda = 0.01$ in each specification’s own units, the median L_1 distance between their unit-gross portfolio weights is 0.58. The two portfolios earn annualized gross Sharpe ratios of 0.69 and 1.07 over the same 594 formation months, June 1975–November 2024 (Table II; Figure 1).

The holdings comparison isolates the mechanism rather than merely documenting two performance numbers. When the factor penalty is expressed in a common set of units, the separately estimated IPCA portfolios coincide to numerical precision: the median one-minus-cosine of their weights is approximately 10^{-11} . The fitted factor covariance matrices remain positive definite without the adjustment, so it is not needed to make those matrices invertible (Internet Appendix IA.D.2). With no factor covariance penalty, both normalizations earn a gross Sharpe ratio of 1.62. The result is not specific to the reference value. Across the six penalties in Figure 1, the two Sharpe ratios agree without the adjustment and separate at larger penalties. At the reference value their difference is 0.38, with a pointwise 95% confidence interval of $[0.09, 0.64]$ conditional on the fitted models

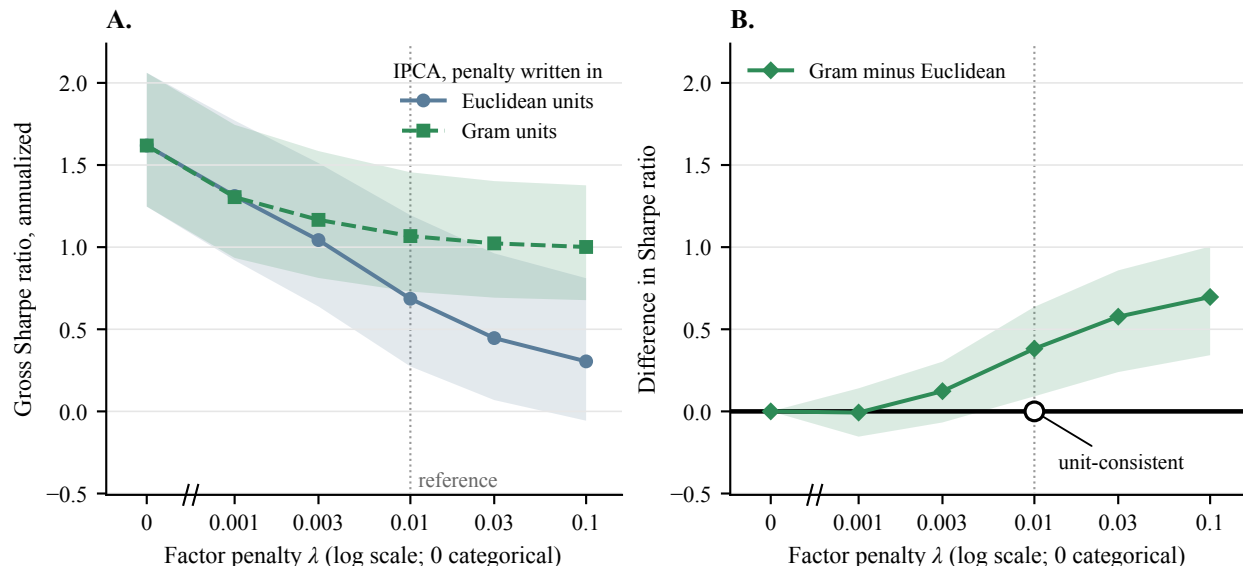


Figure 1. The factor penalty and the Sharpe ratio. Panel A: annualized gross Sharpe ratio at unit gross exposure of IPCA under the Euclidean and the Gram normalization, with the factor penalty λ written in each normalization’s own units. Panel B: Gram minus Euclidean; the open marker is the reference penalty written in one set of units. Bands are pointwise 95% intervals from a 12-month moving-block bootstrap that holds the fitted models fixed; $\lambda = 0$ is a categorical tick on the logarithmic axis. 594 months, June 1975–November 2024.

(Table II, Panel B).¹ Thus nearly identical forecasts need not imply similar positions or Sharpe ratios unless the portfolio’s risk adjustment is specified consistently.

The same concern enters comparisons across model families, although different families need not provide the same forecasts and Proposition 1 does not equate them. Under our common allocation rule, risk-premium PCA (RP-PCA; Lettau and Pelger, 2020) earns a gross Sharpe ratio of 1.49 against 0.69 for Euclidean-normalized IPCA at the reference penalty. The gap of 0.80 becomes -0.08 , statistically indistinguishable from zero in this sample, without the factor covariance penalty; RP-PCA itself changes little along that comparison (Figure 3; Table II). Replacing the fixed adjustment by Ledoit–Wolf linear shrinkage toward a scaled identity does not by itself resolve the units issue: applied separately in the two IPCA normalizations, it yields Sharpe ratios of 1.17 and 1.54 (Table II, Panel C). A ranking by portfolio Sharpe ratios therefore cannot be attributed to

¹Over all 618 formation months the corresponding pointwise interval includes zero (Internet Appendix IA.E); the holdings comparison and the normalization-invariance result are not statistical significance claims.

forecasting differences without also specifying how the risk input is regularized.

The portfolio objective has a second input: which parts of the model’s fitted expected return the investor uses. IPCA allows a characteristic-associated intercept in addition to returns implied by its common factors (Kelly, Pruitt, and Su, 2019). This intercept is the component of the fitted expected return not absorbed by the factor block. Its economic interpretation is not imposed here: it could reflect pricing errors, omitted factors, or conditional dynamics outside that block. We ask what happens when it is included in the expected-return vector passed to the same mean–variance optimizer. This comparison links the return input to the factor-exposure penalty studied above.

In each rolling window we hold one fitted intercept-inclusive IPCA specification and its risk matrix fixed, and construct two portfolios. The *factor-mean portfolio* uses the estimated factor component alone, while the *intercept portfolio* uses the complete fitted expected return. Writing $\widehat{m}_{F,t}$ for the factor means, $Z_t \widehat{\Gamma}_{\alpha,t}$ for the fitted stock-level intercepts, and A_t for the common penalized risk matrix, the two raw positions satisfy

$$w_t^{\text{FM}} \propto A_t^{-1} B_t \widehat{m}_{F,t}, \quad w_t^{\text{IP}} \propto A_t^{-1} (B_t \widehat{m}_{F,t} + Z_t \widehat{\Gamma}_{\alpha,t}).$$

Both are subsequently scaled to unit gross exposure. The comparison changes the mean input, not the estimated model or its risk matrix. The factor-mean portfolio of an intercept-inclusive fit is also distinct from a portfolio estimated from a model that excludes the intercept entirely; we compare with those zero-intercept benchmarks separately.

The factor penalty provides a reason for the two portfolios to respond differently. Lemma 1 shows that the adjustment to a portfolio’s weights depends on the aggregate factor exposures it would hold without the penalty. The factor-mean forecast is carried entirely through the factor block. Adding the fitted intercept can support positions whose expected return is not carried by that same block, making the intercept portfolio less sensitive to a penalty on factor exposures. This is not a factor-neutrality result: orthogonality of coefficient vectors under a characteristic-space

metric does not imply that the resulting stock portfolio has zero factor exposure. Empirically, the intercept portfolio's gross Sharpe ratio ranges only from 2.50 to 2.51 across the full penalty grid, whereas the factor-mean portfolio of the same fit declines from 0.56 to 0.37 (Figure 4; Table III). Re-expressing the full fit in Euclidean units leaves the intercept portfolio identical without the factor penalty and changes its Sharpe ratio by less than 0.01 at positive penalties when the residual risk estimate is held fixed (Internet Appendix IA.F).

We also compare the intercept portfolio with the highest-Sharpe zero-intercept benchmark among those evaluated under the common portfolio rule: unpenalized IPCA, with a gross Sharpe ratio of 1.62. The intercept portfolio earns 2.51 at the reference penalty, a difference of 0.89 with a paired 12-month moving-block bootstrap interval of [0.59, 1.26] (Table III, Panel B).² Under a proportional trading cost of 25 basis points per dollar purchased or sold, the corresponding net Sharpe ratios are 1.60 and 1.03, a difference of 0.57 with an interval of [0.29, 0.89] (Table IV).

The full-sample gain over the best zero-intercept benchmark primarily reflects lower risk per dollar of gross position, not a reliably larger average return. The intercept portfolio has an annualized mean of 7.6% and volatility of 3.0%, against 6.6% and 4.1% for the zero-intercept benchmark; the difference in means is within sampling error (Figure 6). In a time-series regression on that benchmark and the five Fama–French factors, the intercept portfolio has an annualized regression intercept of 4.9% and an R^2 of 0.34 (Table V). The unexplained component is a reduced-form return residual relative to these regressors, not by itself an identified mispricing return or proof that the portfolio is factor-neutral. In line with this interpretation, shuffling fitted intercepts across stocks largely eliminates the portfolio's Sharpe ratio, while removing a group of short-horizon return and trading-activity characteristics reduces its advantage over the zero-intercept benchmark until the difference is no longer statistically distinguishable from zero (Internet Appendix IA.M;

²The common 594-month sample and the plain intercept portfolio as the main-text rule were selected after examining the 618- and 594-month samples and the plain and stabilized intercept rules (Internet Appendix IA.D.1). The pointwise intervals do not adjust for these specification choices.

Section VI.D).

The economic size of the intercept result also depends on when and how the strategy is evaluated. Delaying trades by one month lowers its gross Sharpe ratio from 2.51 to 1.52, against a decline from 1.62 to 1.07 for the delayed zero-intercept benchmark (Table III, Panel C). The portfolio’s net advantage weakens in the later sample: over October 2007–November 2024, its net Sharpe difference relative to the best zero-intercept benchmark is 0.19 with an interval of $[-0.41, 0.75]$ (Internet Appendix IA.I). The cost calculations assume proportional costs and omit stock-specific price impact, financing, short-sale fees, and capacity. Moreover, the processed panel admits formation-month stock observations only when a subsequent return is recorded, and the point-in-time availability of each characteristic has not been independently verified (Section III; Internet Appendix IA.B). The reported bootstrap intervals hold the rolling fits fixed and are pointwise rather than joint across specifications.

Our analysis relates first to covariance estimation and regularization in portfolio choice, including Jagannathan and Ma (2003), Ledoit and Wolf (2004), Ledoit and Wolf (2017), and DeMiguel, Garlappi, and Uppal (2009). It also speaks to the choice of coordinates in penalized asset-pricing models studied by Kozak, Nagel, and Santosh (2020), whose stochastic-discount-factor shrinkage is a distinct estimation problem, and to comparisons of portfolio optimizers and factor models such as Deng et al. (2024). These papers establish the importance of covariance estimation, shrinkage, and the allocation rule. We isolate a different margin: with the fitted latent-factor model and its expected returns held fixed, a numerically identical factor covariance adjustment can impose a different stock-level penalty solely because the factor normalization differs. The invariance condition tells the researcher how to carry an adjustment across equivalent representations, not how to choose its optimal strength.

A second connection is to characteristic-based latent-factor estimation and its interpretation (Kelly, Pruitt, and Su, 2019; Lettau and Pelger, 2020; Gu, Kelly, and Xiu, 2021; Zhang, 2024; Choi and Yuan, 2025). Such models identify fitted stock returns without assigning economically

unique units to every latent factor. Our results show why this freedom matters when factor estimates become inputs to portfolio choice. A third connection is to portfolios that use expected returns outside a fitted factor block, including the principal portfolios of Kelly, Malamud, and Pedersen (2023), hedged characteristic portfolios of Daniel et al. (2020), and arbitrage portfolios of Kim, Korajczyk, and Neuhierl (2021) and Chen, Roussanov, and Wang (2025). We neither introduce the characteristic-associated intercept nor impose an explicit factor hedge. Instead, we compare the full fitted mean with the factor-only component of the same fit, characterize its sensitivity to a factor-exposure penalty, and examine how much of its observed portfolio advantage survives benchmark, cost, timing, and subperiod comparisons. The concerns about estimation risk and trading costs raised by Langlois (2023) remain relevant to the interpretation of our results.

The paper’s contribution is thus to separate two ingredients often bundled into a model’s reported portfolio Sharpe ratio. First, we establish when normalization-dependent factor covariance regularization changes stock allocations despite unchanged fitted returns, and show how a consistently transformed penalty restores invariance. Second, we document the size of this effect within IPCA and its implications for cross-model portfolio comparisons. Third, within an intercept-inclusive IPCA fit, we show how using the full fitted expected return changes sensitivity to the same penalty and quantify the resulting portfolio performance, with explicit limits on its implementation and persistence. A factor-model comparison based on portfolio returns should therefore state both the mean component being used and the units and scale of any factor-exposure regularization.

Section II develops the portfolio objective, the normalization map, and the exposure result. Section III describes the data, rolling estimation, allocation rule, and benchmarks. Section IV examines normalization and penalty strength within and across models. Section V introduces the intercept portfolio, Section VI studies its performance, factor exposures, trading costs, and timing, and Section VII concludes.

II. The Exposure Penalty

This section states the model and its two normalizations, shows that a constant added to factor variances is a penalty on factor exposures written in the units of the loadings, and defines the portfolio that the penalty barely reaches. Proposition 1 gives the map between normalizations of one model and the condition under which a penalized portfolio does not depend on the normalization; Lemma 1 characterizes the intercept portfolio. Internet Appendix IA.A contains the proofs.

A. The Model and Two Normalizations

At formation date t , N_t stocks have L characteristics in $Z_t \in \mathbb{R}^{N_t \times L}$, with row $z'_{i,t}$. Within the training window $\mathcal{W}_t = \{t - 120, \dots, t - 1\}$, next-month excess returns satisfy

$$r_{i,s+1} = z'_{i,s} \Gamma_\beta f_{s+1} + z'_{i,s} \Gamma_\delta g_{s+1} + z'_{i,s} \Gamma_\alpha + \varepsilon_{i,s+1}, \quad s \in \mathcal{W}_t, \quad (1)$$

where $f_{s+1} \in \mathbb{R}^K$ contains latent factors estimated from returns, $g_{s+1} \in \mathbb{R}^{K_o}$ contains observed factors, and $\varepsilon_{i,s+1}$ is the stock-specific disturbance. The maps $\Gamma_\beta \in \mathbb{R}^{L \times K}$ and $\Gamma_\delta \in \mathbb{R}^{L \times K_o}$ turn characteristics into factor loadings, and $\Gamma_\alpha \in \mathbb{R}^L$ turns them into a stock-specific intercept $\alpha_{i,s} = z'_{i,s} \Gamma_\alpha$, the part of the fitted expected return that the factor block does not absorb. The maps are fixed within a window and re-estimated across windows, so exposures move with characteristics. Instrumented principal components analysis (IPCA) sets $\Gamma_\alpha = 0$ (Kelly, Pruitt, and Su, 2019); we call the specification that retains Γ_α , which Kelly, Pruitt, and Su (2019) also permit, the intercept-inclusive specification of IPCA, or IPCA with an intercept.³

³IPCA under the Gram normalization uses the metric of Liu (2026a), and the intercept-inclusive specification is the No-Arbitrage IPCA of Liu (2026b), which also bounds the average squared stock-level intercept by an exposure budget. The budget is slack at every formation date in this paper (largest value across the 618 rolling fits 1.5×10^{-4} , against a budget of 0.035), so no result uses or requires it; Internet Appendix IA.A states it. The intercept-inclusive fit is estimated under the Gram normalization and re-expressed in Euclidean units by the map of Proposition 1.

The complete index $h_{s+1} = \Gamma_\beta f_{s+1} + \Gamma_\delta g_{s+1} + \Gamma_\alpha$ generates the fitted returns $Z_s h_{s+1}$, and its decomposition is not unique: rescaling a latent factor while inversely rescaling its loadings, or moving a component from one block to another, leaves h_{s+1} unchanged. A normalization selects one decomposition. For a symmetric positive-definite metric $Q \in \mathbb{R}^{L \times L}$ we impose

$$\Gamma'_\beta Q \Gamma_\beta = I_K, \quad \Gamma'_\beta Q [\Gamma_\delta, \Gamma_\alpha] = 0. \quad (2)$$

The first restriction fixes the units of the latent loadings up to a rotation; the second separates them from the observed-factor and intercept blocks. The Euclidean normalization sets $Q = I_L$: the latent loadings satisfy the restriction $\Gamma'_\beta \Gamma_\beta = I_K$ of Kelly, Pruitt, and Su (2019), the other two blocks are orthogonal to them in the same metric, and the coefficient vectors that generate the latent loadings are orthonormal. The Gram normalization sets Q to an average of the window's characteristic Gram matrices

$$W_s = \frac{Z'_s Z_s}{N_s}, \quad u' W_s v = \frac{1}{N_s} \sum_{i=1}^{N_s} (z'_{i,s} u)(z'_{i,s} v), \quad (3)$$

whose quadratic form is the average product of the stock exposures that two coefficient vectors u and v generate. Under it the latent loadings have unit average squared stock exposure and zero average exposure products with the other two blocks over the window.⁴ Two coefficient vectors that are orthogonal under I_L can generate overlapping exposures, so the two normalizations scale and separate the latent loadings differently. Within each window the fit minimizes the stock-level sum of squared errors subject to (2); a change of normalization changes the reported blocks but neither this criterion nor the fitted returns (Proposition 1). Internet Appendix IA.B describes the algorithm.

⁴The estimate averages $W_s + \rho_W I_L$ with $\rho_W = 10^{-8}$; Section III.A gives the weights and Internet Appendix IA.L the estimate over time.

B. The Penalty in Loading Units

Write $F_{s+1} = (f'_{s+1}, g'_{s+1})'$ for the $K + K_o$ factors and $B_t = [Z_t \widehat{\Gamma}_{\beta,t}, Z_t \widehat{\Gamma}_{\delta,t}] \in \mathbb{R}^{N_t \times (K+K_o)}$ for the stocks' fitted factor exposures at formation date t . Let $\widehat{m}_{F,t}$ and $\widehat{C}_{F,t}$ be the mean and covariance of the window's 120 factor realizations, $\widehat{D}_t$ a diagonal matrix of residual variances, and $\lambda_\Sigma > 0$ a constant on stock positions, the position penalty. The mean–variance portfolio of a model without an intercept is

$$\Sigma_{F,t} = \widehat{C}_{F,t} + \lambda I_{K+K_o}, \quad \widehat{\mu}_t = B_t \widehat{m}_{F,t}, \quad (4)$$

$$A_t = B_t \Sigma_{F,t} B_t' + \widehat{D}_t + \lambda_\Sigma I_{N_t}, \quad w_t = A_t^{-1} \widehat{\mu}_t. \quad (5)$$

Adding a constant to the factor variances before inverting them is a ridge (Hoerl and Kennard, 1970) on the factor covariance; we call the constant λ the factor penalty. With $A_{0,t} = B_t \widehat{C}_{F,t} B_t' + \widehat{D}_t + \lambda_\Sigma I_{N_t}$, the positive-definite risk matrix without it, w_t is the unique solution of

$$\max_{w \in \mathbb{R}^{N_t}} \left\{ w' \widehat{\mu}_t - \frac{1}{2} w' A_{0,t} w - \frac{\lambda}{2} \|B_t' w\|^2 \right\}. \quad (6)$$

The last term is a penalty on the portfolio's aggregate factor exposures $B_t' w$, written in the units in which the loadings are reported. Doubling a factor's return and halving its loadings leaves every fitted return unchanged and quarters this term for any given position, so the penalty depends on a choice that the estimator leaves free. The next proposition makes the dependence exact.⁵

PROPOSITION 1 (One model, one penalty): *Suppose Γ_β has rank K , $Q > 0$, and B_t has full column rank. Set $C = \Gamma_\beta' Q \Gamma_\beta$, $H = C^{-1} \Gamma_\beta' Q \Gamma_\delta$, and $a = C^{-1} \Gamma_\beta' Q \Gamma_\alpha$, and let $C^{1/2}$ be the symmetric*

⁵Standardizing each factor to unit variance before adding the constant fixes the scale of each factor but not the mixing of latent and observed factors that the map J of Proposition 1 contains, so the standardized penalty still depends on the normalization unless the map between the standardized factors is orthogonal, as it is, for example, when the factors are uncorrelated in both representations (Internet Appendix IA.A.2). The factor covariances here keep the latent–observed cross-covariances (Section III.B).

positive-definite square root of C ; time subscripts and hats are suppressed.

(a) *The blocks*

$$\Gamma_\beta^Q = \Gamma_\beta C^{-1/2}, \quad \Gamma_\delta^Q = \Gamma_\delta - \Gamma_\beta H, \quad \Gamma_\alpha^Q = \Gamma_\alpha - \Gamma_\beta a, \quad f_{s+1}^Q = C^{1/2}(f_{s+1} + Hg_{s+1} + a) \quad (7)$$

satisfy (2) and preserve h_{s+1} at every date, and hence every fitted return and residual. The exposures and the factor moments transform as $B^Q = BJ^{-1}$, $m_F^Q = Jm_F + k$, and $C_F^Q = JC_F J'$, with $J = \begin{pmatrix} C^{1/2} & C^{1/2}H \\ 0 & I_{K_o} \end{pmatrix}$ and $k = \begin{pmatrix} C^{1/2}a \\ 0 \end{pmatrix}$, so that

$$B^Q C_F^Q (B^Q)' = BC_F B', \quad B^Q m_F^Q + Z \Gamma_\alpha^Q = Bm_F + Z \Gamma_\alpha. \quad (8)$$

(b) Let $R \geq 0$ and $R^Q \geq 0$ be factor adjustments added to C_F and C_F^Q . The two stock risk matrices $B(C_F + R)B' + \widehat{D} + \lambda_\Sigma I$ and $B^Q(C_F^Q + R^Q)(B^Q)' + \widehat{D} + \lambda_\Sigma I$ coincide, and with them the portfolios $A^{-1}\mu$ for every expected-return vector μ , if and only if $R^Q = JRJ'$, that is, if and only if the penalty is written in one set of units. For $R = R^Q = \lambda I$ with $\lambda > 0$ this holds if and only if J is orthogonal.

Adding the same constant in each set of units therefore gives

$$B^Q(C_F^Q + \lambda I)(B^Q)' = BC_F B' + \lambda B^Q(B^Q)', \quad (9)$$

whereas the original representation adds $\lambda BB'$. The rotation that (2) leaves free is orthogonal and does not change the penalty, but the map from the Euclidean to the Gram normalization is orthogonal only if $C = I_K$ and $H = 0$, that is, only if the two metrics already agree on the latent loadings. Two consequences follow. Without the factor penalty ($R = R^Q = 0$) the zero-intercept portfolio is the same under every normalization of the fit. At a positive penalty the portfolio depends on the normalization unless the penalty is written in one set of units: chosen in one normalization and

carried to any other as JRJ' , or specified directly on stock positions, as the position penalty λ_Σ already is. Either choice gives the same portfolio under every normalization and at every value of the penalty; neither says what the penalty should be, because the fit is the same under every normalization and cannot select it. Extending the identity to estimators that re-fit the model under each penalty is left for future work; on the rolling fits it holds to numerical precision wherever it is checked (Section IV.B).

A two-factor example. Suppose idiosyncratic risk is diversified away, so that the investor chooses the portfolio's factor exposures $x = B'w$ directly, and let two uncorrelated factors have nonzero means m_1, m_2 and variances σ_1^2, σ_2^2 . The penalized problem $\max_x \{x'm - \frac{1}{2}x'CFx - \frac{\lambda}{2}\|x\|^2\}$ gives $x_k = m_k/(\sigma_k^2 + \lambda)$ against the tangency exposures m_k/σ_k^2 , so for $\lambda > 0$ the penalty leaves the Sharpe ratio at its tangency value if and only if $\sigma_1^2 = \sigma_2^2$, and otherwise tilts the portfolio toward the factor whose variance is larger in the units in which λ is written. Double factor 1 and halve its loadings: every fitted return is unchanged, the variance of factor 1 becomes $4\sigma_1^2$, and the same λ becomes, in the original units, a penalty of $\lambda/4$ on factor 1 and λ on factor 2. Two normalizations of one model, each penalized by the same λ in its own units, therefore hold different portfolios and earn different Sharpe ratios although they fit the data identically, and both regain the tangency value as $\lambda \rightarrow 0$.

A factor penalty adds a matrix BRB' to the risk matrix. Written in each normalization's own units, $R = \lambda I$; written in one set of units, R is carried across a change of normalization $B \rightarrow BJ^{-1}$ as JRJ' . The two ways of writing it differ only in what that change does to the portfolio:

$$\underbrace{(A_0 + \lambda BB')^{-1}\mu}_{\text{changes unless } J \text{ is orthogonal}} \quad \text{and} \quad \underbrace{(A_0 + BRB')^{-1}\mu, \quad R \rightarrow JRJ'}_{\text{the same under every normalization}} \quad (10)$$

the position penalty $\lambda_\Sigma I$ is the case written directly on stock positions. Internet Appendix Table IA.A.1 sets the two side by side.

C. The Intercept Portfolio

IPCA with an intercept supplies a second mean input, the complete fitted mean $\mu_t^{\text{IP}} = B_t \widehat{m}_{F,t} + Z_t \widehat{\Gamma}_{\alpha,t}$: the factor-mean forecast plus the stock-specific fitted intercepts. We call $w_t^{\text{IP}} = A_t^{-1} \mu_t^{\text{IP}}$ the intercept portfolio (IP), and $w_t^{\text{FM}} = A_t^{-1} B_t \widehat{m}_{F,t}$, which uses the factor means of the same fit alone, the factor-mean portfolio. The two share the risk matrix and differ by the positions $A_t^{-1} Z_t \widehat{\Gamma}_{\alpha,t}$. The intercept is orthogonal to the latent loadings in characteristic space by (2), but that does not make these positions factor-neutral, which would require $B_t' A_t^{-1} Z_t \widehat{\Gamma}_{\alpha,t} = 0$. The tangency portfolio (5) of any model without an intercept is a zero-intercept portfolio (ZP). The IPCA zero-intercept portfolio without the factor penalty, which Proposition 1 makes the same under both normalizations of one model and which the two separately estimated fits deliver to numerical precision (Section IV.B), earns the highest Sharpe ratio among the zero-intercept portfolios of Section III.C; we call it the best zero-intercept portfolio.

LEMMA 1 (The intercept portfolio and the factor penalty): *Hold $\widehat{D}_t$ and λ_Σ fixed, suppress t , and let $M = B' A_0^{-1} B$.*

(a) *Without the factor penalty, the intercept portfolio is the same under every normalization of the fit. Without the factor penalty, the factor-mean portfolio changes with the normalization unless $a = 0$ in Proposition 1.*

(b) *For any expected-return vector μ and $\lambda \geq 0$, let $w_\lambda = (A_0 + \lambda B B')^{-1} \mu$. Then*

$$w_\lambda - w_0 = -\lambda A_0^{-1} B (I + \lambda M)^{-1} B' w_0, \quad B' w_\lambda = (I + \lambda M)^{-1} B' w_0, \quad (11)$$

and

$$(w_\lambda - w_0)' A_0 (w_\lambda - w_0) \leq \frac{\lambda}{4} \|B' w_0\|^2. \quad (12)$$

Part (a) follows from (8), which leaves the complete fitted mean and the unpenalized risk matrix unchanged, and (11) from the Woodbury identity.⁶ In words, the factor penalty acts on a portfolio only through its factor exposures: it shrinks them by $(I + \lambda M)^{-1}$, and it moves the portfolio, measured in the norm of the unpenalized risk matrix A_0 , by at most $\frac{1}{2}\sqrt{\lambda}$ times the factor exposure that the portfolio takes without the penalty. Two normalizations, each penalized by λI in its own units, start from the same intercept portfolio at $\lambda = 0$, so at a positive penalty their intercept portfolios differ by at most the sum of the two bounds. The factor-mean portfolio earns its whole forecast through factor exposure, since $B'w_0^{\text{FM}} = M\widehat{m}_F$: the penalty shrinks the positions that carry its mean, and the units decide by how much. The intercept portfolio adds positions that trade on the part of the fitted mean that the factor block does not absorb; to the extent that these positions carry little factor exposure in the units of the penalty, the penalty barely moves the portfolio, and without the penalty the normalization does not move it at all. The lemma bounds how far the penalty can move a portfolio; Figure 4 of Section V measures how far it moves the intercept portfolio.

III. Data, Portfolio Rule, and Benchmarks

Every comparison in the paper holds the stocks, the information dates, the fitted models, and the leverage policy fixed and changes one allocation input. Section IV changes the units and the size of the factor penalty within one model and then across model families; Section V changes the mean input within one fit, from the factor means alone to the complete fitted mean; and Section VI decomposes the resulting gain, measures its factor exposures, and charges it trading costs and a one-month delay. No model is re-estimated when an input changes.

⁶With $u = B'w_0$, the left side of (12) equals $\lambda^2 u'(I + \lambda M)^{-1} M (I + \lambda M)^{-1} u$, and the matrix in this form has eigenvalues $\lambda^2 \nu / (1 + \lambda \nu)^2 \leq \lambda/4$, where $\nu \geq 0$ runs over the eigenvalues of M .

A. Panel and Fits

We use monthly U.S. stock returns and $L = 132$ firm characteristics from Jensen, Kelly, and Pedersen (2023). The universe combines mega-, large-, and small-capitalization stocks and excludes microcaps, and stock-months with more than 30% of their characteristics missing are excluded. Each month the available values of each characteristic are ranked to $[-0.5, 0.5]$ and missing values are set to zero; within each 120-month window the characteristics are centered and scaled with training-sample moments, and the same transformation is applied at formation.⁷ Over the 594 formation months of the main text, June 1975 through November 2024, the panel has 1,210,039 stock-months, with a median monthly cross section of 2,001 stocks (between 1,431 and 3,150); Internet Appendix IA.K defines and summarizes the variables.

Dates denote portfolio formation, and payoffs occur in the following month. At formation date t each model is estimated on the 120 months $\mathcal{W}_t = \{t - 120, \dots, t - 1\}$, using the characteristics Z_s , the returns r_{s+1} , and the observed factors g_{s+1} for $s \in \mathcal{W}_t$, and the current characteristics Z_t give the exposures for the forecast of month $t + 1$. Three rolling fits of IPCA enter the paper: IPCA under the Euclidean normalization, IPCA under the Gram normalization, and IPCA with an intercept under the Gram normalization. Each has $K = 3$ latent factors and $K_o = 5$ observed factors, the five factors of Fama and French (2015).⁸ The Gram metric is an exponentially weighted average of the window's Gram matrices with a half-life of 60 months. Under two alternative averages, with equal weights on all 120 months or on the last 60, the zero-intercept Gram fit earns 1.03 and 1.13 at the

⁷Stock-months enter when the following month's return is available in the source panel, so membership in the formation-month universe requires a recorded next-month return, a mild look-ahead shared by every portfolio; excess returns as low as -100.5% (Internet Appendix IA.K) are consistent with the inclusion of recorded delisting returns. The point-in-time availability of every characteristic in the processed panel is not independently verified here. The convention can affect levels and, through different weights on the affected stocks, comparisons; Internet Appendix IA.B states the conventions.

⁸The specification holds $K = 3$ latent and 5 observed factors and a 120-month window fixed; the fits are held fixed so that every comparison in the paper changes one allocation input at a time.

reference penalty, against 1.07 under the exponentially weighted estimate; Internet Appendix IA.E reports them.

The main text uses the 594 formation months from June 1975 through November 2024. The sample starts in June 1975, the first month in which the stabilized intercept portfolio of Internet Appendix IA.G, which uses the 24 preceding fits, can be formed, so that every portfolio in the paper is compared on the same months. Internet Appendix IA.E reports the results on all 618 formation months from June 1973: the identity of Proposition 1 holds as on the main sample, and the intercept portfolio earns 2.47 (Internet Appendix IA.M.1), against 1.57 for the unpenalized IPCA portfolio and 1.62 for RP-PCA at $\lambda = 0.001$ (Internet Appendix Table IA.E.3).

B. A Common Portfolio Rule

Every model enters the rule (4)–(5) with its own exposures, factor history, and residual variances. The factor mean and covariance are the arithmetic mean and the sample covariance, with denominator 119, of the window’s 120 factor realizations; the covariance keeps the latent–observed cross-covariances, which the orthogonality of the loadings in characteristic space does not remove. The residual variances subtract the factor contributions but not the intercept (Internet Appendix IA.B), and the comparisons within the intercept fit hold them fixed. The factor penalty takes the six values $\lambda \in \{0, 0.001, 0.003, 0.01, 0.03, 0.1\}$ in each fit’s own units, and $\lambda = 0.01$ is the reference penalty. The position penalty is $\lambda_\Sigma = 0.01$ for every model.⁹ No value is tuned to a model or a date.

Raw mean–variance weights carry no risk-aversion coefficient, so their scale is not a position size. For raw weights w_t , define

$$G_t = \sum_{i=1}^{N_t} |w_{i,t}|, \quad w_t^u = \frac{w_t}{G_t}, \quad r_{t+1}^u = (w_t^u)' r_{t+1}. \quad (13)$$

⁹The position penalty is fixed at 0.01 and not varied. It is written in stock units and is therefore invariant to the normalization.

Every portfolio in Table I is held at unit gross exposure.¹⁰ Because G_t varies over time, unit gross exposure is a leverage policy rather than a reporting convention, and it does not equalize volatility across portfolios. Annualized mean, volatility, and Sharpe ratio are $12\bar{r}^u$, $\sqrt{12}s(r^u)$, and $\sqrt{12}\bar{r}^u/s(r^u)$, and maximum drawdown uses compounded returns. Returns are gross of trading costs unless an exhibit says otherwise; Section VI.D charges proportional costs on holdings-based turnover.

C. Benchmarks

Table I lists the benchmarks from the nearest relatives of the intercept portfolio to textbook portfolios. All are evaluated on the same 594 months, and all but the last two are formed by the rule above:

- the factor-mean portfolio of the same fit, which uses the factor means of IPCA with an intercept and omits the fitted intercept;
- the zero-intercept IPCA portfolios under the Euclidean and the Gram normalizations at each factor penalty, and with the penalty written in one set of units; without the penalty the two are one portfolio, the best zero-intercept portfolio;
- RP-PCA (Lettau and Pelger, 2020), eight latent factors extracted from characteristic-managed returns with weight $\gamma_{RP} = 3$ on their means;
- FF5, the five factors of Fama and French (2015) with exposures instrumented by the characteristics;

¹⁰All portfolios are normalized to unit gross exposure each month, so Sharpe ratios are scale-free and means and volatilities are per dollar of gross position. Alphas are annualized from monthly unit-gross returns. Their t -statistics use Newey–West standard errors with six lags. Turnover is the sum of absolute changes in unit-gross weights after returns have drifted the previous month’s positions. It counts purchases and sales, and one-way turnover is half of it. Net returns subtract a proportional cost per dollar traded. The factor penalty is written in each fit’s own units unless an exhibit says “unit-consistent.” All portfolios trade U.S. stocks only.

Table I. Sharpe ratios across the benchmark ladder and the factor penalty. 594 formation months, June 1975–November 2024; annualized Sharpe ratios at unit gross exposure. λ is added to the factor covariance in each fit’s own units; $\lambda = 0.01$ is the reference penalty. Unit-consistent: the Gram fit’s penalty written in Euclidean units ($\lambda = 0.01$ only; by Proposition 1 equal to the Euclidean row at every λ). Net: 25 basis points per dollar traded. “—”: no series (no $\lambda = 0$ holdings for RP-PCA and FF5; no net or delayed market series; delayed series as in Table III). Panel C: IP at $\lambda = 0.01$ minus each portfolio; delayed: weights formed a month earlier; brackets: pointwise 95% intervals, paired 12-month moving-block bootstrap, fitted models held fixed.

Panel A: Gross Sharpe ratio						
Factor penalty λ	0	0.001	0.003	0.01	0.03	0.1
Factor-mean portfolio (same fit)	0.56	0.45	0.42	0.39	0.38	0.37
IPCA, Euclidean units	1.62	1.31	1.04	0.69	0.45	0.30
IPCA, Gram units	1.62	1.30	1.17	1.07	1.02	1.00
IPCA, unit-consistent penalty				0.69		
RP-PCA	1.54	1.61	1.57	1.49	1.40	1.34
FF5 (characteristic-instrumented)	0.93	0.83	0.79	0.75	0.74	0.73
Pooled characteristic regression	0.92	—	—	0.92	—	—
Market (no λ)	0.54	—	—	—	—	—
Equal weight, $1/N$ (no λ)	0.50	—	—	—	—	—
Intercept portfolio (IP)	2.50	2.51	2.51	2.51	2.51	2.51
Panel B: Net Sharpe ratio at 25 basis points per dollar traded						
Factor penalty λ	0			0.01		
Factor-mean portfolio (same fit)	0.03			−0.04		
IPCA, Euclidean units	1.03			0.30		
IPCA, Gram units	1.03			0.59		
RP-PCA	—			0.96		
FF5 (characteristic-instrumented)	—			0.36		
Pooled characteristic regression	0.72			0.71		
Equal weight, $1/N$ (no λ)	0.48					
Intercept portfolio (IP)	1.61			1.60		
Panel C: Intercept portfolio minus each portfolio, $\lambda = 0.01$ (Sharpe difference)						
	Gross	Net, 25 bp	Delayed			
Factor-mean portfolio (same fit)	2.12 [1.64, 2.68]	1.63 [1.18, 2.16]	1.32 [0.87, 1.81]			
IPCA, Euclidean units	1.83 [1.43, 2.24]	1.29 [0.97, 1.64]	1.29 [0.88, 1.69]			
IPCA, Gram units	1.45 [1.09, 1.86]	1.01 [0.67, 1.39]	—			
Best ZP ($\lambda = 0$)	0.89 [0.59, 1.26]	0.57 [0.29, 0.89]	0.45 [0.11, 0.80]			
IPCA, unit-consistent penalty	1.83 [1.43, 2.24]	—	—			
RP-PCA	1.02 [0.71, 1.39]	0.64 [0.33, 0.97]	—			
FF5 (characteristic-instrumented)	1.76 [1.30, 2.30]	1.23 [0.79, 1.73]	—			
Pooled characteristic regression	1.60 [1.14, 2.09]	0.89 [0.46, 1.34]	0.82 [0.40, 1.23]			
Market (no λ)	1.98 [1.44, 2.60]	—	—			
Equal weight, $1/N$ (no λ)	2.01 [1.49, 2.60]	1.12 [0.61, 1.67]	—			

- a pooled characteristic regression, which regresses next-month excess returns on the 132 characteristics and a constant over the training window, a pooled version of the forecasts of Lewellen (2015), and passes the forecast through the risk matrix of the intercept fit;
- the market excess return and the equally weighted portfolio of the same stocks (DeMiguel, Garlappi, and Uppal, 2009).

Internet Appendix IA.C gives each estimator.¹¹

D. Inference

Comparisons are paired. Each bootstrap draw resamples the common return months in 12-month moving blocks, with the same indices for both portfolios, and recomputes the statistic of each; a Sharpe comparison is the difference between two Sharpe ratios, not the Sharpe ratio of the return difference. Every interval is a pointwise 95% percentile interval from 10,000 draws with a fixed seed,¹² and Internet Appendix IA.E repeats the comparisons with 24- and 60-month blocks. Before any input is changed, every portfolio's weights and returns are reproduced from the fitted models to better than 10^{-9} (Internet Appendix IA.D).

IV. One Penalty, One Portfolio

This section holds the rolling fits of IPCA fixed and changes one input of the portfolio rule, the factor penalty. Two normalizations of one model forecast the same returns and hold different stocks (Section IV.A); the difference in holdings is the penalty written in two sets of units (Section IV.B); the scale of the penalty moves the Sharpe ratio by more than its units do (Section IV.C); and a

¹¹A centered-PCA benchmark has a forecast of zero in exact arithmetic under this rule, and its numerical gross exposure is below 4×10^{-16} ; it is excluded from the comparisons and described in Internet Appendix IA.C.

¹²The bootstrap resamples months and holds the fitted models fixed. An interval therefore compares portfolios built from given fits; it is not an estimate of the estimation risk in the fits themselves. Intervals are pointwise, not joint across penalties or portfolios.

comparison across model families at one value of the penalty inherits the same dependence (Section IV.D). The 618-month counterparts of the exhibits in this section are in Internet Appendix IA.E.

A. Same Forecasts, Different Holdings

The Euclidean and the Gram fits are estimated separately, so their agreement on forecasts is measured rather than imposed by Proposition 1. Month by month, the norm of the difference between their stock forecasts, relative to the norm of the Euclidean forecast, has a median of 4.8×10^{-6} and a maximum of 1.4×10^{-4} ; with the reference penalty in each fit's own units, one minus the cosine of their weight vectors has a median of 0.17, and the L_1 distance between their unit-gross weights has a median of 0.58, more than half of the gross position (Table II, Panel A).¹³ The cosine is unchanged when either vector is multiplied by a positive constant, so the difference concerns relative positions across stocks, not leverage. Nearly equal forecasts do not determine the allocation; the risk input does.

B. Written in One Set of Units

Adding λ to every factor variance adds $\frac{\lambda}{2} \|B'_t w\|^2$ to the risk that the optimizer sees, a penalty on the portfolio's factor exposures in the units of the loadings (6). The Euclidean and the Gram loadings are reported in different units, so one number, 0.01, is two different penalties on the same stock position. Proposition 1 gives the map J between the two sets of units, and by its part (b) a penalty λI chosen in one normalization and written in the other as $\lambda J J'$ gives the same portfolio in both. We apply the map to the rolling fits. At each formation date it is built from the Euclidean loadings and the Gram metric, and the rotation of the latent block that the normalization leaves free

¹³The latent loading spans and the residual variances of the two fits also nearly coincide. Internet Appendix IA.D reports these diagnostics together with an exact decomposition of the weight gap into a term that changes the mean at a common risk matrix and a term that changes the risk matrix at a common mean; the mean term is negligible.

is fixed by aligning the two fits' loading matrices, without using any forecast, weight, or return.¹⁴ Only the factor-risk input changes; the forecasts, the residual variances, and the position penalty are those of the fitted models.

Figure 2 and Panel A of Table II report the holdings. Written in one set of units, the reference penalty gives one portfolio under both normalizations. When the Gram fit's penalty is written in Euclidean units, its weights coincide with the Euclidean fit's: one minus the cosine of the two weight vectors has a median of 1.0×10^{-11} . When the Euclidean fit's penalty is written in Gram units, its weights coincide with the Gram fit's, with a median of 4.8×10^{-12} . Without any penalty the two fits coincide with a median of 6.2×10^{-12} . The returns follow the weights: the Gram fit with its penalty written in Euclidean units earns the Euclidean fit's Sharpe ratio of 0.69 (Table I). The holdings gap at the reference penalty is therefore entirely the penalty written in two sets of units.

C. The Sharpe Ratio along the Penalty

Figure 1 plots the gross Sharpe ratio of the two normalizations of one model as the factor penalty takes six values, from zero to ten times the reference value, each written in the normalization's own units. At the reference penalty the Euclidean fit earns 0.69 and the Gram fit 1.07 (Table I). Their difference of 0.38, with interval $[0.09, 0.64]$ (Table II, Panel B), is entirely units: written in one set of units, the two fits hold one portfolio and earn one return.¹⁵

The difference at the reference penalty is one point on a curve. The Gram-minus-Euclidean difference is zero without the penalty, indistinguishable from zero at penalties of 0.003 and below, and significant from the reference penalty upward, where it rises from 0.38 to 0.70 at $\lambda = 0.1$ (Figure 1, Panel B). The ranking of two normalizations of one model is therefore a function of the penalty, and a comparison of the two at one value of the penalty measures the penalty.

¹⁴Across the 618 rolling fits, the map reproduces the Gram fit's stock exposures and factor covariances from the Euclidean fit with median relative errors of 4×10^{-6} and 1×10^{-6} (Internet Appendix IA.D).

¹⁵On all 618 formation months the same difference has an interval that includes zero (Internet Appendix IA.E).

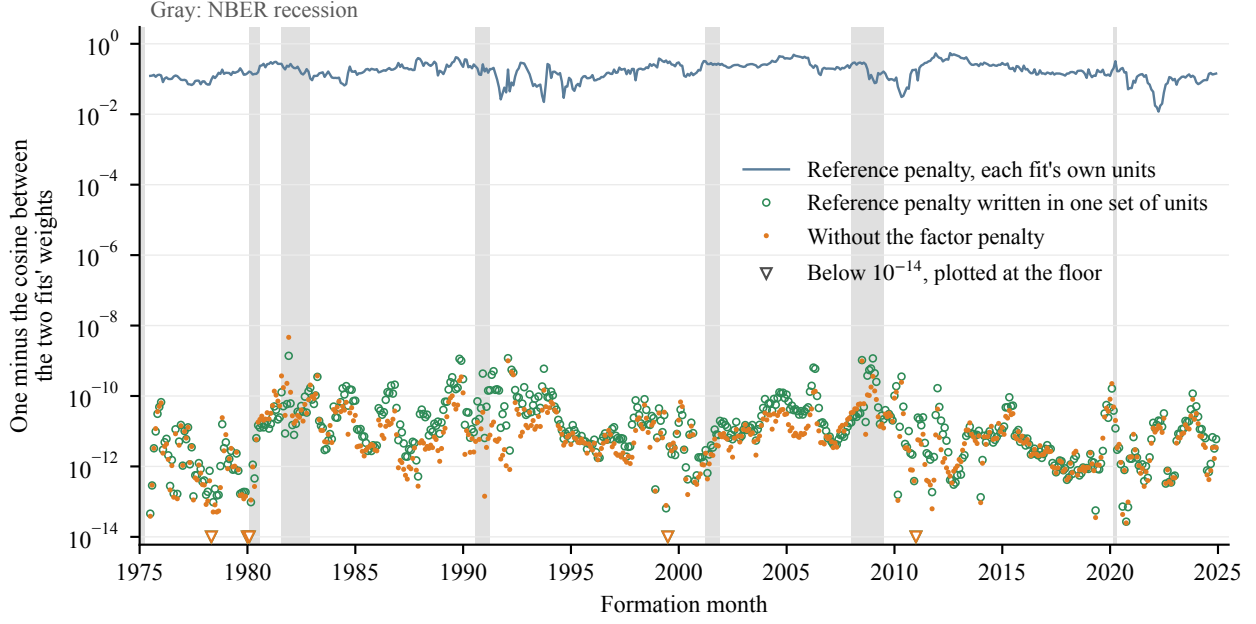


Figure 2. One penalty, one portfolio. One minus the cosine between the stock weights of the Euclidean and the Gram fits of IPCA at each formation month, on a logarithmic scale: with the reference penalty $\lambda = 0.01$ in each fit's own units (line), with that penalty written in one set of units (circles), and without the factor penalty (dots). Values below 10^{-14} are plotted at 10^{-14} (triangles). Gray bands mark NBER recessions (National Bureau of Economic Research, 2023). 594 months, June 1975–November 2024; Panel A of Table II reports the medians.

The scale of the penalty moves the Sharpe ratio by more than its units do. Both curves fall monotonically from 1.62 without the penalty, the Euclidean curve to 0.30 and the Gram curve to 1.00 at $\lambda = 0.1$ (Table I). Removing the reference penalty raises the Euclidean fit's Sharpe ratio by 0.93 and the Gram fit's by 0.55, and their annualized means by 2.6 and 1.9 percentage points, with intervals that exclude zero (Table II, Panel C).¹⁶ Without the penalty the two normalizations are one portfolio by Proposition 1. At 1.62 it is the best zero-intercept portfolio in the paper, and it is the benchmark for the intercept portfolio of Section V.

¹⁶The residual variances are those of the fitted models at every penalty. On the 618 rolling fits, recomputing them without the penalty leaves both removal gains with intervals away from zero (Internet Appendix IA.D). Under a constant-leverage policy in place of unit gross exposure the removal gains are larger, and the size of the Gram-minus-Euclidean difference at the reference penalty depends on the leverage policy as well as on the units (Internet Appendix IA.H).

Table II. The factor penalty within one model and across model families. 594 formation months, June 1975–November 2024; gross returns at unit gross exposure. λ is added to each model’s factor covariance in its own units; “penalty in Euclidean (Gram) units” writes the reference penalty, $\lambda = 0.01$, in the other fit’s units by Proposition 1. Ledoit–Wolf: shrinkage toward a scaled identity in each fit’s units. Panel A: medians and maxima across months; relative gap $\|x_G - x_E\|/\|x_E\|$; L_1 distance between unit-gross weights (maximum two); “—”: not applicable or not computed. Panels B–C: differences in Sharpe ratio and in annualized mean (percentage points); brackets: pointwise 95% intervals, paired 12-month moving-block bootstrap, fitted models held fixed.

Panel A: Distance between the Gram and the Euclidean fits of IPCA						
	One minus cosine		L_1 distance		Relative gap	
	Median	Max.	Median	Max.	Median	Max.
Stock forecasts	—	—	—	—	4.8×10^{-6}	1.4×10^{-4}
<i>Weights</i>						
Reference penalty, own units	0.17	0.55	0.58	1.04	0.61	0.90
Gram fit, penalty in Euclidean units	1.0×10^{-11}	1.4×10^{-9}	4.5×10^{-6}	5.2×10^{-5}	4.9×10^{-6}	5.3×10^{-5}
Euclidean fit, penalty in Gram units	4.8×10^{-12}	2.6×10^{-9}	3.1×10^{-6}	6.9×10^{-5}	3.2×10^{-6}	7.4×10^{-5}
Both unpenalized ($\lambda = 0$)	6.2×10^{-12}	4.6×10^{-9}	3.5×10^{-6}	9.4×10^{-5}	4.0×10^{-6}	9.6×10^{-5}
Ledoit–Wolf, own units	0.045	—	0.30	0.98	—	—

Panel B: Sharpe-ratio differences along the factor penalty					
Factor penalty λ	Gram minus Euclidean fit		RP-PCA minus Euclidean fit		
	Difference	95% interval	Difference	95% interval	
0	0.00	[0.00, 0.00]	−0.08	[−0.44, 0.26]	
0.001	−0.01	[−0.15, 0.14]	0.30	[−0.05, 0.67]	
0.003	0.12	[−0.07, 0.30]	0.53	[0.21, 0.87]	
0.01 (reference)	0.38	[0.09, 0.64]	0.80	[0.44, 1.14]	
0.03	0.58	[0.24, 0.86]	0.95	[0.55, 1.32]	
0.1	0.70	[0.34, 1.00]	1.03	[0.61, 1.43]	

Panel C: Removing the penalty, and a data-driven penalty (b minus a)				
	Sharpe a	Sharpe b	Difference	Mean (pp)
Unpenalized – reference, Euclidean fit	0.69	1.62	0.93 [0.61, 1.25]	2.61 [1.41, 4.01]
Unpenalized – reference, Gram fit	1.07	1.62	0.55 [0.35, 0.77]	1.86 [0.80, 3.09]
Unpenalized – reference, RP-PCA	1.49	1.54	0.05 [−0.20, 0.30]	−0.37 [−1.17, 0.45]
Unpenalized – reference, FF5	0.75	0.93	0.17 [−0.03, 0.39]	1.40 [0.32, 2.55]
Ledoit–Wolf, Gram – Euclidean units	1.17	1.54	0.36 [0.14, 0.64]	0.61 [−0.01, 1.40]
Ledoit–Wolf – reference, Euclidean fit	0.69	1.17	0.49 [0.25, 0.73]	1.71 [0.75, 2.89]
Ledoit–Wolf – reference, Gram fit	1.07	1.54	0.47 [0.30, 0.66]	1.58 [0.76, 2.57]
Ledoit–Wolf – unpenalized, Euclidean fit	1.62	1.17	−0.45 [−0.73, −0.19]	−0.90 [−1.69, −0.25]
Ledoit–Wolf – unpenalized, Gram fit	1.62	1.54	−0.08 [−0.17, 0.02]	−0.28 [−0.63, 0.07]

D. Across Model Families

A comparison across model families at one value of the penalty inherits the same dependence, because each family reports its factors in its own units. Figure 3 re-solves the RP-PCA portfolio of Lettau and Pelger (2020) under the same allocation rule and along the same grid, next to IPCA under

the Euclidean normalization; Table I reports the characteristic-instrumented FF5 portfolio along the same grid. At the reference penalty RP-PCA earns 1.49 against 0.69 for IPCA (Table I), a gap of 0.80. Without the penalty the gap is -0.08 and indistinguishable from zero; it rises monotonically with the penalty to 1.03 at $\lambda = 0.1$, and every interval from $\lambda = 0.003$ upward excludes zero (Table II, Panel B). Removing the reference penalty changes RP-PCA's own Sharpe ratio by 0.05 and FF5's by 0.17, both with intervals that include zero, against 0.93 for IPCA (Panel C). At the reference penalty the four zero-intercept portfolios rank RP-PCA, the Gram fit, FF5, and the Euclidean fit; without it the two IPCA fits earn 1.62, RP-PCA 1.54, and FF5 0.93.

The pattern follows from the exposure penalty and needs no map from one family to the other, which Proposition 1 does not supply because their forecasts differ. A constant added to factor variances is a penalty on factor exposures in the units of each model's loadings, so a comparison of two families at one value of the constant depends on that value. Here the gap is indistinguishable from zero without the penalty and rises with it, while RP-PCA's own Sharpe ratio changes little; the gap at the reference penalty is therefore the penalty acting on the IPCA portfolio rather than an advantage of RP-PCA that survives without it.

A data-driven penalty behaves the same way. The linear shrinkage of Ledoit and Wolf (2004) toward a scaled identity, estimated separately from each normalization's own 120-month factor history, has a median intensity of 0.18 in Euclidean units and 0.14 in Gram units. Because an identity target is itself written in the units of the factors, the two shrunk portfolios differ, with a median L_1 distance of 0.30 between their unit-gross weights (Table II, Panel A), and they earn 1.17 and 1.54, a difference of 0.36 (Panel C).¹⁷ Both lie between the reference penalty and no penalty in their own units. In this application, with eight factors and 120 observations, no shrinkage of the factor covariance is needed, and every identity-target adjustment we examine lowers the Sharpe ratio of the IPCA portfolios; this describes the sample, and the design principle that does not depend

¹⁷On the 618-month sample the Ledoit–Wolf difference is smaller and its 12-month-block interval includes zero; the holdings differ on both samples (Internet Appendix IA.E).

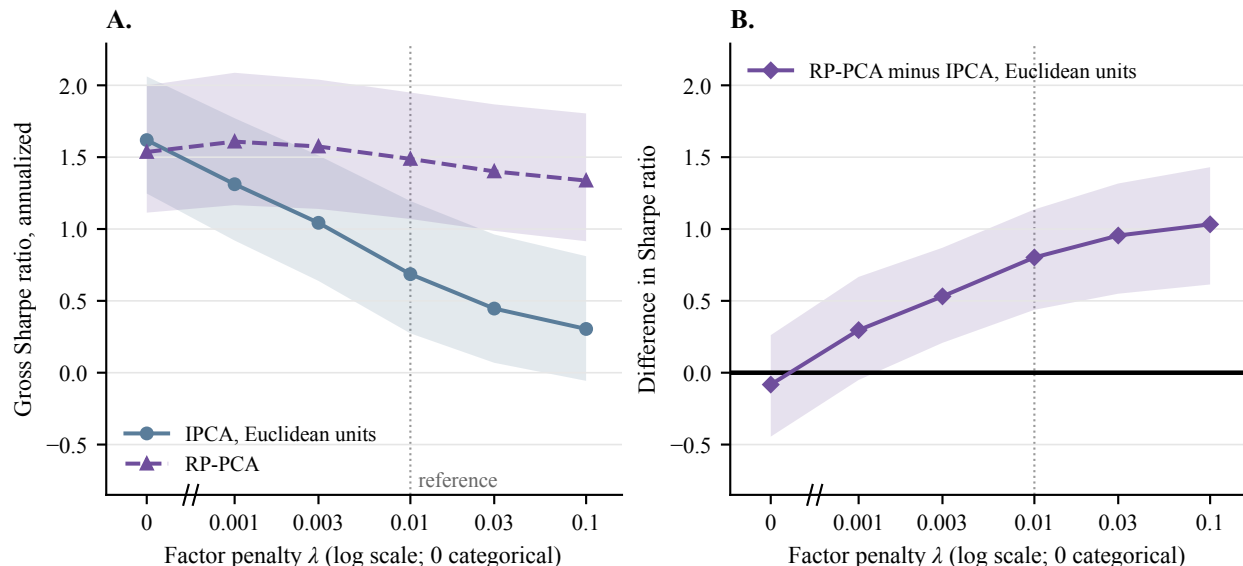


Figure 3. Two model families along the penalty. Panel A: annualized gross Sharpe ratio at unit gross exposure of RP-PCA and of IPCA under the Euclidean normalization, with λ added to each model’s factor covariance in its own units. Panel B: RP-PCA minus IPCA. Bands are pointwise 95% intervals from a 12-month moving-block bootstrap that holds the fitted models fixed; $\lambda = 0$ is a categorical tick on the logarithmic axis. 594 months, June 1975–November 2024.

on it is to write whatever penalty is used in stock units, where it is invariant. The units matter most where they are conventions, as in latent-factor models; Choi and Yuan (2025) document a related dependence on the basis in inference on the intercepts of this model class.

In return space the penalty shows up as negative alpha: regressed on the best zero-intercept portfolio and the five Fama–French factors, both IPCA fits at the reference penalty have negative intercepts, whereas RP-PCA keeps a positive one (Section VI.B and Table V report the regressions). Every zero-intercept portfolio in this section earns its expected return through factor exposure, and factor exposure is what the penalty acts on. Section V turns to a portfolio that earns most of its expected return elsewhere.

V. The Intercept Portfolio

Section IV held the expected-return input fixed and changed the risk input. This section holds one fitted model fixed and changes the expected-return input: whether the characteristic-dependent

intercept that IPCA with an intercept estimates enters the portfolio's expected returns. The two experiments meet in the factor penalty, which acts on factor exposures and therefore acts on one use of the fitted mean and barely reaches the other.

A. Two Uses of One Fitted Mean

IPCA with an intercept supports the two portfolios of Section II.C. Both use the same rolling fit and the same risk matrix A_t , which contains the factor penalty λ and the position penalty. The factor-mean portfolio forms expected returns from the factor block alone, $B_t \widehat{m}_{F,t}$, as every portfolio in Section IV does. The intercept portfolio uses the fitted mean in full,

$$\mu_t^{\text{IP}} = B_t \widehat{m}_{F,t} + Z_t \widehat{\Gamma}_{\alpha,t}, \quad (14)$$

whose second term, the characteristic-dependent intercept, is the part of the fitted mean that the factor block does not absorb.¹⁸ Switching from one portfolio to the other changes a portfolio input, not the fit. The two differ by the positions $A_t^{-1} Z_t \widehat{\Gamma}_{\alpha,t}$, into which the inverse risk matrix turns the intercept, a joint reallocation across stocks that is not factor-neutral in general. Each raw solution is scaled to unit gross exposure, so the returns of the two portfolios are compared directly.

The factor-mean portfolio is a control within this fit. Because the fit assigns part of the fitted mean to the intercept, it is not the same object as a zero-intercept model, and Section V.D shows that it depends on the normalization, whereas the intercept portfolio is nearly invariant to it. The comparison that does not depend on the normalization is with the best zero-intercept portfolio, the unpenalized IPCA portfolio of Section IV.C, which earns 1.62 under either normalization.

¹⁸We use the word intercept for the component of fitted expected return that the factor block does not absorb; whether it is mispricing, an omitted factor, or misspecified dynamics does not affect any portfolio result, which conditions only on the fitted model.

B. The Intercept Portfolio along the Penalty

Figure 4 plots the two portfolios of the same fit along the penalty grid of Section IV, and Table III reports their levels and contrasts. The factor-mean portfolio behaves like the zero-intercept portfolios of Section IV: its Sharpe ratio falls from 0.56 without the penalty to 0.39 at the reference penalty and 0.37 at $\lambda = 0.1$. The intercept portfolio barely moves. It earns 2.50 without the penalty and 2.51 at every positive penalty, with an annualized mean of 7.6% and a volatility of 3.0% at the reference penalty, and its paired gain over the best zero-intercept portfolio has an interval that excludes zero both without the penalty and at the reference penalty (Table III, Panel B). The difference between the two portfolios of the same fit rises from 1.95 to 2.15, and every interval excludes zero (Table III, Panel B). Net of 25 basis points per dollar traded, charged on purchases and sales, the intercept portfolio earns 1.60 at every positive penalty and 1.61 without it, while the factor-mean portfolio earns between -0.05 and 0.03 .

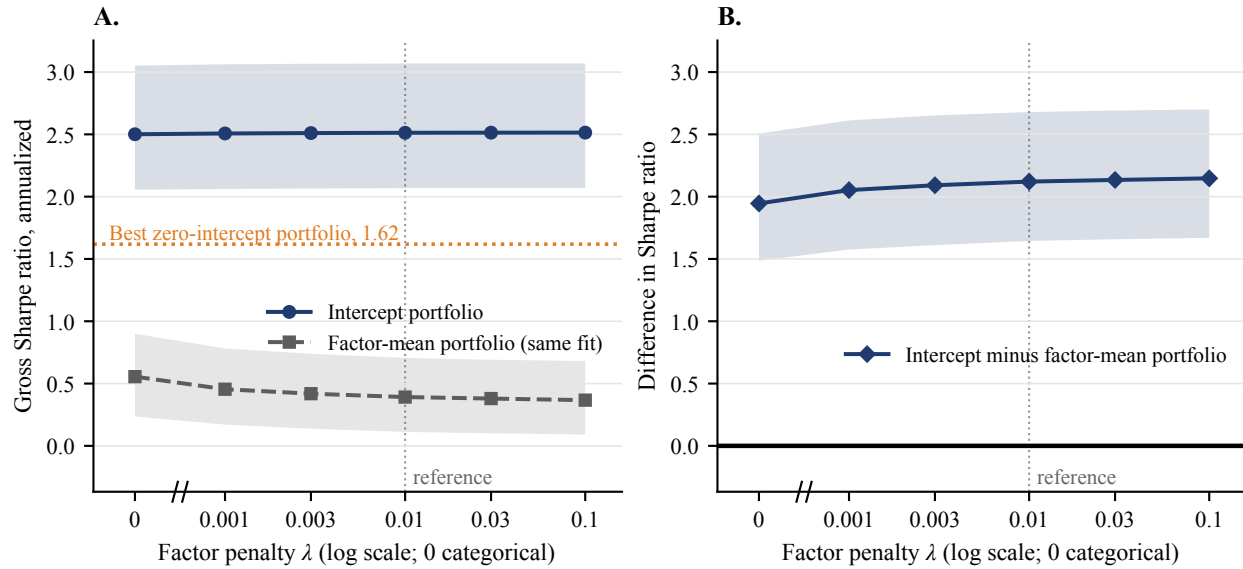


Figure 4. The intercept portfolio and the factor-mean portfolio along the penalty. Panel A: annualized gross Sharpe ratio at unit gross exposure of the intercept portfolio (solid) and the factor-mean portfolio (dashed) of one rolling fit of IPCA with an intercept; the dotted line is the best zero-intercept portfolio. Panel B: their difference. Bands are pointwise 95% intervals from a 12-month moving-block bootstrap that holds the fitted models fixed; $\lambda = 0$ is a categorical tick on the logarithmic axis. 594 months, June 1975–November 2024.

The flat line agrees with Lemma 1, which bounds how far the penalty can move a portfolio by

the factor exposure it takes without the penalty. The penalty acts on a portfolio only through its factor exposures, and the factor-mean portfolio earns its whole forecast through them. Adding the intercept to the same fit raises the annualized mean from 1.7% to 7.6% and lowers the volatility from 4.5% to 3.0% at the reference penalty (Table III, Panel A), and regressed on the best zero-intercept portfolio and the five Fama–French factors, the resulting portfolio loads 0.45 on the former, 0.03 or less in absolute value on each of the latter, and keeps an annualized alpha of 4.9% (Table V). Its return comes mainly from the component that the factor block does not absorb, which a penalty on factor exposures barely reaches; the pattern is consistent with a partial separation of alpha from factor exposure of the kind Kelly, Malamud, and Pedersen (2023) describe.

The demanding comparison is with the best zero-intercept portfolio. The intercept portfolio exceeds it by 0.89 in Sharpe ratio, with interval [0.59, 1.26], at the reference penalty, and by 0.88 when neither portfolio is penalized; net of 25 basis points the gains are 0.57 and 0.58 (Table III, Panel B). That is the smallest of its gains over the portfolios in Table I. Over RP-PCA, the strongest benchmark at the reference penalty, the gain is 1.02; over the pooled characteristic regression, which forecasts returns directly from the 132 characteristics and passes the forecast through the same risk matrix, it is 1.60 gross and 0.89 net. Panel C of Table III delays every trade by one month, and Section VI reports it with the other evidence on timing.

The gain requires the fitted intercept’s assignment to stocks, not only a forecast component of its size and dispersion. Shuffling the fitted intercept across stocks within each formation month, with the risk matrix and the factor-mean forecast held fixed, leaves a gross Sharpe ratio of 0.04 on average and never above 0.30 in 1,000 draws, against 2.51 for the intercept as fitted (Internet Appendix IA.M).

C. The Intercept Test and the Intercept Portfolio

A Sharpe ratio of 2.5 sits beside the finding of Kelly, Pruitt, and Su (2019) that the intercept of this model cannot be distinguished from zero, and the two statements answer different questions.

Table III. The intercept portfolio and the factor-mean portfolio of the same fit. 594 months, June 1975–November 2024; unit gross exposure; means and volatilities annualized, in percent. Both portfolios use one rolling fit of IPCA with an intercept at each factor penalty λ ; the best ZP is IPCA with $\lambda = 0$. “Net, 25 bp” subtracts 25 basis points per dollar traded (purchases plus sales). Panel C holds weights formed one month earlier; delayed series are formed for the five portfolios shown (not for the Gram, RP-PCA, and FF5 portfolios). Intervals: pointwise 95%, 12-month moving-block bootstrap, fitted models held fixed.

Panel A: Levels along the factor penalty				
Portfolio, λ	Mean	Volatility	Sharpe	Net, 25 bp
<i>Intercept portfolio</i>				
$\lambda = 0$	7.8	3.1	2.50	1.61
$\lambda = 0.001$	7.6	3.0	2.51	1.60
$\lambda = 0.003$	7.6	3.0	2.51	1.60
$\lambda = 0.01$ (reference)	7.6	3.0	2.51	1.60
$\lambda = 0.03$	7.6	3.0	2.51	1.60
$\lambda = 0.1$	7.6	3.0	2.51	1.60
<i>Factor-mean portfolio (same fit)</i>				
$\lambda = 0$	2.2	4.0	0.56	0.03
$\lambda = 0.001$	1.9	4.3	0.45	0.00
$\lambda = 0.003$	1.8	4.4	0.42	−0.02
$\lambda = 0.01$ (reference)	1.7	4.5	0.39	−0.04
$\lambda = 0.03$	1.7	4.5	0.38	−0.04
$\lambda = 0.1$	1.7	4.5	0.37	−0.05
Panel B: Paired differences in Sharpe ratio				
	Gross	95% interval	Net, 25 bp	95% interval
<i>IP minus factor-mean portfolio, same λ</i>				
$\lambda = 0$	1.95	[1.49, 2.51]	1.58	[1.14, 2.11]
$\lambda = 0.001$	2.05	[1.58, 2.61]	1.60	[1.15, 2.13]
$\lambda = 0.003$	2.09	[1.61, 2.65]	1.62	[1.16, 2.15]
$\lambda = 0.01$ (reference)	2.12	[1.64, 2.68]	1.63	[1.18, 2.16]
$\lambda = 0.03$	2.13	[1.66, 2.69]	1.64	[1.19, 2.16]
$\lambda = 0.1$	2.15	[1.67, 2.70]	1.65	[1.20, 2.17]
<i>IP minus best ZP</i>				
IP at $\lambda = 0$	0.88	[0.59, 1.23]	0.58	[0.32, 0.88]
IP at $\lambda = 0.01$	0.89	[0.59, 1.26]	0.57	[0.29, 0.89]
<i>IP minus pooled characteristic regression</i>				
IP at $\lambda = 0.01$	1.60	[1.14, 2.09]	0.89	[0.46, 1.34]
Panel C: Trading one month late (Sharpe ratio)				
	Gross	Delayed	IP – row	95% interval
Euclidean fit, $\lambda = 0.01$	0.69	0.22	1.29	[0.88, 1.69]
Best ZP, $\lambda = 0$	1.62	1.07	0.45	[0.11, 0.80]
Pooled regression	0.92	0.70	0.82	[0.40, 1.23]
Factor-mean, $\lambda = 0.01$	0.39	0.20	1.32	[0.87, 1.81]
IP, $\lambda = 0.01$	2.51	1.52	—	—

Their test asks whether the intercept loadings, one per characteristic, are jointly large in characteristic space, and it does not reject that they are zero once several latent factors are included;

Windmüller (2022) and Gu, Kelly, and Xiu (2021) reach similar conclusions with other samples and estimators. The tangency portfolio asks whether the intercept's sum across some two thousand stocks, weighted by the inverse risk matrix, has a mean per unit of risk. The second can be large when the first is not, because an intercept that is small for each characteristic can be systematic across stocks; the estimates here also use 132 characteristics, rolling out-of-sample windows, and observed as well as latent factors. Kelly, Pruitt, and Su (2019) also form pure-alpha portfolios, hedged to zero exposure to their factors by construction, and find them far weaker than their factor portfolios (their Table 7); the intercept portfolio is not hedged, loads 0.45 on the best zero-intercept portfolio, and earns its alpha of 4.9% relative to that portfolio and the five factors (Table V), so what the paper measures is the unhedged mean–variance use of the intercept, not a factor-neutral spread. Their out-of-sample tangency portfolio of five IPCA factors earns a Sharpe ratio of 3.9 in a sample that ends in 2014 (their Table 6), above the 1.62 of the best zero-intercept portfolio here, which is formed at the stock level with idiosyncratic risk in the risk matrix, from 132 characteristics in 120-month rolling windows, and evaluated through 2024; every portfolio in Figure 5 earns its lowest Sharpe ratio in the last third of that sample (Section VI.C), and the comparisons in this paper are made within one rule and one sample. Portfolios built from analogous components earn Sharpe ratios of the same order: Kim, Korajczyk, and Neuhierl (2021) report 1.31 to 1.66 for arbitrage portfolios built from the component of expected returns that characteristic-driven loadings leave unexplained, Chen, Roussanov, and Wang (2023) report arbitrage-portfolio Sharpe ratios above 3 and Chen, Roussanov, and Wang (2025) above 4 out of sample when the factors are combined with orthogonal arbitrage portfolios, and Zhang (2024) and Choi and Yuan (2025) also reject zero pricing errors. The gross 2.5 lies between the figures of Kim, Korajczyk, and Neuhierl (2021) and Chen, Roussanov, and Wang (2023). Langlois (2023) finds significant but declining alphas that do not survive estimation risk and trading costs, a test we do not run; what we report is a net Sharpe

ratio of 1.60 at 25 basis points and a break-even cost of 68 basis points (Table IV).¹⁹

D. Invariance to the Normalization

The intercept is estimated under the Gram normalization, but the intercept portfolio is nearly invariant to it. Without the penalty the invariance is exact: the complete fitted mean and the unpenalized factor risk do not change with the normalization, so at $\lambda = 0$ the intercept portfolio is the same under every normalization of the fit (Lemma 1(a), which follows from Proposition 1); Table IA.F.1 shows that at $\lambda > 0$ the difference is below 0.01 in Sharpe ratio. Re-expressed in Euclidean units by the map of Proposition 1, the same fit delivers the same intercept portfolio without the penalty, with monthly returns that agree to 1.7×10^{-14} , and Sharpe ratios that differ from the Gram-unit values by 0.004 to 0.009 at the positive penalties of the grid (Internet Appendix IA.F).²⁰ Those small differences are the effect of Section IV, a penalty on factor exposures written in two sets of units, acting on a portfolio whose return comes mainly from outside the factor block.

The factor-mean portfolio of the same fit, by contrast, depends on the normalization, because the split of the fitted mean between the factor block and the intercept does: in Euclidean units it earns 0.44 without the penalty and -0.06 at the reference penalty, against 0.56 and 0.39 in Gram units. The Gram normalization therefore flatters the within-fit control rather than the intercept portfolio, and the paper benchmarks the intercept portfolio against the best zero-intercept portfolio, which does not depend on the normalization, rather than against that control.

¹⁹Kozak, Nagel, and Santosh (2018) argue that the absence of near-arbitrage opportunities confines high Sharpe ratios to a few dominant factors; a nearly factor-neutral portfolio with a high gross Sharpe ratio is accordingly a statement about the fitted model and its implementation rather than about equilibrium.

²⁰The Euclidean representation re-expresses the fitted model in other units and holds its residual variances fixed; it is not a separate estimation in Euclidean units with a free intercept, which could select a different fit. The residual variances subtract the factor contributions but not the intercept, so recomputing them in Euclidean units would move them by a median 0.3% and at most 1.2% (Internet Appendix IA.F); the exact invariance at $\lambda = 0$ holds with them fixed.

The intercept portfolio is nearly invariant to the factor penalty, exactly invariant to the normalization of the fit when no penalty is applied, and the portfolio with the highest Sharpe ratio in Table I. The intercept is the alpha the model already estimated, relative to its own factors.

VI. Anatomy of the Intercept Gain

Section V showed that the intercept portfolio earns 2.5 at every value of the factor penalty and exceeds the best zero-intercept portfolio, the unpenalized IPCA portfolio at 1.62, by 0.89. This section takes that gain apart. Figure 5 sets the intercept portfolio against the model-based benchmarks of Table I, gross and net of costs, over the full sample and in each third of it, and Table IV reports what each portfolio trades and what it earns after proportional costs. Section VI.A splits the gain into a mean channel and a volatility channel, Section VI.B measures the factor exposure the intercept portfolio carries, Section VI.C follows the gain through time, and Section VI.D turns to turnover and costs, timing, leverage, and other choices. Unless stated otherwise, the intercept portfolio is the one at the reference penalty, $\lambda = 0.01$, and the comparison portfolio is the best zero-intercept portfolio, the unpenalized IPCA portfolio.

A. The Mean Channel and the Volatility Channel

Per dollar of gross position, the intercept portfolio improves on the best zero-intercept portfolio mainly through the denominator. Figure 6 reports the two portfolios' Sharpe ratios, means, and volatilities. Over the 594 months the intercept portfolio's annualized mean is 7.6%, higher than the best zero-intercept portfolio's 6.6% by an amount within sampling error, and its volatility is 3.0%, about three-quarters of that portfolio's 4.1%. Of the 0.89 gain in Sharpe ratio, 0.66 arrives through the volatility channel and 0.23 through the mean channel (Panel B).²¹

²¹The decomposition writes the Sharpe-ratio difference as $(m_1 - m_0)/v_0 + m_1(1/v_1 - 1/v_0)$, where m and v are the annualized mean and volatility and the subscripts 1 and 0 denote the intercept and the best zero-intercept portfolio: a mean term evaluated at the benchmark's volatility and a volatility term evaluated at the intercept portfolio's mean. The

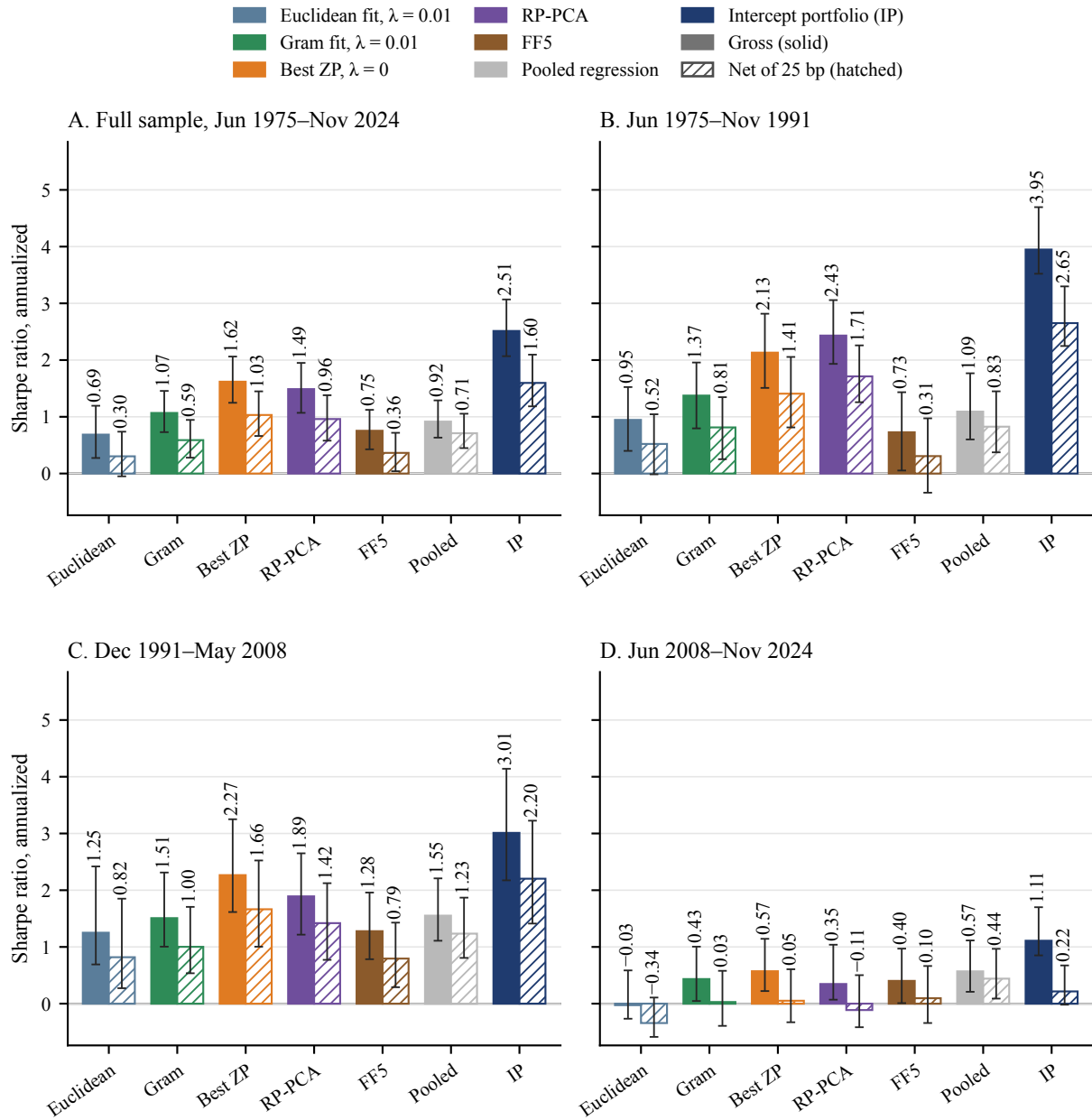


Figure 5. Sharpe ratios by strategy and subperiod, gross and net. Annualized Sharpe ratio at unit gross exposure of seven portfolios in ladder order, penalized portfolios at $\lambda = 0.01$ and the best zero-intercept portfolio (Best ZP) at $\lambda = 0$, over all 594 months (Panel A) and three calendar thirds of 198 months (Panels B–D). Solid bars are gross; hatched bars subtract 25 basis points per dollar of purchases and sales. Whiskers are pointwise 95% percentile intervals from a 12-month moving-block bootstrap (10,000 draws, the same months for every bar in a panel) that holds the fitted models fixed. June 1975–November 2024.

Table IV. Turnover and Sharpe ratios net of proportional costs. 594 months, June 1975–November 2024; unit gross exposure; the factor penalty is $\lambda = 0.01$ in each fit’s own units unless the row states $\lambda = 0$. Turnover is the monthly trade $\sum_i |u_{i,t} - \tilde{u}_{i,t}|$ between the new unit-gross weights and the previous weights after they drift with realized returns; it counts purchases and sales, and one-way turnover is half of it. Net returns subtract a cost per dollar traded from the monthly gross return; the break-even cost sets the mean net return to zero. Panel B: paired differences at 25 basis points; means annualized, in percentage points; pointwise 95% intervals from a 12-month moving-block bootstrap that holds the fitted models fixed.

Panel A: Turnover and Sharpe ratios net of proportional costs						
Portfolio	Turnover	Sharpe ratio at cost (bp)				Break-even cost (bp)
		0	10	25	50	
IPCA, Euclidean units	0.75	0.69	0.53	0.30	−0.08	45
IPCA, Gram units	0.71	1.07	0.88	0.59	0.11	56
Best ZP ($\lambda = 0$)	0.81	1.62	1.38	1.03	0.44	68
RP-PCA	0.69	1.49	1.28	0.96	0.43	70
FF5 (characteristic-instrumented)	0.67	0.75	0.60	0.36	−0.03	48
Factor-mean portfolio (same fit)	0.64	0.39	0.22	−0.04	−0.46	23
Intercept portfolio (IP)	0.94	2.51	2.15	1.60	0.66	68
Panel B: Paired differences net of 25 basis points						
Comparison	Sharpe	95% interval	Mean (pp)	95% interval		
Unpenalized – reference, Euclidean fit	0.73	[0.47, 0.98]	2.4	[1.2, 3.8]		
Gram – Euclidean fit	0.29	[0.04, 0.51]	0.8	[−0.4, 2.2]		
IP – factor-mean portfolio (same fit)	1.63	[1.18, 2.16]	4.9	[3.3, 6.7]		
IP – best ZP	0.57	[0.29, 0.89]	0.6	[−0.5, 1.7]		

The split differs across periods, and the volatility channel is the constant. It is positive in each third of the sample, at 1.20, 0.79, and 0.31, and in each third the intercept portfolio is the less volatile of the two, at 2.4%, 3.3%, and 2.8% against 3.5%, 4.5%, and 3.9%. The mean channel is large in the first third, 0.62, when the intercept portfolio also earns the higher mean, 9.6% against 7.5%. In the middle third it is slightly negative: the intercept portfolio’s mean of 10.0% is below the best zero-intercept portfolio’s 10.2%, and the gain of 0.74 comes through the volatility channel, 0.79, two terms depend on the order of attribution; in the reverse order the split is 0.58 through volatility and 0.32 through the mean. The decomposition is descriptive, not causal. The difference in annualized means is 0.95 percentage points, with a 95% interval of [−0.14, 2.06]. The split also depends on the leverage policy: scaling either portfolio by a constant leaves its Sharpe ratio unchanged but moves the split, and at equal volatility the whole difference is a difference in mean. Under unit gross exposure the split compares the two portfolios per dollar of positions, the scale at which turnover and costs are charged.

against -0.05 through the mean. A lower volatility at a similar mean is what one would expect from a portfolio whose expected return is not earned through factor exposure, since the factor block is where the zero-intercept portfolios take their risk; the exposure regressions of Section VI.B measure how little of that risk the intercept portfolio takes.

B. Factor Exposure

Table V regresses each portfolio's monthly return on the best zero-intercept portfolio and the five factors of Fama and French (2015). The intercept portfolio loads 0.45 on the best zero-intercept portfolio, with a t -statistic of 10.2, and no more than 0.03 in absolute value on any of the five factors. Together with the five factors, IPCA's own best zero-intercept portfolio explains a third of its variance ($R^2 = 0.34$), and the regression leaves an annualized alpha of 4.9% with a t -statistic of 10.9. On the five factors alone, the intercept portfolio's alpha is 7.1% a year with a t -statistic of 12.0, every loading is below 0.06 in absolute value, and the R^2 is 0.02 (Table V, last row).

The penalized portfolios look different. At the reference penalty the Euclidean fit loads 1.03 on the best zero-intercept portfolio and has an alpha of -2.2% a year relative to it ($t = -2.3$), the alpha that the reference penalty leaves that fit relative to the best zero-intercept portfolio; the Gram fit loads 0.74 with an alpha of -0.4% ($t = -0.7$); and the factor-mean portfolio of the same fit as the intercept portfolio loads 0.58 with an alpha of -2.7% ($t = -5.0$). Each behaves like a scaled copy of the best zero-intercept portfolio that earns less than the copy. RP-PCA retains an alpha of 2.6% ($t = 3.7$) on a loading of 0.50, the FF5 portfolio retains none, and the pooled characteristic regression is mostly market and size exposure, with loadings of 0.52 and 0.32 and an R^2 of 0.76. Of the seven portfolios regressed on the best zero-intercept portfolio, the intercept portfolio has the largest alpha and the largest t -statistic.

The intercept portfolio is therefore not a reweighting of the factor exposures the zero-intercept portfolios hold: two-thirds of its variance lies outside the span of the best zero-intercept portfolio and the five factors, and its alpha relative to that span is 4.9% a year. The pattern is consistent

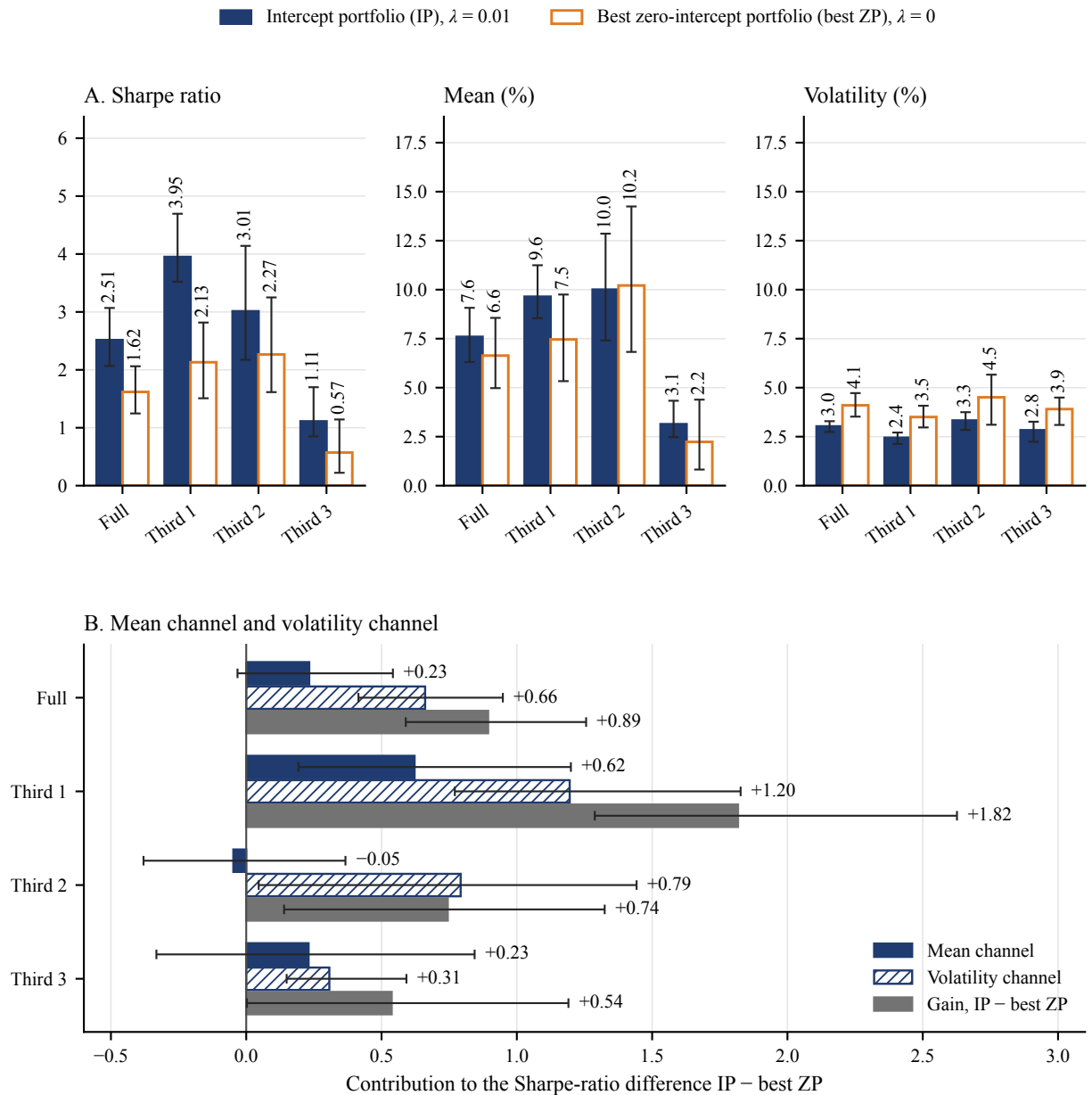


Figure 6. Anatomy of the intercept gain. Panel A: annualized Sharpe ratio, mean, and volatility (percent) of the intercept portfolio (IP, $\lambda = 0.01$; solid) and the best zero-intercept portfolio (best ZP, $\lambda = 0$; hollow), over the full sample and three calendar thirds of 198 months. Panel B: the difference IP – best ZP (gray) and its mean channel $(m_1 - m_0)/v_0$ (solid) and volatility channel $m_1(1/v_1 - 1/v_0)$ (hatched), which sum to it. Means and volatilities are per dollar of gross position; the split in Panel B depends on that normalization. Whiskers are pointwise 95% intervals from a 12-month moving-block bootstrap that holds the fitted models fixed. Gross returns; 594 months, June 1975–November 2024.

Table V. Exposures to the best zero-intercept portfolio and the five Fama–French factors. Each row regresses a portfolio’s monthly gross return at unit gross exposure on the return of the best zero-intercept portfolio (best ZP: IPCA with $\lambda = 0$, the same under the Euclidean and the Gram normalization) and the five factors of Fama and French (2015), dated to the return month; 594 formation months, June 1975–November 2024. All rows except the best ZP row use the reference penalty $\lambda = 0.01$ in each fit’s own units. The best ZP row and the last row use the five factors only; the last row’s loading t -statistics are not reported. Alpha is annualized, in percent; Newey–West t -statistics (six lags, Newey and West, 1987) are in parentheses.

Portfolio	Best ZP	MKT	SMB	HML	RMW	CMA	Alpha	R^2
Euclidean fit	1.03 (13.47)	0.02 (0.61)	−0.04 (−1.65)	−0.02 (−0.55)	−0.09 (−2.88)	−0.07 (−1.41)	−2.21 (−2.30)	0.52
Gram fit	0.74 (12.07)	0.07 (4.24)	0.01 (0.66)	0.06 (2.33)	−0.07 (−2.38)	−0.09 (−2.21)	−0.36 (−0.70)	0.64
Best ZP ($\lambda = 0$)	—	0.08 (5.19)	0.09 (4.77)	−0.03 (−1.13)	0.15 (4.17)	0.10 (2.07)	4.95 (6.59)	0.16
RP-PCA	0.50 (9.36)	0.01 (0.60)	0.00 (−0.21)	−0.06 (−1.32)	−0.02 (−0.55)	0.03 (0.51)	2.55 (3.67)	0.30
FF5	0.57 (7.95)	0.07 (2.84)	0.02 (0.61)	0.11 (3.84)	−0.09 (−2.12)	−0.03 (−0.62)	−0.39 (−0.46)	0.36
Pooled regression	0.27 (2.99)	0.52 (11.22)	0.32 (8.81)	0.09 (1.16)	0.12 (2.71)	−0.07 (−0.94)	2.72 (2.48)	0.76
Factor-mean (same fit)	0.58 (11.89)	0.08 (4.09)	0.02 (1.21)	0.08 (2.56)	−0.04 (−1.15)	−0.07 (−1.52)	−2.73 (−5.04)	0.50
Intercept portfolio (IP)	0.45 (10.21)	−0.03 (−2.83)	−0.01 (−0.52)	0.02 (1.21)	−0.02 (−1.02)	−0.02 (−0.93)	4.92 (10.90)	0.34
IP, five factors only	—	0.01	0.03	0.00	0.05	0.02	7.15 (12.01)	0.02

with a partial separation of alpha from factor exposure of the kind studied by Kelly, Malamud, and Pedersen (2023) and Daniel et al. (2020). Here it does not come from an explicit hedge, and it does not depend on how the fit splits its mean between the factor block and the intercept, since the intercept portfolio is the same under either split when no penalty is applied (Section V.D). The alpha depends on which stock receives which fitted intercept, as the permutation check of Section V.B shows.

C. Over Time

Gross of costs, the intercept portfolio has the highest Sharpe ratio of the seven portfolios in every third of the sample (Figure 5), in both NBER business-cycle regimes (Internet Appendix

Figure IA.I.1), and in every calendar decade except the 59 months from January 2020, in which the FF5 portfolio earns 0.87 against its 0.83 (Internet Appendix Table IA.I.1). Its Sharpe ratio falls from 3.95 in the first third to 1.11 in the last, a 72% decline that matches the 73% decline of the best zero-intercept portfolio, from 2.13 to 0.57, and is smaller than the 86% decline of RP-PCA, from 2.43 to 0.35. Its five-year rolling Sharpe ratio never falls below 0.45, whereas those of the best zero-intercept portfolio and RP-PCA each turn negative at least once (Internet Appendix Figure IA.I.2). Every portfolio in Figure 5 earns its lowest Sharpe ratio in the last third, in line with the attenuation of characteristic premia after their publication,²² and the intercept portfolio's gain over the best zero-intercept portfolio is positive in each third, at 1.82, 0.74, and 0.54 (Figure 6, Panel B; in the last third with an interval whose lower end is 0.003, Internet Appendix IA.I).

Net of 25 basis points per dollar traded, the intercept portfolio is also the highest in the first two thirds, at 2.65 and 2.20. The last third is different. There its net Sharpe ratio is 0.22, with an interval of $[-0.02, 0.67]$ that includes zero and a net mean of 0.6% a year, below the 0.44 of the pooled characteristic regression, whose return is largely market and size exposure (Table V). Over the 206 months from October 2007 through November 2024, the window of Internet Appendix Table IA.I.2, which gives the same signs as the last third, the intercept portfolio's net advantage over the factor-mean portfolio is 0.44 $[-0.22, 1.10]$, over the best zero-intercept portfolio 0.19 $[-0.41, 0.75]$, and over the pooled characteristic regression $-0.19 [-0.71, 0.32]$ (Internet Appendix IA.I), so the net ranking after 2007 is not resolved by this sample.

D. Turnover, Costs, Timing, and Leverage

Turnover and costs. Turnover is high for every portfolio in Table IV, 64–94% of gross exposure per month on purchases and sales, and highest for the intercept portfolio at 94%, which rebalances each month on a signal that moves with the fitted intercept. Costs are charged on purchases and sales at 10, 25, and 50 basis points per dollar traded, with no netting across names and no cost-aware

²²See McLean and Pontiff (2016).

rebalancing.²³ Net of 25 basis points the intercept portfolio earns 1.60, above the gross Sharpe ratio of every benchmark in Table I at the reference penalty and just below the 1.62 that the best zero-intercept portfolio earns before costs, and it earns 1.60 or 1.61 at every penalty on the grid (Table III). Its break-even cost of 68 basis points equals the best zero-intercept portfolio's and is close to RP-PCA's 70, and its net advantage over the best zero-intercept portfolio is 0.57 (Table IV, Panel B). At 50 basis points, the highest cost in the ladder, it earns 0.66, above every zero-intercept portfolio at the same cost. The penalty's effects of Section IV carry over net of costs: removing the penalty raises the Euclidean fit's net Sharpe ratio from 0.30 to 1.03, and at the reference penalty the Gram fit leads the Euclidean fit by 0.29 net (Table IV).

Timing. Every portfolio is formed at the end of its formation month and held over the following month. Delaying every trade by one month, so that the weights formed at $t - 1$ are held through the return of month $t + 1$ and no input is less than a month old, lowers the intercept portfolio's Sharpe ratio from 2.51 to 1.52 and the best zero-intercept portfolio's from 1.62 to 1.07 (Table III, Panel C). The intercept portfolio gives up about two-fifths of its Sharpe ratio to the delay and the best zero-intercept portfolio about one-third. The delayed intercept portfolio still exceeds every benchmark for which a delayed series is formed (the Euclidean fit, the best zero-intercept portfolio, the pooled characteristic regression, and the factor-mean portfolio), each contrast's interval excludes zero, and against the best zero-intercept portfolio the delayed gain is 0.45 (Table III, Panel C). A short-term reversal factor explains little of the intercept portfolio's return: added to the five Fama–French factors, it receives a loading of 0.03 ($t = 2.3$) and leaves the alpha at 7.1% a year. Formed with the intercept loadings of the one-month return and of the 13 characteristics computed from the last 21 days of daily returns set to zero, with the fit unchanged, the intercept portfolio earns 2.16 gross and

²³The cost model is proportional and omits stock-level spreads, price impact, short-sale fees, financing, and capacity. It is the standard device for comparing turnover-heavy portfolios on a common footing (Novy-Marx and Velikov, 2016; DeMiguel et al., 2020), not an implementable net return, which would require stock-level cost models such as those of Detzel, Novy-Marx, and Velikov (2023) and Jensen et al. (2026).

1.44 net of 25 basis points and exceeds the best zero-intercept portfolio by 0.54, with an interval that excludes zero; when the month $t - 12$ return and the two 21-day trading-activity characteristics are also set to zero it earns 1.81, and its gain, 0.19 $[-0.17, 0.53]$, is no longer distinguishable from zero (Internet Appendix IA.M). The intercept thus collects short-horizon premia that the factor block does not absorb, which is also why a one-month delay costs it two-fifths of its Sharpe ratio; a reversal factor does not span that return.

Leverage policy. Every portfolio is held at unit gross exposure (Section III.B). The unnormalized mean–variance solutions carry no risk-aversion coefficient, and their average gross exposure ranges from 3.8 for the Euclidean fit at the reference penalty to 58 for the best zero-intercept portfolio and about 500 for the intercept portfolio; Internet Appendix IA.H reports them, the months in which their raw return falls below -100% , and the penalty-removal and Gram-minus-Euclidean contrasts under two other leverage policies. Under a constant leverage equal to each portfolio’s mean gross exposure, and under a leverage equal to its trailing 12-month mean, the intercept portfolio’s gain over the best zero-intercept portfolio is 0.81 and 0.85, against 0.89 at unit gross exposure, and each interval excludes zero; the gain is therefore a property of the fitted intercept rather than of the leverage policy.

Other choices. Bootstrap blocks of 24 and 60 months leave the intercept portfolio’s gains over the factor-mean portfolio and the best zero-intercept portfolio, and both gains from removing the penalty, with intervals away from zero (Internet Appendix IA.E). On the 618 rolling fits, recomputing the residual variances without the penalty leaves both removal gains away from zero (Internet Appendix IA.D). A stabilized intercept portfolio, which also penalizes positions whose intercept contribution varies across the 24 preceding loading estimates, changes the Sharpe ratio by 0.02 (Internet Appendix IA.G). Applied within capitalization groups, the intercept portfolio exceeds the best zero-intercept portfolio by 0.57 among mega-capitalization stocks, 0.51 among large stocks, and 1.18 among small stocks, each with an interval that excludes zero (Internet Appendix IA.I). Removing each portfolio’s five most extreme months leaves the intercept portfolio’s Sharpe ratio

at 2.51, and its stabilized version's at 2.53 (Internet Appendix IA.M). The results are a property of the penalty and of the fitted intercept, not of these choices.

VII. Conclusion

A constant added to factor variances is a penalty on factor exposures in the units of the loadings, and a latent-factor model leaves those units to a normalization. The intercept a factor model already estimates is the alpha relative to its own factors, and the penalty barely reaches it. The paper establishes three results. First, two normalizations of one IPCA model, given the same number as a penalty, hold different portfolios and earn Sharpe ratios of 0.69 and 1.07; written in one set of units the penalty gives one portfolio, and without the penalty both earn 1.62. Second, the same penalty decides a comparison across model families: under one allocation rule, RP-PCA's advantage over Euclidean-normalized IPCA is 0.80 at the reference penalty, grows with the penalty, and is indistinguishable from zero without it. Third, the intercept portfolio barely moves with the penalty, is exactly invariant to the normalization of the fit when no penalty is applied, and earns 2.5 against 1.62 for the best zero-intercept portfolio.

Quantitatively, the penalty's scale moves the Sharpe ratio of the Euclidean fit from 1.62 to 0.30 across the grid, while the intercept portfolio stays between 2.50 and 2.51. The intercept portfolio's gain of 0.89 over the best zero-intercept portfolio arrives mainly through lower volatility per dollar of gross position, and relative to that portfolio and the five Fama–French factors it carries an alpha of 4.9% a year. It has the highest Sharpe ratio of the portfolios compared in every third of the sample gross of costs and in the first two thirds net of costs, and trading one month late leaves it ahead of the delayed best zero-intercept portfolio, 1.52 against 1.07. Net of 25 basis points per dollar traded it earns 1.60 or 1.61 at every penalty, with a break-even cost of 68 basis points, and over the full sample it exceeds every zero-intercept portfolio net as well as gross. Much of its gain over the best zero-intercept portfolio is a short-horizon premium that the factor block leaves in the intercept's loadings on short-horizon return and trading-activity characteristics.

The design principle that follows is to write a factor penalty in units that do not depend on the normalization, such as stock positions, where it is the same under every normalization, and to use the intercept a factor model already estimates. A comparison of latent-factor models by the Sharpe ratios of their portfolios should state the penalty and the units in which it is written, report the comparison along the penalty's scale or with the penalty written in one set of units, and say which components of the fitted mean the portfolio uses. Reported this way, a comparison of Sharpe ratios compares models under an allocation step that is stated rather than implied.

Data and Code Availability

The characteristics and returns are the U.S. files of the Jensen–Kelly–Pedersen data (Jensen, Kelly, and Pedersen, 2023), obtained through WRDS and built from CRSP and Compustat under an institutional license; the five Fama–French factors, the short-term reversal and momentum factors, and the risk-free rate are from Kenneth French's data library, and NBER recession dates are from National Bureau of Economic Research (2023). The processed panel cannot be redistributed. The replication package contains the code that produces every exhibit in the paper and the Internet Appendix from the rolling fits, the monthly portfolio returns and window-level diagnostics on which the tables and figures are computed, and a design document that describes each computation. The fitted models are available on request to researchers with access to the underlying data.

References

- Chen, Qihui, Nikolai Roussanov, and Xiaoliang Wang, 2023, Semiparametric conditional factor models: Estimation and inference, NBER Working Paper 31817.
- Chen, Qihui, Nikolai Roussanov, and Xiaoliang Wang, 2025, Semiparametric conditional factor models in asset pricing, Working paper, arXiv:2112.07121v5, April 2025.
- Choi, Jungjun, and Ming Yuan, 2025, Inferential theory for pricing errors with latent factors and firm characteristics, Working paper, arXiv:2511.03076, November 2025.
- Cochrane, John H., 2001, *Asset Pricing* (Princeton University Press, Princeton, NJ).
- Daniel, Kent, Lira Mota, Simon Rottke, and Tano Santos, 2020, The cross-section of risk and returns, *Review of Financial Studies* 33, 1927–1979.
- DeMiguel, Victor, Lorenzo Garlappi, and Raman Uppal, 2009, Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?, *Review of Financial Studies* 22, 1915–1953.
- DeMiguel, Victor, Alberto Martín-Utrera, Francisco J. Nogales, and Raman Uppal, 2020, A transaction-cost perspective on the multitude of firm characteristics, *Review of Financial Studies* 33, 2180–2222.
- Deng, Weichuan, Paweł Polak, Abolfazl Safikhani, and Ronakdilip Shah, 2024, A unified framework for fast large-scale portfolio optimization, *Data Science in Science* 3, 2295539.
- Detzel, Andrew, Robert Novy-Marx, and Mihail Velikov, 2023, Model comparison with transaction costs, *Journal of Finance* 78, 1743–1775.
- Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.
- Gu, Shihao, Bryan Kelly, and Dacheng Xiu, 2021, Autoencoder asset pricing models, *Journal of Econometrics* 222, 429–450.
- Hoerl, Arthur E., and Robert W. Kennard, 1970, Ridge regression: Biased estimation for nonorthogonal problems, *Technometrics* 12, 55–67.

- Jagannathan, Ravi, and Tongshu Ma, 2003, Risk reduction in large portfolios: Why imposing the wrong constraints helps, *Journal of Finance* 58, 1651–1683.
- Jensen, Theis Ingerslev, Bryan Kelly, Semyon Malamud, and Lasse Heje Pedersen, 2026, Machine learning and the implementable efficient frontier, *Review of Financial Studies* 39, 3035–3078.
- Jensen, Theis Ingerslev, Bryan Kelly, and Lasse Heje Pedersen, 2023, Is there a replication crisis in finance?, *Journal of Finance* 78, 2465–2518.
- Kelly, Bryan, Semyon Malamud, and Lasse Heje Pedersen, 2023, Principal portfolios, *Journal of Finance* 78, 347–387.
- Kelly, Bryan T., Seth Pruitt, and Yinan Su, 2019, Characteristics are covariances: A unified model of risk and return, *Journal of Financial Economics* 134, 501–524.
- Kim, Soohun, Robert A. Korajczyk, and Andreas Neuhierl, 2021, Arbitrage portfolios, *Review of Financial Studies* 34, 2813–2856.
- Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2018, Interpreting factor models, *Journal of Finance* 73, 1183–1223.
- Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2020, Shrinking the cross-section, *Journal of Financial Economics* 135, 271–292.
- Künsch, Hans R., 1989, The jackknife and the bootstrap for general stationary observations, *Annals of Statistics* 17, 1217–1241.
- Langlois, Hugues, 2023, What matters in a characteristic?, *Journal of Financial Economics* 149, 52–72.
- Ledoit, Olivier, and Michael Wolf, 2004, A well-conditioned estimator for large-dimensional covariance matrices, *Journal of Multivariate Analysis* 88, 365–411.
- Ledoit, Olivier, and Michael Wolf, 2017, Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets Goldilocks, *Review of Financial Studies* 30, 4349–4388.
- Lettau, Martin, and Markus Pelger, 2020, Estimating latent asset-pricing factors, *Journal of Econometrics* 218, 1–31.

- Lewellen, Jonathan, 2015, The cross-section of expected stock returns, *Critical Finance Review* 4, 1–44.
- Liu, Mingyang, 2026a, Characteristic-space metrics in factor models: Identification and inference, Unpublished working paper, August.
- Liu, Mingyang, 2026b, A geometric framework for identification in characteristic-based factor models, Working paper, SSRN 7013178.
- McLean, R. David, and Jeffrey Pontiff, 2016, Does academic research destroy stock return predictability?, *Journal of Finance* 71, 5–32.
- National Bureau of Economic Research, 2023, US business cycle expansions and contractions, Business cycle chronology, last updated March 14, 2023; accessed September 7, 2026.
- Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.
- Novy-Marx, Robert, and Mihail Velikov, 2016, A taxonomy of anomalies and their trading costs, *Review of Financial Studies* 29, 104–147.
- Windmüller, Steffen, 2022, Firm characteristics and global stock returns: A conditional asset pricing model, *Review of Asset Pricing Studies* 12, 447–499.
- Zaffaroni, Paolo, 2025, Factor models for conditional asset pricing, *Journal of Political Economy* 133, 2615–2642.
- Zhang, Chu, 2024, Testing pricing errors of models with latent factors and firm characteristics as covariances, *Management Science* 70, 1706–1728.

Internet Appendix

Same Forecasts, Different Portfolios: Factor Normalization and Regularization

Each section of this appendix answers one question raised by the main text. Section IA.A proves the results of Section II; Sections IA.B and IA.C document the rolling fits and the benchmarks of Section III; Section IA.D shows that every fitted portfolio is reproduced before any input is changed; Section IA.E reports the full penalty grid, other block lengths, the three estimates of the Gram metric, the Ledoit–Wolf shrinkage, and the 618-month sample behind Section IV; Sections IA.F and IA.G support Section V; Sections IA.H to IA.J support Section VI; Sections IA.K and IA.L describe the characteristics and the Gram metric; Section IA.M reports three checks on the intercept signal: a shuffle of the fitted intercept, the removal of extreme months, and a short-term reversal control; and Section IA.N describes what the intercept portfolio holds. The main text uses the 594 formation months from June 1975 through November 2024; every exhibit here names its sample. The replication package contains the code and the monthly series behind every exhibit.

IA.A. Why Does One Constant Impose Two Penalties? Proofs and a Two-Factor Example

This section proves Proposition 1 and Lemma 1 of Section II, adds a two-factor example with idiosyncratic risk, and states the exposure budget of the estimator. Notation follows Section II: B collects the stocks' factor exposures, m_F and C_F are the factor mean and covariance, $\widehat{D}$ is the diagonal of residual variances, λ_Σ the position penalty, λ the factor penalty, and $A_0 = BC_F B' + \widehat{D} + \lambda_\Sigma I$ the risk matrix without the factor penalty; time subscripts are suppressed.

IA.A.1. Matched representations: Proposition 1(a)

Let Γ_β have rank K and $Q > 0$, and set $C = \Gamma_\beta' Q \Gamma_\beta$, $H = C^{-1} \Gamma_\beta' Q \Gamma_\delta$, and $a = C^{-1} \Gamma_\beta' Q \Gamma_\alpha$. Full column rank of Γ_β and $Q > 0$ make C positive definite, so its symmetric square root and inverse

exist. The transformed blocks

$$\Gamma_\beta^Q = \Gamma_\beta C^{-1/2}, \quad \Gamma_\delta^Q = \Gamma_\delta - \Gamma_\beta H, \quad \Gamma_\alpha^Q = \Gamma_\alpha - \Gamma_\beta a, \quad f_{s+1}^Q = C^{1/2}(f_{s+1} + Hg_{s+1} + a) \quad (\text{IA1})$$

satisfy the normalization, because $(\Gamma_\beta^Q)' Q \Gamma_\beta^Q = C^{-1/2} C C^{-1/2} = I_K$ and $(\Gamma_\beta^Q)' Q [\Gamma_\delta^Q, \Gamma_\alpha^Q] = C^{-1/2} \{ \Gamma_\beta' Q [\Gamma_\delta, \Gamma_\alpha] - C[H, a] \} = 0$. They preserve the complete index, because

$$\Gamma_\beta^Q f_{s+1}^Q + \Gamma_\delta^Q g_{s+1} + \Gamma_\alpha^Q = \Gamma_\beta (f_{s+1} + Hg_{s+1} + a) + (\Gamma_\delta - \Gamma_\beta H) g_{s+1} + \Gamma_\alpha - \Gamma_\beta a = \Gamma_\beta f_{s+1} + \Gamma_\delta g_{s+1} + \Gamma_\alpha,$$

and multiplying by Z_s preserves every fitted return and residual.

In terms of exposures $B = Z[\Gamma_\beta, \Gamma_\delta]$ and factors $F = (f', g')'$ the change is $F^Q = JF + k$ with

$$J = \begin{pmatrix} C^{1/2} & C^{1/2}H \\ 0 & I_{K_o} \end{pmatrix}, \quad k = \begin{pmatrix} C^{1/2}a \\ 0 \end{pmatrix}, \quad J^{-1} = \begin{pmatrix} C^{-1/2} & -H \\ 0 & I_{K_o} \end{pmatrix}. \quad (\text{IA2})$$

Then $BJ^{-1} = Z[\Gamma_\beta C^{-1/2}, \Gamma_\delta - \Gamma_\beta H] = B^Q$; the factor mean and covariance of the transformed history are $m_F^Q = Jm_F + k$ and $C_F^Q = JC_F J'$, since the constant k shifts the mean and not the covariance. Hence $B^Q C_F^Q (B^Q)' = BJ^{-1} JC_F J' J^{-1'} B' = BC_F B'$, and

$$B^Q m_F^Q = Bm_F + BJ^{-1}k = Bm_F + Z\Gamma_\beta a, \quad Z\Gamma_\alpha^Q = Z\Gamma_\alpha - Z\Gamma_\beta a, \quad (\text{IA3})$$

so the complete mean $B^Q m_F^Q + Z\Gamma_\alpha^Q = Bm_F + Z\Gamma_\alpha$ is preserved, while its split between the factor block and the intercept moves by $Z\Gamma_\beta a$. With $\Gamma_\alpha = 0$, $a = 0$ and the factor-mean forecast is itself preserved. $\square$

IA.A.2. One penalty, one portfolio: Proposition 1(b)

Let $R \geq 0$ and $R^Q \geq 0$ be factor adjustments in the two representations, with common $\widehat{D}$ and λ_Σ . By part (a), $B^Q C_F^Q (B^Q)' = BC_F B'$, so the two stock risk matrices differ only through

BRB' and $B^Q R^Q (B^Q)' = BJ^{-1} R^Q J^{-1'} B'$. If $R^Q = JRJ'$ the two terms are equal. Conversely, if $BJ^{-1} R^Q J^{-1'} B' = BRB'$ and B has full column rank, multiplying on the left by $(B'B)^{-1} B'$ and on the right by its transpose gives $J^{-1} R^Q J^{-1'} = R$, that is, $R^Q = JRJ'$. The risk matrices are positive definite, so the portfolios $A^{-1}\mu$ coincide for every μ if and only if the risk matrices coincide. For $R = R^Q = \lambda I$ with $\lambda > 0$ the condition is $JJ' = I$, which for the square nonsingular J means that J is orthogonal; since

$$JJ' = \begin{pmatrix} C^{1/2}(I + HH')C^{1/2} & C^{1/2}H \\ H'C^{1/2} & I_{K_o} \end{pmatrix},$$

this holds if and only if $H = 0$ and $C = I_K$, that is, only if the two metrics already agree on the latent loadings. The rotation that the normalization leaves free is orthogonal and does not change the penalty. Expanding the risk term shows what each penalty charges,

$$w'Aw = (B'w)'C_F(B'w) + w'\widehat{D}w + (B'w)'R(B'w) + \lambda_\Sigma \|w\|_2^2, \quad (\text{IA4})$$

so with $R = \lambda I$ the factor penalty is $\frac{\lambda}{2} \|B'w\|^2$ in the objective, a charge on the portfolio's factor exposures in the units of the loadings, whereas the position penalty is written in stock units and does not depend on the normalization. $\square$

Standardized factors. Dividing each factor by its standard deviation before adding λ replaces $R = \lambda I$ by $R = \lambda S$, with $S = \text{diag}(C_F)$, and $R^Q = \lambda I$ by $R^Q = \lambda S_Q$, with $S_Q = \text{diag}(C_F^Q)$. By part (b) the two portfolios coincide if and only if $JSJ' = S_Q$, that is, if and only if $\tilde{J} = S_Q^{-1/2}JS^{1/2}$ is orthogonal, since $\tilde{J}\tilde{J}' = S_Q^{-1/2}JSJ'S_Q^{-1/2}$. This holds, for example, when the factors are uncorrelated in both representations: then $C_F = S$ and $C_F^Q = S_Q$, so $JSJ' = JC_FJ' = C_F^Q = S_Q$. In general J mixes the latent and the observed factors, and standardization fixes the scale of each factor but not that mixing. Table IA.A.1 sets the two ways of writing a factor penalty side by side.

Table IA.A.1. Two ways to write a factor penalty. B : the stocks' factor exposures; A_0 : the risk matrix without the factor penalty; J : the map of Proposition 1 between two normalizations, $B \rightarrow BJ^{-1}$; $\lambda_\Sigma I$: the position penalty.

	A constant on factor variances, in each normalization's units	A penalty in stock-position units
Objective	$w'\mu - \frac{1}{2}w'A_0w - \frac{\lambda}{2}\ B'w\ ^2$	$w'\mu - \frac{1}{2}w'A_0w - \frac{1}{2}w'Pw$, with $P = BRB'$
Change of normalization, $B \rightarrow BJ^{-1}$	Term becomes $\frac{\lambda}{2}\ (J^{-1})'B'w\ ^2$; it changes unless J is orthogonal	Term unchanged, with R carried across as JRJ'
Optimal weights	$(A_0 + \lambda BB')^{-1}\mu$, which depend on the normalization	$(A_0 + P)^{-1}\mu$, the same under every normalization
What it penalizes	Factor exposures in the units in which the loadings are reported	Factor exposures in one fixed set of units, or positions directly ($P = \lambda_\Sigma I$)
Without a penalty	$A_0^{-1}\mu$	$A_0^{-1}\mu$

IA.A.3. The intercept portfolio: Lemma 1

Part (a). At $\lambda = 0$ the risk matrix is A_0 in both representations, by Proposition 1(a), and the complete mean is preserved by (IA3); the intercept portfolio $A_0^{-1}(Bm_F + Z\Gamma_\alpha)$ is therefore the same under every normalization. The factor-mean portfolio uses Bm_F alone, which moves by $Z\Gamma_\beta a$; when $Z\Gamma_\beta$ has full column rank this is zero only if $a = 0$. $\square$

Part (b). Let $w_0 = A_0^{-1}\mu$, $w_\lambda = (A_0 + \lambda BB')^{-1}\mu$, $M = B'A_0^{-1}B \geq 0$, and $u = B'w_0$. The Woodbury identity gives $(A_0 + \lambda BB')^{-1} = A_0^{-1} - \lambda A_0^{-1}B(I + \lambda M)^{-1}B'A_0^{-1}$, so

$$w_\lambda - w_0 = -\lambda A_0^{-1}B(I + \lambda M)^{-1}u, \quad B'w_\lambda = u - \lambda M(I + \lambda M)^{-1}u = (I + \lambda M)^{-1}u. \quad (\text{IA5})$$

Since M and $(I + \lambda M)^{-1}$ commute,

$$(w_\lambda - w_0)'A_0(w_\lambda - w_0) = \lambda^2 u'(I + \lambda M)^{-1}M(I + \lambda M)^{-1}u.$$

On each eigenvector of M with eigenvalue $\nu \geq 0$ the matrix in this form has eigenvalue $\lambda^2 \nu / (1 + \lambda \nu)^2 = \lambda \cdot \lambda \nu / (1 + \lambda \nu)^2 \leq \lambda/4$, because $x/(1+x)^2 \leq 1/4$ for $x \geq 0$. Hence

$$(w_\lambda - w_0)' A_0 (w_\lambda - w_0) \leq \frac{\lambda}{4} \|B' w_0\|^2, \quad \|B' w_\lambda\| \leq \|B' w_0\|, \quad (\text{IA6})$$

the second because the eigenvalues of $(I + \lambda M)^{-1}$ lie in $(0, 1]$. $\square$

Two consequences are used in the text. First, write $\|v\|_{A_0} = (v' A_0 v)^{1/2}$. Two normalizations, each penalized by λI in its own units, share A_0 and, by part (a), the unpenalized intercept portfolio w_0 ; by the triangle inequality their penalized intercept portfolios satisfy $\|w_\lambda^I - w_\lambda^Q\|_{A_0} \leq \frac{1}{2} \sqrt{\lambda} (\|B' w_0\| + \|B^Q w_0\|)$, a bound set by the factor exposure of one portfolio written in the two sets of units. Second, for the factor-mean portfolio $w_0^{\text{FM}} = A_0^{-1} B m_F$ the exposure is $B' w_0^{\text{FM}} = M m_F$, so its whole forecast is carried by the exposures that the penalty shrinks. The bound concerns weights; the Sharpe ratios along the penalty grid are reported in Section V and Table IA.F.1.

IA.A.4. A two-factor example with idiosyncratic risk

The example of Section II lets the investor choose factor exposures directly. With idiosyncratic risk, let two factors have orthogonal exposure vectors $b_1, b_2 \in \mathbb{R}^N$, variances σ_1^2 and σ_2^2 , zero covariance, and means m_1 and m_2 ; let the residual risk be dI_N , with the position penalty absorbed into d ; and let $R = \lambda I_2$. Because $A^{-1} b_k = b_k / \{d + (\sigma_k^2 + \lambda) \|b_k\|^2\}$, the portfolio is

$$w = \sum_{k=1}^2 \frac{m_k b_k}{d + (\sigma_k^2 + \lambda) \|b_k\|^2}. \quad (\text{IA7})$$

Rescale the first factor by $c > 0$: $f_1 \mapsto c f_1$, $b_1 \mapsto b_1/c$, $m_1 \mapsto c m_1$, $\sigma_1^2 \mapsto c^2 \sigma_1^2$. Fitted returns are unchanged, and the first term of (IA7) becomes $m_1 b_1 / \{d + (\sigma_1^2 + \lambda/c^2) \|b_1\|^2\}$. Without the penalty the portfolio is unchanged; with it, the same number λ acts as λ/c^2 on the rescaled factor, so doubling a factor and halving its loadings quarters the penalty on every position exposed to it,

whatever the idiosyncratic risk.

IA.A.5. *The exposure budget of the estimator*

The intercept-inclusive specification of IPCA used in the paper, introduced as No-Arbitrage IPCA by Liu (2026b), is estimated under the Gram normalization subject to the exposure budget

$$\Gamma'_\alpha Q \Gamma_\alpha \leq \delta_\alpha, \quad \delta_\alpha = 0.035. \quad (\text{IA8})$$

With $Q = \sum_s \omega_s W_s$, $\omega_s \geq 0$, $\sum_s \omega_s = 1$, the left side equals $\sum_s \omega_s N_s^{-1} \sum_i \alpha_{i,s}^2$ with $\alpha_{i,s} = z'_{i,s} \Gamma_\alpha$: the budget bounds a weighted average of squared stock-level intercepts rather than their signed average. For any position w , Cauchy–Schwarz gives

$$|w' Z \Gamma_\alpha| \leq (w' Z Q^{-1} Z' w)^{1/2} (\Gamma'_\alpha Q \Gamma_\alpha)^{1/2} \leq \sqrt{\delta_\alpha} (w' Z Q^{-1} Z' w)^{1/2}, \quad (\text{IA9})$$

so the budget bounds a portfolio's intercept contribution relative to its characteristic exposure, measured by $Z Q^{-1} Z'$, which is not a return covariance. Liu (2026b) reads $\sqrt{\delta_\alpha}$ as a maximum alpha Sharpe ratio; that reading treats $Z Q^{-1} Z'$ in (IA9) as the residual return covariance, which it is not here, so the budget does not bound the intercept portfolio's Sharpe ratio. The motivation is the square-summability of cross-sectional pricing errors that Zaffaroni (2025) obtains under asymptotic no-arbitrage and further regularity conditions; the finite-sample budget is not that condition. Taking conditional expectations in the model, expected returns are $z'_{i,s} \Gamma_\beta m_{f,s} + z'_{i,s} \Gamma_\delta m_{g,s} + \alpha_{i,s}$; the intercept is a pricing error only under a factor-pricing specification (Cochrane, 2001, Chapters 1 and 4), and omitted factors or misspecified conditional dynamics produce the same term.

The budget is slack at every formation date: across the 618 final estimates the largest value of $\Gamma'_\alpha Q \Gamma_\alpha$ is 1.51×10^{-4} , against 0.035 (Section IA.B.3), so no result uses or requires a no-arbitrage bound. A fixed numerical budget measures coefficient size under I_L but stock-level exposure under a Gram metric, so after a change of normalization it is evaluated on $(\Gamma_\alpha^Q)' Q \Gamma_\alpha^Q$.

IA.B. How Are the Rolling Fits Estimated?

This section documents the estimation behind the rolling fits of Section III: the criterion, its numerical regularization, the normalization and stopping rules, the portfolio solve, and the information used at each formation date. Every fit uses $L = 132$ characteristics, $K = 3$ latent and $K_o = 5$ observed factors, and a 120-month window.

IA.B.1. The least-squares criterion

Within a formation window write the complete index as $h_{s+1} = \Gamma x_{s+1}$, with $\Gamma = [\Gamma_\beta, \Gamma_\delta, \Gamma_\alpha]$ and $x_{s+1} = (f'_{s+1}, g'_{s+1}, 1)'$; the final column and regressor are omitted when the intercept is excluded. With $W_s = Z'_s Z_s / N_s$ and the characteristic-managed returns $X_s = Z'_s r_{s+1} / N_s$, define $m_s = X_s - W_s h_{s+1}$. Since X_s lies in the range of W_s , the Moore–Penrose identities give

$$\begin{aligned} m'_s W_s^+ m_s &= X'_s W_s^+ X_s - 2X'_s h_{s+1} + h'_{s+1} W_s h_{s+1}, \\ \frac{\|r_{s+1} - Z_s h_{s+1}\|^2}{N_s} &= \frac{r'_{s+1} r_{s+1}}{N_s} - 2X'_s h_{s+1} + h'_{s+1} W_s h_{s+1}. \end{aligned} \quad (\text{IA10})$$

The two criteria differ only in their constant term, so the moment criterion $\sum_s N_s m'_s W_s^+ m_s$ weights every stock-month equally, as the underlying least-squares problem does; equal weighting of the normalized moment criterion would instead weight dates equally.

Unique factor profiling requires $\Gamma'_\beta W_s \Gamma_\beta > 0$, not full rank of W_s . With loadings held fixed, the factor normal equation is

$$(\Gamma'_\beta W_s \Gamma_\beta) f_{s+1} = \Gamma'_\beta [X_s - W_s (\Gamma_\delta g_{s+1} + \Gamma_\alpha)]. \quad (\text{IA11})$$

With factors fixed, differentiating the objective in Γ gives

$$\left\{ \sum_s N_s (x_{s+1} x'_{s+1} \otimes W_s) \right\} \text{vec}(\Gamma) = \sum_s N_s \text{vec}(X_s x'_{s+1}), \quad (\text{IA12})$$

where vec stacks columns and $\text{vec}(W\Gamma xx') = (xx' \otimes W) \text{vec}(\Gamma)$. These are unconstrained conditional normal equations; the normalization and cross-block restrictions are restored after each update (Section IA.B.3).

IA.B.2. Numerical regularization in estimation

The estimator replaces W_s by $W_s^{\text{imp}} = W_s + \rho_W I_L$, with $\rho_W = 10^{-8}$. Expanding the inverse-weighted criterion,

$$\begin{aligned} & \sum_s N_s (X_s - W_s^{\text{imp}} h_{s+1})' (W_s^{\text{imp}})^{-1} (X_s - W_s^{\text{imp}} h_{s+1}) \\ &= \sum_s \|r_{s+1} - Z_s h_{s+1}\|^2 + \rho_W \sum_s N_s \|h_{s+1}\|^2 + C^{\text{imp}}, \end{aligned} \quad (\text{IA13})$$

where C^{imp} does not depend on the parameters. The adjustment is a small quadratic penalty on the complete index h_{s+1} , which a matched change of normalization preserves (Section IA.A.1), so it does not depend on the normalization. Separate jitters may be added to the linear systems (IA11)–(IA12), and no penalty on the loading blocks is used. This moment regularization, the jitters, the factor penalty λ , and the position penalty λ_Σ are four distinct operations; only the last two enter the portfolio step.

IA.B.3. Normalization, the budget, and stopping

After each update the normalization and cross-block restrictions (2) are restored by matched transformations of the loading blocks and the factor path, as in Section IA.A.1. Cholesky normalization, for example, uses $R'R = \Gamma'_\beta Q \Gamma_\beta$ and replaces Γ_β by $\Gamma_\beta R^{-1}$ and the latent factors by Rf ; removing a loading component in the latent span adds the compensating observed-factor or constant term to f . Both operations preserve the fitted index. The intercept projection also weakly reduces

its exposure: with $c = \Gamma'_\beta Q \Gamma_\alpha$ and $\Gamma'_\beta Q \Gamma_\beta = I_K$,

$$(\Gamma_\alpha - \Gamma_\beta c)' Q (\Gamma_\alpha - \Gamma_\beta c) = \Gamma'_\alpha Q \Gamma_\alpha - c' c, \quad (\text{IA14})$$

so a second budget rescale after the projection is redundant in exact Q geometry. Radial rescaling of an intercept that violates the budget is not in general fit-preserving, and an unconstrained update followed by a projection need not be the exact constrained minimizer; slackness of the budget at the final estimates therefore describes the solution reached, not the whole numerical path.

The iteration monitor is the change in the unweighted managed-moment reconstruction norm. Every conditional-model window uses a tolerance of 10^{-6} and a cap of 1,000 cycles, and the rolling fits of IPCA with an intercept warm-start from the preceding window. Across the 618 final estimates of that fit, the largest intercept exposure is 1.51×10^{-4} against the budget of 0.035, the largest latent-normalization residual is 1.74×10^{-12} , and the largest cross-block residual is 5.27×10^{-15} . All 618 windows report convergence, one of them at the iteration cap. In 404 windows the monitor rises at least once by more than 10^{-10} , which is not a test of descent in the differently weighted objective, and the fits do not record budget-enforcement events. A small reconstruction change and a stationary constrained solution are different properties, and the convergence report concerns the first.

Table IA.B.1 starts the Euclidean and the Gram fits from one common singular-vector initialization. The two fits again span nearly the same latent space and give nearly equal forecasts, and their portfolios again differ at the reference penalty and coincide without it, so the difference in holdings does not come from starting values. One window, June 1999, reaches the iteration cap in both fits and remains in every diagnostic.

Table IA.B.1. Common-initialization diagnostic, Gram versus Euclidean fit. A separate estimation over the 618 all-stock windows, June 1973–November 2024, starts both fits from one common singular-vector initialization, with zero block-loading penalty, moment regularization 10^{-8} , tolerance 10^{-6} , and at most 1,000 iterations. Both fits meet the stopping rule in 617 windows and reach the cap in June 1999, which remains in every row. Both weight comparisons hold the Euclidean fit’s residual diagonal common. The last row removes the factor penalty from the re-estimated factor covariances and keeps the position penalty and a numerical eigenvalue floor. The rows compare spans, forecasts, and weights.

Diagnostic	Median	Maximum
Largest latent-span angle (degrees)	4.3×10^{-4}	1.1×10^{-2}
Relative forecast gap	4.7×10^{-6}	1.4×10^{-4}
Weight cosine, reference penalty $\lambda = 0.01$	0.830	0.988
1–cosine, no factor penalty	5.8×10^{-12}	4.6×10^{-9}

IA.B.4. The portfolio solve

The stock risk matrix has the form $D + VCV'$ with $D > 0$ diagonal and $C \geq 0$, where D includes $\lambda_{\Sigma}I$. With the symmetric square root of C and $U = VC^{1/2}$, the Woodbury identity gives

$$(D + UU')^{-1}\mu = D^{-1}\mu - D^{-1}U(I + U'D^{-1}U)^{-1}U'D^{-1}\mu. \quad (\text{IA15})$$

Only the small middle system is inverted, and C itself is never inverted, so the formula also applies when C is singular. It is the same portfolio solve as $A_t^{-1}\mu_t$ up to numerical precision.

The residual diagonal $\widehat{D}_t$ holds each stock’s training-window variance of $r - Bf$: it subtracts factor contributions, and in the intercept fit it does not subtract the fitted intercept. Every same-fit comparison holds this diagonal fixed, so a change of input changes nothing else in the risk matrix. Section IA.D.4 reports the alternative convention that subtracts the penalized factor variance.

IA.B.5. Information used at formation and the 594-month sample

At month t the fit uses the pairs $\{Z_s, r_{s+1}, g_{s+1}\}_{s \in \mathcal{W}_t}$; the last pair is (Z_{t-1}, r_t, g_t) , and the current cross section Z_t forms the forecast of r_{t+1} . Writing $\widehat{F}_{s+1}^{(t)} = (\widehat{f}_{s+1}^{(t)}, g'_{s+1})'$ for the factor realizations

of the window- t fit, the factor mean and covariance inputs are

$$\begin{aligned}\widehat{m}_{F,t} &= \frac{1}{120} \sum_{s \in \mathcal{W}_t} \widehat{F}_{s+1}^{(t)}, \\ \widehat{C}_{F,t} &= \frac{1}{119} \sum_{s \in \mathcal{W}_t} (\widehat{F}_{s+1}^{(t)} - \widehat{m}_{F,t})(\widehat{F}_{s+1}^{(t)} - \widehat{m}_{F,t})'.\end{aligned}\tag{IA16}$$

These are moments of monthly factor returns within the current fit, distinct from the dispersion of past loading estimates that the stabilized intercept portfolio of Section IA.G uses.

The first rolling fit is dated June 1973. The stabilized intercept portfolio needs the 24 preceding intercept-loading estimates, and the twenty-fifth fit, dated June 1975, is the first with 24 predecessors (June 1973 through May 1975); its training characteristics run from June 1965 through May 1975, and its payoff is realized in July 1975. The 618-month sample therefore runs from June 1973 and the 594-month sample from June 1975, both through November 2024. The 594-month sample changes the evaluation dates, not the 120-month estimation window, and every portfolio in the main text is evaluated on it. Adjacent fits share 119 training months, so the 24 prior estimates are not independent. Figure IA.B.1 places both histories relative to formation and payoff.

IA.B.6. Preprocessing and information dates

The estimator uses a processed rank-based panel: characteristics are ranked cross-sectionally each month to $[-0.5, 0.5]$, nonfinite entries are set to zero, entries are clipped, and each window then standardizes the characteristics with its training-sample moments and applies the same transformation at formation. In the processed panel none of the 132 characteristic columns or the return column contains a nonfinite value, and no entry lies outside the clipping threshold. The panel holds 1,362,000 stock-months of mega-, large-, and small-capitalization stocks, of which 1,244,002 fall in the 618-month evaluation sample. A stock-month enters when its next-month return is available in the source panel, and no evaluated cross section has a missing next-month return. Four evaluation

A. Two histories feed one formation-date portfolio

Current model fit

120 pairs: (Z_s, r_{s+1}, g_{s+1})

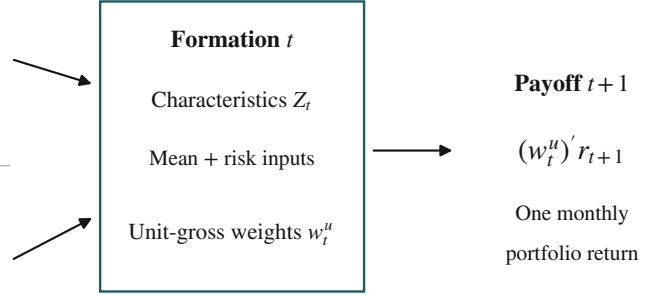
$s = t - 120, \dots, t - 1$; last payoff r_t

Current loadings; factor mean and covariance

Prior loading estimates

24 vectors: $\hat{\Gamma}_{\alpha, t-24}, \dots, \hat{\Gamma}_{\alpha, t-1}$

Instability input; current estimate excluded



B. June 1975: first formation with 24 prior estimates

Training characteristics	June 1965 - May 1975	120 months
Paired returns and factors	July 1965 - June 1975	120 months
Prior intercept-loading estimates	June 1973 - May 1975	24 vectors
Formation / next-month payoff	June 1975 / July 1975	1 portfolio return

Figure IA.B.1. Training, prior loading estimates, and portfolio formation. Panel A distinguishes the 120 training pairs from the 24 prior monthly intercept-loading estimates used by the stabilized intercept portfolio; both feed the formation-date portfolio, whose return is realized one month later. Panel B gives the first formation with 24 prior estimates, June 1975, the start of the 594-month sample. The current loading enters the mean forecast and is excluded from that month's loading-dispersion input.

excess returns are below -1 ; adding the next-month risk-free rate leaves one implied total return below -1 , by 1.83 basis points, and the observation is kept as supplied.

Across the 4,944 rolling fits of the eight conditional-model sequences, every estimation span is 120 months, the information cutoff equals the training end, and evaluation dates run consecutively from June 1973 to November 2024. Observed factors are dated to the return month, one month after the characteristics paired with them. These conventions apply to every portfolio in the paper in the same way.

IA.C. How Are the Benchmarks Built?

This section defines the benchmarks of the ladder in Section III and Table I, and Table IA.C.1 reports the levels of every portfolio on the 594-month sample. Every fitted benchmark enters the allocation rule of Section IA.A.2 with its own exposures, factor history, and residual variances, and with the same penalties and unit-gross policy. Throughout, t is the formation month and $\mathcal{W}_t = \{t - 120, \dots, t - 1\}$ its training window, with $n = 120$.

IA.C.1. The characteristic-instrumented FF5 benchmark

The FF5 benchmark uses the market, size, value, profitability, and investment factors of Fama and French (2015), $g_{s+1} \in \mathbb{R}^5$, with exposures instrumented by characteristics and no latent factor or intercept:

$$r_{i,s+1} = z'_{i,s} \Gamma_{\delta,t} g_{s+1} + \varepsilon_{i,s+1}, \quad \beta_{i,s}^{\text{FF5}} = \Gamma'_{\delta,t} z_{i,s}. \quad (\text{IA17})$$

The map $\Gamma_{\delta,t} \in \mathbb{R}^{132 \times 5}$ has 660 coefficients and solves the pooled least-squares problem

$$\widehat{\Gamma}_{\delta,t} \in \arg \min_{G \in \mathbb{R}^{L \times 5}} \sum_{s \in \mathcal{W}_t} \|r_{s+1} - Z_s G g_{s+1}\|^2, \quad (\text{IA18})$$

which weights stock-months equally; its normal equation is (IA12) with $x_{s+1} = g_{s+1}$ and the moment regularization of Section IA.B.2. At formation,

$$B_t^{\text{FF5}} = Z_t \widehat{\Gamma}_{\delta,t}, \quad \widehat{m}_{g,t} = \frac{1}{n} \sum_{s \in \mathcal{W}_t} g_{s+1}, \quad \widehat{\mu}_t^{\text{FF5}} = B_t^{\text{FF5}} \widehat{m}_{g,t}, \quad (\text{IA19})$$

and the factor covariance is the centered sample covariance of the same factor history. The benchmark estimates how characteristics map into observed-factor exposures; it is not a stock-by-stock beta regression or a Fama–MacBeth second pass, and it has five factors rather than eight.

IA.C.2. RP-PCA on characteristic-managed returns

Stack the managed returns $X_s = Z'_s r_{s+1} / N_s$ into $M_t \in \mathbb{R}^{n \times L}$, a balanced panel of characteristic-managed returns, with mean $\bar{x}_t = M'_t \mathbf{1}_n / n$ and centered version $M_{c,t} = M_t - \mathbf{1}_n \bar{x}'_t$. The RP-PCA estimator of Lettau and Pelger (2020) selects the leading eight eigenvectors of

$$S_{\gamma,t} = \frac{1}{n} M'_t \left(I_n + \frac{\gamma_{\text{RP}}}{n} \mathbf{1}_n \mathbf{1}'_n \right) M_t = \frac{1}{n} M'_{c,t} M_{c,t} + (1 + \gamma_{\text{RP}}) \bar{x}_t \bar{x}'_t, \quad \gamma_{\text{RP}} = 3, \quad (\text{IA20})$$

so that expected-return information shares in selecting directions with covariance variation. Let $V_t \in \mathbb{R}^{L \times 8}$ collect these orthonormal eigenvectors, with signs oriented by the training factor means. No additional managed-return standardization, factor-variance normalization, or factor orthogonalization is applied. The loadings, factor history, and factor mean are

$$\widehat{\Gamma}_{\beta,t}^{\text{RP}} = V_t, \quad \widehat{F}_t^{\text{RP}} = M_t V_t, \quad \widehat{m}_{f,t}^{\text{RP}} = V'_t \bar{x}_t, \quad (\text{IA21})$$

and the portfolio uses $B_t^{\text{RP}} = Z_t V_t$, $\widehat{\mu}_t^{\text{RP}} = B_t^{\text{RP}} \widehat{m}_{f,t}^{\text{RP}}$, and the centered sample covariance of $\widehat{F}_t^{\text{RP}}$. The estimator's auxiliary regression intercepts are not passed to the portfolio, so RP-PCA is a zero-intercept benchmark. The extraction matrix weights training months equally, whereas the FF5 criterion aggregates stock-level errors across the unbalanced panel.

IA.C.3. The pooled characteristic regression

At each formation date the pooled characteristic regression regresses next-month excess returns on the 132 standardized characteristics and a constant over the training window, a pooled version of the characteristic forecasts of Lewellen (2015), solved from the window's moment matrices. Its fitted forecast $Z_t \widehat{b}_t$ is passed through the risk matrix of the intercept fit, with the factor penalty at $\lambda = 0.01$ or $\lambda = 0$, the same residual diagonal, and the same position penalty. The benchmark therefore forms a mean without passing through a factor model, and holds the risk input fixed at

that of the intercept portfolio.

IA.C.4. The market and equal weights

The market benchmark is the Fama–French excess market return dated to the return month; it has no holdings, so it has no turnover, net return, or delayed version. The equal-weight benchmark $(1/N)$ holds every stock of the evaluated cross section with equal weight, so it uses the same universe and next-month excess returns as every fitted portfolio; its turnover is computed from return-drifted weights exactly as in Section IA.J.1. Neither benchmark depends on the factor penalty.

IA.C.5. Why centered PCA is not a benchmark

A PCA configuration that selects covariance eigenvectors V_t^{PCA} but returns the centered history $M_{c,t}V_t^{\text{PCA}}$ has a factor mean of zero by construction. With no separate mean block, its portfolio forecast is zero in exact arithmetic; in the rolling fits its forecast is numerically zero and its raw gross exposure is below 4×10^{-16} , so it supplies no portfolio and is excluded from every ranking. RP-PCA is the mean-sensitive eight-factor alternative.

Table IA.C.1. Levels of every portfolio on the 594-month sample. The sample is the 594 formation months from June 1975 through November 2024. Returns are gross of costs at unit gross exposure; means and volatilities are annualized and in percent; drawdowns are compounded. λ is the factor penalty in each fit's own units; the three Gram rows use the three estimates of the Gram metric of Section IA.E.5. Market is Mkt-RF and $1/N$ weights the evaluated stocks equally; neither uses a fitted model or λ . The last five rows use one rolling fit of IPCA with an intercept. Panel B: paired differences with pointwise 95% intervals from the 12-month moving-block bootstrap, holding the fitted models fixed; mean differences are in percentage points a year.

Panel A: Gross performance				
Portfolio	Mean	Volatility	Sharpe	Drawdown
IPCA, Euclidean units, $\lambda = 0.01$	4.0	5.9	0.69	-34.6
IPCA, Gram units (EWMA estimate), $\lambda = 0.01$	4.8	4.5	1.07	-14.2
IPCA, Gram units (equal-120 estimate), $\lambda = 0.01$	4.7	4.6	1.03	-17.3
IPCA, Gram units (rolling-60 estimate), $\lambda = 0.01$	4.9	4.3	1.13	-11.1
Best ZP, $\lambda = 0$	6.6	4.1	1.62	-8.4
RP-PCA, $\lambda = 0.01$	5.8	3.9	1.49	-13.4
FF5 (characteristic-instrumented), $\lambda = 0.01$	3.9	5.2	0.75	-20.7
Pooled characteristic regression, $\lambda = 0.01$	10.0	10.9	0.92	-29.4
Pooled characteristic regression, $\lambda = 0$	10.2	11.1	0.92	-29.4
Market	8.3	15.5	0.54	-54.1
Equal weight ($1/N$)	9.5	18.9	0.50	-55.9
Factor-mean portfolio (same fit), $\lambda = 0.01$	1.7	4.5	0.39	-17.7
Factor-mean portfolio (same fit), $\lambda = 0$	2.2	4.0	0.56	-12.0
Intercept portfolio (IP), $\lambda = 0.01$	7.6	3.0	2.51	-3.9
Intercept portfolio (IP), $\lambda = 0$	7.8	3.1	2.50	-4.5
IP, stabilized, $\lambda = 0.01$	7.6	3.0	2.53	-3.9
Panel B: Paired differences				
Comparison	Sharpe	95% interval	Mean (pp)	95% interval
IP minus factor-mean portfolio, $\lambda = 0.01$	2.12	[1.64, 2.68]	5.84	[4.16, 7.60]
IP minus factor-mean portfolio, $\lambda = 0$	1.95	[1.49, 2.51]	5.55	[3.94, 7.28]
IP ($\lambda = 0.01$) minus best ZP	0.89	[0.59, 1.26]	0.95	[-0.14, 2.06]
IP ($\lambda = 0$) minus best ZP	0.88	[0.59, 1.23]	1.16	[0.12, 2.21]
IP ($\lambda = 0.01$) minus RP-PCA	1.02	[0.71, 1.39]	1.79	[0.74, 2.93]

IA.D. Are the Fitted Portfolios Reproduced Exactly?

Every experiment in Sections IV to VI changes one portfolio input while the rolling fits stay fixed. This section describes that design, shows that every fitted portfolio is reproduced before any input changes (Table IA.D.1), reports the diagnostics behind Section IV.A, and examines the convention for the residual diagonal.

IA.D.1. Design

For each of the 618 formation windows the fitted objects of each model are the loading matrices $\widehat{\Gamma}_\beta$ and $\widehat{\Gamma}_\delta$ (and $\widehat{\Gamma}_\alpha$ for IPCA with an intercept), the 120-month factor history, the factor covariance including the reference penalty, the stock residual variances, the stock forecast, and the raw portfolio weights. Stock exposures B_t are rebuilt from the standardized characteristics and the fitted loadings, and the residual diagonal is each stock’s training-window residual variance; for the zero-intercept fits it is recovered by subtracting the penalized common-factor variance from the fitted total variance, which removes the penalty exactly. The factor penalty is then replaced: by nothing ($\lambda = 0$); by the same constant written in the other fit’s units, through the map J of Proposition 1; by λI for $\lambda \in \{0, 0.001, 0.003, 0.01, 0.03, 0.1\}$; or by the Ledoit–Wolf estimate of Section IA.E.6. The portfolio is re-solved with the same forecast, residual diagonal, and position penalty and normalized to unit gross exposure. The intercept portfolio and the factor-mean portfolio change the mean input of one fit in the same way, with everything else held fixed.

The penalty grid, the paired contrasts, and the bootstrap seed were fixed before any return was computed; the intercept grid, the cost calculation, the Ledoit–Wolf check, and the choice of the best zero-intercept portfolio as the benchmark for the intercept portfolio were specified afterwards, each before it was run. The main-text sample of 594 months and the plain intercept portfolio as the main-text rule were chosen after both samples and both rules had been computed; Sections IA.E and IA.G report the alternatives in full. All variants and contrasts are reported.

IA.D.2. The map between the two normalizations

The map J of (IA2) is built from the fitted Euclidean loadings and the Gram metric of the same window. The rotation of the normalized latent block, which the normalization leaves free, is fixed by the orthogonal polar factor of the alignment matrix between the transformed Euclidean loadings and the fitted Gram loadings; no forecast, weight, or return enters the construction. Table IA.D.1 shows that the map reproduces the Gram exposures and factor covariances with median relative

errors below 10^{-5} , and that every weight and monthly return is reproduced to better than 10^{-9} before an input is changed. The fitted factor covariances are themselves reproduced from the fitted factor histories, with observed factors dated to the return month, to a relative precision of 10^{-15} , and they are positive definite without the factor penalty: the smallest eigenvalue of a fitted covariance less $0.01I$ is 3.1×10^{-5} .

Table IA.D.1. Reproduction of the fitted portfolios. The sample is all 618 formation months from June 1973 through November 2024. Panel A compares each portfolio rebuilt from the fitted models with the portfolio of the fit itself, before any input is changed: raw weights by relative L_2 error and monthly unit-gross returns by absolute error. Panel B checks the map J of Proposition 1 between the Euclidean and the Gram fits, built from the fitted loadings and the Gram metric without using any forecast, weight, or return. The last row is the smallest eigenvalue of a fitted factor covariance less the reference penalty $0.01I$. Blank cells: only the maximum is reported.

Check	Median	Maximum
Panel A: Fitted portfolios rebuilt		
Raw weights, IPCA in Euclidean units (relative L_2 error)	3.2×10^{-13}	3.4×10^{-10}
Raw weights, IPCA in Gram units (relative L_2 error)	3.9×10^{-13}	1.7×10^{-10}
Monthly returns, IPCA in Euclidean units (absolute error)		2.6×10^{-12}
Monthly returns, IPCA in Gram units (absolute error)		9.4×10^{-12}
Monthly returns, factor-mean portfolio, same fit (absolute error)		7.7×10^{-12}
Monthly returns, intercept portfolio (absolute error)		1.4×10^{-14}
Monthly returns, RP-PCA and FF5 (absolute error)		4.2×10^{-16}
Panel B: The map between the two normalizations		
Loadings: $\ B_Q - B_I J^{-1}\ _F / \ B_Q\ _F$	3.8×10^{-6}	8.9×10^{-5}
Factor covariance: $\ C_Q - J C_I J'\ _F / \ C_Q\ _F$	1.1×10^{-6}	2.7×10^{-5}
Latent rotation: distance of the alignment matrix from orthogonality	2.5×10^{-11}	1.0×10^{-8}
Smallest eigenvalue of a fitted factor covariance less $0.01I$		3.1×10^{-5}

IA.D.3. Same forecasts, different holdings

At a common formation date let A_I and A_Q be the stock risk matrices and μ_I and μ_Q the forecasts of the Euclidean and the Gram fits. Their raw weights satisfy $A_I w_I = \mu_I$ and $A_Q w_Q = \mu_Q$, so

$$w_Q - w_I = A_Q^{-1}(\mu_Q - \mu_I) + A_Q^{-1}(A_I - A_Q)w_I, \quad (\text{IA22})$$

an exact split into a mean term at the Gram risk matrix and a risk term at the Euclidean forecast. For the common stock set the relative forecast gap and the weight cosine are

$$d_{\mu,t} = \frac{\|\widehat{\mu}_{Q,t} - \widehat{\mu}_{I,t}\|_2}{\|\widehat{\mu}_{I,t}\|_2}, \quad c_{w,t} = \frac{w'_{Q,t} w_{I,t}}{\|w_{Q,t}\|_2 \|w_{I,t}\|_2}. \quad (\text{IA23})$$

The relative weight gap replaces the forecasts in $d_{\mu,t}$ by raw weights and so measures scale as well as direction; the cosine measures direction only.

The two fits are separately estimated, so their agreement is measured rather than assumed. Across the 618 windows the relative forecast gap has a median of 4.7×10^{-6} , a 95th percentile of 1.8×10^{-5} , and a maximum of 1.4×10^{-4} ; the latent loading spans differ by a largest principal angle of median 4.3×10^{-4} degrees; and the residual diagonals differ by a median relative gap of 4.2×10^{-7} . The mean term of (IA22) is a median 0.07% of the weight gap. At the reference penalty the weight cosine has a median of 0.83 and an interquartile range of [0.753, 0.878]; without the factor penalty one minus the cosine has a median of 6×10^{-12} (Table IA.D.2).

Table IA.D.2. Forecasts, risk inputs, and holdings of the Euclidean and the Gram fits. The sample is the 618 formation months from June 1973 through November 2024, with the Gram metric estimated by the EWMA construction and the factor penalty at the reference value $\lambda = 0.01$ unless a row says otherwise. Norm ratios are diagnostic magnitudes, not additive shares; cosines are unchanged by positive rescaling of either weight vector. The rows without the factor penalty delete only λI from the fitted risk matrices. Dashes: the statistic is a single summary.

Diagnostic	Median	Maximum or minimum
Mean term / total weight-gap norm (%)	0.07	—
Relative residual-variance gap	4.2×10^{-7}	—
Largest latent-span angle (degrees)	4.3×10^{-4}	Max. 0.011
Weight cosine, reference penalty	0.830	—
Relative raw-weight gap, reference penalty	0.607	Max. 0.895
1–cosine without the factor penalty	6.0×10^{-12}	Max. 4.6×10^{-9}
Relative raw-weight gap without the factor penalty	3.9×10^{-6}	Max. 9.6×10^{-5}
Rebuilt factor covariance, relative gap, both fits	—	$< 2 \times 10^{-15}$

The covariance row compares the sample covariance of the fitted latent factors and the observed factors, paired with the same return dates, with the fitted factor covariance less $0.01I$, over 1,236 fit-window matrices. The first three rows measure the mean and risk terms of equation (IA22); none is a contribution to realized returns.

IA.D.4. The residual diagonal

Every same-fit comparison holds the fitted residual diagonal D fixed, so that a change of input changes nothing else in the risk matrix. An alternative diagonal subtracts the penalized factor variance from each stock's total variance, $D' = D + \lambda \|b_i\|^2$ at $\lambda = 0.01$; it moves part of the penalty into idiosyncratic risk rather than removing it, it is a median 2.4 times D in Euclidean units and 3.8 times in Gram units, and, unlike D , it depends on the units of the loadings. Recomputing the unpenalized portfolio with D' moves its Sharpe ratio from 1.57 to 1.52 in Euclidean units and to 1.46 in Gram units on the 618-month sample, and the gains from removing the penalty become 0.88 and 0.55, with intervals that still exclude zero (Table IA.D.3). The gain from removing the penalty is therefore a property of the penalty rather than of the diagonal convention.

Table IA.D.3. The residual diagonal and the unpenalized portfolio. The sample is the 618 formation months from June 1973 through November 2024; the main text uses 594 months. The fitted diagonal holds each stock's training-window residual variance; the alternative D' subtracts the penalized factor variance from each stock's total variance, $D' = D + 0.01 \|b_i\|^2$, so it depends on the units of the loadings. Sharpe ratios are annualized, gross, at unit gross exposure. Panels B and C: the unpenalized portfolio minus the reference-penalty portfolio of the same fit, with pointwise 95% intervals from the 12-month moving-block bootstrap; mean differences in percentage points a year.

Residual diagonal	Euclidean units	Gram units
Panel A: Sharpe ratio of the unpenalized portfolio		
Fitted diagonal D	1.57	1.57
Alternative diagonal D'	1.52	1.46
Panel B: Unpenalized minus reference penalty, Sharpe ratio		
Fitted diagonal D	0.93 [0.62, 1.22]	0.66 [0.39, 0.88]
Alternative diagonal D'	0.88 [0.57, 1.16]	0.55 [0.26, 0.80]
Panel C: Unpenalized minus reference penalty, annualized mean (pp)		
Fitted diagonal D	2.69 [1.50, 4.05]	1.93 [0.72, 3.20]
Alternative diagonal D'	2.49 [1.38, 3.75]	1.55 [0.43, 2.66]

IA.E. What Do the Full Penalty Grid, Other Block Lengths, and the 618-Month Sample Show?

The main text uses the 594 formation months from June 1975 through November 2024. This section reports the penalty grid and the separately fitted benchmarks on the 618 formation months from June 1973, together with the bootstrap, other block lengths, the three estimates of the Gram metric, and the Ledoit–Wolf shrinkage behind Section IV.

IA.E.1. *The paired bootstrap*

Every interval in the paper resamples common return months in moving blocks, following the dependent-data bootstrap of Künsch (1989): block starts are drawn uniformly, $\lceil n/b \rceil$ blocks of b months are concatenated and truncated to the evaluation length n , and the same indices are applied to both portfolios of a comparison, with gross returns and turnover carried jointly when net returns are recomputed. Each comparison uses 10,000 draws and seed 20260714, reset for each contrast and block length. A Sharpe comparison recomputes each portfolio’s Sharpe ratio in every draw before taking the difference,

$$\Delta \text{SR}^{(b)} = \sqrt{12} \left\{ \frac{\bar{r}_1^{u,(b)}}{s_1^{u,(b)}} - \frac{\bar{r}_0^{u,(b)}}{s_0^{u,(b)}} \right\}, \quad (\text{IA24})$$

which is not the Sharpe ratio of the return difference. Intervals are pointwise 95% percentile intervals. Recentered two-sided mean tests resample $d_{t+1} - \bar{d}$ with seed 20260713 and report $(1 + N_{\text{extreme}})/(10,000 + 1)$; where a Holm adjustment is reported, it covers the four capitalization groups of one panel. The bootstrap holds the fitted models fixed and does not re-estimate them, so it compares portfolios built from given fits. Blocks of 12 months are the primary length; 24- and 60-month blocks retain more within-block dependence with fewer blocks, and Table IA.E.5 reports all three for the main comparisons. Each statistic is computed from unrounded monthly data and rounded once, so a printed difference can differ in the last digit from the difference of two printed

levels; the Gram (rolling-60) minus Euclidean difference on 618 months, for example, is 0.368, printed 0.37, whereas $1.00 - 0.64 = 0.36$.

IA.E.2. How large is the reference penalty?

Across the 618 windows the median eigenvalues of the fitted 8×8 factor covariances, without the penalty, are 8.9×10^{-5} (smallest), 8.8×10^{-4} (median), and 9.5×10^{-3} (largest) in Euclidean units, and 8.7×10^{-5} , 3.8×10^{-4} , and 3.1×10^{-3} in Gram units. The reference value $\lambda = 0.01$ is therefore about eleven times the median factor variance in Euclidean units and twenty-seven times in Gram units, and of the same order as the largest eigenvalue, which it exceeds in about half of the windows in Euclidean units and in nine of ten in Gram units; the seven smaller eigenvalues are below 0.01 in every window. Eigenvalue by eigenvalue, the median ratio of the penalty to the factor variance runs from 1 to 115 in Euclidean units and from 3 to 115 in Gram units. The smallest eigenvalue of any fitted factor covariance is 3.1×10^{-5} , so the eight-factor covariance estimated from 120 months is positive definite without the penalty. Written in Gram units, the Euclidean reference penalty $0.01I$ becomes $0.01JJ'$, whose median eigenvalues along the three latent directions are 0.067, 0.136, and 0.267 times 0.01: the same number is a different penalty in the two sets of units.

IA.E.3. The penalty grid on 618 months

Table IA.E.1 replaces the reference penalty by λI for the six values of the grid, in Euclidean and in Gram units, on the 618-month sample. Both curves fall monotonically from 1.57 without the penalty, to 0.28 in Euclidean units and to 0.85 in Gram units at $\lambda = 0.1$, and every step away from the reference value has an interval that excludes zero. The Gram minus Euclidean difference is 0.00 without the penalty, -0.07 at 0.001, 0.27 at the reference value, with interval $[-0.02, 0.58]$, and 0.57 at 0.1. Figure IA.E.1 plots the two curves and their difference with pointwise bands. Table IA.E.2 shows that on these months the Gram fit with the Euclidean penalty written in Gram units earns exactly the Euclidean fit's 0.64, and the Euclidean fit with the Gram penalty written in

Euclidean units earns exactly the Gram fit's 0.91; without the penalty the two fits are one portfolio, earning 1.57, and removing the penalty adds 0.93 and 0.66 to their Sharpe ratios and 2.69 and 1.93 percentage points a year to their means.

Table IA.E.1. Gross performance along the factor penalty, 618 months. The sample is the 618 formation months from June 1973 through November 2024; the main text uses the 594 months from June 1975. Each row replaces the reference penalty by λI in the indicated loading units and leaves the fitted forecast, residual variances, and position penalty unchanged; all six values of the grid are reported. Differences are relative to $\lambda = 0.01$ in the same units, with pointwise 95% intervals from the 12-month moving-block bootstrap. Means and volatilities are annualized and in percent; drawdowns are compounded; returns are gross at unit gross exposure.

Factor penalty	Mean	Volatility	Sharpe	Drawdown	Sharpe diff.	95% interval
<i>Euclidean units</i>						
$\lambda = 0$	6.57	4.20	1.57	-8.4	0.93	[0.62, 1.22]
$\lambda = 0.001$	5.73	4.71	1.22	-11.0	0.58	[0.38, 0.83]
$\lambda = 0.003$	4.96	5.16	0.96	-18.8	0.33	[0.21, 0.47]
$\lambda = 0.01$	3.88	6.11	0.64	-34.6	reference	
$\lambda = 0.03$	2.96	7.14	0.42	-49.4	-0.22	[-0.33, -0.14]
$\lambda = 0.1$	2.20	7.88	0.28	-62.9	-0.36	[-0.56, -0.22]
<i>Gram units</i>						
$\lambda = 0$	6.57	4.20	1.57	-8.4	0.66	[0.39, 0.88]
$\lambda = 0.001$	5.45	4.74	1.15	-13.8	0.24	[0.16, 0.34]
$\lambda = 0.003$	4.97	4.96	1.00	-16.6	0.09	[0.06, 0.14]
$\lambda = 0.01$	4.64	5.11	0.91	-18.0	reference	
$\lambda = 0.03$	4.48	5.16	0.87	-18.5	-0.04	[-0.07, -0.02]
$\lambda = 0.1$	4.37	5.16	0.85	-18.5	-0.06	[-0.10, -0.03]
<i>Gram units minus Euclidean units at the same λ</i>						
$\lambda = 0$					0.00	[0.00, 0.00]
$\lambda = 0.001$					-0.07	[-0.24, 0.11]
$\lambda = 0.003$					0.04	[-0.18, 0.26]
$\lambda = 0.01$					0.27	[-0.02, 0.58]
$\lambda = 0.03$					0.45	[0.13, 0.79]
$\lambda = 0.1$					0.57	[0.23, 0.92]

The same constant decides comparisons across model families. Table IA.E.3 re-solves the fitted FF5 and RP-PCA portfolios across the grid on 618 months. RP-PCA minus IPCA in Euclidean units is 0.00 without the penalty, with interval [-0.37, 0.38], and rises with the penalty to 0.85 at the reference value, with interval [0.48, 1.15], and 1.05 at 0.1; at 0.001 its interval excludes zero with 12- and 24-month blocks but not with 60-month blocks. On these months RP-PCA at

Table IA.E.2. One penalty, one portfolio: the 618-month sample. 618 formation months, June 1973–November 2024; returns gross of trading costs at unit gross exposure; means and volatilities annualized, in percent. Every row uses the forecasts, residual variances, and position penalty of the indicated rolling fit of IPCA and changes only the factor-risk input; the reference penalty is $\lambda = 0.01$. J is the map of Proposition 1 between the Euclidean and the Gram loading units. Panel B: one minus the cosine and the L_1 distance between unit-gross weights (maximum two), month by month. Panel C: paired differences with pointwise 95% intervals from a 12-month moving-block bootstrap that holds the fitted models fixed.

Panel A: Gross performance				
Portfolio	Mean	Volatility	Sharpe	Drawdown
Euclidean fit, λI in Euclidean units (reference)	3.88	6.11	0.64	−34.6
Gram fit, λI in Gram units (reference)	4.64	5.11	0.91	−18.0
Gram fit, Euclidean-unit penalty $\lambda JJ'$	3.88	6.11	0.64	−34.6
Euclidean fit, Gram-unit penalty $\lambda J^{-1} J^{-1'}$	4.64	5.11	0.91	−18.0
Euclidean fit, unpenalized	6.57	4.20	1.57	−8.4
Gram fit, unpenalized	6.57	4.20	1.57	−8.4
Panel B: Distance between stock weights				
Comparison	One minus cosine		Unit-gross L_1 distance	
	Median	Maximum	Median	Maximum
Reference penalty, Gram vs. Euclidean fit	0.17	0.55	0.57	1.04
Gram fit with Euclidean-unit penalty vs. Euclidean fit	9.4×10^{-12}	1.4×10^{-9}	4.3×10^{-6}	5.2×10^{-5}
Euclidean fit with Gram-unit penalty vs. Gram fit	4.6×10^{-12}	2.6×10^{-9}	3.0×10^{-6}	6.9×10^{-5}
Unpenalized, Gram vs. Euclidean fit	5.8×10^{-12}	4.6×10^{-9}	3.4×10^{-6}	9.4×10^{-5}
Panel C: Paired differences				
Comparison	Sharpe	95% interval	Mean (pp)	95% interval
Gram minus Euclidean fit, reference penalty	0.27	[−0.02, 0.58]	0.76	[−0.45, 2.12]
Gram fit with Euclidean-unit penalty minus Euclidean fit	0.00	[0.00, 0.00]	0.00	[0.00, 0.00]
Unpenalized minus reference, Euclidean fit	0.93	[0.62, 1.22]	2.69	[1.50, 4.05]
Unpenalized minus reference, Gram fit	0.66	[0.39, 0.88]	1.93	[0.72, 3.20]

$\lambda = 0.001$ earns 1.62, above the 1.57 of the unpenalized IPCA portfolio; on the 594 months of the main text the order is reversed, 1.61 against 1.62. FF5 earns 0.98 without the penalty and 0.68 at the reference value, a difference of 0.30 with interval [0.05, 0.56]; RP-PCA moves by 0.08, with interval [−0.15, 0.36]. Figure IA.E.2 plots the families with bands. Table IA.E.4 reports the five-factor exposures of the zero-intercept fits on these months.

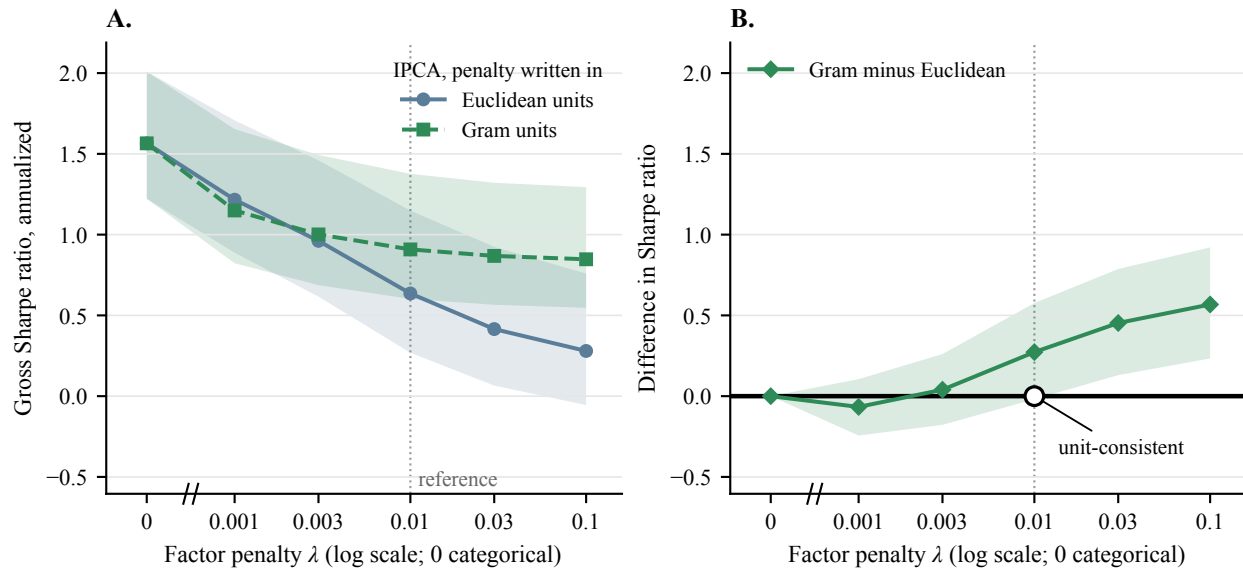


Figure IA.E.1. The factor penalty and the Sharpe ratio, 618 months. Annualized gross Sharpe ratio at unit gross exposure of IPCA when the factor penalty λI is written in Euclidean or in Gram units, for $\lambda \in \{0, 0.001, 0.003, 0.01, 0.03, 0.1\}$; forecasts, residual variances, and the position penalty are those of the rolling fits. Panel A: the two levels. Panel B: Gram minus Euclidean; the open marker on the zero line at $\lambda = 0.01$ is the Gram fit with the Euclidean penalty written in Gram units. Bands are pointwise 95% intervals from a 12-month moving-block bootstrap (10,000 draws) holding the fitted models fixed. The leftmost tick, $\lambda = 0$, is categorical on the log axis, and the dotted line marks the reference penalty. 618 months, June 1973–November 2024; Figure 1 shows the 594-month version.

Table IA.E.3. Separately fitted portfolios along the factor penalty, 618 months. Gross Sharpe ratios at unit gross exposure over the 618 formation months from June 1973 through November 2024; the main text uses the 594 months from June 1975. FF5 and RP-PCA are the fitted benchmarks of Section IA.C, re-solved with λI in their own factor units and with forecasts, residual variances, and the position penalty unchanged; the IPCA columns are those of Table IA.E.1. The last two columns report RP-PCA minus IPCA at the same λ , with pointwise 95% intervals from the 12-month moving-block bootstrap. $\lambda = 0.01$ is the reference penalty of every portfolio.

Factor penalty	FF5	RP-PCA	IPCA		RP-PCA minus IPCA	
			Euclidean	Gram	Euclidean units	Gram units
$\lambda = 0$	0.98	1.57	1.57	1.57	0.00 [−0.37, 0.38]	0.00 [−0.37, 0.38]
$\lambda = 0.001$	0.79	1.62	1.22	1.15	0.41 [0.01, 0.76]	0.47 [0.05, 0.86]
$\lambda = 0.003$	0.73	1.58	0.96	1.00	0.62 [0.26, 0.92]	0.58 [0.16, 0.95]
$\lambda = 0.01$	0.68	1.48	0.64	0.91	0.85 [0.48, 1.15]	0.58 [0.15, 0.94]
$\lambda = 0.03$	0.66	1.40	0.42	0.87	0.98 [0.58, 1.32]	0.53 [0.08, 0.90]
$\lambda = 0.1$	0.65	1.33	0.28	0.85	1.05 [0.64, 1.42]	0.49 [0.02, 0.87]

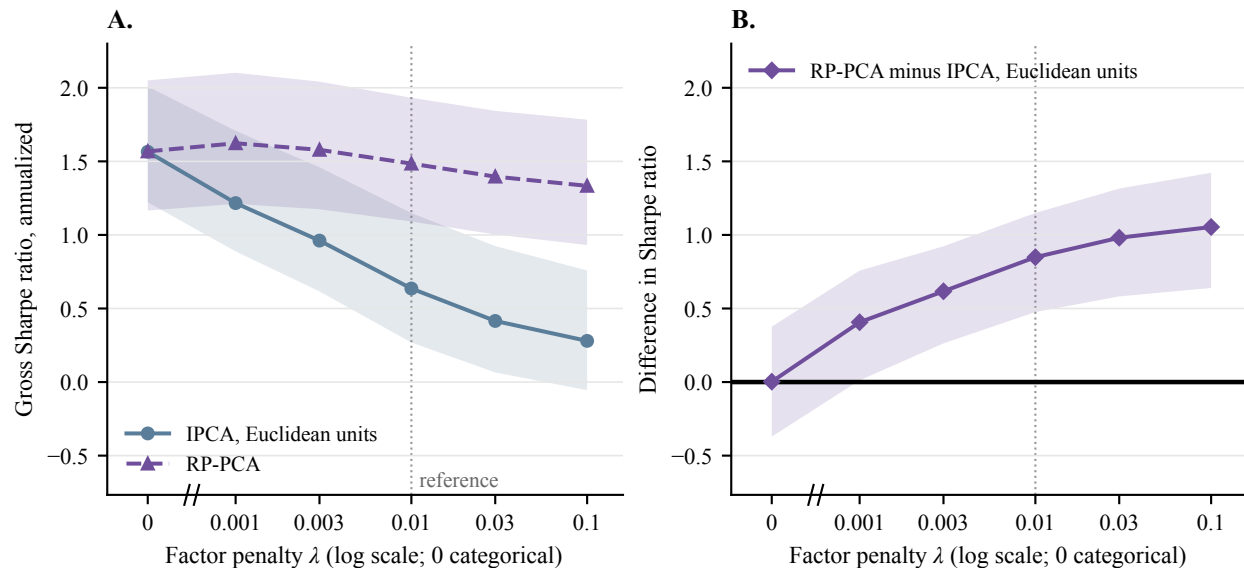


Figure IA.E.2. Two model families along the penalty, 618 months. Annualized gross Sharpe ratio at unit gross exposure of RP-PCA and IPCA in Euclidean units when λI replaces the reference penalty in each fit's own units, for $\lambda \in \{0, 0.001, 0.003, 0.01, 0.03, 0.1\}$; forecasts, residual variances, and the position penalty are those of the rolling fits. Panel A: levels. Panel B: RP-PCA minus IPCA in Euclidean units. Bands are pointwise 95% intervals from a 12-month moving-block bootstrap (10,000 draws) holding the fitted models fixed. The leftmost tick, $\lambda = 0$, is categorical on the log axis, and the dotted line marks the reference penalty. 618 months, June 1973–November 2024; Figure 3 shows the 594-month version.

Table IA.E.4. Exposures of the zero-intercept fits to the five Fama–French factors, 618 months. Each row regresses a portfolio's monthly gross return at unit gross exposure on the five factors of Fama and French (2015), dated to the return month, over the 618 formation months from June 1973 through November 2024; the main text uses 594 months (Table V and Table IA.I.5). Alpha is annualized and in percent, with its Newey–West t -statistic (six lags, Newey and West, 1987) beneath; t -statistics of the factor loadings are not reported for this sample. The regressions describe exposure and are not tests of a pricing model.

Portfolio	Alpha	MKT	SMB	HML	RMW	CMA	R^2
IPCA, Euclidean units, $\lambda = 0.01$	3.27 (3.00)	0.07	0.00	−0.05	0.01	0.06	0.04
IPCA, Gram units, $\lambda = 0.01$	3.82 (4.82)	0.08	0.02	0.03	−0.03	0.03	0.08
Best ZP (Euclidean = Gram), $\lambda = 0$	5.13 (7.06)	0.07	0.06	−0.04	0.12	0.13	0.11

IA.E.4. Block lengths

Table IA.E.5 repeats the main Sharpe-ratio differences with 12-, 24-, and 60-month blocks. The gains from removing the penalty and every intercept-portfolio contrast exclude zero at all three

lengths. The Gram minus Euclidean differences at the reference penalty on 618 months include zero at every length for the EWMA and equal-120 estimates; the rolling-60 estimate's interval excludes zero with 12-month blocks and includes it with 24- and 60-month blocks. The stabilized intercept portfolio exceeds the intercept portfolio by 0.022 at every length, and the corresponding monthly mean intervals include zero at every length.

Table IA.E.5. Block-length sensitivity of the Sharpe-ratio differences. Pointwise 95% percentile intervals from the paired moving-block bootstrap with 12-, 24-, and 60-month blocks, 10,000 draws each, and seed 20260714; the two portfolios of each row are resampled on the same blocks and the fitted models are held fixed. Panel A uses the 618 formation months from June 1973 and Panel B the 594 from June 1975, both through November 2024. Returns are gross at unit gross exposure; penalized portfolios use the reference penalty $\lambda = 0.01$, and the factor-mean portfolio is that of the same fit as the intercept portfolio (IP). The last row is printed to three decimals because the difference is at the second decimal.

Comparison	Difference	95% interval by block length		
		12 months	24 months	60 months
Panel A: 618 months				
Gram (EWMA) minus Euclidean	0.27	[−0.02, 0.58]	[−0.05, 0.59]	[−0.03, 0.64]
Gram (equal-120) minus Euclidean	0.21	[−0.06, 0.51]	[−0.09, 0.53]	[−0.06, 0.57]
Gram (rolling-60) minus Euclidean	0.37	[0.02, 0.68]	[−0.01, 0.69]	[−0.01, 0.76]
Unpenalized minus reference, Euclidean	0.93	[0.62, 1.22]	[0.61, 1.23]	[0.58, 1.31]
Unpenalized minus reference, Gram	0.66	[0.39, 0.88]	[0.42, 0.87]	[0.39, 0.86]
Panel B: 594 months				
IP minus factor-mean, $\lambda = 0.01$	2.12	[1.64, 2.68]	[1.58, 2.84]	[1.42, 3.18]
IP minus factor-mean, $\lambda = 0$	1.95	[1.49, 2.51]	[1.41, 2.65]	[1.23, 2.98]
IP ($\lambda = 0.01$) minus best ZP	0.89	[0.59, 1.26]	[0.60, 1.30]	[0.53, 1.41]
IP, stabilized, minus RP-PCA	1.05	[0.73, 1.41]	[0.77, 1.40]	[0.83, 1.50]
IP, stabilized, minus Gram fit	1.47	[1.11, 1.88]	[1.09, 1.97]	[1.03, 2.15]
IP, stabilized, minus IP	0.022	[0.010, 0.035]	[0.011, 0.036]	[0.012, 0.035]

IA.E.5. Three estimates of the Gram metric

Number the 120 training months from oldest to newest by $j = 0, \dots, 119$ and write $W_{t,j}^{\text{imp}}$ for the regularized second-moment matrix of month j . The EWMA estimate is

$$\widehat{Q}_t^{\text{EWMA}} = \sum_{j=0}^{119} \frac{d^{119-j}}{\sum_{k=0}^{119} d^k} W_{t,j}^{\text{imp}}, \quad d = 2^{-1/60}; \quad (\text{IA25})$$

the equal-120 estimate averages all 120 matrices equally and the rolling-60 estimate the final 60 ($j = 60, \dots, 119$). All three weight dates, not stock counts, and pass the 132-dimensional feature block to the estimator. The main text uses the EWMA estimate. Table IA.E.6 reports the three side by side at the reference penalty. On 618 months they earn 0.91, 0.85, and 1.00 against 0.64 for the Euclidean fit; on 594 months they earn 1.07, 1.03, and 1.13 against 0.69, with Gram minus Euclidean differences of 0.38, 0.34, and 0.44 whose intervals exclude zero and mean differences whose intervals include zero. Over three 206-month periods of the 618-month sample, the rolling-60 estimate earns 0.93 against 0.69 for the Euclidean fit from June 1973 to July 1990, 1.88 against 1.75 from August 1990 to September 2007, and 0.41 against -0.14 from October 2007 to November 2024. Each comparison is made at the reference penalty written in each fit's own units; for the EWMA estimate, Section IV.B shows that the whole difference is that penalty written in two sets of units.

IA.E.6. Ledoit–Wolf shrinkage in both sets of units

For each fit and window the estimator of Ledoit and Wolf (2004) shrinks the sample covariance S of the 120-month factor history toward mI , with $m = \text{tr}(S)/p$, intensity $\min(\bar{b}^2, d^2)/d^2$, $d^2 = \|S - mI\|_F^2/p$, and $\bar{b}^2 = n^{-2} \sum_k \|x_k x_k' - S\|_F^2/p$, where x_k are the demeaned factor observations and $p = 8$. The sample covariance reproduces the fitted factor covariance less the reference penalty to a relative precision of 10^{-15} . On 618 months the median intensity is 0.17 in Euclidean units and 0.13 in Gram units, with median targets m of 2.1×10^{-3} and 7.9×10^{-4} . Because the target mI is written in the units of the factors, the two shrunk portfolios differ: their median unit-gross L_1 distance is 0.30 on both samples (Table IA.E.7). Their Sharpe ratios are 1.06 and 1.28 on 618 months, a difference of 0.22 with interval $[-0.04, 0.55]$ ($[0.01, 0.65]$ with 60-month blocks), and 1.17 and 1.54 on 594 months, a difference of 0.36 with interval $[0.14, 0.64]$. On both samples both shrunk portfolios earn more than the reference penalty and less than the unpenalized portfolio, except that in Gram units on 594 months the shortfall from the unpenalized portfolio, -0.08 , has

Table IA.E.6. Three estimates of the Gram metric against the Euclidean fit. IPCA at the reference penalty $\lambda = 0.01$, with the Gram metric estimated as an exponentially weighted average of the window's characteristic second-moment matrices (EWMA, 60-month half-life), their equal-weighted average over all 120 months (equal-120), or over the final 60 months (rolling-60). Panels A and B use the 618 formation months from June 1973 and Panel C the 594 from June 1975, both through November 2024; Table IA.C.1 gives the 594-month levels. Returns are gross at unit gross exposure; means, volatilities, and drawdowns are annualized or compounded, in percent. Differences are Gram minus Euclidean, with pointwise 95% intervals from the 12-month moving-block bootstrap beneath; p is the p -value of a recentered two-sided mean test (seed 20260713).

		IPCA, Gram units		
	Euclidean	EWMA	Equal-120	Rolling-60
Panel A: Levels, 618 months				
Sharpe ratio	0.64	0.91	0.85	1.00
Annualized mean (%)	3.9	4.6	4.5	4.8
Annualized volatility (%)	6.1	5.1	5.3	4.8
Maximum drawdown (%)	−34.6	−18.0	−19.8	−16.5
Panel B: Gram minus Euclidean, 618 months				
Sharpe ratio		0.27	0.21	0.37
		[−0.02, 0.58]	[−0.06, 0.51]	[0.02, 0.68]
Annualized mean (percentage points)		0.76	0.64	0.95
		[−0.45, 2.12]	[−0.54, 1.96]	[−0.38, 2.36]
Annualized volatility (percentage points)		−1.00	−0.76	−1.29
		[−2.32, 0.18]	[−2.06, 0.37]	[−2.62, 0.04]
Recentered mean test, p		0.25	0.32	0.17
Panel C: Gram minus Euclidean, 594 months				
Sharpe ratio		0.38	0.34	0.44
		[0.09, 0.64]	[0.08, 0.57]	[0.11, 0.75]
Mean (basis points a month)		6.19	5.57	7.18
		[−4.12, 17.68]	[−4.41, 16.48]	[−3.82, 19.76]

an interval that includes zero. With eight factors and 120 months, the factor covariance is positive definite without any shrinkage.

Table IA.E.7. Ledoit–Wolf shrinkage of the factor covariance in two sets of units. Each fit’s 120-month factor covariance S is shrunk toward mI , $m = \text{tr}(S)/8$, with the intensity of Ledoit and Wolf (2004), estimated separately in Euclidean and in Gram units; forecasts, residual variances, and the position penalty are those of the fits. The two columns use the 594 formation months from June 1975 and the 618 from June 1973, both through November 2024. Sharpe ratios are annualized, gross, at unit gross exposure. Panel C: paired differences with pointwise 95% intervals from the 12-month moving-block bootstrap; “reference” is $\lambda = 0.01$ and “unpenalized” $\lambda = 0$ in the same units.

	594 months	618 months
Panel A: Shrinkage and holdings (medians across formation months)		
Shrinkage intensity, Euclidean units	0.18	0.17
Shrinkage intensity, Gram units	0.14	0.13
Unit-gross L_1 distance between the two shrunk portfolios	0.30	0.30
One minus cosine between the two shrunk portfolios	0.045	0.046
Panel B: Sharpe ratio		
Shrunk covariance, Euclidean units	1.17	1.06
Shrunk covariance, Gram units	1.54	1.28
Panel C: Paired differences in Sharpe ratio		
Gram units minus Euclidean units	0.36 [0.14, 0.64]	0.22 [−0.04, 0.55]
Shrunk minus reference, Euclidean units	0.49 [0.25, 0.73]	0.42 [0.22, 0.69]
Shrunk minus reference, Gram units	0.47 [0.30, 0.66]	0.37 [0.22, 0.60]
Shrunk minus unpenalized, Euclidean units	−0.45 [−0.73, −0.19]	−0.51 [−0.73, −0.23]
Shrunk minus unpenalized, Gram units	−0.08 [−0.17, 0.02]	−0.29 [−0.50, −0.01]

IA.F. Is the Intercept Portfolio a Property of the Normalization?

Section V.D states that the intercept portfolio is the same under every normalization of the fit without a penalty and within 0.01 in Sharpe ratio at every penalty. This section reports the computation behind that statement (Table IA.F.1).

At every formation date the rolling fit of IPCA with an intercept, estimated under the Gram normalization, is re-expressed in Euclidean loading units by the matched transformation of Section IA.A.1 with $Q = I_L$: the latent loadings become $\Gamma_\beta C^{-1/2}$ with $C = \Gamma'_\beta \Gamma_\beta$, the observed-factor and intercept blocks are re-orthogonalized, and the latent factors, their forecast means, and their covariance follow the same map. The complete fitted mean and every fitted return are unchanged, with a relative difference below 10^{-8} at every date. Residual variances, the position penalty, and

the forecasts are those of the fit, and λI is added to the factor covariance in the units of each representation. A random orthogonal rotation of the re-expressed latent block changes the intercept portfolio's weights by at most 10^{-11} , so the re-expressed portfolio does not depend on a choice of rotation.

Without a penalty the two representations deliver the same intercept portfolio month by month, as Section IA.A.3 requires: the largest monthly return difference is 1.7×10^{-14} , and the median unit-gross L_1 distance between the two portfolios is 1.9×10^{-15} . With the penalty the two differ by -0.004 to -0.009 in Sharpe ratio across the grid, below 0.01 at every value; the intervals at $\lambda \geq 0.003$ exclude zero, and the distance between the portfolios at the reference penalty is 0.0027 . On the 618-month sample the intercept portfolio earns 2.458 – 2.470 in Gram units and 2.458 – 2.466 in Euclidean units across the grid.

The factor-mean portfolio of the same fit depends on the normalization at every value of λ , including zero, because the split of the fitted mean between the factor block and the intercept does. In Euclidean units the projection of the intercept loadings on the latent loading span, a median 29% of the intercept's stock-level norm (5th to 95th percentile, 13% to 74%), is reassigned to the factor means, and the factor-mean forecast changes by a median 49% in relative L_2 norm. The factor-mean portfolio then earns 0.44 rather than 0.56 without the penalty and -0.06 rather than 0.39 at the reference penalty, and its unit-gross L_1 distance from its Gram-unit version is 0.24 and 0.54 . On 618 months it falls from 0.53 to 0.29 across the grid in Gram units and from 0.39 to -0.24 in Euclidean units. The difference between the intercept portfolio and the factor-mean portfolio is correspondingly larger in Euclidean units, 2.57 rather than 2.12 at the reference penalty and 2.06 rather than 1.95 without it. The Gram normalization therefore favors the factor-mean portfolio rather than the intercept portfolio, and the normalization-free benchmark for the intercept portfolio is the unpenalized zero-intercept portfolio, which is one portfolio under either normalization.

The residual diagonal is held fixed in both representations; recomputing the factor-only diagonal in Euclidean units would move it by a median 0.3% and at most 1.2% . The exercise re-expresses

one fit; a separate estimation of the model in Euclidean units with a free intercept could reach a different fitted index, and the exposure budget, defined in Gram units, is slack at every date.

Table IA.F.1. The intercept portfolio and the factor-mean portfolio under two normalizations of one fit. 594 formation months, June 1975–November 2024; gross returns at unit gross exposure; means annualized, in percent. At every formation date the rolling fit of IPCA with an intercept, estimated under the Gram normalization, is re-expressed in Euclidean units by the map of Proposition 1 with $Q = I_L$, which leaves every fitted return unchanged; residual variances, the position penalty, and the forecasts are those of the fit, and λI is added to the factor covariance in each representation’s own units. Differences are Euclidean minus Gram units, with pointwise 95% intervals from a 12-month moving-block bootstrap; Panel A has three decimals because the differences are at the third decimal.

Factor penalty	Mean		Sharpe		Sharpe diff.	95% interval
	Gram	Euclidean	Gram	Euclidean		
Panel A: Intercept portfolio (factor means and fitted intercept)						
$\lambda = 0$	7.79	7.79	2.501	2.501	0.000	[0.000, 0.000]
$\lambda = 0.001$	7.63	7.63	2.507	2.503	−0.004	[−0.010, 0.002]
$\lambda = 0.003$	7.60	7.60	2.510	2.503	−0.008	[−0.015, −0.001]
$\lambda = 0.01$ (reference)	7.59	7.58	2.513	2.504	−0.009	[−0.017, −0.002]
$\lambda = 0.03$	7.59	7.58	2.514	2.506	−0.007	[−0.014, −0.002]
$\lambda = 0.1$	7.58	7.58	2.514	2.509	−0.005	[−0.009, −0.001]
Panel B: Factor-mean portfolio (same fit)						
$\lambda = 0$	2.24	1.71	0.56	0.44	−0.12	[−0.28, 0.02]
$\lambda = 0.001$	1.95	1.07	0.45	0.24	−0.22	[−0.38, −0.06]
$\lambda = 0.003$	1.83	0.49	0.42	0.10	−0.32	[−0.52, −0.13]
$\lambda = 0.01$ (reference)	1.75	−0.39	0.39	−0.06	−0.46	[−0.69, −0.22]
$\lambda = 0.03$	1.72	−1.12	0.38	−0.16	−0.54	[−0.80, −0.28]
$\lambda = 0.1$	1.67	−1.70	0.37	−0.22	−0.59	[−0.88, −0.32]

IA.G. What Does Stabilizing the Intercept Portfolio Change?

Section V.A uses the plain intercept portfolio. A stabilized version penalizes positions whose intercept contribution varies across past loading estimates. This section defines it and shows that it changes the Sharpe ratio by 0.02.

IA.G.1. Definition

For m_t prior intercept-loading estimates, the dispersion input is

$$V_{\Gamma_{\alpha,t}} = \frac{1}{m_t - 1} \sum_{j=1}^{m_t} (\widehat{\Gamma}_{\alpha,t-j} - \bar{\Gamma}_{\alpha,t})(\widehat{\Gamma}_{\alpha,t-j} - \bar{\Gamma}_{\alpha,t})', \quad (\text{IA26})$$

where $\bar{\Gamma}_{\alpha,t}$ is their average. The rule uses the 24 most recent estimates, excluding the current one, so $m_t = 24$ throughout the 594-month sample. The stabilized intercept portfolio keeps the mean of the intercept portfolio and adds a term to its risk matrix:

$$w_t^{\text{stab}} = (B_t \Sigma_{F,t} B_t' + \widehat{D}_t + \kappa_U Z_t V_{\Gamma_{\alpha,t}} Z_t' + \lambda_{\Sigma} I_{N_t})^{-1} (B_t \widehat{m}_{F,t} + Z_t \widehat{\Gamma}_{\alpha,t}), \quad \kappa_U = 1, \quad (\text{IA27})$$

with $\Sigma_{F,t} = \widehat{C}_{F,t} + \lambda I$. For a position v the added term charges $(\kappa_U/2)(Z_t' v)' V_{\Gamma_{\alpha,t}} (Z_t' v)$, which is $\kappa_U/2$ times the sample variance of the position's intercept contribution across the 24 past estimates. The matrix has rank at most 23 in 132 dimensions, adjacent estimates share 119 of their 120 training months, and the past coefficients are expressed in their own windows' standardizations, so their dispersion combines coefficient movement with changes in preprocessing. It is a loading-dispersion term rather than a calibrated forecast-error covariance, and κ_U is set, not optimized. The requirement of 24 prior estimates defines the start of the 594-month sample (Section IA.B.5).

IA.G.2. What it changes

Table IA.G.1 decomposes the gain over the factor-mean portfolio of the same fit at the reference penalty. Adding the fitted intercept to the mean raises the Sharpe ratio from 0.3917 to 2.5126; adding the stabilizing term raises it further to 2.5347, a difference of 0.0222 with interval $[0.0100, 0.0355]$, and leaves the mean unchanged: the monthly mean difference is 0.08 basis points, with interval $[-0.15, 0.26]$. The gain of the stabilized portfolio over the plain one is a small reduction in volatility, from 3.02% to 3.00%, with the same mean. Its five-factor alpha is 7.14%

a year ($t = 12.05$) against 7.15% ($t = 12.01$) for the intercept portfolio (Table IA.G.2). Over the three 198-month thirds of the sample it earns 3.97, 3.03, and 1.13, against 3.95, 3.01, and 1.11 for the intercept portfolio. Its turnover is 0.94 per month (purchases plus sales), its net Sharpe ratios are 2.17, 1.61, and 0.67 at 10, 25, and 50 basis points per dollar traded, and its break-even cost is 68 basis points, as for the intercept portfolio (Table IA.G.3).

On the 618-month sample, Table IA.G.4 varies the weight κ_U with the fit and the mean held fixed: the Sharpe ratio rises from 2.4686 without the term to 2.4926 at $\kappa_U = 1$ and 2.5096 at $\kappa_U = 2$. A full-rank alternative, built from window residuals with equal date weights, gives 2.4758 at $\kappa_U = 1$; its residuals omit the fitted intercept and it does not account for the restriction $\Gamma'_\beta Q \Gamma_\alpha = 0$, which leaves $L - K$ admissible intercept directions, so it is a sensitivity calculation rather than the covariance of the constrained estimator. Across nine conventions for the loading history, holding the mean and the factor risk fixed, the Sharpe ratio at $\kappa_U = 1$ ranges from 2.4895 to 2.4938 on 618 months; two of the nine are equivalent EWMA specifications. The stabilizing term is a modest change to the risk input without a separate mean effect, and the main text reports the plain intercept portfolio.

Table IA.G.1. What the fitted intercept and the stabilizing term each add. All entries use one rolling fit of IPCA with an intercept (Gram normalization, EWMA estimate), the reference penalty $\lambda = 0.01$, the 594 formation months from June 1975 through November 2024, and unit gross exposure. The factor-mean portfolio uses the factor means alone; the intercept portfolio adds the fitted intercept to the mean; the stabilized intercept portfolio also adds the loading-dispersion term with $\kappa_U = 1$. Panel B: pointwise 95% intervals from the paired 12-month moving-block bootstrap (10,000 draws, seed 20260714); recentered mean tests use seed 20260713. The columns change inputs of one fit; they are not separately fitted models.

Panel A: Gross performance			
Statistic	Factor-mean (same fit)	IP	IP, stabilized
Annualized mean (%)	1.75	7.59	7.60
Annualized volatility (%)	4.46	3.02	3.00
Sharpe ratio	0.3917	2.5126	2.5347
Maximum drawdown (%)	−17.73	−3.93	−3.87
Panel B: Incremental input effects			
Comparison	Monthly mean (bp)	95% interval	<i>p</i>
IP minus factor-mean portfolio	48.69	[34.66, 63.36]	< 0.001
IP, stabilized, minus IP	0.08	[−0.15, 0.26]	0.437
Comparison	Sharpe difference	95% interval	
IP minus factor-mean portfolio	2.1209	[1.6445, 2.6791]	
IP, stabilized, minus IP	0.0222	[0.0100, 0.0355]	

Table IA.G.2. Five-factor exposures of the intercept portfolio with and without the stabilizing term. Each row regresses a portfolio's monthly gross return at unit gross exposure on the five factors of Fama and French (2015), dated to the return month, over the 594 formation months from June 1975 through November 2024, at the reference penalty $\lambda = 0.01$. Alpha is annualized and in percent, with its Newey–West *t*-statistic (six lags, Newey and West, 1987) beneath; *t*-statistics of the loadings are not reported. Table IA.I.5 gives the same regression for the other portfolios. The regressions describe exposure and are not tests of a pricing model.

Portfolio	Alpha	MKT	SMB	HML	RMW	CMA	R^2
Intercept portfolio (IP)	7.15 (12.01)	0.01	0.03	0.00	0.05	0.02	0.02
IP, stabilized	7.14 (12.05)	0.01	0.03	0.00	0.05	0.02	0.03

Table IA.G.3. Turnover and net Sharpe ratios of the stabilized intercept portfolio. 594 months, June 1975–November 2024; unit gross exposure; factor penalty $\lambda = 0.01$. The stabilized intercept portfolio adds the stabilizing term of Internet Appendix IA.G to the intercept portfolio of Table IV; the intercept portfolio row repeats that table. Turnover counts purchases and sales per dollar of gross exposure (one-way turnover is half of it); net returns subtract a cost per dollar traded; the break-even cost sets the mean net return to zero.

Portfolio	Turnover	Sharpe ratio at cost (bp)				Break-even cost (bp)
		0	10	25	50	
Intercept portfolio (IP)	0.94	2.51	2.15	1.60	0.66	68
Stabilized intercept portfolio	0.94	2.53	2.17	1.61	0.67	68

Table IA.G.4. Sensitivity to the weight on the stabilizing term, 618 months. Annualized gross Sharpe ratios from 618 unit-gross monthly returns, June 1973–November 2024; the main text uses 594 months. The fit, the reference penalty, and the mean input are fixed while κ_U varies. The dispersion input uses up to 24 prior intercept-loading estimates; the alternative is a full-rank residual-based matrix, a sensitivity calculation rather than the covariance of the constrained loading estimator.

Portfolio input	Prior-estimate dispersion	Full-rank alternative
Factor-mean portfolio (same fit), no stabilizing term	0.32	0.32
Intercept portfolio, $\kappa_U = 0$	2.4686	2.4686
Intercept portfolio, $\kappa_U = 0.25$	2.4755	2.4705
Intercept portfolio, $\kappa_U = 0.5$	2.4818	2.4723
Intercept portfolio, $\kappa_U = 1$ (stabilized)	2.4926	2.4758
Intercept portfolio, $\kappa_U = 2$	2.5096	2.4825

The two $\kappa_U = 0$ entries agree by construction because the stabilizing term is absent. These re-solves of one fit carry no sampling intervals.

IA.H. What Do the Unnormalized Solutions and Other Leverage Policies

Show?

Every portfolio in the paper is held at unit gross exposure. This section reports the unnormalized mean–variance solutions behind that policy and the main contrasts under two other leverage policies, supporting Section VI.D.

IA.H.1. The unit-gross policy

For raw weights w_t , define

$$G_t = \sum_{i=1}^{N_t} |w_{i,t}|, \quad w_t^u = \frac{w_t}{G_t}, \quad r_{t+1}^u = (w_t^u)' r_{t+1}. \quad (\text{IA28})$$

Absolute weights sum to one each month, so Sharpe ratios are scale-free and means and volatilities are per dollar of gross position; annualized mean, volatility, and Sharpe ratio are $12\bar{r}^u$, $\sqrt{12} s(r^u)$, and $\sqrt{12} \bar{r}^u / s(r^u)$, and drawdowns compound monthly returns. Because G_t varies over time, unit gross exposure is a leverage policy rather than a reporting convention, and it does not equalize volatility across portfolios. The raw solution $A_t^{-1} \mu_t$ carries no risk-aversion coefficient, so its scale is not a position size.

IA.H.2. The unnormalized solutions

Panel A of Table IA.H.1 reports the mean gross exposure of each raw solution on the 594-month sample: 3.8 for the Euclidean fit and 2.4 for the Gram fit at the reference penalty, 58 for the best zero-intercept portfolio, and about 500 for the intercept portfolios, 507 for the intercept portfolio and 501 for its stabilized version; across the penalty grid the intercept portfolio's mean raw exposure is 507 to 509. At that scale the raw returns are arithmetic, not wealth: 87 of the 594 raw monthly returns of the intercept portfolio, 83 of its stabilized version, and 16 of the best zero-intercept portfolio are below -100% , and their raw Sharpe ratios of 2.45, 2.48, and 1.64 are moment ratios of those series.

No raw return of the Euclidean fit, the Gram fit, or the factor-mean portfolio is below -100% , and no portfolio's unit-gross return is: in no month does any portfolio lose more than its gross position. Normalizing each month's weights to unit gross exposure removes the scale month by month. On the 618-month sample the correlation between a portfolio's gross exposure and its unit-gross return is -0.091 for the Euclidean fit, -0.008 for the Gram fit, and 0.165 for the unpenalized portfolio.

IA.H.3. Other leverage policies

Panel B of Table IA.H.1 replaces the unit-gross policy on the 618-month sample by a constant leverage, which scales each portfolio's raw weights by its full-sample mean gross exposure, and by a leverage equal to the trailing 12-month mean of gross exposure, which uses only past exposures and starts after 12 months. Under constant leverage the Sharpe ratios are 0.46 for the Euclidean fit and 0.90 for the Gram fit at the reference penalty and 1.60 without the penalty, against 0.64 , 0.91 , and 1.57 at unit gross exposure. The gain from removing the penalty is larger under constant leverage, 1.14 in Euclidean units and 0.70 in Gram units, and both gains exclude zero under all three policies; the gain is therefore a property of the penalty rather than of the leverage policy. The Gram minus Euclidean difference at the reference penalty is 0.27 at unit gross exposure, 0.43 under constant leverage, and 0.16 under the trailing rule, so its size depends on the leverage policy as well as on the units of the penalty.

Panel C reports the intercept portfolio's gain over the best zero-intercept portfolio on the 594-month sample under the same three policies: 0.89 at unit gross exposure, 0.81 under constant leverage, and 0.85 under the trailing rule. Each interval excludes zero, at 12-month blocks and also at 24- and 60-month blocks. Unlike the gains from removing the penalty, this gain is slightly smaller under constant leverage than under unit gross exposure; under constant leverage the intercept portfolio's volatility is lower by 1.16 percentage points, with an interval that excludes zero, and its mean is higher by 0.94 points, with an interval that includes zero, the same pattern as at unit gross exposure. The gain is therefore a property of the fitted intercept rather than of the leverage policy.

Constant leverage uses the full-sample mean exposure and cannot be implemented in real time; the trailing rule can.

Table IA.H.1. Unnormalized solutions and the unit-gross policy. Panel A, 594 formation months from June 1975: mean gross exposure $\sum_i |w_{i,t}|$ of the raw mean–variance solution, Sharpe ratios of the raw and the unit-gross series, and the number of months with a raw return below -100% ; gross of costs, reference penalty unless stated; “—”: not computed. Panels B (618 formation months from June 1973; 605 for the trailing rule, which needs 12 prior months) and C (594 months): Sharpe differences under unit gross exposure, a constant leverage equal to the full-sample mean gross exposure, and the trailing 12-month mean, Panel B to three decimals; pointwise 95% intervals, 12-month moving-block bootstrap.

Panel A: Raw and unit-gross portfolios, 594 months				
Portfolio	Mean gross	Raw Sharpe	Raw < -100%	Unit-gross Sharpe
IPCA, Euclidean units	3.79	0.49	0	0.69
IPCA, Gram units (EWMA estimate)	2.37	0.97	0	1.07
IPCA, Gram units (rolling-60 estimate)	2.37	1.00	—	1.13
Best ZP, $\lambda = 0$	58.3	1.64	16	1.62
RP-PCA	0.53	1.53	—	1.49
FF5	1.98	0.77	—	0.75
Factor-mean portfolio (same fit)	2.27	0.38	0	0.39
Intercept portfolio (IP)	507	2.45	87	2.51
IP, stabilized	501	2.48	83	2.53

Panel B: Sharpe differences under three leverage policies, 618 months			
Comparison	Unit gross	Constant leverage	Trailing 12-month
Unpenalized minus reference, Euclidean units	0.929 [0.620, 1.216]	1.138 [0.684, 1.515]	0.877 [0.607, 1.180]
Unpenalized minus reference, Gram units	0.657 [0.389, 0.883]	0.705 [0.419, 0.977]	0.720 [0.406, 0.986]
Gram minus Euclidean, reference penalty	0.273 [−0.019, 0.577]	0.433 [0.047, 0.752]	0.156 [−0.089, 0.488]

Panel C: The intercept gain under three leverage policies, 594 months			
Comparison	Unit gross	Constant leverage	Trailing 12-month
IP minus best ZP	0.89 [0.59, 1.26]	0.81 [0.53, 1.13]	0.85 [0.54, 1.21]

IA.I. How Does Performance Vary over Time and Across Size Groups?

Section VI.C reports the portfolios over three calendar thirds of the 594-month sample (Figure 5). This section adds decades, NBER business-cycle regimes, five-year rolling windows, the

net comparisons in the last third and in the months after October 2007, size groups, alphas on each benchmark, and cumulative returns. All exhibits use the 594 formation months from June 1975 through November 2024 and, unless stated, the seven portfolios of Figure 5 in the same order.

IA.I.1. Thirds and decades

Gross of costs, the intercept portfolio earns 3.95, 3.01, and 1.11 in the three 198-month thirds (June 1975–November 1991, December 1991–May 2008, and June 2008–November 2024), the highest Sharpe ratio in each; the next highest are, in turn, 2.43 for RP-PCA, 2.27 for the best zero-intercept portfolio, and 0.57 for that portfolio and the pooled regression. From the first third to the last the intercept portfolio's Sharpe ratio falls by 72%, against 73% for the best zero-intercept portfolio (2.13 to 0.57) and 86% for RP-PCA (2.43 to 0.35). The gross difference between the intercept portfolio and the best zero-intercept portfolio is positive in every third, 1.82, 0.74, and 0.54, and in the last third its 12-month interval is [0.003, 1.19].

Table IA.I.1 reports Sharpe ratios by decade and by third, the paired difference between the intercept portfolio and the best zero-intercept portfolio in every period, and the percentage change from the first third to the last. Gross of costs, the intercept portfolio has the highest Sharpe ratio of the seven portfolios in every decade except the last, the 59 months from January 2020 through November 2024, in which the FF5 portfolio earns 0.87 against its 0.83; the intercept portfolio's difference from the FF5 portfolio in those months is -0.04 , with interval $[-1.00, 1.30]$. Its gain over the best zero-intercept portfolio is positive in every decade and largest in the first two, 2.03 over the 55 months from June 1975 and 1.64 in the 1980s; the interval excludes zero in four of the six decades and includes it in the 2000s and from 2020 (Panel B). Panel C repeats the thirds: from the first to the last, the Sharpe ratio changes by -71.9% for the intercept portfolio, -73.2% for the best zero-intercept portfolio, and -85.8% for RP-PCA.

Table IA.I.1. Sharpe ratios by decade. Annualized Sharpe ratios of monthly gross returns at unit gross exposure; 594 formation months, June 1975–November 2024; periods by formation month (the first decade starts in June 1975, the last ends in November 2024). All portfolios except the best ZP (IPCA with $\lambda = 0$) use the reference penalty $\lambda = 0.01$. Panel B: IP minus best ZP, with a pointwise 95% interval (paired 12-month moving-block bootstrap, 10,000 draws). Panel C: calendar thirds of 198 months (June 1975–November 1991, December 1991–May 2008, June 2008–November 2024).

	1975–79	1980s	1990s	2000s	2010s	2020–24	Full
<i>Panel A: Gross Sharpe ratio</i>							
Euclidean fit	1.10	0.94	2.00	0.16	0.27	0.33	0.69
Gram fit	1.18	1.57	2.08	0.73	0.46	0.63	1.07
Best ZP ($\lambda = 0$)	1.70	2.37	3.31	1.41	0.72	0.60	1.62
RP-PCA	2.69	2.41	2.94	0.67	0.66	0.45	1.49
FF5	1.14	0.58	1.37	0.74	0.23	0.87	0.75
Pooled regression	1.67	0.96	1.87	1.03	0.82	0.35	0.92
Intercept portfolio (IP)	3.73	4.01	4.68	1.72	1.69	0.83	2.51
Months	55	120	120	120	120	59	594
<i>Panel B: Contrast, IP minus best ZP</i>							
	Sharpe	95% interval					
1975–79	2.03	[1.07, 3.94]					
1980s	1.64	[0.93, 2.44]					
1990s	1.36	[0.39, 1.97]					
2000s	0.30	[−0.05, 0.78]					
2010s	0.97	[0.31, 1.96]					
2020–24	0.24	[−0.81, 1.25]					
Full	0.89	[0.59, 1.26]					
<i>Panel C: Gross Sharpe ratio by third, and change from the first to the last third (percent)</i>							
	First	Second	Last	Change			
Intercept portfolio (IP)	3.95	3.01	1.11	−71.9			
Best ZP ($\lambda = 0$)	2.13	2.27	0.57	−73.2			
RP-PCA	2.43	1.89	0.35	−85.8			

IA.I.2. Net of costs, and the months after October 2007

Net of 25 basis points per dollar traded, the last third differs: the intercept portfolio earns 0.22, with interval $[-0.02, 0.67]$, against 0.44 for the pooled characteristic regression, and its net differences from the factor-mean portfolio, the best zero-intercept portfolio, and the pooled regression in that third are 0.38 $[-0.24, 1.16]$, 0.16 $[-0.41, 0.80]$, and -0.23 $[-0.71, 0.35]$; only its difference from the Euclidean fit, 0.56 $[0.11, 1.04]$, excludes zero. The last-third net means are small: 0.61% a year for the intercept portfolio and 0.20% for the best zero-intercept portfolio, against 6.47% for

the pooled regression with a volatility of 14.7%.

Table IA.I.2 reports the same comparisons over the 206 months from October 2007 through November 2024. Net of 25 basis points the intercept portfolio's advantage is 0.44 $[-0.22, 1.10]$ over the factor-mean portfolio, 0.19 $[-0.41, 0.75]$ over the best zero-intercept portfolio, and -0.19 $[-0.71, 0.32]$ over the pooled regression; gross of costs the three differences are 0.96 $[0.32, 1.60]$, 0.57 $[-0.00, 1.12]$, and 0.56 $[0.07, 1.13]$. The gain from removing the penalty from the Euclidean fit is 0.45 $[0.08, 0.79]$ net over the same months. The two windows give the same signs, and every net interval for the intercept portfolio against these three benchmarks includes zero in both. The table also reports the one-month delay: trading every portfolio a month late lowers the intercept portfolio to 1.52 and the best zero-intercept portfolio to 1.07, and every delayed difference excludes zero.

IA.I.3. Business-cycle regimes and rolling windows

Figure IA.I.1 splits the 594 return months by the NBER chronology into 530 expansion months and 64 recession months. The intercept portfolio has the highest Sharpe ratio of the seven portfolios in both regimes, 2.50 in expansions and 2.69 in recessions, and the best zero-intercept portfolio is second in both, at 1.65 and 1.47. The intercept portfolio's gain over it is 0.85 $[0.49, 1.24]$ in expansions and 1.22 $[0.28, 2.43]$ in recessions, with bootstrap blocks drawn within each regime; drawn on the full calendar, with each month keeping its regime label, the recession intervals are $[1.44, 4.84]$ for the intercept portfolio and $[0.31, 2.73]$ for its gain. Its annualized mean is 7.3% in expansions and 9.7% in recessions, with volatilities of 2.9% and 3.6%. The intercept portfolio is also the highest in both regimes when months are classified by formation month or when a recession is dated from the month after the peak.

Figure IA.I.2 plots Sharpe ratios over the 535 overlapping 60-month windows. The intercept portfolio's Sharpe ratio never falls below 0.45, its value in the window ending in March 2012, whereas the best zero-intercept portfolio's reaches -0.24 (September 2018) and is negative in 20

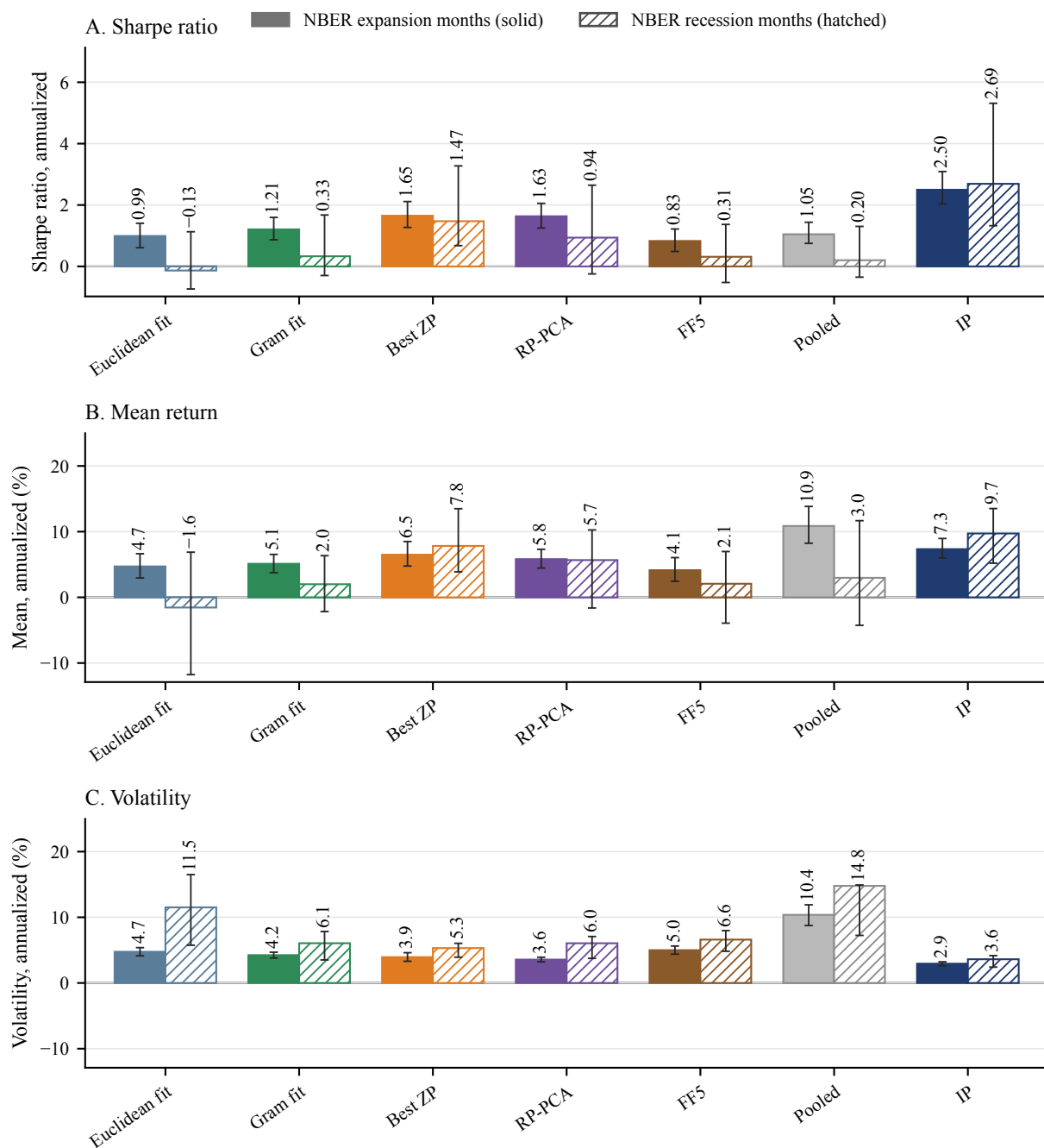


Figure IA.I.1. Sharpe ratio, mean, and volatility by business-cycle regime. Annualized Sharpe ratio (A), mean in percent (B), and volatility in percent (C) of the seven portfolios of Figure 5, gross of costs at unit gross exposure, in expansion months (solid) and recession months (hatched): the peak-to-trough months of the NBER chronology (National Bureau of Economic Research, 2023), endpoints included, by return month (64 of 594). Whiskers are pointwise 95% intervals from a 12-month moving-block bootstrap on the months of each regime in calendar order (10,000 draws), fitted models held fixed. Panels B and C share one y-range. June 1975–November 2024.

Table IA.I.2. The pooled regression, trading one month late, and the months after October 2007. The sample is the 594 formation months from June 1975 through November 2024. The pooled characteristic regression of Section IA.C.3 uses the risk input of the intercept fit. “Delayed” holds the unit-gross weights formed one month earlier through the return month, so every input is at least one month old. Net returns subtract 25 basis points per dollar traded (purchases plus sales). Panel B: paired Sharpe-ratio differences with pointwise 95% intervals from the 12-month moving-block bootstrap; the last column, net of costs, uses the 206 months from October 2007 through November 2024. Means and volatilities are annualized and in percent.

Panel A: Levels					
Portfolio	Mean	Volatility	Sharpe	Net Sharpe, 25 bp	Sharpe, delayed
IPCA, Euclidean units, $\lambda = 0.01$	4.03	5.87	0.69	0.30	0.22
Best ZP, $\lambda = 0$	6.64	4.10	1.62	1.03	1.07
Factor-mean (same fit), $\lambda = 0.01$	1.75	4.46	0.39	−0.04	0.20
IP, $\lambda = 0.01$	7.59	3.02	2.51	1.60	1.52
Pooled regression, $\lambda = 0.01$	10.01	10.94	0.92	0.71	0.70
Pooled regression, $\lambda = 0$	10.25	11.12	0.92	0.72	0.71

Panel B: Paired Sharpe differences, 95% intervals beneath				
Comparison	Gross	Net, 25 bp	Delayed	Net, 2007–2024
Unpenalized minus reference, Euclidean	0.93 [0.61, 1.25]	0.73 [0.47, 0.98]	0.84 [0.55, 1.11]	0.45 [0.08, 0.79]
IP minus factor-mean (same fit)	2.12 [1.64, 2.68]	1.63 [1.18, 2.16]	1.32 [0.87, 1.81]	0.44 [−0.22, 1.10]
IP minus best ZP	0.89 [0.59, 1.26]	0.57 [0.29, 0.89]	0.45 [0.11, 0.80]	0.19 [−0.41, 0.75]
IP minus pooled regression	1.60 [1.14, 2.09]	0.89 [0.46, 1.34]	0.82 [0.40, 1.23]	−0.19 [−0.71, 0.32]
Pooled regression minus factor-mean	0.52 [0.24, 0.86]	0.75 [0.47, 1.09]	0.51 [0.20, 0.85]	0.63 [0.17, 1.11]

windows, and RP-PCA’s reaches −0.08 (April 2012) and is negative in 7. The intercept portfolio is above the best zero-intercept portfolio in 496 of the 535 windows (93%), above RP-PCA in all of them, and the highest of the seven portfolios in 472. The paired band of its difference from the best zero-intercept portfolio (Panel B) lies above zero in 273 windows and below zero in 4, and the lower end of the intercept portfolio’s own band reaches −0.04. Overlapping windows are not independent, and these counts are descriptive.

For the stabilized intercept portfolio against the factor-mean portfolio of the same fit, the rolling 120-month all-stock Sharpe-ratio difference is positive in all 475 windows. Overlapping windows are not independent replications, and Table IA.I.3 gives the corresponding mean increments by size group with intervals: 0.488% a month [0.348, 0.634] over the full sample and 0.202%

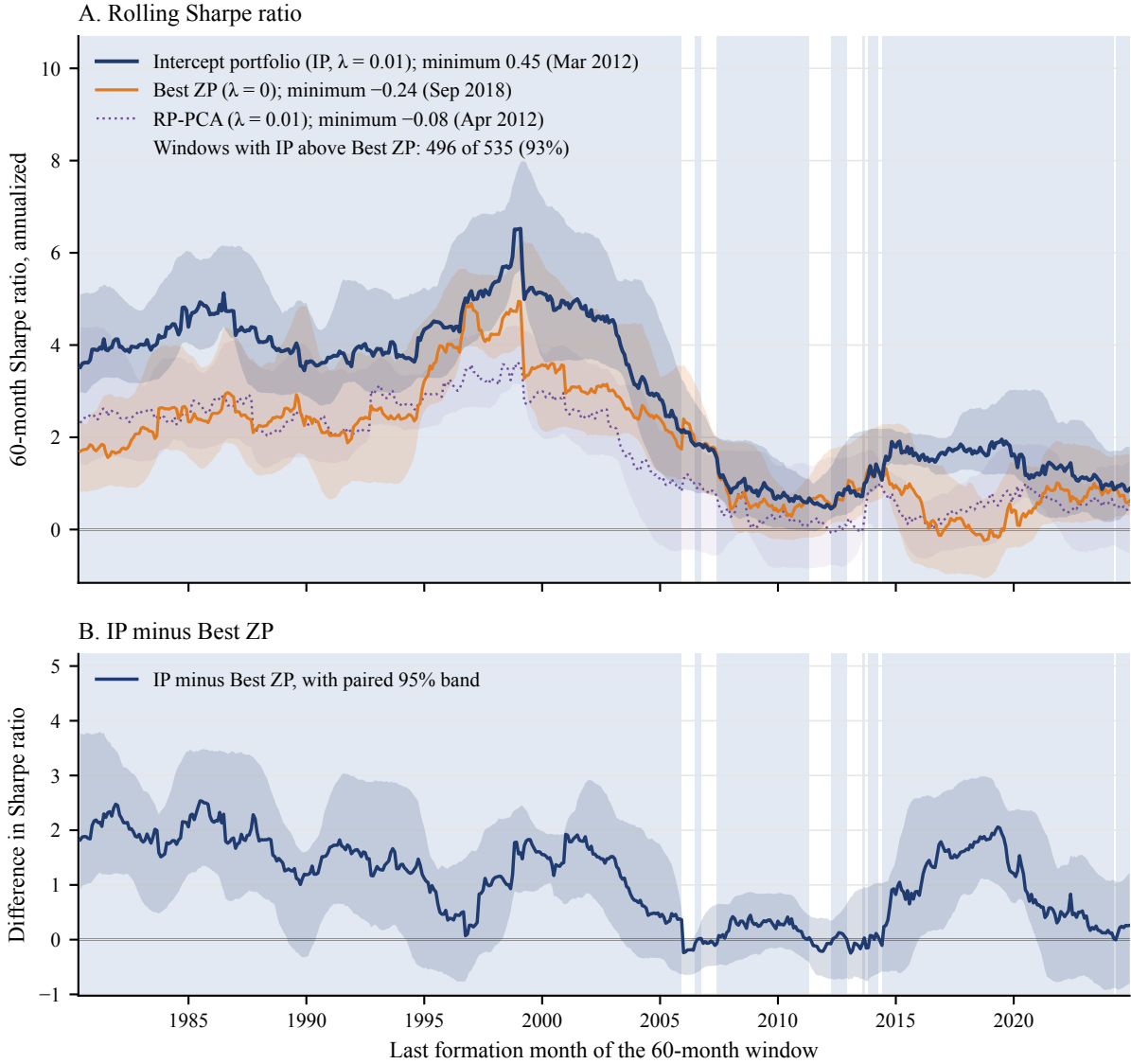


Figure IA.I.2. Five-year rolling Sharpe ratios. Panel A: annualized gross Sharpe ratio over the 60 formation months ending in each month from May 1980 to November 2024 (535 windows) of the intercept portfolio (IP), the best zero-intercept portfolio (best ZP, $\lambda = 0$), and RP-PCA (dotted, without band), with pointwise 95% bands from a 12-month moving-block bootstrap within each window (10,000 draws) that holds the fitted models fixed; shading marks the windows in which the IP is above the best ZP, and the legend gives each line's minimum. Panel B: IP minus best ZP, with the paired band. 594-month sample, June 1975–November 2024.

$[-0.071, 0.409]$ over the 206 months after October 2007, where no group's interval excludes zero.

Table IA.I.3. Paired mean-return increment of the stabilized intercept portfolio over the factor-mean portfolio of the same fit. Returns use unit-gross weights and are gross of costs; means and intervals are in percent per month, at the reference penalty $\lambda = 0.01$. The 12-month moving-block bootstrap uses 10,000 draws, seed 20260714, and a recentered two-sided mean test; the Holm adjustment covers the four universes within each panel. All rows use the rolling fit of IPCA with an intercept estimated on all stocks, applied and re-solved within each indicated stock group with its own unit-gross normalization.

Universe	Mean	95% interval	p	Holm p
Full sample: June 1975–November 2024 (594 months)				
All stocks	0.488	[0.348, 0.634]	< 0.001	< 0.001
Mega	0.239	[0.105, 0.409]	0.003	0.003
Large	0.347	[0.194, 0.508]	< 0.001	< 0.001
Small	0.666	[0.487, 0.840]	< 0.001	< 0.001
Recent period: October 2007–November 2024 (206 months)				
All stocks	0.202	[−0.071, 0.409]	0.11	0.32
Mega	0.206	[−0.035, 0.507]	0.14	0.32
Large	0.174	[−0.120, 0.389]	0.19	0.32
Small	0.251	[−0.061, 0.464]	0.08	0.32

IA.I.4. Size groups

Table IA.I.4 applies the all-stock fits within each capitalization group, solving and normalizing each portfolio there, so the rows are dependent applications of the same fitted models. The intercept portfolio earns 1.25, 1.53, and 2.76 in the mega, large, and small groups, against 0.67, 1.02, and 1.58 for the best zero-intercept portfolio, and its gains, 0.57, 0.51, and 1.18, each have an interval that excludes zero at 12-, 24-, and 60-month blocks. Within each group the gain comes through lower volatility, by 1.75, 0.94, and 2.06 percentage points, each with an interval that excludes zero, while the differences in mean, 0.74, 0.87, and 0.90 points, have intervals that include zero. Against the factor-mean portfolio of the same fit, which earns 0.23 to 0.30, the intercept portfolio and its stabilized version (1.25, 1.54, and 2.77) gain 1.02 to 2.47, and every difference has an interval that excludes zero.

IA.I.5. Alphas on each benchmark

Table IA.I.5 regresses each portfolio on the five Fama–French factors alone. The intercept portfolio has an annualized alpha of 7.15% with a Newey–West t -statistic of 12.01, loadings below

Table IA.I.4. Sharpe ratios within capitalization groups. The all-stock fits are applied to the stocks of one capitalization group, and each portfolio is solved and normalized to unit gross exposure within that group, so the rows are dependent applications of the same fitted models. The intercept portfolio (IP), its stabilized version and the factor-mean portfolio use the reference penalty $\lambda = 0.01$; the best zero-intercept portfolio (best ZP) is IPCA with $\lambda = 0$. Gross returns; 594 formation months, June 1975–November 2024. Panel B reports pointwise 95% intervals from the paired 12-month moving-block bootstrap.

Panel A: Sharpe ratios				
Universe	IP	Best ZP ($\lambda = 0$)	Factor-mean (same fit)	IP, stabilized
All stocks	2.51	1.62	0.39	2.53
Mega	1.25	0.67	0.23	1.25
Large	1.53	1.02	0.25	1.54
Small	2.76	1.58	0.30	2.77

Panel B: Differences in Sharpe ratio			
Universe	IP minus best ZP	IP minus factor-mean	IP, stabilized minus factor-mean
All stocks	0.89	2.12	2.14
	[0.59, 1.26]	[1.64, 2.68]	[1.66, 2.70]
Mega	0.57	1.02	1.02
	[0.34, 0.85]	[0.70, 1.44]	[0.70, 1.44]
Large	0.51	1.28	1.29
	[0.21, 0.85]	[0.83, 1.77]	[0.84, 1.77]
Small	1.18	2.46	2.47
	[0.86, 1.55]	[1.94, 3.06]	[1.95, 3.07]

0.06 in absolute value, and an R^2 of 0.02; the factor-mean portfolio of the same fit has an alpha of 0.15% ($t = 0.25$) and an R^2 of 0.26. Table V adds the best zero-intercept portfolio as a regressor.

Figure IA.I.3 regresses the intercept portfolio on each of the six other portfolios in turn together with the five factors, over the full sample and its two halves of 300 and 294 formation months. Over the full sample the alphas run from 4.9% a year, relative to the best zero-intercept portfolio, to 6.7%, relative to the FF5 portfolio, with Newey–West t -statistics of 10.3 to 13.2. They are 6.5% to 9.6% from June 1975 through May 2000 and 3.0% to 3.7% from June 2000 through November 2024, with t -statistics of 8.7 to 15.7 and 5.4 to 6.7, so every whisker excludes zero. On the five factors alone the alpha is 10.4% and 3.7% in the two halves.

Table IA.I.5. Five-factor alphas of every fitted portfolio. Each row regresses a portfolio's monthly gross return at unit gross exposure on the five factors of Fama and French (2015), dated to the return month, over the 594 formation months from June 1975 through November 2024. Penalized portfolios use the reference penalty $\lambda = 0.01$ in their own units; the best ZP is the unpenalized IPCA portfolio. Alpha is annualized and in percent, with its Newey–West t -statistic (six lags, Newey and West, 1987) beneath; t -statistics of the loadings are in Table V for the best ZP and are not reported for the other rows. Table V adds the best ZP as a regressor. The regressions describe exposure and are not tests of a pricing model.

Portfolio	Alpha	MKT	SMB	HML	RMW	CMA	R^2
IPCA, Euclidean units	2.87 (2.71)	0.10	0.05	−0.05	0.06	0.04	0.09
IPCA, Gram units	3.32 (5.01)	0.13	0.08	0.04	0.04	−0.01	0.25
Best ZP ($\lambda = 0$)	4.95 (6.59)	0.08	0.09	−0.03	0.15	0.10	0.16
RP-PCA	5.01 (6.63)	0.06	0.04	−0.07	0.06	0.08	0.07
FF5 portfolio	2.43 (2.94)	0.12	0.07	0.10	−0.01	0.03	0.19
Factor-mean portfolio (same fit)	0.15 (0.25)	0.13	0.08	0.06	0.05	−0.01	0.26
Intercept portfolio (IP)	7.15 (12.01)	0.01	0.03	0.00	0.05	0.02	0.02

IA.I.6. Cumulative returns

Figure IA.I.4 plots cumulative log returns of the seven portfolios, each scaled by a constant so that its full-sample volatility is 10% a year; the scaling is ex post and leaves Sharpe ratios unchanged. The intercept portfolio's path ends at 12.1, against 7.7 for the best zero-intercept portfolio, 7.1 for RP-PCA, 5.0 for the Gram fit, 4.3 for the pooled regression, 3.5 for FF5, and 3.1 for the Euclidean fit. Every path is flatter after mid-2008, the start of the last third, in which every portfolio earns its lowest Sharpe ratio (Figure 5).

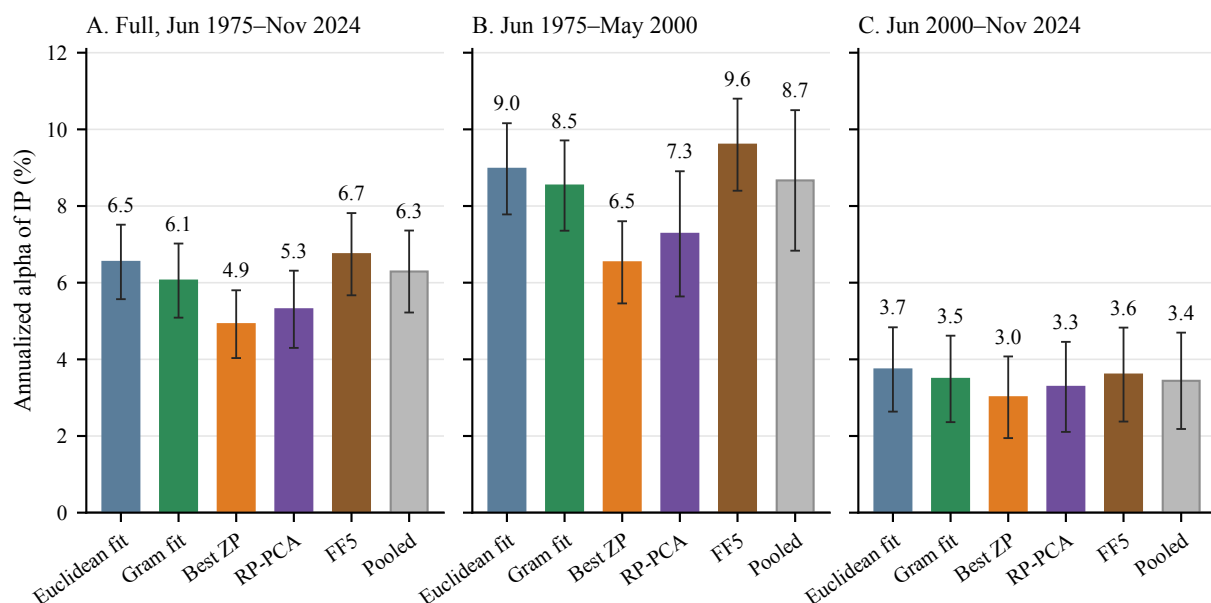


Figure IA.I.3. Alpha of the intercept portfolio on each benchmark. Each bar is the annualized intercept, in percent, of a regression of the intercept portfolio’s monthly gross return on the return of one other portfolio of Figure 5 and the five Fama–French factors, dated to the return month; whiskers are ± 1.96 Newey–West standard errors (six lags), and values are printed above them. Panel A: 594 formation months, June 1975–November 2024; Panel B: June 1975–May 2000 (300 months); Panel C: June 2000–November 2024 (294 months). One y-range; returns at unit gross exposure.

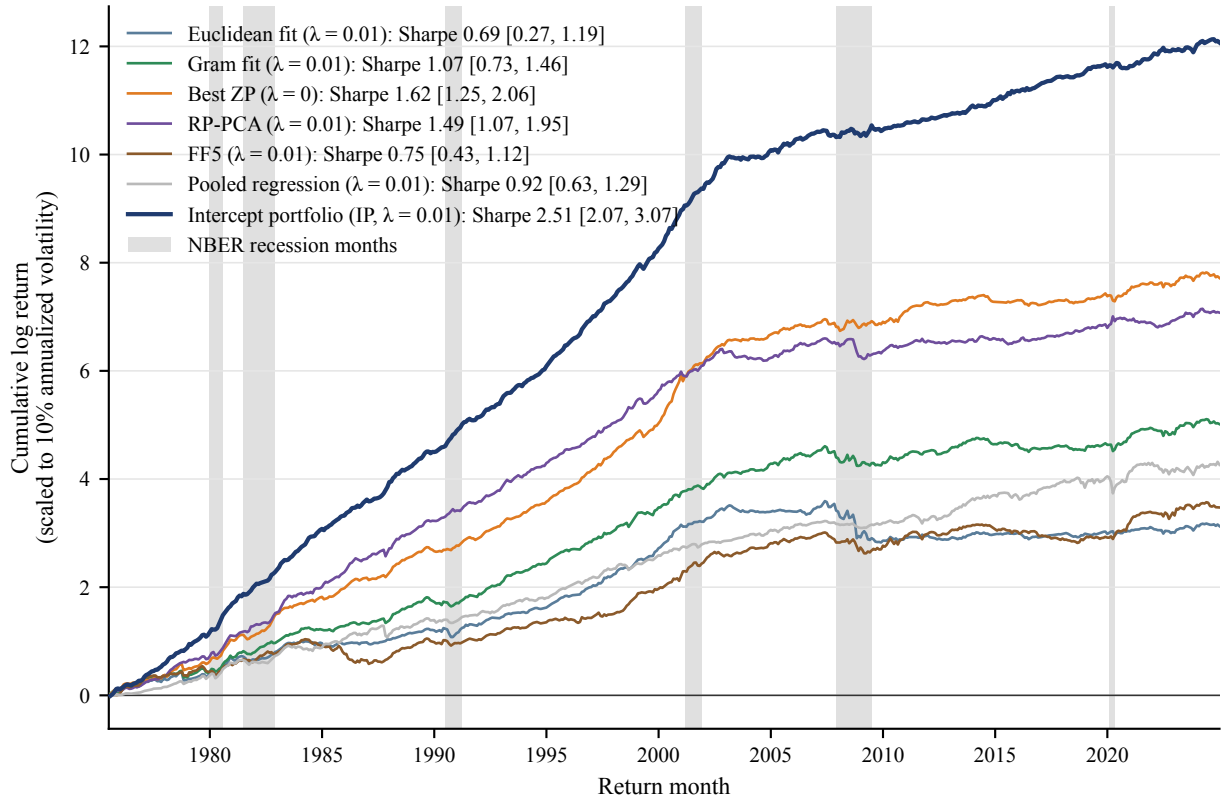


Figure IA.I.4. Cumulative log returns at 10% volatility. Each portfolio's monthly gross return r_t at unit gross exposure is scaled by $k = 0.10/(\sqrt{12}s)$, with s its full-sample monthly standard deviation, and cumulated as $\sum \log(1 + k r_t)$ by return month; the scaling is ex post and descriptive. NBER recession months are shaded (National Bureau of Economic Research, 2023). The legend prints each portfolio's full-sample Sharpe ratio with its pointwise 95% interval (12-month moving-block bootstrap, 10,000 draws); the paths are realized and carry no band. 594 formation months, June 1975–November 2024.

IA.J. How Much Do the Portfolios Trade, and How Accurate Are Their Forecasts?

This section defines the turnover and cost calculation behind Table IV and Section VI.D, describes turnover over time, and reports the accuracy of the stock-level forecasts.

IA.J.1. Turnover and proportional costs

Turnover is computed from holdings. Let u_t be the unit-gross weights at formation t and $R_{i,t+1}$ the realized total return of stock i over the following month, its excess return plus the risk-free rate. The drifted position before the next rebalance is $\tilde{u}_{i,t+1} = u_{i,t}(1 + R_{i,t+1})$; stocks that leave the universe are liquidated and entrants are bought from zero. The turnover of the rebalance at $t + 1$ is

$$q_{t+1} = \sum_i |u_{i,t+1} - \tilde{u}_{i,t+1}| \quad (\text{IA29})$$

over the union of names. It counts purchases and sales together, so one-way turnover is half of it, and capital is reset to unit gross exposure at each rebalance; the initial position is not charged. The net return subtracts c times the turnover of the trade that established the month's holdings, for c of 10, 25, and 50 basis points per dollar traded. With $k = c/10^4$,

$$\text{Var}(r^u - kq) = \text{Var}(r^u) + k^2 \text{Var}(q) - 2k \text{Cov}(r^u, q), \quad (\text{IA30})$$

so costs change volatility as well as the mean, and every net Sharpe ratio is recomputed from the net series. The break-even cost is the value $c^* = 10^4 \bar{r}^u / \bar{q}$ at which the mean net return is zero.

The cost model is proportional: it omits stock-level spreads, price impact, short-sale fees, financing, and capacity, and it is the standard device for comparing turnover-heavy portfolios on a common footing (Novy-Marx and Velikov, 2016; DeMiguel et al., 2020). Monthly turnover is 0.64 to 0.94 of gross exposure for the portfolios of Table IV, or 0.32 to 0.47 one way, and it is highest for

the intercept portfolio, 0.94. Across the penalty grid the intercept portfolio's turnover is 0.935 to 0.940 and its break-even cost 68 to 69 basis points, against 22 to 27 basis points for the factor-mean portfolio of the same fit; its net Sharpe ratio at 25 basis points lies between 1.5968 and 1.6093 at every penalty value (Table III). In the last third of the sample turnover is lower for every portfolio: 0.84 for the intercept portfolio, 0.68 for the best zero-intercept portfolio, and 0.60 to 0.66 for the others. Figure IA.J.1 plots the trailing 12-month mean turnover of the intercept portfolio and the best zero-intercept portfolio: it ranges from 0.73 to 1.19 for the intercept portfolio and from 0.54 to 1.03 for the best zero-intercept portfolio, is higher for the intercept portfolio in 489 of the 583 windows, and averages 0.94 and 0.81 over the full sample.

A proxy that scales each month's raw one-way turnover by the contemporaneous raw gross exposure, instead of reconstructing trades from return-drifted holdings, gives a one-way turnover of 0.47 a month for the intercept portfolio, consistent with half the holdings-based figure, and break-even costs about twice as large: about 136 basis points for the stabilized intercept portfolio, 142 for RP-PCA, and 111 to 115 for the three Gram estimates on 594 months, with 50-basis-point Sharpe ratios of 0.27 and 0.57 for the Euclidean fit and the rolling-60 Gram fit on 618 months and 1.62 for the stabilized intercept portfolio on 594 months. No exhibit uses the proxy; every exhibit uses the holdings-based calculation.

IA.J.2. Forecast accuracy and investment value

Table IA.J.1 evaluates the stock-level forecasts on a common set of 1,210,039 stock-months of the 594-month sample, a different count from the 1,244,002 stock-months of the 618-month evaluation. A negative predictive R^2 means larger pooled squared error than a zero forecast; a positive information coefficient can coexist with it because correlation does not assess scale. These statistics cannot rank portfolios. Hold the risk matrix A_t fixed and multiply a nonzero forecast $\hat{\mu}_t$

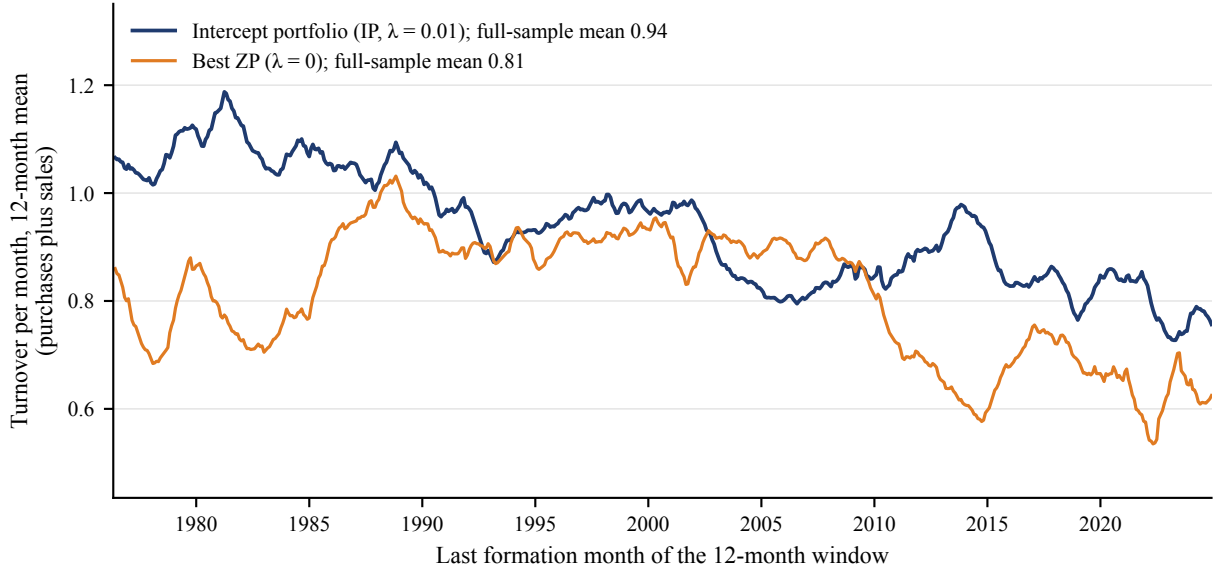


Figure IA.J.1. Turnover over time. Trailing 12-month mean of monthly turnover, purchases plus sales per dollar of gross exposure (one-way turnover is half of it), of the intercept portfolio ($\lambda = 0.01$) and the best zero-intercept portfolio ($\lambda = 0$), for the 12 formation months ending in each month from May 1976 to November 2024; the legend prints the full-sample means. Turnover is a realized trading quantity of the fitted portfolios, not an estimated performance statistic, and no band is drawn. 594-month sample, June 1975–November 2024.

by $c > 0$; the unit-gross allocation satisfies

$$\frac{A_t^{-1}(c\hat{\mu}_t)}{\|A_t^{-1}(c\hat{\mu}_t)\|_1} = \frac{c w_t}{c \|w_t\|_1} = w_t^u, \quad (\text{IA31})$$

so positions and returns are unchanged while squared forecast errors change. The Euclidean and the Gram fits agree to four decimals in all three statistics, and RP-PCA has the highest rank correlation together with a predictive R^2 of about -0.26 and the highest Sharpe ratio among the separately fitted benchmarks.

Table IA.J.1. Forecast accuracy on the 594-month sample. The sample is June 1975–November 2024, with 1,210,039 stock-months in the common forecast evaluation set. IC is the mean monthly cross-sectional Pearson correlation between forecasts and next-month excess returns; rank IC uses Spearman correlations. Predictive R^2 is $1 - \sum (r - \hat{r})^2 / \sum r^2$, pooled over stock-months, with a zero forecast as the benchmark. All entries are in decimal units. The Gram row uses the rolling-60 estimate of the metric. The last row is the factor-mean forecast of the intercept fit; the complete forecast of the intercept portfolio is not evaluated here.

Forecast	IC	Rank IC	Predictive R^2
FF5	0.0159	0.0170	−0.0004
IPCA, Euclidean units	0.0254	0.0249	−0.0188
IPCA, Gram units (rolling-60 estimate)	0.0254	0.0249	−0.0188
RP-PCA	0.0364	0.0415	−0.2568
Factor-mean forecast, IPCA with an intercept	−0.0021	−0.0046	−0.0174

IA.K. Which Characteristics Enter the Fits?

Every fit in the paper uses the same 132 characteristics, described in Section III. Table IA.K.1 lists them with their coverage, mean, standard deviation, minimum, median, and maximum over the 1,244,002 stock-months of the 618-month sample, together with the next-month excess return. Over the 618 formation months from June 1973 through November 2024 the panel has 1,244,002 stock-months from 12,703 stocks, with a median monthly cross section of 1,990 stocks (between 1,335 and 3,150); over the 594 formation months of the main text it has 1,210,039 stock-months, with a median cross section of 2,001 stocks (between 1,431 and 3,150). The set is the Jensen–Kelly–Pedersen library (Jensen, Kelly, and Pedersen, 2023) as supplied by the data provider at the extraction date, less the variables unavailable in the processed panel; the selection rule is the data provider’s. Stock-months with more than 30% missing characteristics are excluded before ranking, and the remaining missing values are set to zero after ranking.

Table IA.K.1. Summary statistics of the 132 characteristics and the return. The sample is all mega-, large-, and small-capitalization stocks over the 618 formation months from June 1973 through November 2024, 1,244,002 stock-months; the main text uses the 594 months from June 1975. Characteristics are the Jensen–Kelly–Pedersen variables, ranked cross-sectionally each month to $[-0.5, 0.5]$. The processed panel already sets missing values to zero (Section IA.B.6), so coverage is 100% by construction and the means and standard deviations mix ranks with zeros; the source documentation reports raw coverage. The return is the next-month excess return in decimal form. Variable definitions follow Jensen, Kelly, and Pedersen (2023) and the documentation of their data library; the names are the library’s identifiers.

Variable	Coverage (%)	Mean	SD	Min	Median	Max
cowc_gr1a	100.0	-0.005	0.249	-0.500	0.000	0.500
oaccruals_at	100.0	0.003	0.256	-0.500	0.000	0.500
oaccruals_ni	100.0	-0.015	0.269	-0.500	0.000	0.500
taccruals_at	100.0	0.002	0.263	-0.500	0.000	0.500
taccruals_ni	100.0	-0.013	0.272	-0.500	0.000	0.500
debt_gr3	100.0	0.015	0.249	-0.500	0.000	0.500
fnl_gr1a	100.0	0.007	0.280	-0.500	0.006	0.500
ncol_gr1a	100.0	0.034	0.278	-0.500	0.031	0.500
nfna_gr1a	100.0	-0.005	0.276	-0.500	-0.015	0.500
noa_at	100.0	0.014	0.262	-0.500	0.000	0.500
aliq_at	100.0	-0.021	0.270	-0.500	0.000	0.500
at_gr1	100.0	0.024	0.265	-0.500	0.026	0.500
be_gr1a	100.0	0.027	0.264	-0.500	0.021	0.500
capx_gr1	100.0	0.009	0.245	-0.500	0.000	0.500
capx_gr2	100.0	0.010	0.243	-0.500	0.000	0.500
capx_gr3	100.0	0.013	0.238	-0.500	0.000	0.500
coa_gr1a	100.0	0.005	0.254	-0.500	0.000	0.500
col_gr1a	100.0	0.012	0.255	-0.500	0.000	0.500
emp_gr1	100.0	0.007	0.255	-0.500	0.000	0.500

Variable	Coverage (%)	Mean	SD	Min	Median	Max
inv_gr1	100.0	0.006	0.241	-0.500	0.000	0.500
inv_gr1a	100.0	-0.013	0.281	-0.500	0.000	0.500
lnoa_gr1a	100.0	0.024	0.257	-0.500	0.000	0.500
mispricing_mgmt	100.0	-0.002	0.284	-0.500	-0.002	0.500
ncoa_gr1a	100.0	0.032	0.262	-0.500	0.012	0.500
nncoa_gr1a	100.0	0.024	0.260	-0.500	0.000	0.500
noa_gr1a	100.0	0.012	0.256	-0.500	0.000	0.500
ppeinv_gr1a	100.0	0.015	0.258	-0.500	0.000	0.500
sale_gr1	100.0	0.013	0.266	-0.500	0.008	0.500
sale_gr3	100.0	0.016	0.257	-0.500	0.000	0.500
saleq_gr1	100.0	0.015	0.264	-0.500	0.003	0.500
age	100.0	0.067	0.298	-0.500	0.114	0.491
aliq_mat	100.0	-0.040	0.254	-0.500	0.000	0.500
at_be	100.0	0.017	0.270	-0.500	0.015	0.500
bidaskhl_21d	100.0	-0.134	0.238	-0.500	-0.161	0.500
cash_at	100.0	-0.024	0.279	-0.500	-0.026	0.500
netdebt_me	100.0	0.019	0.262	-0.500	0.025	0.500
tangibility	100.0	-0.029	0.265	-0.500	0.000	0.500
z_score	100.0	0.013	0.264	-0.500	0.000	0.500
beta_60m	100.0	-0.000	0.254	-0.500	0.000	0.500
beta_dimson_21d	100.0	0.022	0.260	-0.500	0.020	0.500
betabab_1260d	100.0	0.019	0.250	-0.500	0.000	0.500
betadown_252d	100.0	0.026	0.259	-0.500	0.018	0.500
chcsho_12m	100.0	-0.023	0.288	-0.500	-0.007	0.500

Variable	Coverage (%)	Mean	SD	Min	Median	Max
eqnetis_at	100.0	-0.041	0.287	-0.500	-0.015	0.500
fcf_me	100.0	0.040	0.248	-0.500	0.033	0.500
ivol_capm_21d	100.0	-0.109	0.258	-0.500	-0.140	0.500
ivol_capm_252d	100.0	-0.130	0.248	-0.500	-0.163	0.500
ivol_ff3_21d	100.0	-0.110	0.257	-0.500	-0.140	0.500
ivol_hxz4_21d	100.0	-0.110	0.257	-0.500	-0.141	0.500
netis_at	100.0	-0.027	0.266	-0.500	-0.018	0.500
rmax1_21d	100.0	-0.084	0.265	-0.500	-0.108	0.500
rmax5_21d	100.0	-0.087	0.262	-0.500	-0.110	0.500
rvol_21d	100.0	-0.096	0.260	-0.500	-0.123	0.500
turnover_126d	100.0	0.077	0.264	-0.500	0.093	0.500
zero_trades_126d	100.0	-0.106	0.241	-0.500	-0.118	0.500
zero_trades_21d	100.0	-0.094	0.248	-0.500	-0.105	0.500
zero_trades_252d	100.0	-0.110	0.239	-0.500	-0.121	0.500
prc_highprc_252d	100.0	0.073	0.261	-0.500	0.091	0.500
resff3_12_1	100.0	0.004	0.274	-0.500	0.000	0.500
resff3_6_1	100.0	0.006	0.275	-0.500	0.000	0.500
ret_12_1	100.0	0.042	0.260	-0.500	0.044	0.500
ret_3_1	100.0	0.022	0.266	-0.500	0.030	0.500
ret_6_1	100.0	0.032	0.264	-0.500	0.038	0.500
ret_9_1	100.0	0.038	0.262	-0.500	0.043	0.500
seas_1_1na	100.0	0.025	0.259	-0.500	0.019	0.500
dsale_drec	100.0	-0.002	0.263	-0.500	0.000	0.500
niq_at_chg1	100.0	0.008	0.254	-0.500	0.000	0.500

Variable	Coverage (%)	Mean	SD	Min	Median	Max
niq_be_chg1	100.0	0.012	0.254	-0.500	0.000	0.500
niq_su	100.0	0.006	0.277	-0.500	0.000	0.500
ocf_at_chg1	100.0	0.006	0.259	-0.500	0.000	0.500
qmj_safety	100.0	0.028	0.270	-0.500	0.036	0.500
ret_12_7	100.0	0.032	0.263	-0.500	0.032	0.500
sale_emp_gr1	100.0	0.002	0.256	-0.500	0.000	0.500
saleq_su	100.0	0.008	0.276	-0.500	0.000	0.500
seas_1_1an	100.0	0.018	0.264	-0.500	0.012	0.500
sti_gr1a	100.0	-0.069	0.288	-0.500	-0.212	0.500
dolvol_var_126d	100.0	-0.129	0.264	-0.500	-0.175	0.500
ebit_bev	100.0	0.070	0.262	-0.500	0.081	0.500
ebit_sale	100.0	0.078	0.261	-0.500	0.100	0.500
f_score	100.0	-0.035	0.256	-0.500	0.000	0.499
ival_me	100.0	-0.017	0.249	-0.500	0.000	0.500
ni_be	100.0	0.077	0.264	-0.500	0.095	0.500
niq_be	100.0	0.068	0.261	-0.500	0.077	0.500
o_score	100.0	-0.080	0.247	-0.500	-0.063	0.500
ocf_at	100.0	0.069	0.261	-0.500	0.082	0.500
ope_be	100.0	0.068	0.246	-0.500	0.036	0.500
ope_bel1	100.0	0.063	0.242	-0.500	0.026	0.500
turnover_var_126d	100.0	-0.125	0.266	-0.500	-0.169	0.500
at_turnover	100.0	-0.000	0.276	-0.500	-0.006	0.500
cop_at	100.0	0.050	0.259	-0.500	0.043	0.500
cop_atl1	100.0	0.053	0.261	-0.500	0.048	0.500

Variable	Coverage (%)	Mean	SD	Min	Median	Max
dgp_dsale	100.0	0.011	0.265	-0.500	0.000	0.500
gp_at	100.0	0.010	0.275	-0.500	0.006	0.500
gp_atl1	100.0	0.013	0.276	-0.500	0.010	0.500
mispricing_perf	100.0	0.073	0.269	-0.500	0.094	0.500
ni_inc8q	100.0	-0.077	0.312	-0.500	0.000	0.429
niq_at	100.0	0.068	0.263	-0.500	0.084	0.500
op_at	100.0	0.067	0.267	-0.500	0.084	0.500
op_atl1	100.0	0.065	0.266	-0.500	0.083	0.500
opex_at	100.0	-0.031	0.280	-0.500	-0.049	0.500
qmj_prof	100.0	0.065	0.276	-0.500	0.087	0.500
sale_bev	100.0	-0.007	0.276	-0.500	-0.003	0.500
tax_gr1a	100.0	0.017	0.284	-0.500	0.039	0.500
corr_1260d	100.0	0.122	0.235	-0.500	0.118	0.500
coskew_21d	100.0	0.004	0.286	-0.500	0.000	0.500
dbnetis_at	100.0	-0.009	0.277	-0.500	-0.020	0.500
kz_index	100.0	-0.016	0.263	-0.500	0.000	0.500
lti_gr1a	100.0	-0.084	0.306	-0.500	-0.257	0.500
pi_nix	100.0	0.007	0.257	-0.500	0.000	0.500
ami_126d	100.0	-0.201	0.195	-0.500	-0.222	0.500
dolvol_126d	100.0	0.201	0.200	-0.500	0.226	0.500
market_equity	100.0	0.236	0.158	-0.235	0.241	0.500
prc	100.0	0.179	0.225	-0.500	0.217	0.500
iskew_capm_21d	100.0	-0.013	0.285	-0.500	-0.007	0.500
iskew_ff3_21d	100.0	-0.011	0.285	-0.500	-0.006	0.500

Variable	Coverage (%)	Mean	SD	Min	Median	Max
iskew_hxz4_21d	100.0	-0.011	0.285	-0.500	-0.005	0.500
ret_1_0	100.0	0.017	0.267	-0.500	0.024	0.500
rmax5_rvol_21d	100.0	0.007	0.273	-0.500	0.000	0.500
rskew_21d	100.0	-0.015	0.282	-0.500	-0.011	0.500
at_me	100.0	-0.029	0.282	-0.500	-0.041	0.500
be_me	100.0	-0.050	0.268	-0.500	-0.062	0.500
bev_mev	100.0	-0.042	0.263	-0.500	-0.042	0.500
debt_me	100.0	0.001	0.278	-0.500	0.011	0.500
div12m_me	100.0	-0.038	0.379	-0.500	0.104	0.500
ebitda_mev	100.0	0.028	0.250	-0.500	0.024	0.500
eq_dur	100.0	0.010	0.254	-0.500	0.002	0.500
eqnpo_12m	100.0	0.039	0.285	-0.500	0.060	0.500
eqnpo_me	100.0	0.052	0.272	-0.500	0.064	0.500
eqpo_me	100.0	0.032	0.298	-0.500	0.046	0.500
ni_me	100.0	0.046	0.251	-0.500	0.051	0.500
ocf_me	100.0	0.036	0.250	-0.500	0.031	0.500
sale_me	100.0	-0.027	0.265	-0.500	-0.037	0.500
next-month excess return	100.0	0.008	0.129	-1.005	0.005	16.250

IA.L. How Does the Gram Metric Evolve?

The Gram normalization of Section II measures loadings by the stock exposures they generate, through a trailing average of the characteristic second-moment matrices. This section describes the 618 EWMA estimates of that metric used by the rolling fits, from June 1973 through November

2024.

IA.L.1. Spectral concentration and drift

Let $\ell_{1,t} \geq \dots \geq \ell_{L,t} > 0$ be the eigenvalues of the 132×132 feature block $\widehat{Q}_t$ passed to each fit, and $p_{j,t} = \ell_{j,t} / \sum_k \ell_{k,t}$. The spectral participation dimension and the adjacent-window spectral drift are

$$d_t^{\text{part}} = \frac{1}{\sum_{j=1}^L p_{j,t}^2} = \frac{(\text{tr } \widehat{Q}_t)^2}{\text{tr}(\widehat{Q}_t^2)}, \quad v_t^{\text{spec}} = \frac{1}{2} \sum_{j=1}^L |p_{j,t} - p_{j,t-1}|. \quad (\text{IA32})$$

By Cauchy–Schwarz $1 \leq d_t^{\text{part}} \leq L$, with equality at L when the eigenvalues are equal; the drift compares distributions of second-moment mass across ordered spectral positions, lies in $[0, 1]$, and is not defined for the first month. Both statistics are unchanged by a common positive scaling and by permutations of the characteristics; they are not invariant to characteristic-specific rescaling, and the drift does not track rotations of the eigenvectors.

The participation dimension has a median of 19.8 and a range of 17.4 to 22.6, against 132 for equal eigenvalues, and the leading eigenvalue accounts for a median 14.2% of the trace. Every matrix has full rank, with condition numbers from 1,735 to 13,747. The spectrum moves gradually: the median monthly drift is 0.063 percentage points and the largest 0.115 (Table IA.L.1, Figure IA.L.1). Each window’s preprocessing differs from the previous window’s in all 617 transitions, so the paths combine changing standardization and stock samples with changing relationships among characteristics; no crisis date, threshold, or return enters the calculation. Neither statistic estimates the number of priced risks. The metric sets the units in which the Gram fit reports its loadings, and Section IV shows that those units affect portfolios only through the factor penalty.

Table IA.L.1. The Gram metric over time. The sample is all 618 EWMA feature-block Gram matrices, June 1973–November 2024, with 132 characteristics throughout; the main text uses the 594 months from June 1975. The participation dimension summarizes the dispersion of eigenvalue shares and is not a number of priced factors. Spectral drift is defined in equation (IA32) and has 617 observations. The condition number is the ratio of the largest to the smallest eigenvalue.

Diagnostic	Minimum	Median	Maximum
Participation dimension	17.36	19.83	22.63
Largest eigenvalue share (%)	11.76	14.24	15.21
Spectral drift (percentage points)	0.033	0.063	0.115
Condition number	1,735	5,922	13,747

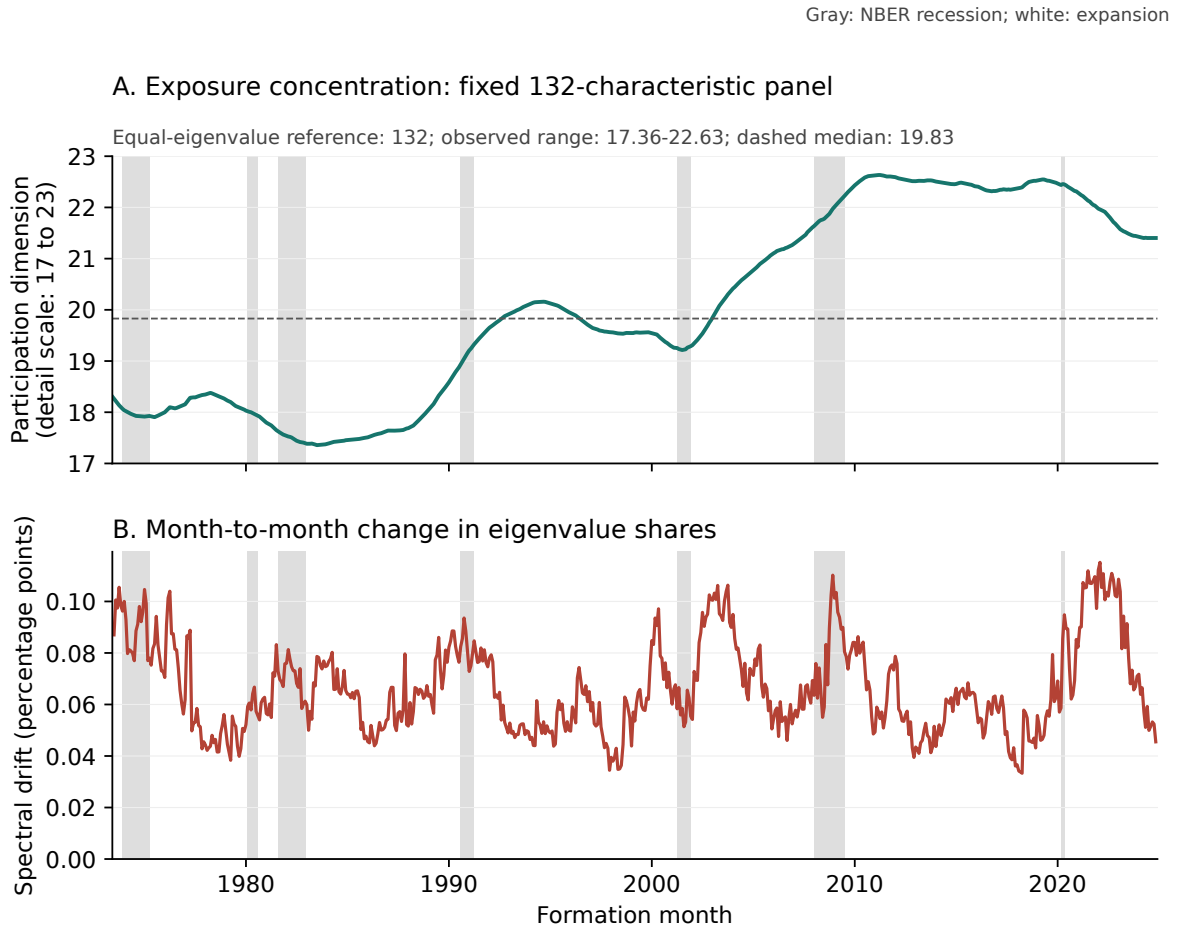


Figure IA.L.1. The Gram metric over time. Panel A: spectral participation dimension of the 132-characteristic EWMA Gram matrix in each of the 618 formation months, June 1973–November 2024; the detail axis spans 17–23, the dashed line is the median, and the equal-eigenvalue value of 132 lies outside the axis. Panel B: one half of the absolute change in the normalized ordered eigenvalue shares, summed over all 132 eigenvalues, in percentage points. The matrices incorporate each window’s preprocessing and stock sample. Gray bands mark NBER recession months on the formation-date axis (National Bureau of Economic Research, 2023).

IA.M. Does the Intercept Signal Survive a Shuffle, the Removal of Extreme Months, and a Reversal Control?

Section VI attributes the intercept portfolio's performance to the fitted intercept. This section reports three checks on that signal: a within-month permutation of the intercept across stocks, the removal of each portfolio's most extreme months, and a control for short-term reversal.

IA.M.1. A within-month shuffle of the fitted intercept

At every formation month the risk matrix $A_t = B_t \Sigma_{F,t} B_t' + \widehat{D}_t + \lambda_{\Sigma} I$ at the reference penalty, the factor-mean forecast $B_t \widehat{m}_{F,t}$, the universe, and the realized returns are held fixed, and the stock-level fitted intercept $a_t = Z_t \widehat{\Gamma}_{\alpha,t}$ is replaced by a random permutation Pa_t of its N_t entries across that month's stocks. The portfolio $A_t^{-1}(B_t \widehat{m}_{F,t} + Pa_t)$ is normalized to unit gross exposure and its next-month gross return computed. Each of 1,000 draws permutes every month independently, from one generator with seed 20260714; the identity permutation reproduces the intercept portfolio to 1.7×10^{-14} a month. The shuffle keeps the intercept's within-month distribution, including any common component, and removes its pairing with each stock's characteristics and next return.

Shuffling the fitted intercept leaves a gross Sharpe ratio of 0.04 on average, with a standard deviation of 0.08, and never above 0.30 in 1,000 draws, against 2.51 for the intercept as fitted (Table IA.M.1, Figure IA.M.1); the actual value lies 27 null standard deviations above the largest draw, and the permutation p -value is at its floor of $1/1,001$. On the 618-month sample the null mean is -0.02 and its maximum 0.22 , against 2.47 . The shuffled portfolios have a mean return of 0.07% a year on average and a volatility of 1.76% , against 7.59% and 3.02% for the intercept portfolio, so the gap is a gap in mean. The factor-mean portfolio of the same fit, at 0.39 , also exceeds every draw: the allocator scales the intercept up relative to the factor-mean forecast, so a shuffled intercept dominates the raw weights and the shuffled portfolio is close to cross-sectional noise rather than to the factor-mean portfolio. In the data the cross-sectional correlation between the fitted intercept

and the next month's return averages 0.045 and is positive in 73% of the months.

Table IA.M.1. The intercept portfolio under a within-month shuffle of the fitted intercept. In every formation month the fitted stock-level intercept of the rolling fit of IPCA with an intercept is permuted at random across that month's stocks, holding the risk matrix (reference penalty $\lambda = 0.01$), the factor-mean forecast, the universe, and the realized returns fixed; the portfolio is re-solved and held at unit gross exposure. Entries are annualized gross Sharpe ratios over 1,000 independent permutations (seed 20260714) on the 594 formation months from June 1975 and the 618 from June 1973, both through November 2024. The first two rows are the portfolios with the intercept as fitted and without it.

	594 months	618 months
Intercept portfolio, intercept as fitted	2.51	2.47
Factor-mean portfolio (same fit)	0.39	0.32
Shuffled intercept, 1,000 draws		
Mean	0.04	-0.02
Standard deviation	0.08	0.08
Minimum	-0.22	-0.24
1st percentile	-0.15	-0.20
5th percentile	-0.10	-0.14
Median	0.04	-0.02
95th percentile	0.16	0.10
99th percentile	0.21	0.15
Maximum	0.30	0.22
Draws at or above the intercept portfolio	0	0

IA.M.2. Extreme months

Removing each portfolio's five largest absolute-return months leaves the intercept portfolio's Sharpe ratio at 2.51 (2.513 with every month and 2.506 without the five) and the stabilized intercept portfolio's at 2.53. The best zero-intercept portfolio rises from 1.62 to 1.72, RP-PCA from 1.49 to 1.68, and the Gram fit from 1.07 to 1.23, while the ordering of the Euclidean fit and the FF5 portfolio at the reference penalty, 0.69 and 0.75 with every month, reverses to 1.00 and 0.69. The intercept portfolio's most extreme months are monthly gains of 2.8% to 3.4% per dollar of gross position. Because the removed months differ across portfolios, this is a check on the influence of extreme observations within each series rather than a paired comparison.

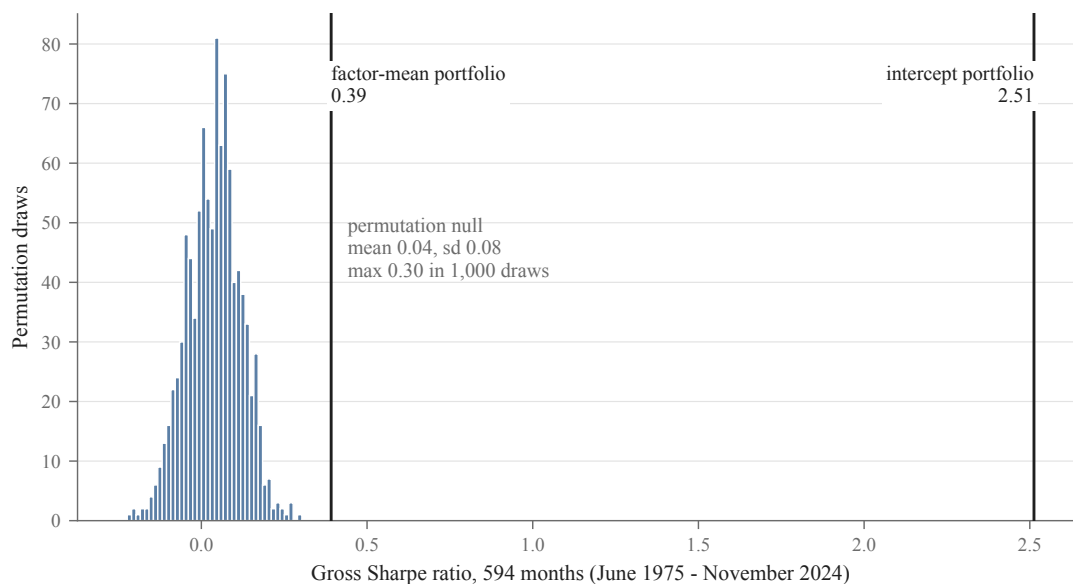


Figure IA.M.1. The intercept portfolio under a within-month shuffle of the fitted intercept. Histogram of the annualized gross Sharpe ratios, at unit gross exposure, of 1,000 portfolios in which the fitted stock-level intercept is permuted across the stocks of each formation month, with the risk matrix at the reference penalty, the factor-mean forecast, the universe, and the realized returns held fixed. The left vertical line marks the factor-mean portfolio of the same fit and the right one the intercept portfolio with the intercept as fitted; the annotation gives the mean, standard deviation, and maximum of the draws. 594 months, June 1975–November 2024.

IA.M.3. A short-term reversal control

Trading one month late lowers the intercept portfolio’s Sharpe ratio from 2.51 to 1.52 (Section VI.D), and the 132 characteristics include the one-month return, `ret_1_0`, and 13 characteristics computed from the daily returns of the last 21 trading days (Table IA.K.1). Table IA.M.2 asks in two ways how much of the portfolio is a short-term reversal position. Panel A adds the short-term reversal factor and the momentum factor of Kenneth French’s data library to the five Fama–French factors. The intercept portfolio’s alpha moves from 7.15% a year to 7.06% ($t = 11.73$), with a reversal loading of 0.03 ($t = 2.27$) and an R^2 of 0.04; adding momentum leaves an alpha of 6.73% ($t = 11.53$). The best zero-intercept portfolio does not load on the reversal factor (-0.01 , $t = -0.40$).

Panels B and C form the intercept portfolio with the fitted intercept loadings of nested sets

of short-horizon characteristics set to zero at every formation date, with the factor block, the risk matrix, and the allocation rule unchanged. Without the one-month return it earns 2.37 gross and 1.55 net of 25 basis points. Without the one-month return and the 13 daily-return characteristics, it earns 2.16 gross and 1.44 net, and its gain over the best zero-intercept portfolio is 0.54 [0.18, 0.92] gross and 0.41 [0.08, 0.76] net; adding the month $t - 12$ return, `seas_1_1an`, changes little. When the two 21-day trading-activity characteristics, `bidaskhl_21d` and `zero_trades_21d`, are also set to zero, the portfolio earns 1.81 gross and 1.18 net, and its gain over the best zero-intercept portfolio, 0.19 [−0.17, 0.53] gross and 0.15 [−0.19, 0.47] net, is no longer distinguishable from zero. The part removed from the intercept has, on average, 29% of the norm of the fitted intercept for the one-month return alone and 79% for all 17 characteristics, and each modified portfolio is below the intercept portfolio, by 0.14 to 0.71 in gross Sharpe ratio, with intervals that exclude zero. After the removal, the loadings on the reversal factor added to the five factors lie between −0.03 and −0.01 and the alphas on the seven factors between 5.1% and 6.3% a year (t of 8.8 to 11.2).

A reversal factor therefore explains little of the intercept portfolio's return, but the intercept's loadings on short-horizon return and trading-activity characteristics carry much of its gain over the best zero-intercept portfolio. The check holds the fit fixed: the remaining loadings were estimated with the removed characteristics present, so correlated characteristics may carry part of their contribution, and a re-estimation without them would require new rolling fits.

Table IA.M.2. A short-term reversal control for the intercept portfolio. Panel A regresses the monthly gross return of the intercept portfolio (IP, $\lambda = 0.01$) and the best zero-intercept portfolio (best ZP, $\lambda = 0$) on the five Fama–French factors, the short-term reversal factor (ST Rev), and the momentum factor (Mom); annualized alpha in percent, Newey–West t -statistics (six lags) beneath, five loadings not shown. Panels B and C zero the fitted intercept loadings of the listed characteristics, fit unchanged; share: mean norm of the removed part relative to the fitted intercept; net: 25 basis points per dollar traded; Panel C: pointwise 95% intervals, paired 12-month moving-block bootstrap. 594 months, June 1975–November 2024.

Panel A: Regressions on factor returns

Portfolio and factors	Alpha	ST Rev	Mom	R^2
IP, FF5	7.15 (12.01)			0.02
IP, FF5 + ST Rev	7.06 (11.73)	0.03 (2.27)		0.04
IP, FF5 + ST Rev + Mom	6.73 (11.53)	0.04 (2.97)	0.04 (3.57)	0.07
Best ZP ($\lambda = 0$), FF5	4.95 (6.59)			0.16
Best ZP ($\lambda = 0$), FF5 + ST Rev	4.97 (6.49)	−0.01 (−0.40)		0.16
Best ZP ($\lambda = 0$), FF5 + ST Rev + Mom	4.37 (5.83)	0.02 (0.71)	0.08 (4.36)	0.22

Panel B: The IP with intercept loadings set to zero (cumulative sets)

Characteristics set to zero	Names	Gross	Net	Turnover	Share
None (the IP)	0	2.51	1.60	0.94	0.00
ret_1_0	1	2.37	1.55	0.83	0.29
and 13 daily-return, 21-day characteristics	14	2.16	1.44	0.71	0.57
and seas_1_1an	15	2.16	1.43	0.72	0.58
and bidaskhl_21d, zero_trades_21d	17	1.81	1.18	0.62	0.79

Panel C: Differences in Sharpe ratio

Characteristics set to zero	Minus best ZP, gross	Minus best ZP, net	Minus IP, gross
None (the IP)	0.89 [0.59, 1.26]	0.57 [0.29, 0.89]	—
ret_1_0	0.75 [0.44, 1.11]	0.52 [0.24, 0.85]	−0.14 [−0.27, −0.03]
and 13 daily-return, 21-day characteristics	0.54 [0.18, 0.92]	0.41 [0.08, 0.76]	−0.35 [−0.55, −0.19]
and seas_1_1an	0.54 [0.18, 0.91]	0.40 [0.07, 0.75]	−0.36 [−0.56, −0.20]
and bidaskhl_21d, zero_trades_21d	0.19 [−0.17, 0.53]	0.15 [−0.19, 0.47]	−0.71 [−0.99, −0.51]

IA.N. What Does the Intercept Portfolio Hold?

Section V describes the intercept portfolio through its returns. Figure IA.N.1 describes it through its inputs and positions over the 594 formation months from June 1975 through November 2024, next to the best zero-intercept portfolio. The largest average intercept loadings of the rolling fits are on the 21-day zero-trading-days measure (`zero_trades_21d`), 21-day idiosyncratic volatility (`ivol_capm_21d`), gross profitability (`gp_at` and `gp_atl1`), and 21-day return volatility (`rvol_21d`); the loading on the one-month return (`ret_1_0`) is negative in 98% of the fits (Panel A). The two portfolios are equally broad: the effective number of positions, the inverse of the sum of squared unit-gross weights, averages 1,244 for the intercept portfolio and 1,247 for the best zero-intercept portfolio, the largest 1% of positions hold 4.1% of gross exposure in both, and both hold 51% of their gross exposure in small stocks, which are 49% of the names (Panel C). Their long sides hold 52% and 54% of gross exposure. The intercept portfolio's one distinct tilt is toward the previous month's losers: the average rank of its long positions on the one-month return minus that of its short positions is -0.21 , against -0.12 for the best zero-intercept portfolio, a difference of -0.09 with interval $[-0.12, -0.07]$. On size and on illiquidity the two tilts are close to each other and to zero, -0.02 and 0.03 for both (Panel B). Section IA.M.3 asks how much of the intercept portfolio's return this tilt explains.

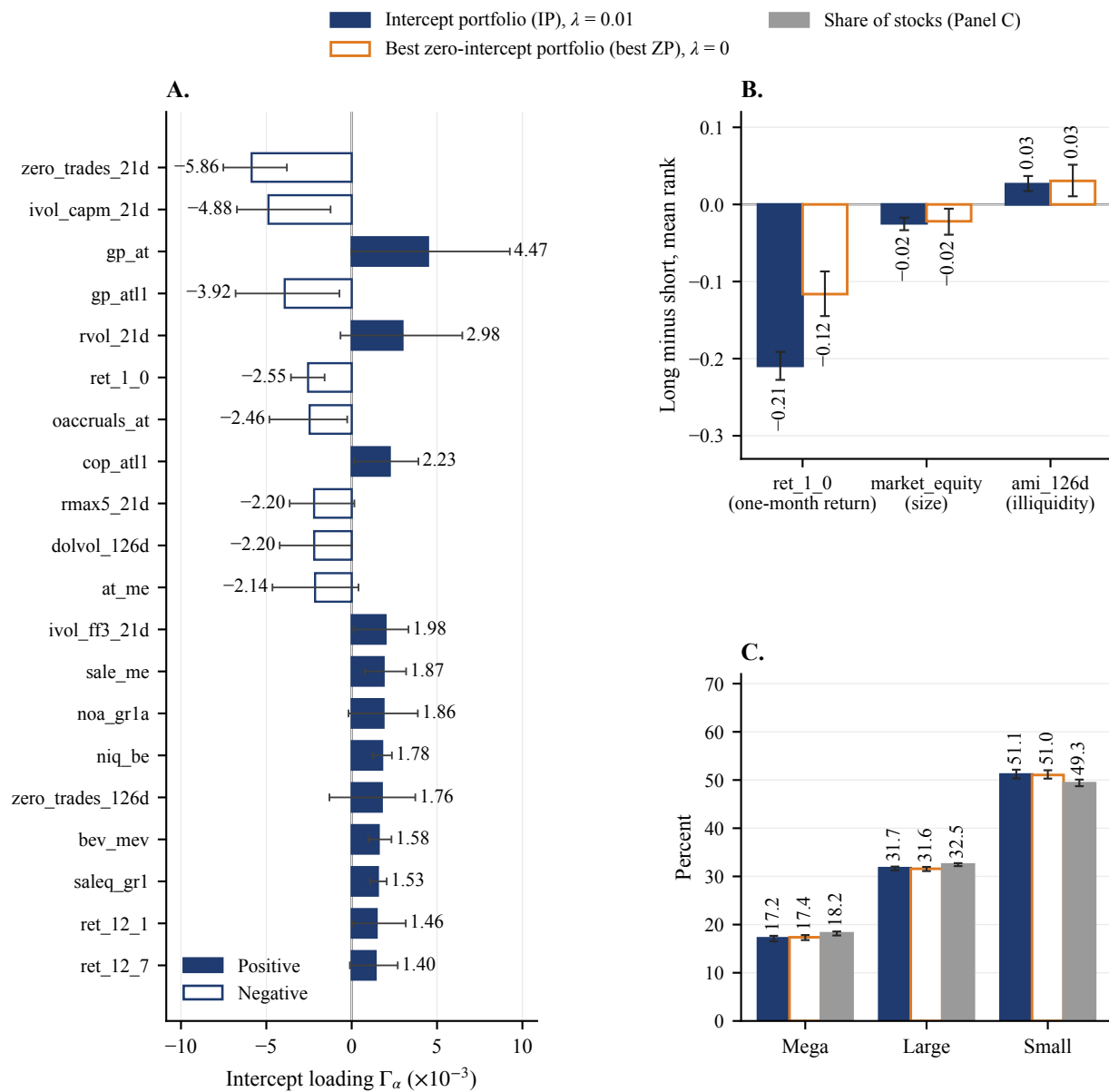


Figure IA.N.1. What the intercept portfolio holds. Panel A: the 20 largest time-series mean intercept loadings $\hat{\Gamma}_\alpha$ in absolute value ($\times 10^{-3}$, standardized characteristics) of the rolling fits of IPCA with an intercept; hollow bars are negative; whiskers: 25th and 75th percentiles across the 594 fits, not a confidence interval. Panel B: mean rank (-0.5 to 0.5) on each characteristic shown of the long minus the short positions, each side weighted by its unit-gross weights; IP ($\lambda = 0.01$) solid, best ZP ($\lambda = 0$) hollow. Panel C: percent of gross exposure in mega-, large-, and small-capitalization stocks; gray: each group's share of the stocks. Whiskers in Panels B and C: pointwise 95% intervals of the time-series mean, 12-month moving-block bootstrap, fitted models held fixed. 594 formation months, June 1975–November 2024.