When Complexity Outpaces Factor Models *

Carter Davis[1] and Alejandro Lopez-Lira[2]

[1] *Indiana University*

[2] *University of Florida*

This Version: March 2025

Abstract

We propose a solution to several asset pricing puzzles by investigating how complexity—the dimensionality of the parameter space relative to the available sample size—affects pricing errors when rational, risk-averse agents have imperfect knowledge of the data-generating process. Our theory yields three implications as complexity increases: (1) pricing errors grow, (2) the optimal portfolio increasingly exploits estimation error components rather than fundamental risk, and (3) multiple strategies can achieve higher Sharpe ratios while maintaining low cross-correlations. Our model explains the limited pricing power of parsimonious factor models, the weak relationship between betas and average returns, and the proliferation of anomalies. Empirically, we document substantial complexity in return predictability and covariance structures. Consistent with our model, systematic hedge funds' active positions exhibit near-zero correlationss, suggesting they exploit distinct dimensions of estimation errors.

*We thank multiple seminar participants at the Chicago Booth Asset Pricing Conference, Carnegie Mellon University, the University of Florida, CUNEF University, ITAM, Seoul National University, NOVA University, the Theory-based Empirical Asset Pricing Research Workshop, the Korean University and the Korea Advanced Institute of Science & Technology for helpful comments and suggestions. Correspondence: ckd1@iu.edu, alejandro.lopez-lira@warrington.ufl.edu. There are no financial disclosures or conflicts of interest.

For decades, asset pricing theory has been driven by the search for parsimonious models that link expected returns to systematic risk factors. Despite their theoretical elegance, the capital asset pricing model (CAPM), the Arbitrage Pricing Theory (APT), and their multi-factor extensions have consistently fallen short in empirical tests. Meanwhile, a puzzling disconnect has emerged: machine-learning strategies generate high Sharpe ratios with little correlation to each other, and hundreds of predictive signals show limited connection to traditional risk factors. This persistent gap between theoretical simplicity and empirical complexity raises a fundamental question: Why do factor models struggle to explain asset prices, and what drives the proliferation of diverse, profitable anomalies?

We argue that complexity—defined as the dimensionality of the parameter space relative to the available sample size—provides the key to resolving this puzzle. Our equilibrium model examines how rational, risk-averse agents navigate high-dimensional markets when faced with uncertainty about expected returns and covariances. Without resorting to market inefficiencies or behavioral biases, we demonstrate that estimation uncertainty alone can explain the empirical patterns observed in asset markets. As complexity increases through either more assets or predictive signals, pricing errors systematically grow, and high Sharpe ratio portfolios increasingly derive their performance from estimation errors rather than fundamental risk. This framework offers a unified explanation for factor model failures, the weak beta-return relationship, and the emergence of diverse, high-Sharpe-ratio strategies—all as natural consequences of rational optimization in complex environments.

Our theory yields three main implications. First, we show that pricing errors of factor models increase linearly with complexity. As complexity increases, the Sharpe ratio gap between the optimal portfolio under the true data-generating process (the portfolio formed using the true inverse covariance matrix and expected returns) and any feasible portfolio investors can construct in equilibrium widens. The Gibbons, Ross, and Shanken (1989) framework links this divergence to larger pricing errors. In equilibrium, the market portfolio's Sharpe ratio, set by risk aversion and market risk, remains stable despite complexity, in contrast to the Sharpe ratio of the optimal portfolio that grows by loading on agents' estimation errors. This result contradicts a key prediction

of APT, which suggests that pricing errors should decrease as the number of assets increases.

Second, we show that optimal portfolio weights capitalize on these opportunities, increasingly deriving profitability from estimation errors rather than fundamental risk premia. Our model yields a novel decomposition of expected returns into three components: mean estimation error, covariance estimation error, and true risk compensation. Hence, the optimal portfolio, the inverse covariance multiplied by excess returns, can be decomposed into three components: (1) the market portfolio, scaled by risk aversion, with a Sharpe ratio largely independent of complexity; (2) weights exploiting errors in expected payoffs; and (3) weights leveraging covariance estimation errors. As complexity grows, the Sharpe ratios of the latter two rise relative to the market's, fueled by increasing estimation uncertainty, while fundamental risk compensation stays relatively flat.

Third, we show that as complexity increases, there is a natural raise of the number of anomalies. In high-dimensional markets, the increasing number of estimation errors creates multiple independent dimensions for exploitation. Strategies focusing on distinct error components naturally show low correlations and high Sharpe ratios, remaining unspanned by one another. As complexity rises, more strategies can attain a given Sharpe ratio while staying weakly correlated, alternatively, a fixed number of low-correlation strategies can achieve higher Sharpe ratios. Intuitively, rising complexity allows more ways to split the increasing number of estimation errors, letting strategies exploit distinct opportunities absent in simpler markets without overlap.

Together, our model proposes a unified solution to several asset pricing puzzles. The limited power of parsimonious factor models arises as complexity increases pricing errors, outpacing what achievable factor models can span. This disconnect manifests temporally: while agents may make optimal decisions given their ex ante information sets, it is only clear ex post, for example when an econometrician analyzes return patterns under the objective data-generating process and identifies apparent inefficiencies. These apparent anomalies reflect estimation uncertainty rather than irrationality. The weak link between betas and average returns follows from these errors, which grow with market dimensionality and obscure systematic risk's role (Daniel and Titman (1997), Lopez-Lira and Roussanov (2022)). Likewise, our framework suggests that sophisticated quanti-

tative strategies achieve high Sharpe ratios while remaining largely unspanned by one another, as they target orthogonal error dimensions (Baba Yara, Boyer, and Davis (2021)). Far from indicating inefficiency or irrationality, these patterns may reflect rational responses to high dimensionality.

Moreover, we document pervasive complexity in the stock market, suggesting we operate in a high-dimensional market. From the 17,517 characteristic-sorted portfolios from Chen, Lopez-Lira, and Zimmermann (2023), 347 portfolios (2%) exhibit absolute t-statistics above 3 over our sample period, with 197 retaining such significance even after 2000, despite evidence of declining return predictability.[1] The covariance structure further reveals this dimensionality: the first ten principal components account for just 20% of return variance, while 75 factors are needed to capture 75% of these diversified portfolios. This high-dimensional environment complicates estimating expected returns and covariances, supporting our model's assumptions.

Beyond supporting our model's high-dimensional assumptions, we also provide empirical evidence for its implications, particularly the emergence of diverse trading strategies in complex markets. Quantitative strategies show little overlap in their return structures: a diverse set of proposed factor models from the literature yields an average pairwise R^2 of just 7%, while the active components of systematic hedge fund portfolios, based on 13F holdings data, show an even lower pairwise average R^2 below 1%. This weak explanatory power among sophisticated investors' positions confirms our model's prediction that high-dimensional markets allow strategies to exploit distinct estimation error dimensions.

Related Literature

Our paper contributes to the literature on economies with high-dimensional properties, building on models that explore complexity in asset pricing. The closest work is Da, Nagel, and Xiu (2022), which adopts a partial equilibrium approach to examine the statistical limits of arbitrage. In their framework, alphas are assumed exogenous, focusing on characterizing attainable Sharpe ratios

1. See Harvey, Liu, and Zhu (2016), McLean and Pontiff (2016), Yan and Zheng (2017), Chen (2021), Chen and Zimmermann (2022).

for investors under uncertainty. By contrast, we develop a general equilibrium model where Sharpe ratios, asset prices, and pricing errors emerge endogenously. Our equilibrium reveals that risk compensation remains largely insensitive to the economy's complexity while pricing errors increase markedly.

Martin and Nagel (2022) model an economy with numerous potential predictors, where risk-neutral Bayesian investors learn about expected returns. We advance their framework by incorporating risk-averse agents and uncertainty in the covariance matrix, enabling us to analyze risk pricing and the relationships among trading strategies. Similarly, Kozak and Nagel (2024) demonstrate that characteristics-based factors can span the stochastic discount factor (SDF) under covariance conditions more readily met with many characteristics. We build on their theoretical spanning results by illustrating why, in practice, such spanning becomes increasingly elusive as complexity rises, driven by estimation errors.

Andrei, Cujean, and Wilson (2023) illustrate how informational disparities between investors and empiricists can bias beta estimates, flattening the Security Market Line. We extend their framework by demonstrating that market complexity alone generates pricing errors, even without dispersed information, as the estimation challenge impacts all participants. Similarly, Merton (1987) employs incomplete information in asset pricing, emphasizing heterogeneous investor recognition as a driver of market outcomes. We build on this line of inquiry by exploring how market complexity alone can sustain persistent pricing anomalies, even among rational, homogeneous investors with shared information. In such high-dimensional settings, anomalies endure despite uniform rationality and access to data.

Finally, we advance the literature on high-dimensional asset pricing by offering an equilibrium-based explanation for the efficacy of complex factor models, such as those in Kelly, Malamud, and Zhou (2022), Bryzgalova, Pelger, and Zhu (2020a), and Didisheim et al. (2024). Our model rests on the assumption of return predictability, supported by a robust body of research, including Fama and MacBeth (1973), Jacobs and Levy (1988), Lewellen (2015), Gu, Kelly, and Xiu (2020), and

Freyberger, Neuhierl, and Weber (2020a), among others.[2]

Next, we introduce our general equilibrium model to explore how rising complexity shapes market outcomes. Featuring rational agents under uncertainty, the model begins with a homogeneous agent economy to analyze pricing errors and optimal portfolio behavior, then extends to heterogeneous information settings to explain the emergence of diverse, uncorrelated trading strategies.

1 Model

We work with economies with an increasing number of assets (N). Let d_t represent the N dimensional column vector of dividends in period t. Similarly, p_t denotes the N-dimensional column vector of asset prices at time t. Share excess returns, dollar excess returns, or excess dollar payoffs are defined as $r_{t+1} = d_{t+1} - (1 + r_f)p_t$ where r_f is a risk-free exogenous rate.[3]

Dividends serve as the primary driver of asset payoffs in our model. Hence, the covariance matrix of returns, denoted as $\Lambda_t = \mathbb{V}\text{ar}_t[d_{t+1}] = \mathbb{V}\text{ar}_t[r_{t+1}]$, remains independent of endogenous prices in the subsequent period, enhancing the model's tractability. We define the expected dividends under the objective data-generating process as $a_t = E_t[d_{t+1}]$ and the expected excess returns as $\mu_t = E_t[r_{t+1}]$. For notational simplicity, we omit time subscripts where the context permits.

Initially, we assume a unit measure of identical agents in the economy, each endowed with constant absolute risk aversion (CARA) utility with risk aversion γ facing a per-asset supply of $\frac{1}{N}$ as in Kozak, Nagel, and Santosh (2018). Additionally, let ι denote the N-dimensional column vector of ones.[4] Later, we extend the framework to accommodate heterogeneous agents.

All proofs are in the Appendix.

2. See Alti and Titman (2019) for an alternative model justifying return predictability.

3. The terms share excess returns, dollar excess returns, and excess dollar payoffs are all used in the literature. Following convention, we adopt this formulation rather than percentage returns.

4. This assumption ensures that under perfect information and common dividend variation, the maximum Sharpe ratio of the remains well-behaved as N increases. Specifically, with a unit asset supply, $SR(w^*)^2 = \mu'\Lambda_t^{-1}\mu = \gamma^2\iota'\Lambda_t\Lambda_t^{-1}\Lambda_t\iota = \gamma^2\iota'\Lambda_t\iota \in O(N^2)$. For uncorrelated assets, $SR(w^*)^2 \in O(N)$.

1.1 Statistical Assumptions

For each asset i, the dividend process is given by:

$$d_{i,t+1} = b'_{i,t} x_{t+1} + c_{i,t} z_{i,t+1} \tag{1}$$

where x_{t+1} is an S-dimensional column vector of state variables, and $b_{i,t}$ is a vector of loadings. These loadings $b_{i,t}$ can be interpreted as firm-specific characteristics— such as size, age, capital structure, productivity, or management quality—that influence dividend outcomes when interacting with productivity shocks captured by x_{t+1}.[5]

Each element of $b_{i,t}$ represents the sensitivity or weight of asset i's dividends to these systematic shocks. For example, $b'_{i,t} x_{t+1}$ can be considered a reduced form specification of the production function, where firm production potentially covaries with each other via common shocks. Moreover, there is a firm-specific state variable $z_{i,t+1}$ with loading $c_{i,t}$.[6]

Stacking the previous equation across all assets, we express the dividend vector as:

$$d_{t+1} = B_t X_{t+1} \tag{2}$$

where d_{t+1} is the N dimensional vector of dividends, and X_{t+1} is the $S + N$ dimensional column vector defined as $X_{t+1} = (x_{1,t+1}, ..., x_{S,t+1}, z_{1,t+1}, ..., z_{N,t+1})'$ and B_t is the $N \times (S+N)$ matrix defined as

$$B_t = \begin{bmatrix} b'_{1,t} & c_{1,t} & \dots & 0 \\ b'_{2,t} & 0 & c_{2,t} \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ b'_{N,t} & 0 & 0 & \dots & c_{N,t} \end{bmatrix}. \tag{3}$$

Under the objective data-generating process, conditional on the full information set at time t,

5. Alternatively, x_{t+1} can represent the coefficients of OLS regressions as in Martin and Nagel (2022).
6. The idiosyncratic component ensures covariance estimation uncertainty persists as N increases.

denoted $\mathcal{F}_t$, the state variables follow:

$$X_{t+1} \mid \mathcal{F}_t \sim N(\pi, V), \tag{4}$$

where π is the mean vector and (V) is the covariance matrix of (X_{t+1}). Consequently, dividends d_{t+1} follow a multivariate normal distribution with mean $B_t\pi$ and covariance B_tVB_t'.

Throughout the paper, we refer to quantities derived from the data-generating process as "objective" or "true" indistinctively. These terms indicate values based on the actual parameters of the economy rather than agents' estimates or beliefs.

We assume investors perfectly observe B_t at time t. For the general case, they lack exact knowledge of π and V, relying instead on prior beliefs and Bayesian updates using observed dividends up to time t. In subsequent sections, we analyze specific scenarios where investors may have perfect knowledge of either π, V, or both to better understand how different types of estimation uncertainty affect equilibrium outcomes.

1.2 Agents' Estimation

We begin with the general case where agents observe neither the mean vector π nor the covariance matrix V of the state variables. Given the dividend process and state variable structure, agents must form beliefs about future dividends based on historical observations. With t observations of the state variables, agents update their beliefs using Bayesian inference.

1.2.1 Prior Distributions

For joint estimation of both the mean vector and covariance matrix, agents use a normal-inverse-Wishart prior, which is the conjugate prior for multivariate normal distributions with unknown mean

and covariance. Specifically, the prior is:

$$V \sim \text{Inverse-Wishart}(\nu_0, V_0^{-1}) \tag{5}$$

$$\pi | V \sim N(\pi_0, V/\kappa_0) \tag{6}$$

where:

- π_0 is the prior mean vector for π
- $\kappa_0 > 0$ represents the strength of the prior belief about π
- $\nu_0 > S + N - 1$ is the degrees of freedom parameter for the inverse-Wishart distribution
- V_0 is the prior scale matrix for V

To ensure the posterior variance remains well-defined as N increases, we set $\nu_0 = N + S + \delta$ for some $\delta > 1$. This guarantees that when updating with data, the resulting degrees of freedom ν_t will exceed 2, which is necessary for the variance of the t-distribution to be defined.

1.2.2 Posterior Distributions

Let $\mathcal{I}_t$ denote the investors' information set at time t, consisting of observed state variables up to t. Variables denoted with a hat (e.g., $\hat{a}_t$, $\hat{\Lambda}_t$) represent agents' subjective estimates based on $\mathcal{I}_t$. The posterior mean of π is:

$$\hat{\pi}_t = \frac{\kappa_0 \pi_0 + t \bar{X}_t}{t + \kappa_0}, \tag{7}$$

where $\bar{X}_t = \frac{1}{t} \sum_{\tau=1}^{t} X_\tau$ is the sample mean. The scale matrix for the predictive t-distribution is defined as:

$$\hat{\Sigma}_t = \frac{\kappa_0 + t + 1}{(\kappa_0 + t)(\nu_0 + t - (S+N) + 1)} \left(V_0 + \sum_{\tau=1}^{t} (X_\tau - \bar{X}_t)(X_\tau - \bar{X}_t)' + \frac{\kappa_0 t}{\kappa_0 + t} (\bar{X}_t - \pi_0)(\bar{X}_t - \pi_0)' \right), \tag{8}$$

with degrees of freedom updated as $\nu_t = \nu_0 + t - (S + N) + 1$.[7] To ensure the posterior variance remains well-defined as N increases, we set $\nu_0 = N + S + \delta$ for some $\delta > 1$, which guarantees that $\nu_t = t + \delta + 1 > 2$ for all $t \geq 0$. Together, these imply a posterior predictive distribution for X_{t+1} given by:

$$X_{t+1} \mid \mathcal{I}_t \sim t_{\nu_t}(\hat{\pi}_t, \hat{\Sigma}_t), \tag{9}$$

where t_{ν_t} denotes a multivariate t-distribution with ν_t degrees of freedom, location $\hat{\pi}_t$, and scale matrix $\hat{\Sigma}_t$.

Due to the computational complexity of the t-distribution, agents approximate it as:

$$X_{t+1} \mid \mathcal{I}_t \sim N(\hat{\pi}_t, \hat{V}_t), \tag{10}$$

where the covariance matrix is adjusted as $\hat{V}_t = \frac{\nu_t}{\nu_t - 2}\hat{\Sigma}_t$ to match the variance of the t-distribution under this normal approximation.

This belief about state variables translates directly to beliefs about dividends. Since $d_{t+1} = B_t X_{t+1}$, investors perceive dividends as normally distributed:

$$d_{t+1} \mid \mathcal{I}_t \sim N(B_t\hat{\pi}_t, B_t\hat{V}_t B_t'), \tag{11}$$

with expected dividends $\hat{a}_t = B_t\hat{\pi}_t$ and covariance matrix $\hat{\Lambda}_t = B_t\hat{V}_t B_t'$.

Consequently, investors perceive excess returns as:

$$r_{t+1} \mid \mathcal{I}_t \sim N(\hat{\mu}_t, \hat{\Lambda}_t), \tag{12}$$

where $\hat{\mu}_t = \hat{a}_t - (1 + r_f)p_t$ represents their expected excess returns and $\hat{\Lambda}_t = B_t\hat{V}_t B_t'$ is the perceived covariance matrix of returns. These subjective beliefs shape agents' demands and, in

7. Note that $\hat{\Sigma}_t$ as defined here is not simply the posterior mean of V from the Inverse-Wishart distribution. Rather, it is the scale parameter appearing in the predictive Student-t distribution for X_{t+1}. These differ by factors involving $\nu_0 + t$ and $\kappa_0 + t$. The additional factor $\frac{\kappa_0+t+1}{(\kappa_0+t)}$ is precisely the standard correction that appears in predictive distribution formulas when integrating out both π and V.

turn, equilibrium prices, as explored next.

For the special cases where agents know either π, V, or both, the estimation problem simplifies accordingly. When π is known, agents use the true mean directly ($\hat{\pi}_t = \pi$) with no estimation error. When V is known, the covariance estimation becomes exact ($\hat{V}_t = V$). [8]

1.2.3 Special Case: Known Covariance Matrix

When the covariance matrix V is known, agents only estimate the mean vector π. The posterior mean remains:

$$\hat{\pi}_t = \frac{\kappa_0 \pi_0 + t\bar{X}_t}{t + \kappa_0}, \tag{13}$$

but the predictive distribution simplifies to:

$$X_{t+1} \mid \mathcal{I}_t \sim N\left(\hat{\pi}_t, V + \frac{V}{t + \kappa_0}\right), \tag{14}$$

where the second term represents parameter uncertainty.

1.2.4 Special Case: Known Mean Vector

With known mean vector π, the posterior scale matrix becomes:

$$\hat{\Sigma}_t = \frac{1}{\nu_0 + t - (S+N) + 1}\left(V_0 + \sum_{\tau=1}^{t}(X_\tau - \pi)(X_\tau - \pi)'\right), \tag{15}$$

with degrees of freedom $\nu_t = \nu_0 + t$. Using a normal approximation to the t-distribution:

$$X_{t+1} \mid \mathcal{I}_t \sim N(\pi, \hat{V}_t), \tag{16}$$

where $\hat{V}_t = \frac{\nu_t}{\nu_t - 2}\hat{\Sigma}_t$ for $\nu_t > 2$.

8. These complete information scenarios can be formally represented as limiting cases of our Bayesian framework. For known π, set the prior mean $\pi_0 = \pi$ with $\kappa_0 \to \infty$, which forces $\hat{\pi}_t = \pi$ for all t. For known V, set $V_0 = \nu_0 V$ with $\nu_0 \to \infty$, ensuring $\hat{V}_t = V$. When both parameters are known, apply both adjustments simultaneously to eliminate all estimation uncertainty, resulting in agents' beliefs matching the true distribution: $d_{t+1} \mid \mathcal{I}_t \sim N(B_t\pi, B_t V B_t')$.

1.3 Agents' Demand

Given that CARA agents perceive the world as normally distributed, and there are no transaction costs, their demand for assets takes the familiar mean-variance form:

$$x_t = \frac{1}{\gamma}\hat{\Lambda}_t^{-1}\hat{\mu}_t = \frac{1}{\gamma}\hat{\Lambda}_t^{-1}(\hat{a}_t - (1+r_f)p_t). \tag{17}$$

Where $\hat{\Lambda}_t$ is the agents' estimated covariance matrix of returns, $\hat{\mu}_t$ is their expected excess returns, $\hat{a}_t$ is their expected dividend forecast, and p_t is the asset price vector. This demand reflects agents' optimal portfolios under their subjective beliefs, which typically deviate from the true optimal portfolio due to imperfect knowledge of the actual covariance Λ_t and expected dividends a_t. In the limiting case of perfect information, agents' estimates align with reality: $\hat{\Lambda}_t = \Lambda_t$ and $\hat{a}_t = a_t$.[9]

1.4 Equilibrium

Given a per-asset supply of $\frac{1}{N}$, market clearing requires $x_t = \frac{1}{N}\iota$, where ι is the N-dimensional vector of ones. Substituting the demand function into this condition and solving for prices yields the equilibrium price vector:

$$p_t = \frac{1}{1+r_f}\left(\hat{a}_t - \frac{\gamma}{N}\hat{\Lambda}_t\iota\right). \tag{18}$$

For convenience, define $\gamma_N = \frac{\gamma}{N}$ as the scaled risk aversion. The true expected excess returns under the data-generating process are then:

$$\text{Expected Returns} = \mu_t \equiv \mathbb{E}_t[d_{t+1}] - (1+r_f)p_t \tag{19}$$

$$= a_t - (1+r_f)p_t = (a_t - \hat{a}_t) + \gamma_N\hat{\Lambda}_t\iota \tag{20}$$

9. Given the CARA utility structure with normally distributed returns, price taking, and absence of transaction costs, the demand function is the same if we consider single-period agents that observe the entire history of dividends or infitely lived agents.

This expression can be decomposed into three components:

$$\mu_t = \underbrace{(a_t - \hat{a}_t)}_{\text{mean estimation error}} + \underbrace{\gamma_N(\hat{\Lambda}_t - \Lambda_t)\iota}_{\text{covariance estimation error}} + \underbrace{\gamma_N \Lambda_t \iota}_{\text{risk compensation}} \tag{21}$$

Here, the mean estimation error captures discrepancies between true and perceived expected dividends, while the covariance estimation error reflects inaccuracies in agents' risk assessments. The risk compensation term, $\gamma_N \Lambda_t \iota$, represents the premium demanded for bearing systematic risk, scaled by the economy's risk aversion and asset supply.

In contrast, agents' subjective expected returns are simpler: $\hat{\mu}_t = \hat{a}_t - (1 + r_f)p_t = \gamma_N \hat{\Lambda}_t \iota$. This reflects their perceived risk premia, driven solely by their estimated covariance.

If taken as a primitive, the demand function in equation 17 can accommodate various estimation methods beyond the Bayesian approach detailed in the previous section. Alternative preference structures or belief formation mechanisms—including regression-based forecasts, or even sophisticated machine-learning algorithms—can be viewed as different ways to generate $\hat{\Lambda}_t$ and $\hat{a}_t$. Regardless of how beliefs are formed, equilibrium prices and expected returns will reflect both estimation errors and risk premia in a consistent manner, as long as agents act as feasible mean-variance optimizers using their subjective estimates of means and covariances.

1.5 Objective Optimal Portfolio

In equilibrium, market clearing requires that agents are satisfied holding the market portfolio given their subjective beliefs. That is, prices adjust until agents' optimal portfolios (based on their subjective estimates of means and covariances) align with the available supply. Hence, the market is subjectively optimal. However, the market portfolio will generally be different from the optimal portfolio under complete knowledge of the economy's parameters.

Given the vector of expected returns in equilibrium, the weights of this objective optimal

portfolio are given by

$$w_t^* = \Lambda_t^{-1}\mu_t = \underbrace{\Lambda_t^{-1}(a_t - \hat{a}_t)}_{\text{errors in means}} + \underbrace{\gamma_N \Lambda_t^{-1}(\hat{\Lambda}_t - \Lambda_t)\iota}_{\text{errors in covariances}} + \underbrace{\gamma_N \iota}_{\text{market portfolio}} \tag{22}$$

Intuitively, given a vector of prices formed with imperfect information beliefs, the optimal strategy is to form a portfolio that exploits these estimation errors arising from both the covariance and expected returns while also having a market portfolio component exploiting the positive risk premium. Hence, the stochastic discount factor (SDF) must load on agents' estimation errors.

1.6 Pricing Errors

Let $\sigma^2_{m,t} = \iota'\Lambda_t\iota$ denote the market variance and $\hat{\sigma}^2_{m,t} = \iota'\hat{\Lambda}_t\iota$ the perceived variance. Further, let $\varepsilon = a_t - \hat{a}_t$ denote the mean estimation error. In contrast to its expected return, the market variance does not depend on agents' expectations since prices adjust so that investors hold the same share of each asset, and the variance depends only on the dividends.

Proposition 1 (Decomposition of Pricing Errors)**.** *The weighted sum of squared pricing errors in the economy, relative to the market portfolio, is given by the difference in squared Sharpe ratios (SR) of the optimal portfolio and the market portfolio. This sum can be decomposed into mean and covariance estimation errors plus an interaction term.*

$$\alpha'\Lambda_t^{-1}\alpha = SR(w^*)^2 - SR(\frac{1}{N}\iota)^2 \tag{23}$$

$$= \underbrace{\left[\varepsilon'\,\Lambda_t^{-1}\,\varepsilon \;-\; \frac{(\iota'\varepsilon)^2}{\sigma^2_{m,t}}\right]}_{\textit{Mean estimation errors}} + \underbrace{2\,\gamma_N\left[\varepsilon'\,\Lambda_t^{-1}\,\hat{\Lambda}_t\,\iota - \frac{(\iota'\varepsilon)\,\hat{\sigma}^2_{m,t}}{\sigma^2_{m,t}}\right]}_{\textit{Interaction term}} + \underbrace{(\gamma_N)^2\left[\iota'\,\hat{\Lambda}_t\,\Lambda_t^{-1}\,\hat{\Lambda}_t\,\iota \;-\; \frac{(\hat{\sigma}^2_{m,t})^2}{\sigma^2_{m,t}}\right]}_{\textit{Covariance estimation errors}} \tag{24}$$

This decomposition builds on the insight from Gibbons, Ross, and Shanken (1989) that pricing errors can be expressed as differences in squared Sharpe ratios, but extends their framework to the equilibrium case where agents have uncertainty about means and covariances. The first term

captures errors in estimating means, comparing how the correct weighting exploits mean estimation errors versus market weights. The second term represents the interaction between mean estimation errors and differences in perceived versus true covariance weighting. The third term isolates the pure covariance mispricing effect, measuring how differences between estimated and true covariances affect the Sharpe ratio.

Importantly, the market portfolio in this economy is the correct factor model from agents' perspective–they perceive it as the mean-variance efficient portfolio given their beliefs about expected returns and covariances. An econometrician analyzing historical returns, however, would observe that the market portfolio is not truly mean-variance efficient under the objective data-generating process. The econometrician would detect apparent pricing errors and might conclude that additional factors beyond the market are needed to explain returns. These "factors" would effectively capture the estimation errors made by agents.

Equation 23 characterizes the pricing errors given the statistical properties of the dividends and agents' estimates about them. In what follows, we consider specific cases with perfect information, and when means or covariances are known.

1.7 Perfect Information

The case of perfect information, where agents know the true mean and covariance, is useful as a benchmark.

Proposition 2. *When agents know the economy's parameters, CAPM and Arbitrage Pricing Theory (APT) hold. In particular,*

1. *Mean returns are given by* $\mu_t = \gamma_N \Lambda_t \iota$.
2. *The mean-variance efficient portfolio is the market portfolio with weights of* $\frac{1}{N}$.
3. *Pricing errors with respect to the market portfolio are zero,* $\alpha_t = 0$.
4. *For* $N \to \infty$*, and rotating the factors to make them orthogonal, the price of risk for factor*

k is given by $\gamma \bar{b}_{t,k} \sigma^2_{t,k}$, where $\bar{b}_{t,k}$ is the average loading on factor k and $\sigma_{t,k}$ is the factor's volatility.

The intuition for CAPM is straightforward: mean-variance optimizers will only choose optimal portfolios. Hence, prices adjust so that agents are satisfied with holding the market portfolio by making the market portfolio mean-variance efficient.

For the APT, consider a setting with uncorrelated normal factors[10]. Further, assume N going to infinity, so that idiosyncratic risk is completely diversifiable. For asset i, the risk compensation simplifies to:

$$(\gamma_N \Lambda_t \iota)_i = \gamma \sum_{k=1}^{S} b_{i,t,k} \bar{b}_{t,k} \sigma^2_{t,k}, \tag{25}$$

Hence, APT holds with a price of risk given by:

$$\gamma \bar{b}_{t,k} \sigma^2_{t,k}, \tag{26}$$

where $\bar{b}_{t,k} = \frac{1}{N} \sum_{j=1}^{N} b_{j,t,k}$ is the average loading (or beta) on factor k, $b_{i,t,k}$ is asset i's loading on factor k, and $\sigma_{t,k}$ is the factor's volatility.

For instance, a pure-play portfolio with a beta of 1 on factor m and zero on all others has an expected excess return of:

$$\gamma \bar{b}_{t,m} \sigma^2_{t,m}, \tag{27}$$

where $\gamma \bar{b}_{t,m} \sigma^2_{t,m}$ is the price of risk for factor m. This term quantifies the compensation per unit of exposure to factor m, determined by the market's average loading, factor volatility, and scaled risk aversion. Moreover, a factor's exposure is only priced if its average beta across the market is non-zero. When the average loading is zero, positive and negative exposures across assets cancel out in the aggregate, leaving no systematic risk for the market to price.

For example, consider decompositions such as Kozak, Nagel, and Santosh (2018), where the first factor is an equal-weighted portfolio (i.e., $q_1 = \sqrt{\frac{1}{N}} \iota$) and the rest are long-short portfolios

10. Achievable via orthogonal rotation, albeit the interpretation of betas would change.

orthogonal to it ($q_k'\iota = 0$ for $k > 1$). Factors with $\bar{b}_{t,k} = 0$, akin to their long-short PCs, should command no premium, as they lack systematic impact. In the case of uncorrelated factors, the optimal weights $w_f = Cov(F)^{-1}E(f)$ should have zero weight on all the factors but the first. This also partially explains why in Lopez-Lira and Roussanov (2022), it is optimal to hedge most factors. If the factors betas are net zero across the economy, then the Sharpe ratio will be increasing in the number of hedged factors.

1.8 Known Means and Unknown Covariances

Now consider the case, where agents know the true expected value of the fundamental but not the covariance.

Proposition 3. *When agents know the expected value of dividends, a, but have uncertainty about the covariance, Λ, neither CAPM nor APT hold. Further, pricing errors are, on expectation, increasing in complexity defined as the ratio of the number of state variables plus the number of firms $N + S$ over t, the number of available observations. In particular:*

1. *Mean returns are given by $\mu_t = \gamma_N(\hat{\Lambda}_t - \Lambda_t)\iota + \gamma_N\Lambda_t\iota$, and hence APT fails.*
2. *CAPM fails, and the mean-variance efficient portfolio weights are given by $\gamma_N\Lambda_t^{-1}(\hat{\Lambda}_t - \Lambda_t)\iota + \gamma_N\iota$.*
3. *Pricing errors are given by $\alpha'\Lambda_t^{-1}\alpha = (\gamma_N)^2\left[\iota'\,\hat{\Lambda}_t\,\Lambda_t^{-1}\,\hat{\Lambda}_t\,\iota\; -\; \frac{(\hat{\sigma}^2_{m,t})^2}{\sigma^2_{m,t}}\right]$.*

Moreover, we have that APT fails and the opposite intuition holds: for a fixed number of observations, price errors with respect to the market portfolio (agents' best factor model) are increasing in the number of assets:

$$\mathbb{E}[\alpha_t'\Lambda^{-1}\alpha_t] \geq \frac{\gamma^2}{N^2}\left(\frac{(S+N)+t}{(S+N)+t-2}\right)^2\left(\frac{1}{t+1}\right)^2 t(S+N-1)(S+N)^2\bar{\varrho} \sim O\left(\frac{N+S}{t}\right), \tag{28}$$

where ϱ is

$$\bar{\varrho} = \frac{1}{((S+N))^2} \sum_{i=1}^{(S+N)} \sum_{j=1}^{(S+N)} b_{i,t} b_{j,t} V_{i,j} \tag{29}$$

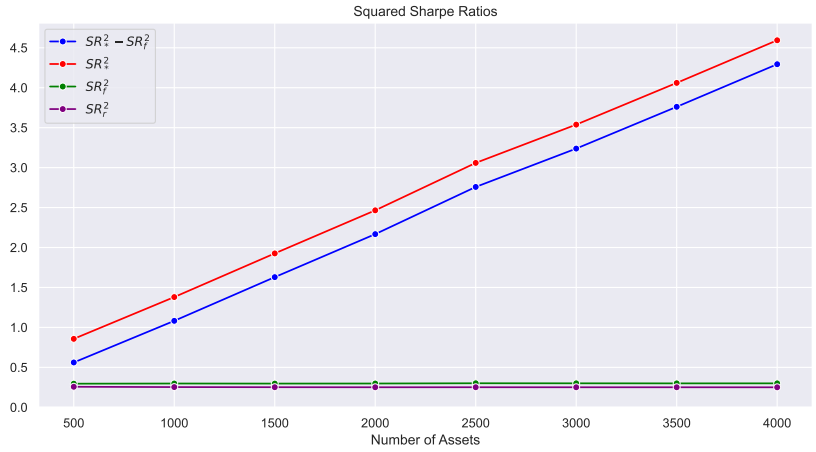
This asymptotic behavior assumes the number of systematic factors S is small relative to the number of assets N (i.e., $S \ll N$), thereby simplifying $(S+N)^3/N^2 \approx N + S$ for large N.

To illustrate the implications of this result, we provide a simple calibration. Figure 1 depicts the components of the weighted squared pricing errors for a simulated economy calibrated to target an equilibrium Sharpe ratio of 0.5 when there is perfect information. We assume that asset payoffs are driven by a multi-factor model with 5 factors. The first factor has a dominant influence, with a variance of 0.04, while the remaining factors each contribute a smaller total variance of 0.01. The loadings (Betas) for the first factor are centered around 1, so most assets have significant exposure to the primary market factor, while the other factors have Betas centered around 0. The total systematic variance from all factors is 0.05, and we adjust the idiosyncratic variance to 0.16 to achieve a target total variance of 0.25 for each stock.

Agents in the economy know the expected dividend payoffs and have to estimate the covariance Λ_t from sample realizations. SR_*^2, SR_f^2, SR_r^2 denote the squared Sharpe ratios implied by the optimal portfolio under the true data-generating process, the feasible optimal weights, and the weights in a rational expectation equilibrium. In figure 1a, agents regularize their estimation using one factor extracted from PCA, while in figure 1b, agents shrink their estimate towards the sample diagonal matrix.

Neither the feasible nor the rational equilibrium squared Sharpe ratios depend on N. However, the optimal portfolio under the true data-generating process Sharpe ratio is, on expectation, a linear function of N. Intuitively, as the number of assets grows, the estimation gets more challenging, and the number of mispriced assets increases. Hence, there are better optimal combinations. The squared pricing errors increase because the optimal portfolio under the true data-generating process Sharpe ratio increases but not the feasible one.

Figure 1: Squared Sharpe Ratios as a Function of the Number of Assets



(a) Agents Using Sample Covariance Matrix



(b) Agents Using Regularized Sample Covariance Matrix

The figure shows the Sharpe ratios of a simulated economy with different numbers of assets. The economy has a one-factor model structure in assets' payoffs with $\beta \sim N(1, 0.25^2)$, variance of aggregate factor = 0.2^2, idiosyncratic variance = 0.45^2, $\gamma = 2.5$, and $t = 252$. The blue line denotes the difference in squared Sharpe ratios between the optimal portfolio under the true data-generating process and feasible squared Sharpe ratios. The red line denotes the optimal portfolio under the true data-generating process squared Sharpe ratio. The green line denotes the feasible squared Sharpe ratio. The purple line denotes the perfect information squared Sharpe ratio.

1.8.1 Unknown Means and Known Covariances

Similarly to the previous case, the pricing errors are generally increasing in the number of assets or states. This result demonstrates how estimation uncertainty in expected returns leads to systematic mispricing even when agents have perfect knowledge of the covariance structure.

Proposition 4. *If investors have perfect information about covariances but only the correct prior about the expectation, neither CAPM nor APT hold. In particular:*

1. *Mean returns are given by* $\mu_t = (a_t - \hat{a}_t) + \gamma_N \Lambda_t \iota$*, and hence APT fails.*
2. *CAPM fails, and the mean-variance efficient portfolio weights are given by* $\Lambda_t^{-1}(a_t - \hat{a}_t) + \gamma_N \iota$*.*
3. *Pricing errors are given by* $\alpha' \Lambda_t^{-1} \alpha = \varepsilon' \Lambda_t^{-1} \varepsilon - \frac{1}{\sigma_m^2}(\iota' \varepsilon)^2$*, where* $\varepsilon = a_t - \hat{a}_t$*.*

Moreover, the average pricing errors relative to the feasible optimal portfolio increase in the number of assets and states at a rate of $\frac{N+S}{t}$*:*

$$\mathbb{E}[\alpha_t' \Lambda_t^{-1} \alpha_t] \geq \frac{t}{(t+\kappa_0)^2}(N+S) - \frac{t}{(t+\kappa_0)^2} = \frac{t}{(t+\kappa_0)^2}(N+S-1) \sim O(\frac{N+S}{t}) \tag{30}$$

1.8.2 Unknown Means and Covariances

The last case of interest is considering unknown means and covariances. For tractability, we consider the case where the estimation error is approximately zero on average, $E[\iota' \varepsilon] \approx 0$, and a slightly stronger approximation imposing that $E[\iota' A \varepsilon] \approx 0$ if A is non-random and positive definite. In that case, the total pricing errors can be approximated by the Effectively, sum of both estimation errors about expected values and covariances.

$$E[\alpha'\Lambda_t^{-1}\alpha] \approx E[\varepsilon'\Lambda_t^{-1}\varepsilon] + E[\frac{\gamma^2}{N^2}\iota'\hat{\Lambda}_t\Lambda_t^{-1}\hat{\Lambda}_t\iota] - E[\frac{1}{\sigma_m^2}(\frac{\gamma^2}{N^2}(\hat{\sigma}_m^2)^2)] \tag{31}$$

$$= \underbrace{E[\varepsilon'\Lambda_t^{-1}\varepsilon] - E[\frac{1}{\sigma_m^2}(\iota'\varepsilon)^2]}_{\alpha \text{ from unknown means}} + \underbrace{E[\frac{\gamma^2}{N^2}(\iota'\hat{\Lambda}_t\Lambda_t^{-1}\hat{\Lambda}_t\iota] - E[\frac{(\iota'\hat{\Lambda}_t\iota)^2}{\iota'\Lambda_t\iota})]}_{\alpha \text{ from unknown means covariances}}. \tag{32}$$

$$\tag{33}$$

2 Heterogeneous Beliefs

The homogeneous agent setting can be interpreted as a special case of our framework where only an atomistic (measure zero) set of agents might have correct beliefs about the data-generating process. However, if nobody has perfect information about the means and covariances, a natural question arises: why do we care about pricing errors? There are two key reasons. First, an econometrician testing asset pricing models ex-post will find large pricing errors relative to standard factor models. Specifically, while agents at time t make optimal decisions given their beliefs, an econometrician analyzing the realized returns at time $t+1$ observes that the market portfolio wasn't truly mean-variance efficient under the objective data-generating process. These apparent anomalies arise not from irrationality but from the fundamental challenge of estimating high-dimensional parameters with limited data. Second, in real markets, there may be a subset of agents that have partial or complete information about the data-generating process, yet mispricing persists.

To understand how mispricing persists in equilibrium even when a significant fraction of investors knows the true data-generating process, we extend the model to include heterogeneous beliefs. Throughout this section, we maintain standard assumptions that agents agree to disagree, are price takers, and do not learn from prices. While the model could be extended to relax these assumptions, we leave such extensions for future research.

For analytical tractability, we first consider a case where investors learn about the mean of X_t and know covariances. This setup makes the problem more intuitive, but it is straightforward to

adapt it to unknown means and unknown covariances at the cost of more algebra. Investors use the t observations to learn about X_t. We consider I different types of investors with different priors, with θ_i mass for each investor type i.

Each investor i has a prior belief about π:

$$\pi \sim N(\pi_i, V_i) \tag{34}$$

where $V_i = k_i V$, and $k_i > 0$ is a scalar representing investor i's confidence in their prior. A larger k_i indicates more uncertainty about the prior mean.

After observing t periods of data $\{X_t\}_{t=1}^{t}$, investor i updates their belief to:

$$\pi_{\text{post},i} = \alpha_i \pi_i + (1 - \alpha_i)\bar{X} \tag{35}$$

where

$$\alpha_i = \frac{1}{1 + k_i t}, \quad \bar{X} = \frac{1}{t}\sum_{\tau=1}^{t} X_\tau \tag{36}$$

and the posterior variance of the mean is:

$$V_{\text{post},i} = (1 - \alpha_i)V \tag{37}$$

In the Appendix, we show that the variance-covariance matrix perceived by investor i combines fundamental uncertainty with parameter uncertainty:

$$\Lambda_i = B_t V \left(\frac{1 + k_i(t+1)}{1 + k_i t} \right) B_t^\top = \Lambda_t \varphi_i \tag{38}$$

where $\Lambda_t = B_t V B_t^\top$ is the variance of dividends and $\varphi_i = \frac{1+k_i(t+1)}{1+k_i t} > 1$ adjusts for uncertainty about means. This additional uncertainty leads investors to take more conservative positions.

Given their beliefs, each investor's demand follows a mean-variance optimization:

$$x_{i,t} = \frac{1}{\gamma}\Lambda_t^{-1}\varphi_i^{-1}\left(B_t\pi_{\text{post},i} - (1+r_f)p_t\right). \tag{39}$$

This demand combines the standard mean-variance portfolio Λ_t^{-1}(expected excess returns) with an adjustment factor φ_i^{-1} that reduces positions based on parameter uncertainty.

Market clearing requires that the aggregate demand equals the per capita supply of risky assets:

$$\frac{\iota}{N} = \sum_{i=1}^{I}\theta_i x_{i,t}. \tag{40}$$

Solving for equilibrium prices yields:

$$p_t = \frac{1}{1+r_f}\left(\hat{a}_E - \frac{\gamma}{NW}\Lambda_t\iota\right), \tag{41}$$

where $\hat{a}_E = B_t\left(\frac{1}{W}\sum i = 1^I w_i\pi_{\text{post},i}\right)$ represents the economy's weighted average expectation. The weights $w_i = \theta_i\varphi_i^{-1}$ and total weight $W = \sum_{i=1}^{I} w_i$ $(0 < W < 1)$ reflect both the mass of each investor type and their uncertainty-adjusted demand.

The resulting vector of true expected returns mirrors the homogeneous case but with aggregate beliefs:

$$\mu_t \equiv \mathbb{E}_t[d_{t+1}] - (1+r_f)p_t = a_t - \hat{a}_E + \frac{\gamma}{NW}\Lambda_t\iota. \tag{42}$$

The expected returns compensate for both fundamental risk ($\frac{\gamma}{NW}\Lambda_t\iota$) and loads on the estimation errors ($a_t - \hat{a}_E$). Each investor type perceives different expected returns based on their beliefs:

$$\hat{\mu}_{i,t} \equiv \mathbb{E}_{i,t}[d_{t+1}] - (1+r_f)p_t = a_i - \hat{a}_E + \frac{\gamma}{NW}\Lambda_t\iota. \tag{43}$$

This heterogeneity in beliefs leads each agent to perceive mispricing relative to their own valuation. Defining $\varepsilon_i = a_i - \hat{a}_E$ as the perceived mispricing, each investor's optimal portfolio combines a

speculative component trading on perceived mispricing with a market component:

$$x_i = \frac{1}{\gamma\varphi_i}\Lambda_t^{-1}\varepsilon_i + \frac{1}{N\varphi_i W}\iota. \tag{44}$$

The pricing error properties from the homogeneous case persist in a straightforward way through the weighted average expectation $\hat{a}_E$. Since $\hat{a}_E$ is a weighted average of each investor type's beliefs, and each investor faces growing estimation challenges as complexity increases (higher N or S), the estimation errors of $\hat{a}_E$ itself grow with market complexity. In other words, the market's aggregate expectation inherits the same fundamental relationship between complexity and estimation error that we found in the homogeneous case. Even informed investors do not eliminate these pricing errors due to risk aversion limiting the size of their positions.

2.1 Two Types of Investors

To better understand how mispricing persists in equilibrium even with informed investors, consider a special case with two types of investors. Type 1 investors are fully informed about the true mean μ and have complete confidence in their knowledge ($k_1 = 0$). Type 2 investors are less informed, with a prior belief $\mu_2 \neq \mu$ and confidence level $k_2 > 0$.

The posterior beliefs for Type 1 (informed) investors are straightforward: $\pi_{\text{post},1} = \mu$, $V_{\text{post},1} = 0$, and $\varphi_1 = 1$.

For Type 2 (uninformed) investors, the posterior beliefs reflect both their prior and the data:

$$\pi_{\text{post},2} = \alpha_2\mu_2 + (1-\alpha_2)\bar{X} \tag{45}$$

$$V_{\text{post},2} = \frac{k_2 V}{1+k_2 T} \tag{46}$$

$$\varphi_2 = \frac{1+k_2(T+1)}{1+k_2 T} > 1 \tag{47}$$

where $\alpha_2 = \frac{1}{1+k_2 T}$.

The weighted expectations of the economy combine these beliefs according to:

$$\hat{a}_E = B_t \left(\frac{\theta_1 \mu + \theta_2 \varphi_2^{-1} (\alpha_2 \mu_2 + (1 - \alpha_2) \bar{X})}{\theta_1 + \theta_2 \varphi_2^{-1}} \right) \tag{48}$$

where $w_1 = \theta_1$ is the weight of informed investors, $w_2 = \theta_2 \varphi_2^{-1}$ is the (effective) weight of unformed investors, and $W = w_1 + w_2$ is the total weight.

The estimation error in this two-type model reveals why mispricing persists:

$$\begin{aligned} \varepsilon &= a - \hat{a}_E && (49) \\ &= B_t \mu - B_t \left(\frac{1}{W} (w_1 \mu + w_2 \pi_{\text{post},2}) \right) && (50) \\ &= B_t \left(\frac{w_2}{W} \left[\alpha_2 (\mu - \mu_2) + (1 - \alpha_2)(\mu - \bar{X}) \right] \right) && (51) \end{aligned}$$

This decomposition shows that mispricing persists even with informed investors because of two key factors. First, risk aversion limits the size of positions that informed investors can take, preventing them from completely eliminating mispricing. Second, the presence of less informed investors with different beliefs creates a weighted average expectation that differs from the true mean, leading to systematic deviations in prices. The magnitude of these deviations depends on both the relative mass of informed versus uninformed investors (θ_1 vs θ_2) and the confidence level of the uninformed investors (k_2). As before, the estimation errors increases with complexity.

2.2 Sophisticated Investors with Partial Information

The previous sections demonstrated how market complexity impacts pricing errors and the limitations of parsimonious factor models. We now extend this analysis by examining how sophisticated investors with partial information can generate profitable trading strategies that maintain low correlations with each other, even as market complexity increases. This finding helps explain the empirical observation that different quantitative strategies can simultaneously achieve high Sharpe ratios while remaining largely uncorrelated.

Consider sophisticated investors who have perfect information about a subset of the available assets. We model this partial information structure by assuming each investor i has access to a random subset of assets, represented by a diagonal selection matrix S_i where each diagonal element equals 1 with probability q (indicating the asset is known/accessible to the investor) and 0 otherwise. This captures the realistic constraint that even sophisticated investors cannot perfectly analyze and trade all available securities.

We consider the case where atomistic informed investors trade in an otherwise homogeneous agents economy as in the first section. For tractability, all agents have full information about the covariance matrix Λ_t, and we assume the homogeneous agents have the correct prior about expected payoffs.

For tractability, we assume the payoffs' covariance matrix follows a one-factor structure:

$$\Lambda = \beta\beta'\sigma_m^2 + \sigma_\epsilon^2 I \tag{52}$$

where β captures aggregate factor exposures, σ_m^2 is the aggregate variance, and σ_ϵ^2 represents idiosyncratic risk. While this simplified structure might appear restrictive, it captures the essential feature needed for our analysis: the tension between diversifiable idiosyncratic risk and undiversifiable systematic risk. The key insights about correlation scaling emerge from the interplay between estimation errors and risk structure, not the specific number of factors.

Given their partial information set, investors form portfolios with weights:

$$w_i = \frac{1}{N}\Lambda^{-1} S_i (a - \hat{a}_E) \tag{53}$$

where Λ is the return covariance matrix, a represents true expected payoffs, and $\hat{a}_E$ captures the market's aggregate estimate of expected payoffs. The estimation error $(a - \hat{a}_E)$ follows a normal distribution with mean zero and covariance matrix Λ_a that reflects both fundamental uncertainty and parameter estimation error:

$$(a - \hat{a}_E) \sim N(0, \Lambda_a) \tag{54}$$

where

$$\Lambda_a = \left(\frac{1}{\kappa} + t\right)^{-1} \left(\beta\beta'\sigma_m^2 + \sigma_\epsilon^2 I\right) \tag{55}$$

The expected return of this strategy scales with the fraction of assets known as $O(q)$.

$$\mathrm{E}(R_w) = \frac{1}{N}(\Sigma^{-1} S_i (a - \hat{a}_E))'(a - \hat{a}_E) + \gamma_N \Sigma\beta) \sim O(q) \tag{56}$$

The quadratic form $(a - \hat{a}_e)'\Sigma^{-1} S_i (a - \hat{a}_e)/N$, which corresponds to the expected return component arising from estimation errors, has expected value q because S_i selects q fraction of a chi-squared random variable and the $1/N$ scaling in portfolio weights gives us final expectation q. The component corresponding to the risk premium, $a(a - \hat{a}_e)' S_i \beta$, has zero expected value because the portfolio is zero-cost and estimation errors are centered at zero. Economically, this means strategies' expected returns depend only on the fraction of assets an investor understands deeply (q), not market size (N), as any benefits from more trading opportunities are exactly offset by the need to reduce position sizes to manage estimation risk.

Using the Sherman-Morrison formula, we can decompose the inverse of the assets variance as:

$$\Lambda^{-1} = \frac{1}{\sigma_\epsilon^2} I - \frac{\sigma_m^2}{\sigma_\epsilon^2(\sigma_\epsilon^2 + \sigma_m^2 \beta'\beta)} \beta\beta' \tag{57}$$

The variance of each investor's portfolio contains two components. The first comes from idiosyncratic risk and scales as $O(q/N)$, declining with the number of assets but increasing with the fraction of assets known. The second component reflects systematic market risk and scales as $O(q^2/N)$. Intuitively, as investors have access to more assets (higher q), they can better diversify idiosyncratic risk but remain exposed to systematic risk that cannot be diversified away. Thus, the total variance of the portfolio scales as:

$$\mathrm{Var}(R_w) = \frac{1}{N^2}(a - \hat{a}_E)' S_i \Lambda^{-1} S_i (a - \hat{a}_E) \sim O\left(\frac{q}{N}\right).$$

The Sharpe ratio of the strategy scales as $O(q\sqrt{N})$ because expected returns scale with q while

the standard deviation scales as $O(\sqrt{q/N})$. This scaling has a powerful economic interpretation: as markets become more complex (higher N), investors who understand a fixed fraction of assets (constant q) can achieve higher risk-adjusted returns by taking smaller positions in more assets, benefiting from improved diversification without sacrificing the profitability that comes from their understanding of individual securities.

The covariance between two strategies scales as $O(q^2/N)$ and each strategy's variance scales as $O(q/N)$, leading to correlations that scale as $O(q)$ independent of market size. The economic intuition is compelling: while both covariance and variance benefit from diversification (the $1/N$ terms), the correlation depends only on the overlap in known assets (q). This explains why even sophisticated investors pursuing similar strategies can maintain low correlations in complex markets - as long as they focus on different subsets of assets, the benefits of a larger investment universe (N) reduce both risk and comovement proportionally, leaving the correlation dependent only on the fraction of shared expertise (q).

$$\text{Cov}(R_{w_i}, R_{w_j}) = \frac{1}{N^2}(a - \hat{a}_E)' S_i \Lambda^{-1} S_j (a - \hat{a}_E) \sim O\left(\frac{q^2}{N}\right) \tag{58}$$

Hence, the correlation between different investors' portfolios scales as $O(q)$:

$$\text{Corr}(R_{w_i}, R_{w_j}) = \frac{\text{Cov}(R_{w_i}, R_{w_j})}{\sqrt{\text{Var}(R_{w_i})\text{Var}(R_{w_j})}} \sim O(q) \tag{59}$$

This means that as markets become more complex (larger N), investors who each analyze different subsets of assets will naturally form portfolios with low correlations, even when pursuing similar statistical arbitrage strategies. In high-dimensional markets, there are many different ways to construct portfolios with attractive risk-return tradeoffs by focusing on different subsets of assets.

These theoretical predictions align with our empirical findings about the low correlations between systematic hedge fund strategies and the challenges faced by parsimonious factor models. They also yield several testable implications about how strategy relationships should evolve as markets become more complex. As investment universes expand (higher N), funds face a choice

between maintaining deep knowledge of a fixed fraction of assets (constant q) or spreading resources more thinly (declining q). Our model predicts that successful funds will maintain stable q by scaling research capabilities with N, explaining why large quantitative funds invest heavily in research infrastructure. Funds that fail to scale research appropriately should see declining Sharpe ratios relative to competitors.

The model also makes specific predictions about cross-fund correlation patterns. Correlations between funds should depend primarily on research overlap (shared q) rather than strategy similarity. Funds using similar techniques but focusing on different market segments should maintain low correlations. This explains why even "copycat" strategies can succeed without eroding the profits of the original strategy. Furthermore, increasing market complexity should lead to more, not fewer, profitable trading strategies. However, these strategies should become increasingly specialized and research-intensive. The "factor zoo" should continue to expand with market complexity, but new factors should show less overlap with existing ones.

These predictions help rationalize why we observe a proliferation of seemingly unrelated anomalies and suggest that this pattern is a natural outcome of rational investors optimally responding to the opportunities and constraints created by market complexity. Rather than indicating market inefficiency, the existence of numerous successful but uncorrelated strategies reflects the inherent challenge of spanning high-dimensional markets with limited information processing capacity.

3 Empirical Study

We first look for empirical backing for the assumption that we are in a high-dimensional setting. We provide evidence of the high dimensionality of stock returns. Then, we look for empirical support for the model's predictions

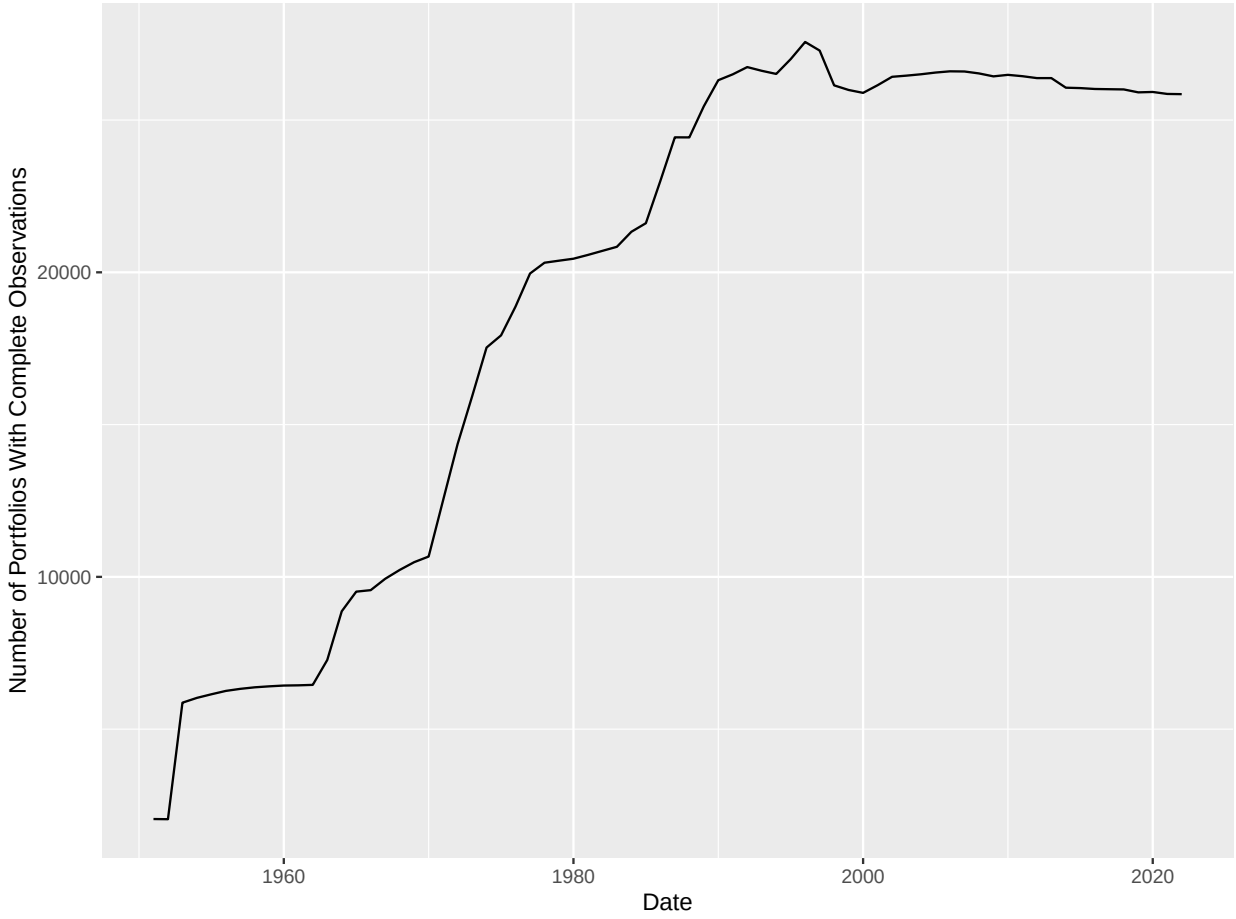
3.1 Data

We use the data from Chen, Lopez-Lira, and Zimmermann (2023). They generate 29,315 firm-level signals as follows. Let X be one of 242 Compustat accounting variables + CRSP market equity and Y be one of the 65 variables observed and positive for $> 25\%$ of firms in 1963 with matched CRSP data. They form signals by combining all combinations of ratios (X/Y) and scaled first differences $(\Delta X/\text{lag}(Y))$. This procedure would lead to $242 \times 65 \times 2 = 31,460$ signals, but there are 2,145 signals redundant in the "unsigned" portfolio sorts. They lag each signal by six months and then form long-short decile strategies by sorting stocks on the lagged signals each June. Delisting returns and other data handling methods follow Chen and Zimmermann (2022) to ensure that the published and data-mined strategies are comparable.

We consider the in-sample period to be $[1980, 2000]$ and the out-of-sample period to be $[2001, 2021]$. We restrict our analysis to the 17517 long-short portfolios that, due to data availability, have complete observations for the sample of 1980-2021. We refer to these signals simply as the 'portfolios.' There are no firm survivorship bias issues since the delisting returns of individual firms are handled appropriately. Rather, the data for some firm characteristics was unavailable until later in the sample.

These portfolios map nicely into our theoretical framework since each is constructed using one signal at a time, and hence, they proxy for the different characteristics that map to states. Moreover, being diversified portfolios implies little if any, idiosyncratic variation left, and we can focus on the systematic components.

Figure 2: Number of Portfolios with Complete Observations by Year



3.2 High-Dimensionality

We first motivate the high-dimensional nature of stock returns by exploring these portfolios' cross-sectional and time-series variations. First, we focus on the time-series variation. Even though these portfolios are very diversified, they retain a high-dimensional structure as depicted in Figure 3. About 20% of the overall variance of one-year daily return variance is explained by the first ten greatest principal components (in-sample). To account for 50% of the variation, we need about 50 factors. We need 75 factors to explain roughly 75% of the variation. The prior result is still obtained even though PCA is known to yield the (in-sample) solution for which the (time-series) variance explained is the highest for a given number of factors.

Then, we explore the cross-sectional variation and show that consistent with the factor zoo literature, there are too many signals with out-of-sample predictability. First, we collect the top 100 portfolios with the highest t-stats (we flip the portfolio returns if the t-stat is negative). Then, we look at the out-of-sample distribution of t-statistics. We find that there are way too many signals to be due to chance. Figure 4 shows the distribution of out-of-sample t-statistics and a fitted normal distribution. Panel 4a looks at the long-short returns while Panel 4b looks at the CAPM alphas. In both cases, we find many signals retain the out-of-sample predictability. The result holds despite the well-known fact that most of the predictability coming from characteristics declines very fast after the year 2000 (Lopez-Lira and Roussanov (2022)).

Figure 3: Cumulative Percentage of Variance Explained by the Principal Components

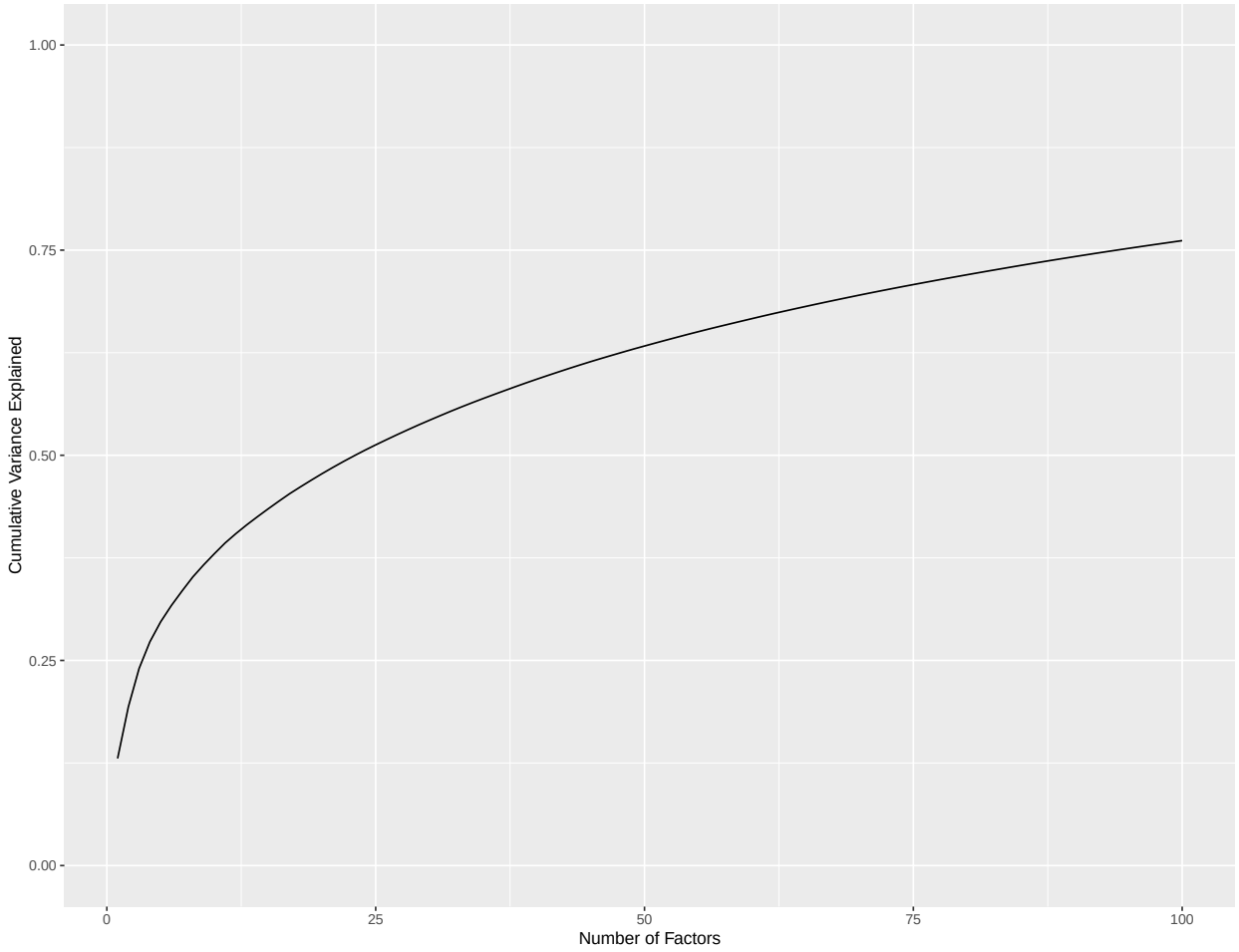
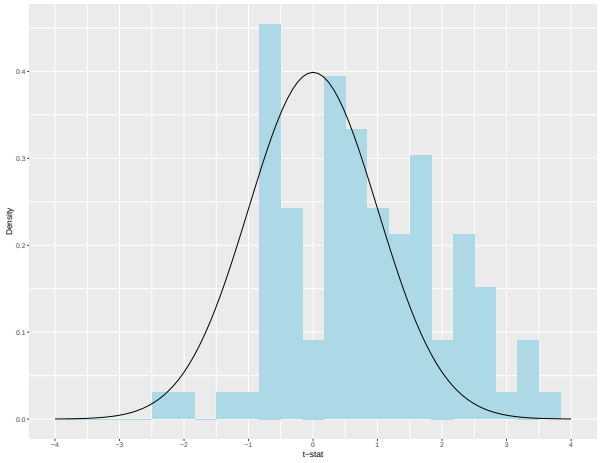
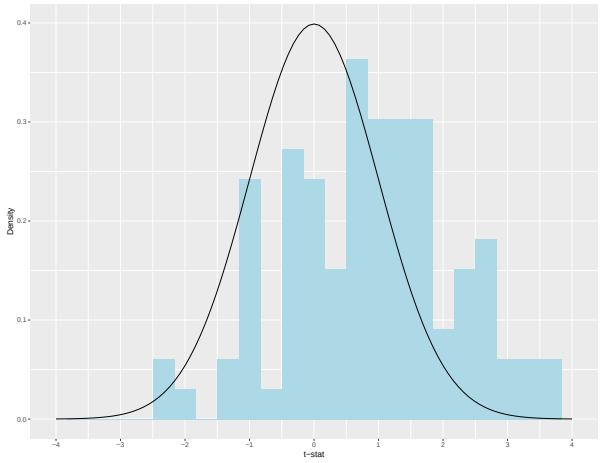


Figure 4: Comparative distributions of Out-of-Sample t-statistics



(a) Distribution of Out-of-Sample t-statistics from the top-100 In-Sample Portfolios



(b) Distribution of Out-of-Sample CAPM Alpha t-statistics from the top-100 In-Sample Portfolios

3.3 Low Correlations Between Strategies

An important implication from the model is that in a high-dimensional world, portfolios with high Sharpe ratios can have low correlations. We explore if this is the case empirically. We first consider portfolio correlations of investment strategies taken from the literature.

It is well known that a factor model prices the cross-section if and only if the factor model can be collapsed down to a mean-variance efficient portfolio (Huberman and Kandel 1987). Table 1 shows the variety of models from the literature that are used here, and Davis 2024 has more details about the implementation of these models. The Freyberger, Neuhierl, and Weber 2020b dataset of 62 asset characteristics is used. We follow Kozak, Nagel, and Santosh 2020 and Kelly, Pruitt, and Su 2019 by converting characteristics to cross-sectional percentiles (between 0 and 1) and then subtracting 0.5 to create long-short characteristic-managed portfolios. The data is monthly. The data from July 1962 up until the end of 1989 is used to train the model, and the parameters of each model are refitted every year. The results come from January 1990 to December of 2023, which are trained using strictly ex-ante data.

Panel (a) of Figure 6 shows the distribution of pair-wise correlations across all strategies. The average pair-wise correlation is about 0.26 which corresponds to an R^2 of a little bit more than 5%.

While these results already show that different methods to produce expected return estimates result in very low correlation, we further consider how is the correlation of actual strategies that are traded in the market using 13F holdings data. We follow the KY methodology to correct investors types in the 13F data, and define a systematic hedge fund as an "investment advisor" that holds at least 500 stocks. There are many hedge funds that hold only 10 or 20 stocks, instead of investing in systematic strategies, and these do not correspond to the sophisticated and diversified strategies we consider. With this filter, we capture some well-known hedge funds that use quantitative models to pursue systematic equities trading strategies, including AQR, Renaissance, Tower Research, Citadel, Two Sigma, S.A.C Capital, D.E. Shaw, Roll and Ross Asset Management, Paradigm Asset Management, Quantitative Systematic Strategies, Bridgewater Associates, Cubist Systematic Strategies, Martingale Asset Management, and Quantitative Systematic Strategies.

We consider only the active part of the portfolio weights, meaning we take the portfolio weights of the systematic hedge funds, and subtract off the fund's market beta times the market portfolio weights. In particular, if the weights of fund i are given by w_i, and its beta given by β_i, we define the market-neutral weights as

$$w_{i|M} = w_i - \beta_i w_M. \tag{60}$$

This result in the correlation of the active part of funds i and j given by:

$$\text{corr}_{i,j} = \frac{(w_{i|M})'\Sigma(w_{j|M})}{\sqrt{(w_{i|M})'\Sigma(w_{i|M})}\sqrt{(w_{j|M})'\Sigma(w_{j|M})}} \tag{61}$$

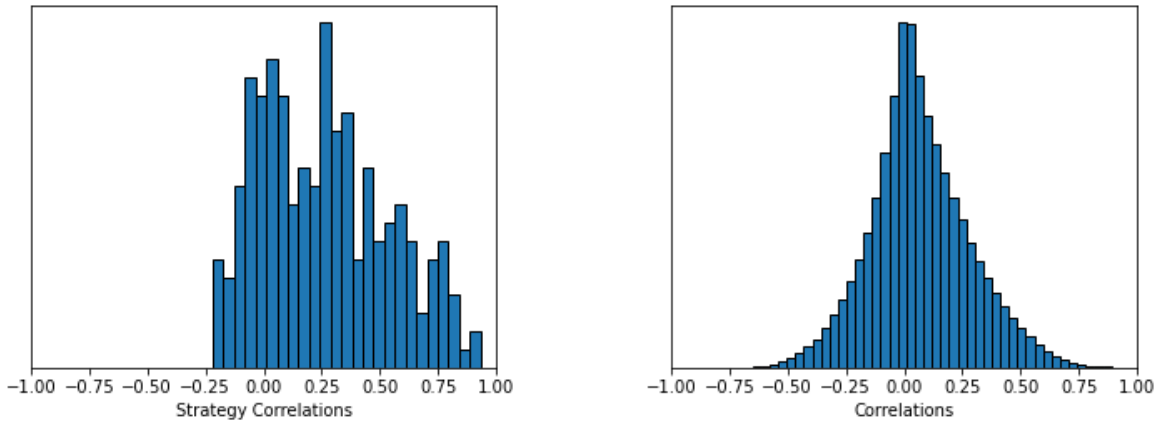
Panel (b) of Figure 6 shows the pair-wise correlations of these strategies. The average pair-wise correlation is about 0.07. This is in many ways similar to Koijen and Yogo 2019, who show that much of demand is not explained by classic characteristics. We go one step further by showing that much of the active component of hedge fund portfolios is not even explained by even other hedge fund's active portfolios.

Initialism	Factor Model
BPZF	Bryzgalova, Pelger, and Zhu 2020b
L	linear characteristic-weighted portfolios
CRW	Chen, Roussanov, and Wang 2021 linear portfolios
FF3	Fama and French 1993
FF6	Fama and French 2015 with Carhart 1997 momentum
GKX	Gu, Kelly, and Xiu 2021
HXZ	Hou, Xue, and Zhang 2014
KPS	Kelly, Pruitt, and Su 2019
NN	neural network
RF	random forest

Initialism	Collapse Method
BPZ	Bryzgalova, Pelger, and Zhu 2020b
KNS	Kozak, Nagel, and Santosh 2020
BSV	Brandt and Santa-Clara 2006
DGU	DeMiguel, Garlappi, and Uppal 2007

Table 1: **Statistical Arbitrageur Demand Models** This table lists the initialisms and sources for the investment strategy models.

Figure 5: Correlations of Portfolios



(a) Trading Strategy Correlations

(b) Systematic Hedge Fund Cross Correlations

Figure 6: This figure shows the correlations across trading strategies in Panel (a), and correlations of the active component of Systematic Hedge Funds from the 13F data in Panel (b).

Panel A: Weber CAPM Regressions Part 1

	NN	L-BPZ	KPS-BPZ	CRW-KNS	BPZF-BPZ	GKX-BPZ	GKX-BSV	FF6-DGU	HXZ-BSV	L-DGU	FF6-KNS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
α	51.10***	45.26***	45.01***	40.80***	27.12***	28.22***	27.29***	18.35***	22.79***	23.48***	20.42***
	(2.61)	(2.67)	(2.67)	(2.68)	(2.53)	(2.68)	(2.67)	(2.21)	(2.68)	(2.63)	(2.66)
β	-0.22***	0.08*	-0.06	-0.01	0.33***	-0.05	-0.07	0.57***	0.02	-0.18***	0.10**
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)
Observations	408	408	408	408	408	408	408	408	408	408	408
R^2	0.05	0.01	0.00	0.00	0.11	0.00	0.01	0.32	0.00	0.03	0.01
Sharpe Ratio	3.20	2.98	2.89	2.64	1.94	1.81	1.73	1.50	1.49	1.42	1.38
Note:											*p<0.1; **p<0.05; ***p<0.01

Panel B: Weber CAPM Regressions Part 2

	GKX-KNS	KPS-DGU	BPZF-DGU	HXZ-KNS	FF3-KNS	RF	FF3-DGU	CAPM	CRW-BPZ	GKX-DGU	BPZF-BSV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
α	21.58***	16.49***	10.74***	9.34***	12.93***	8.72***	5.18***		7.25***	4.00	-4.24
	(2.68)	(2.38)	(2.07)	(1.74)	(2.56)	(2.04)	(1.30)		(2.67)	(2.67)	(2.68)
β	-0.04	0.46***	0.63***	0.76***	0.29***	0.65***	0.87***		0.06	0.08*	-0.02
	(0.05)	(0.04)	(0.04)	(0.03)	(0.05)	(0.04)	(0.02)		(0.05)	(0.05)	(0.05)
Observations	408	408	408	408	408	408	408	408	408	408	408
R^2	0.00	0.21	0.40	0.58	0.08	0.42	0.76		0.00	0.01	0.00
Sharpe Ratio	1.38	1.32	1.04	1.02	1.00	0.92	0.82	0.55	0.50	0.31	-0.29
Note:											*p<0.1; **p<0.05; ***p<0.01

Table 2: **Weber CAPM Regressions.** This table presents the CAPM regression results for the Weber dataset, divided into two panels. Panel A contains results from the first part of the data, and Panel B contains results from the second part. The table includes the relevant coefficients and standard errors, calculated over the specified period. Detailed model descriptions are available in Table 1.

Panel A: Weber Cross Regressions Beta Part 1

	Rows: Test Assets, Columns: Right-hand-side Portfolios										
	NN	L-BPZ	KPS-BPZ	CRW-KNS	BPZF-BPZ	GKX-BPZ	GKX-BSV	FF6-DGU	HXZ-BSV	L-DGU	FF6-KNS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
NN		0.62***	0.61***	0.65***	0.43***	0.42***	0.44***	0.04	0.21**	0.25**	0.12
		(0.05)	(0.08)	(0.07)	(0.08)	(0.07)	(0.08)	(0.08)	(0.10)	(0.11)	(0.14)
L-BPZ	0.62***		0.54***	0.55***	0.41***	0.36***	0.35***	0.27***	0.16	0.22	0.19
	(0.08)		(0.06)	(0.10)	(0.11)	(0.11)	(0.13)	(0.09)	(0.14)	(0.14)	(0.21)
KPS-BPZ	0.61***	0.54***		0.74***	0.31***	0.56***	0.44***	0.18***	0.16**	0.30***	0.23**
	(0.05)	(0.09)		(0.07)	(0.05)	(0.06)	(0.07)	(0.07)	(0.06)	(0.08)	(0.09)
CRW-KNS	0.65***	0.55***	0.74***		0.50***	0.65***	0.66***	0.03	-0.02	0.05	-0.07
	(0.06)	(0.07)	(0.07)		(0.09)	(0.06)	(0.06)	(0.08)	(0.10)	(0.11)	(0.14)
BPZF-BPZ	0.43***	0.41***	0.31***	0.50***		0.40***	0.44***	0.27***	-0.05	-0.09	-0.02
	(0.11)	(0.11)	(0.07)	(0.10)		(0.10)	(0.11)	(0.09)	(0.11)	(0.12)	(0.16)
GKX-BPZ	0.42***	0.36***	0.56***	0.65***	0.40***		0.57***	-0.07	-0.03	-0.09	-0.19*
	(0.07)	(0.08)	(0.06)	(0.07)	(0.07)		(0.06)	(0.08)	(0.09)	(0.10)	(0.11)
GKX-BSV	0.44***	0.35***	0.44***	0.66***	0.44***	0.57***		-0.10	-0.05	-0.12	-0.21
	(0.09)	(0.12)	(0.08)	(0.08)	(0.10)	(0.09)		(0.09)	(0.11)	(0.12)	(0.14)
FF6-DGU	0.04	0.27**	0.18***	0.03	0.27**	-0.07	-0.10		0.57***	0.62***	0.81***
	(0.08)	(0.13)	(0.06)	(0.08)	(0.11)	(0.07)	(0.08)		(0.07)	(0.05)	(0.07)
HXZ-BSV	0.21**	0.16	0.16***	-0.02	-0.05	-0.03	-0.05	0.57***		0.60***	0.77***
	(0.10)	(0.15)	(0.06)	(0.10)	(0.11)	(0.09)	(0.11)	(0.06)		(0.06)	(0.06)
L-DGU	0.25**	0.22	0.30***	0.05	-0.09	-0.09	-0.12	0.62***	0.60***		0.83***
	(0.11)	(0.17)	(0.06)	(0.12)	(0.12)	(0.10)	(0.11)	(0.06)	(0.08)		(0.04)
FF6-KNS	0.12	0.19	0.23***	-0.07	-0.02	-0.19	-0.21	0.81***	0.77***	0.83***	
	(0.13)	(0.21)	(0.07)	(0.14)	(0.16)	(0.12)	(0.14)	(0.06)	(0.07)	(0.06)	
GKX-KNS	0.51***	0.35***	0.51***	0.66***	0.46***	0.56***	0.56***	-0.06	-0.05	-0.09	-0.17
	(0.06)	(0.08)	(0.07)	(0.05)	(0.07)	(0.06)	(0.06)	(0.08)	(0.09)	(0.10)	(0.11)
KPS-DGU	0.27***	0.37***	0.46***	0.51***	0.69***	0.44***	0.42***	0.33***	-0.07	-0.08	0.02
	(0.07)	(0.06)	(0.06)	(0.05)	(0.05)	(0.06)	(0.05)	(0.07)	(0.08)	(0.09)	(0.10)
BPZF-DGU	0.14	0.30***	0.15**	0.32***	0.77***	0.26***	0.26**	0.45***	-0.08	-0.17*	0.03
	(0.12)	(0.10)	(0.08)	(0.11)	(0.04)	(0.08)	(0.11)	(0.08)	(0.09)	(0.10)	(0.14)
HXZ-KNS	-0.03	0.23***	0.03	0.11	0.55***	0.06	0.06	0.76***	0.27***	0.09	0.40***
	(0.09)	(0.06)	(0.07)	(0.08)	(0.06)	(0.06)	(0.08)	(0.05)	(0.08)	(0.07)	(0.12)
FF3-KNS	0.05	0.16	0.08	-0.03	0.16	0.01	-0.02	0.74***	0.71***	0.50***	0.68***
	(0.08)	(0.13)	(0.06)	(0.08)	(0.11)	(0.08)	(0.10)	(0.04)	(0.05)	(0.06)	(0.06)
RF	0.18	0.36***	0.14*	0.35**	0.75***	0.23**	0.26*	0.47***	-0.09	-0.13	0.02
	(0.15)	(0.13)	(0.08)	(0.14)	(0.07)	(0.11)	(0.15)	(0.10)	(0.12)	(0.13)	(0.18)
FF3-DGU	-0.10	0.20***	0.01	0.08	0.51***	0.08	0.05	0.72***	0.19**	-0.03	0.26***
	(0.08)	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)	(0.07)	(0.06)	(0.08)	(0.07)	(0.10)
CAPM	-0.22***	0.08	-0.06	-0.01	0.33***	-0.05	-0.07	0.57***	0.02	-0.18***	0.10
	(0.06)	(0.05)	(0.05)	(0.05)	(0.07)	(0.05)	(0.05)	(0.06)	(0.07)	(0.06)	(0.07)
CRW-BPZ	0.17**	0.22***	0.24***	0.24***	0.31***	0.34***	0.26***	0.15**	0.19**	-0.02	0.07
	(0.08)	(0.07)	(0.07)	(0.07)	(0.06)	(0.06)	(0.07)	(0.07)	(0.08)	(0.09)	(0.10)
GKX-DGU	0.24***	0.32***	0.27***	0.36***	0.53***	0.28***	0.36***	0.12	-0.05	-0.01	0.00
	(0.08)	(0.09)	(0.08)	(0.08)	(0.05)	(0.09)	(0.08)	(0.08)	(0.09)	(0.11)	(0.13)
BPZF-BSV	0.12*	0.11*	0.03	0.00	0.09	0.05	-0.02	-0.10*	-0.20***	-0.06	-0.16***
	(0.07)	(0.06)	(0.06)	(0.07)	(0.07)	(0.07)	(0.08)	(0.05)	(0.05)	(0.07)	(0.06)
Observations	408	408	408	408	408	408	408	408	408	408	408
F-test	6.94	7.13	7.93	10.96	18.42	21.33	22.45	19.06	18.81	17.95	17.55
p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Note: *p<0.1; **p<0.05; ***p<0.01

Panel B: Weber Cross Regressions Beta Part 2

	Rows: Test Assets, Columns: Right-hand-side Portfolios										
	GKX-KNS	KPS-DGU	BPZF-DGU	HXZ-KNS	FF3-KNS	RF	FF3-DGU	CAPM	CRW-BPZ	GKX-DGU	BPZF-BSV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
NN	0.51***	0.27***	0.14	-0.03	0.05	0.18	-0.10	-0.22***	0.17*	0.24***	0.12*
	(0.08)	(0.08)	(0.12)	(0.09)	(0.08)	(0.13)	(0.07)	(0.05)	(0.09)	(0.09)	(0.07)
L-BPZ	0.35***	0.37***	0.30***	0.23***	0.16	0.36***	0.20***	0.08	0.22**	0.32***	0.11
	(0.11)	(0.09)	(0.11)	(0.07)	(0.11)	(0.12)	(0.06)	(0.05)	(0.10)	(0.11)	(0.07)
KPS-BPZ	0.51***	0.46***	0.15**	0.03	0.08	0.14**	0.01	-0.06	0.24***	0.27***	0.03
	(0.05)	(0.06)	(0.07)	(0.07)	(0.06)	(0.07)	(0.06)	(0.05)	(0.07)	(0.08)	(0.06)
CRW-KNS	0.66***	0.51***	0.32***	0.11	-0.03	0.35***	0.08	-0.01	0.24**	0.36***	0.00
	(0.07)	(0.08)	(0.11)	(0.09)	(0.09)	(0.11)	(0.07)	(0.05)	(0.10)	(0.10)	(0.07)
BPZF-BPZ	0.46***	0.69***	0.77***	0.55***	0.16	0.75***	0.51***	0.33***	0.31***	0.53***	0.09
	(0.10)	(0.07)	(0.07)	(0.07)	(0.10)	(0.05)	(0.06)	(0.04)	(0.10)	(0.09)	(0.08)
GKX-BPZ	0.56***	0.44***	0.26***	0.06	0.01	0.23**	0.08	-0.05	0.34***	0.28***	0.05
	(0.06)	(0.07)	(0.08)	(0.06)	(0.08)	(0.09)	(0.06)	(0.05)	(0.08)	(0.09)	(0.07)
GKX-BSV	0.56***	0.42***	0.26**	0.06	-0.02	0.26*	0.05	-0.07	0.26***	0.36***	-0.02
	(0.09)	(0.09)	(0.11)	(0.08)	(0.10)	(0.14)	(0.07)	(0.05)	(0.10)	(0.11)	(0.07)
FF6-DGU	-0.06	0.33***	0.45***	0.76***	0.74***	0.47***	0.72***	0.57***	0.15**	0.12	-0.10
	(0.07)	(0.08)	(0.10)	(0.05)	(0.04)	(0.15)	(0.05)	(0.06)	(0.07)	(0.08)	(0.06)
HXZ-BSV	-0.05	-0.07	-0.08	0.27***	0.71***	-0.09	0.19***	0.02	0.19**	-0.05	-0.20***
	(0.09)	(0.08)	(0.10)	(0.06)	(0.05)	(0.11)	(0.06)	(0.07)	(0.07)	(0.09)	(0.08)
L-DGU	-0.09	-0.08	-0.17	0.09	0.50***	-0.13	-0.03	-0.18***	-0.02	-0.01	-0.06
	(0.10)	(0.09)	(0.11)	(0.07)	(0.07)	(0.12)	(0.07)	(0.07)	(0.09)	(0.11)	(0.08)
FF6-KNS	-0.17	0.02	0.03	0.40***	0.68***	0.02	0.26***	0.10	0.07	0.00	-0.16*
	(0.12)	(0.10)	(0.13)	(0.06)	(0.07)	(0.18)	(0.07)	(0.06)	(0.10)	(0.13)	(0.08)
GKX-KNS		0.43***	0.32***	0.10	-0.07	0.32***	0.07	-0.04	0.30***	0.36***	0.04
		(0.08)	(0.09)	(0.08)	(0.08)	(0.09)	(0.06)	(0.05)	(0.09)	(0.10)	(0.07)
KPS-DGU	0.43***		0.76***	0.61***	0.21***	0.71***	0.60***	0.46***	0.34***	0.59***	0.04
	(0.06)		(0.04)	(0.05)	(0.08)	(0.07)	(0.05)	(0.05)	(0.07)	(0.05)	(0.07)
BPZF-DGU	0.32***	0.76***		0.82***	0.27***	0.93***	0.80***	0.63***	0.33***	0.50***	0.03
	(0.10)	(0.06)		(0.05)	(0.09)	(0.05)	(0.04)	(0.03)	(0.09)	(0.08)	(0.08)
HXZ-KNS	0.10	0.61***	0.82***		0.52***	0.79***	0.92***	0.76***	0.29***	0.36***	-0.08
	(0.07)	(0.05)	(0.06)		(0.07)	(0.11)	(0.03)	(0.03)	(0.07)	(0.06)	(0.06)
FF3-KNS	-0.07	0.21**	0.27***	0.52***		0.25*	0.59***	0.29***	0.23***	0.11	-0.15**
	(0.08)	(0.08)	(0.10)	(0.06)		(0.14)	(0.06)	(0.07)	(0.07)	(0.09)	(0.07)
RF	0.32**	0.71***	0.93***	0.79***	0.25**		0.79***	0.65***	0.29***	0.45***	0.08
	(0.13)	(0.08)	(0.06)	(0.06)	(0.11)		(0.05)	(0.03)	(0.11)	(0.11)	(0.08)
FF3-DGU	0.07	0.60***	0.80***	0.92***	0.59***	0.79***		0.87***	0.24***	0.29***	-0.05
	(0.06)	(0.06)	(0.07)	(0.04)	(0.07)	(0.12)		(0.03)	(0.06)	(0.06)	(0.06)
CAPM	-0.04	0.46***	0.63***	0.76***	0.29***	0.65***	0.87***		0.06	0.08	-0.02
	(0.05)	(0.05)	(0.09)	(0.05)	(0.07)	(0.14)	(0.03)		(0.05)	(0.07)	(0.05)
CRW-BPZ	0.30***	0.34***	0.33***	0.29***	0.23***	0.29***	0.24***	0.06		0.35***	-0.16**
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.06)	(0.05)		(0.08)	(0.07)
GKX-DGU	0.36***	0.59***	0.50***	0.36***	0.11	0.45***	0.29***	0.08	0.35***		0.07
	(0.08)	(0.07)	(0.07)	(0.07)	(0.08)	(0.08)	(0.07)	(0.07)	(0.07)		(0.08)
BPZF-BSV	0.04	0.04	0.03	-0.08	-0.15***	0.08	-0.05	-0.02	-0.16**	0.07	
	(0.07)	(0.07)	(0.08)	(0.06)	(0.05)	(0.08)	(0.06)	(0.05)	(0.08)	(0.07)	
Observations	408	408	408	408	408	408	408	408	408	408	408
F-test	21.64	26.59	26.19	28.01	24.18	25.40	28.38	26.82	30.61	29.93	27.40
p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3: **Weber Unconditional Cross Regressions (Betas).** Panel B presents the regression results for the Weber dataset, focusing on beta coefficients from the second part of the data.

4 Conclusion

Our framework resolves several enduring asset pricing mysteries by reframing them as natural outcomes of market complexity. In high-dimensional environments, where the parameter space outstrips available data, rational optimization under uncertainty generates systematic pricing errors, dilutes the explanatory power of risk premiums, and generates diverse anomalies. Far from being robust universal solutions, traditional factor models like the CAPM and its multi-factor successors falter—not due to market inefficiencies but because their parsimony cannot keep pace with the estimation challenges posed by modern markets.

This perspective recasts the proliferation of anomalies and the breakdown of parsimonious models as evidence of rational adaptation rather than market failure. Sophisticated investors can exploit distinct estimation errors to construct high-Sharpe-ratio strategies that remain little correlated. Our theory reveals this divergence as a feature of high-dimensional markets, where abundant error dimensions allow multiple strategies to thrive without overlapping.

Empirical evidence reinforces these insights. Analyzing thousands of characteristic-sorted portfolios, we uncover a stark high-dimensional reality: the first ten principal components explain just 20% of the return variance of diversified portfolios, and even 75 factors leave 25% unaccounted for. Similarly, the active positions of systematic hedge funds show an average R^2 below 1%, highlighting their ability to target orthogonal facets of market complexity.

References

Alti, Aydogan, and Sheridan Titman. 2019. "A Dynamic Model of Characteristic-Based Return Predictability." *Journal of Finance* 74, no. 6 (December): 3187–3216. ISSN: 15406261. https://doi.org/10.1111/jofi.12839.

Andrei, Daniel, Julien Cujean, and Mungo Wilson. 2023. "The Lost Capital Asset Pricing Model." *The Review of Economic Studies* 90, no. 6 (November): 2703–2762. ISSN: 0034-6527. https://doi.org/10.1093/RESTUD/RDAD013.

Baba Yara, Fahiz, Brian H. Boyer, and Carter Davis. 2021. "Messy Asset Pricing: Can AI Models Lead to a Consensus?" *SSRN Electronic Journal* (November). https://doi.org/10.2139/SSRN.3967588.

Brandt, Michael W, and Pedro Santa-Clara. 2006. "Dynamic portfolio selection by augmenting the asset space." *The Journal of Finance* 61 (5): 2187–2217.

Bryzgalova, Svetlana, Markus Pelger, and Jason Zhu. 2020a. "Forest Through the Trees: Building Cross-Sections of Stock Returns." *working paper* (September). https://doi.org/10.2139/SSRN.3493458.

———. 2020b. "Forest through the trees: Building cross-sections of stock returns." *Available at SSRN 3493458.*

Carhart, Mark M. 1997. "On Persistence in Mutual Fund Performance." *The Journal of Finance* 52 (1): 57–82.

Chen, Andrew Y. 2021. "Most Claimed Statistical Findings in Cross-Sectional Return Predictability Are Likely True." *SSRN Electronic Journal* (October). https://doi.org/10.2139/SSRN.3912915.

Chen, Andrew Y., Alejandro Lopez-Lira, and Tom Zimmermann. 2023. "Peer-Reviewed Theory Does Not Help Predict the Cross-section of Stock Returns." *SSRN Electronic Journal* (March). https://doi.org/10.2139/SSRN.4308069.

Chen, Andrew Y., and Tom Zimmermann. 2022. "Open Source Cross-Sectional Asset Pricing." *Critical Finance Review* 11, no. 2 (May): 207–264. ISSN: 2164-5744. https://doi.org/10.1561/104.00000112.

Chen, Qihui, Nikolai L. Roussanov, and Xiaoliang Wang. 2021. "Semiparametric Conditional Factor Models: Estimation and Inference." *SSRN Electronic Journal* (July). ISSN: 1556-5068. https://doi.org/10.2139/ssrn.3984633.

Da, Rui, Stefan Nagel, and Dacheng Xiu. 2022. "The Statistical Limit of Arbitrage."

Daniel, Kent, and Sheridan Titman. 1997. "Evidence on the Characteristics of Cross Sectional Variation in Stock Returns." *The Journal of Finance* 52 (1): 1–33. https://doi.org/https://doi.org/10.1111/j.1540-6261.1997.tb03806.x.

Davis, Carter. 2024. "The Elasticity of Quantitative Investment." *working paper*.

DeMiguel, Victor, Lorenzo Garlappi, and Raman Uppal. 2007. "Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy?" *Review of Financial Studies* 22 (5): 1915–1953. ISSN: 08939454.

Didisheim, Antoine, Shikun (Barry) Ke, Bryan T. Kelly, and Semyon Malamud. 2024. "APT or "AIPT"? The Surprising Dominance of Large Factor Models." (Cambridge, MA) (September). https://doi.org/10.3386/W33012.

Fama, Eugene F., and Kenneth R. French. 1993. "Common risk factors in the returns on stocks and bonds." *Journal of Financial Economics* 33 (1): 3–56. ISSN: 0304-405X.

———. 2015. "A five-factor asset pricing model." *Journal of Financial Economics* 116:1–22. ISSN: 0304-405X.

Fama, Eugene F., and James D. MacBeth. 1973. "Risk, Return, and Equilibrium: Empirical Tests." *Journal of Political Economy* 81 (3): 607–636. ISSN: 00223808, 1537534X. http://www.jstor.org/stable/1831028.

Freyberger, Joachim, Andreas Neuhierl, and Michael Weber. 2020a. "Dissecting Characteristics Nonparametrically." *The Review of Financial Studies* 33 (5): 2326–2377. ISSN: 0893-9454. https://doi.org/10.1093/rfs/hhz123.

———. 2020b. "Dissecting Characteristics Nonparametrically." *The Review of Financial Studies* 33, no. 5 (April): 2326–2377. ISSN: 0893-9454. https://doi.org/10.1093/rfs/hhz123.

Gibbons, Michael R, Stephen A Ross, and Jay Shanken. 1989. "A Test of the Efficiency of a Given Portfolio." *Econometrica* 57 (5): 1121–1152. ISSN: 00129682, 14680262. http://www.jstor.org/stable/1913625.

Gu, Shihao, Bryan Kelly, and Dacheng Xiu. 2020. "Empirical Asset Pricing via Machine Learning." *The Review of Financial Studies* 33 (5): 2223–2273. ISSN: 0893-9454. https://doi.org/10.1093/rfs/hhaa009.

———. 2021. "Autoencoder asset pricing models." Annals Issue:Financial Econometrics in the Age of the Digital Economy, *Journal of Econometrics* 222 (1, Part B): 429–450. ISSN: 0304-4076.

Harvey, Campbell R, Yan Liu, and Heqing Zhu. 2016. "... and the Cross-Section of Expected Returns." *The Review of Financial Studies* 29 (1): 5–68. https://doi.org/10.1093/rfs/hhv059.

Hou, Kewei, Chen Xue, and Lu Zhang. 2014. "Digesting Anomalies: An Investment Approach." *The Review of Financial Studies* 28, no. 3 (September): 650–705. ISSN: 0893-9454.

Huberman, Gur, and Shmuel Kandel. 1987. "Mean-Variance Spanning." *The Journal of Finance* 42 (4): 873–888.

Jacobs, Bruce I, and Kenneth N Levy. 1988. "Disentangling Equity Return Regularities: New Insights and Investment Opportunities." *Financial Analysts Journal* 44 (3). ISSN: 0015-198X. https://doi.org/10.2469/faj.v44.n3.18.

Kelly, Bryan, Seth Pruitt, and Yinan Su. 2019. "Characteristics Are Covariances: A Unified Model of Risk and Return." *Journal of Financial Economics* 134 (3): 501–524.

Kelly, Bryan T., Semyon Malamud, and Kangying Zhou. 2022. "The Virtue of Complexity Everywhere." *SSRN Electronic Journal* (July). https://doi.org/10.2139/SSRN.4166368.

Koijen, Ralph S. J., and Motohiro Yogo. 2019. "A Demand System Approach to Asset Pricing." *Journal of Political Economy* 127 (4): 1475–1515. https://doi.org/10.1086/701683.

Kozak, Serhiy, and Stefan Nagel. 2024. "When Do Cross-Sectional Asset Pricing Factors Span the Stochastic Discount Factor?" *SSRN Electronic Journal* (May). https://doi.org/10.2139/SSRN.4443643.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh. 2018. "Interpreting Factor Models." *The Journal of Finance* 73 (3): 1183–1223. ISSN: 15406261. https://doi.org/10.1111/jofi.12612.

———. 2020. "Shrinking the Cross Section." *Journal of Financial Economics* 135 (2): 271–292.

Lewellen, Jonathan. 2015. "The Cross-section of Expected Stock Returns." *Critical Finance Review* 4 (1): 1–44. ISSN: 2164-5744. https://doi.org/10.1561/104.00000024.

Lopez-Lira, Alejandro, and Nikolai L. Roussanov. 2022. "Do Common Factors Really Explain the Cross-Section of Stock Returns?" *SSRN Electronic Journal* (July). https://doi.org/10.2139/ssrn.3628120.

Martin, Ian W.R., and Stefan Nagel. 2022. "Market efficiency in the age of big data." *Journal of Financial Economics* 145, no. 1 (July): 154–177. ISSN: 0304-405X. https://doi.org/10.1016/J.JFINECO.2021.10.006.

McLean, R David, and Jeffrey Pontiff. 2016. "Does Academic Research Destroy Stock Return Predictability?" *The Journal of Finance* 71 (1): 5–32. https://doi.org/10.1111/jofi.12365.

Merton, Robert C. 1987. "A simple model of capital market equilibrium with incomplete information." *The journal of finance* 42, no. 3 (July): 483–510. ISSN: 1540-6261. https://doi.org/10.1111/J.1540-6261.1987.TB04565.X.

Yan, Xuemin Sterling, and Lingling Zheng. 2017. "Fundamental Analysis and the Cross-Section of Stock Returns: A Data-Mining Approach." *The Review of Financial Studies* 30, no. 4 (April): 1382–1423. ISSN: 0893-9454. https://doi.org/10.1093/RFS/HHX001.

Appendix

4.1 Decomposition of Pricing Errors

Proposition. *The weighted sum of squared pricing errors in the economy, relative to the market portfolio, is given by the difference in squared Sharpe ratios (SR) of the optimal portfolio and the market portfolio. This sum can be decomposed into mean and covariance estimation errors plus an interaction term.*

$$\alpha' \Lambda_t^{-1} \alpha = SR(w^*)^2 - SR(\frac{1}{N}\iota)^2 \tag{62}$$

$$= \underbrace{\left[\varepsilon' \, \Lambda_t^{-1} \, \varepsilon \; - \; \frac{(\iota'\varepsilon)^2}{\sigma_{m,t}^2}\right]}_{\textit{Mean estimation errors}} + \underbrace{2\,\gamma_N \left[\varepsilon' \, \Lambda_t^{-1} \, \hat{\Lambda}_t \, \iota - \; \frac{(\iota'\varepsilon)\,\hat{\sigma}_{m,t}^2}{\sigma_{m,t}^2}\right]}_{\textit{Interaction term}} + \underbrace{(\gamma_N)^2 \left[\iota' \, \hat{\Lambda}_t \, \Lambda_t^{-1} \, \hat{\Lambda}_t \, \iota \; - \; \frac{(\hat{\sigma}_{m,t}^2)^2}{\sigma_{m,t}^2}\right]}_{\textit{Covariance estimation errors}} \tag{63}$$

Proof. The squared Sharpe ratio for a portfolio with weights w is defined as:

$$SR(w)^2 = \frac{(w'\mu_t)^2}{w'\Lambda_t w}$$

Derivation for the Optimal Portfolio For the optimal portfolio with weights $w^* = \Lambda_t^{-1}\mu_t$, we can simplify the squared Sharpe ratio as follows:

$$SR(w^*)^2 = \frac{(w^{*\prime}\mu_t)^2}{w^{*\prime}\Lambda_t w^*} \tag{64}$$

$$= \frac{(\Lambda_t^{-1}\mu_t)'\mu_t}{(\Lambda_t^{-1}\mu_t)'\Lambda_t(\Lambda_t^{-1}\mu_t)} \tag{65}$$

$$= \frac{\mu_t'\Lambda_t^{-1}\mu_t}{\mu_t'\Lambda_t^{-1}\Lambda_t\Lambda_t^{-1}\mu_t} \tag{66}$$

$$= \frac{\mu_t'\Lambda_t^{-1}\mu_t}{\mu_t'\Lambda_t^{-1}\mu_t} \tag{67}$$

$$= \mu_t'\Lambda_t^{-1}\mu_t \tag{68}$$

Substituting our expression for $\mu_t = \varepsilon + \gamma_N \hat{\Lambda}_t \, \iota$, we get:

$$SR(w^*)^2 = \mu_t' \Lambda_t^{-1} \mu_t = (\varepsilon + \gamma_N \hat{\Lambda}_t \, \iota)' \, \Lambda_t^{-1} \, (\varepsilon + \gamma_N \hat{\Lambda}_t \, \iota)$$

Derivation for the Market Portfolio For the market portfolio with weights $\frac{1}{N}\iota$, we have:

$$SR\left(\tfrac{1}{N}\,\iota\right)^2 = \frac{\left(\frac{1}{N}\iota'\mu_t\right)^2}{\frac{1}{N^2}\iota'\Lambda_t\iota} \tag{69}$$

$$= \frac{\frac{1}{N^2}(\iota'\mu_t)^2}{\frac{1}{N^2}\iota'\Lambda_t\iota} \tag{70}$$

$$= \frac{(\iota'\mu_t)^2}{\iota'\Lambda_t\iota} \tag{71}$$

Now, substituting $\mu_t = \varepsilon + \gamma_N \hat{\Lambda}_t \, \iota$ and noting that $\sigma^2_{m,t} = \frac{1}{N^2}\iota'\Lambda_t\iota$ and $\hat{\sigma}^2_{m,t} = \frac{1}{N^2}\iota'\hat{\Lambda}_t\iota$:

$$SR\left(\tfrac{1}{N}\,\iota\right)^2 = \frac{(\iota'(\varepsilon + \gamma_N \hat{\Lambda}_t \, \iota))^2}{\sigma^2_{m,t}} \tag{72}$$

$$= \frac{(\iota'\varepsilon + \gamma_N \iota'\hat{\Lambda}_t \, \iota)^2}{\sigma^2_{m,t}} \tag{73}$$

$$= \frac{(\iota'\varepsilon + \gamma_N \hat{\sigma}^2_{m,t})^2}{\sigma^2_{m,t}} \tag{74}$$

where:

- $\varepsilon = a_t - \hat{a}_t$ is the mean estimation error. - $\sigma^2_{m,t} = \frac{1}{N^2}\iota'\Lambda_t\iota$ is the (true) market variance. - $\hat{\sigma}^2_{m,t} = \frac{1}{N^2}\iota'\hat{\Lambda}_t\iota$ is the *perceived* market variance. - γ_N is the representative investor's risk-aversion scaled by the number of assets (or by wealth, depending on the normalization).

Hence the difference in squared Sharpe ratios can be written as:

$$\alpha'\Lambda_t^{-1}\alpha \;=\; SR(w^*)^2 \;-\; SR\left(\tfrac{1}{N}\,\iota\right)^2 \;=\; (\varepsilon + \gamma_N \hat{\Lambda}_t \, \iota)' \, \Lambda_t^{-1} \, (\varepsilon + \gamma_N \hat{\Lambda}_t \, \iota) \;-\; \frac{\left(\iota'\varepsilon \;+\; \gamma_N \, \hat{\sigma}^2_{m,t}\right)^2}{\sigma^2_{m,t}}.$$

Now, we expand each part:

1. **Expand** $\left(\varepsilon + \gamma_N \hat{\Lambda}_t \iota\right)' \Lambda_t^{-1} \left(\varepsilon + \gamma_N \hat{\Lambda}_t \iota\right)$:

$$(\varepsilon + \gamma_N \hat{\Lambda}_t \iota)' \Lambda_t^{-1} (\varepsilon + \gamma_N \hat{\Lambda}_t \iota) = \varepsilon' \Lambda_t^{-1} \varepsilon + 2\gamma_N \varepsilon' \Lambda_t^{-1} \hat{\Lambda}_t \iota + (\gamma_N)^2 \iota' \hat{\Lambda}_t \Lambda_t^{-1} \hat{\Lambda}_t \iota.$$

2. **Expand** $\dfrac{\left(\iota'\varepsilon + \gamma_N \hat{\sigma}^2_{m,t}\right)^2}{\sigma^2_{m,t}}$:

$$\frac{\left(\iota'\varepsilon + \gamma_N \hat{\sigma}^2_{m,t}\right)^2}{\sigma^2_{m,t}} = \frac{(\iota'\varepsilon)^2}{\sigma^2_{m,t}} + \frac{2\gamma_N (\iota'\varepsilon)\, \hat{\sigma}^2_{m,t}}{\sigma^2_{m,t}} + (\gamma_N)^2 \frac{(\hat{\sigma}^2_{m,t})^2}{\sigma^2_{m,t}}.$$

Subtracting the second from the first gives:

$$(\varepsilon + \gamma_N \hat{\Lambda}_t \iota)' \Lambda_t^{-1} (\varepsilon + \gamma_N \hat{\Lambda}_t \iota) - \frac{\left(\iota'\varepsilon + \gamma_N \hat{\sigma}^2_{m,t}\right)^2}{\sigma^2_{m,t}}$$

$$= \left[\varepsilon' \Lambda_t^{-1} \varepsilon - \frac{(\iota'\varepsilon)^2}{\sigma^2_{m,t}}\right] + 2\gamma_N \left[\varepsilon' \Lambda_t^{-1} \hat{\Lambda}_t \iota - \frac{(\iota'\varepsilon)\, \hat{\sigma}^2_{m,t}}{\sigma^2_{m,t}}\right] + (\gamma_N)^2 \left[\iota' \hat{\Lambda}_t \Lambda_t^{-1} \hat{\Lambda}_t \iota - \frac{(\hat{\sigma}^2_{m,t})^2}{\sigma^2_{m,t}}\right].$$

Thus we can write the final result as the **sum of three terms**:

$$\alpha' \Lambda_t^{-1} \alpha = \underbrace{\left[\varepsilon' \Lambda_t^{-1} \varepsilon - \frac{(\iota'\varepsilon)^2}{\sigma^2_{m,t}}\right]}_{\text{Term 1 (mean errors)}} + \underbrace{2\gamma_N \left[\varepsilon' \Lambda_t^{-1} \hat{\Lambda}_t \iota - \frac{(\iota'\varepsilon)\, \hat{\sigma}^2_{m,t}}{\sigma^2_{m,t}}\right]}_{\text{Term 2 (cross term)}} + \underbrace{(\gamma_N)^2 \left[\iota' \hat{\Lambda}_t \Lambda_t^{-1} \hat{\Lambda}_t \iota - \frac{(\hat{\sigma}^2_{m,t})^2}{\sigma^2_{m,t}}\right]}_{\text{Term 3 (covariance errors)}}$$

□

4.2 Perfect Information

Proposition 5. *When agents know the economy's parameters, CAPM and Arbitrage Pricing Theory (APT) hold. In particular,*

1. *Mean returns are given by* $\mu_t = \gamma_N \Lambda_t \iota$.

2. *The mean-variance efficient portfolio is the market portfolio with weights of* $\frac{1}{N}$.

3. *Pricing errors with respect to the market portfolio are zero, $\alpha_t = 0$.*

4. *For $N \to \infty$, and rotating the factors to make them orthogonal, the price of risk for factor k is given by $\gamma \bar{b}_{t,k} \sigma^2_{t,k}$, where $\bar{b}_{t,k}$ is the average loading on factor k and $\sigma_{t,k}$ is the factor's volatility.*

Proof. When agents have perfect information, their subjective estimates match the true values: $\hat{a}_t = a_t = B_t \pi$ and $\hat{\Lambda}_t = \Lambda_t = B_t V B_t'$. Recall that the true expected excess returns are:

$$\mu_t = \gamma_N \Lambda_t \iota \tag{75}$$

CAPM Verification. To verify that CAPM holds, we need to show that expected excess returns are proportional to market betas. The market portfolio in our setting has weights $w_m = \frac{1}{N}\iota$, reflecting equal investment across all assets.

The excess return on the market portfolio is $r_{m,t+1} = \frac{1}{N}\iota' r_{t+1}$, with expected excess return:

$$\mu_{m,t} = \frac{1}{N}\iota' \mu_t = \gamma_N \frac{1}{N}\iota' \Lambda_t \iota = \frac{\gamma}{N^2}\iota' \Lambda_t \iota \tag{76}$$

For any asset i, its beta with respect to the market is:

$$\beta_{i,t} = \frac{\mathrm{Cov}_t(r_{i,t+1}, r_{m,t+1})}{\mathrm{Var}_t(r_{m,t+1})} = \frac{\frac{1}{N}(\Lambda_t \iota)_i}{\frac{1}{N^2}\iota' \Lambda_t \iota} = \frac{N(\Lambda_t \iota)_i}{\iota' \Lambda_t \iota} \tag{77}$$

The expected excess return for asset i is:

$$\mu_{i,t} = (\gamma_N \Lambda_t \iota)_i = \frac{\gamma}{N}(\Lambda_t \iota)_i \tag{78}$$

Multiplying the market's expected excess return by asset i's beta:

$$\beta_{i,t}\mu_{m,t} = \frac{N(\Lambda_t \iota)_i}{\iota' \Lambda_t \iota} \cdot \frac{\gamma}{N^2}\iota' \Lambda_t \iota = \frac{\gamma}{N}(\Lambda_t \iota)_i = \mu_{i,t} \tag{79}$$

Thus, $\mu_{i,t} = \beta_{i,t}\mu_{m,t}$, confirming that CAPM holds under perfect information.

APT Verification. For APT, we need to show that expected excess returns are a linear function of factor loadings when idiosyncratic risks are diversified away as N grows large.

Assuming uncorrelated factors (achievable via orthogonal rotation), the covariance matrix can be expressed as:

$$\Lambda_t = \sum_{k=1}^{S} \sigma_k^2 b_{.,k} b'_{.,k} + \text{diag}(c_{1,t}^2\sigma_{z1}^2, \ldots, c_{N,t}^2\sigma_{zN}^2) \tag{80}$$

where $b_{.,k} = (b_{1,t,k}, \ldots, b_{N,t,k})'$ is the vector of loadings on factor k, σ_k^2 is the variance of factor k, and $c_{i,t}^2\sigma_{zi}^2$ represents idiosyncratic variance.

For asset i, the risk compensation term becomes:

$$(\Lambda_t \iota)_i = N \sum_{k=1}^{S} \sigma_k^2 \bar{b}_{t,k} b_{i,t,k} + c_{i,t}^2\sigma_{zi}^2 \tag{81}$$

where $\bar{b}_{t,k} = \frac{1}{N}\sum_{i=1}^{N} b_{i,t,k}$ is the average loading on factor k.

The expected excess return is therefore:

$$\mu_{i,t} = \gamma_N(\Lambda_t \iota)_i = \gamma \sum_{k=1}^{S} \sigma_k^2 \bar{b}_{t,k} b_{i,t,k} + \frac{\gamma}{N} c_{i,t}^2\sigma_{zi}^2 \tag{82}$$

As $N \to \infty$, if idiosyncratic variances remain bounded, the term $\frac{\gamma}{N} c_{i,t}^2\sigma_{zi}^2 \to 0$, and:

$$\mu_{i,t} \approx \gamma \sum_{k=1}^{S} \sigma_k^2 \bar{b}_{t,k} b_{i,t,k} = \sum_{k=1}^{S} \lambda_k b_{i,t,k} \tag{83}$$

where $\lambda_k = \gamma\sigma_k^2\bar{b}_{t,k}$ is the risk premium for factor k.

This confirms that APT holds in the limit as N increases: expected returns are a linear function of factor loadings, with risk premia determined by factor variances, average exposures, and risk aversion. □

4.3 Known Means and Unknown Covariances

Proposition 6. *When agents know the expected value of dividends, a, but have uncertainty about the covariance, Λ, neither CAPM nor APT hold. Further, pricing errors are, on expectation, increasing in complexity defined as the ratio of the number of state variables plus the number of firms $N + S$ over t, the number of available observations. In particular:*

1. *Mean returns are given by $\mu_t = \gamma_N(\hat{\Lambda}_t - \Lambda_t)\iota + \gamma_N \Lambda_t \iota$, and hence APT fails.*
2. *CAPM fails, and the mean-variance efficient portfolio weights are given by $\gamma_N \Lambda_t^{-1}(\hat{\Lambda}_t - \Lambda_t)\iota + \gamma_N \iota$.*
3. *Pricing errors are given by $\alpha' \Lambda_t^{-1} \alpha = (\gamma_N)^2 \left[\iota' \hat{\Lambda}_t \Lambda_t^{-1} \hat{\Lambda}_t \iota - \frac{(\hat{\sigma}^2_{m,t})^2}{\sigma^2_{m,t}} \right]$.*

Moreover, we have that APT fails and the opposite intuition holds: for a fixed number of observations, price errors with respect to the market portfolio (agents' best factor model) are increasing in the number of assets:

$$\mathbb{E}[\alpha_t' \Lambda^{-1} \alpha_t] = \frac{\gamma^2}{N^2} \left(\frac{(S+N)+t}{(S+N)+t-2} \right)^2 \left(\frac{1}{t+1} \right)^2 t(S+N-1)(S+N)^2 \bar{\varrho} \sim O\left(\frac{N+S}{t} \right), \tag{84}$$

where ϱ is

$$\bar{\varrho} = \frac{1}{((S+N))^2} \sum_{i=1}^{(S+N)} \sum_{j=1}^{(S+N)} b_{i,t} b_{j,t} V_{i,j} \tag{85}$$

Proof. Recall that the sum of squared pricing errors, when the mean is known, is given by:

$$= \frac{\gamma^2}{N^2} \left(\iota' \hat{\Lambda}_t \Lambda^{-1} \hat{\Lambda}_t \iota - \frac{(\iota' \hat{\Lambda}_t \iota)^2}{\iota' \Lambda \iota} \right). \tag{86}$$

Note that $\hat{\Lambda}_t = B_t \hat{V}_t B_t'$, where $\hat{V}_t$ is the agent's estimate of the covariance matrix of state variables. In the case of known mean but unknown covariance, we first define the scale matrix:

$$\hat{\Sigma}_t = \frac{1}{t + \nu_0 - (S+N) + 1} \left(V_0 + \sum_{\tau=1}^{t} (X_\tau - \pi)(X_\tau - \pi)' \right) \tag{87}$$

Then, using the normal approximation to the t-distribution with $\nu_t = \nu_0 + t$, we have:

$$\hat{V}_t = \frac{\nu_t}{\nu_t - 2}\hat{\Sigma}_t = \frac{\nu_0 + t}{\nu_0 + t - 2}\hat{\Sigma}_t \tag{88}$$

Simplifying by parts:

$$\mathbb{E}[\iota'\hat{\Lambda}_t\Lambda^{-1}\hat{\Lambda}_t\iota] = \tag{89}$$

$$= \left(\frac{\nu_0 + t}{\nu_0 + t - 2}\right)^2 \left(\frac{1}{t + \nu_0 - (S+N) + 1}\right)^2 \left(\mathbb{E}\Big[\iota' B_t \hat{V}_t B_t' \Lambda^{-1} B_t \hat{V}_t B_t'\iota\Big]\right) \tag{90}$$

$$= \left(\frac{\nu_0 + t}{\nu_0 + t - 2}\right)^2 \left(\frac{1}{t + \nu_0 - (S+N) + 1}\right)^2 \Bigg(\iota' B_t V_0 B_t' \Lambda^{-1} B_t V_0 B_t'\iota \tag{91}$$

$$+ \iota' B_t V_0 B_t' \Lambda^{-1} B_t \mathbb{E}\Big[\sum_{\tau=1}^{t}(X_t - \pi)(X_t - \pi)'\Big] B_t'\iota \tag{92}$$

$$+ \iota' B_t \mathbb{E}\Big[\sum_{\tau=1}^{t}(X_t - \pi)(X_t - \pi)'\Big] B_t' \Lambda^{-1} B_t V_0 B_t'\iota \tag{93}$$

$$+ \iota' B_t \mathbb{E}\Big[\Big(\sum_{\tau=1}^{t}(X_t - \pi)(X_t - \pi)'\Big) B_t' \Lambda^{-1} B_t \times \Big(\sum_{\tau=1}^{t}(X_t - \pi)(X_t - \pi)'\Big)\Big] B_t'\iota\Bigg) \tag{94}$$

We use the iid nature of X_t to write:

$$\mathbb{E}\left[\sum_{\tau=1}^{t}(X_t - \pi)(X_t - \pi)'\right] . = tV \tag{95}$$

Moreover,

$$\iota' B_t \mathbb{E}\left[\left(\sum_{\tau=1}^{t}(X_t - \pi)(X_t - \pi)'\right) B_t' \Lambda^{-1} B_t \left(\sum_{\tau=1}^{t}(X_t - \pi)(X_t - \pi)'\right)\right] B_t'\iota \tag{96}$$

$$= \iota' B_t \mathbb{E}\left[\sum_{\tau=1}^{t}\left((X_t - \pi)(X_t - \pi)'\right) B_t' \Lambda^{-1} B_t \left((X_t - \pi)(X_t - \pi)'\right)\right] B_t'\iota \tag{97}$$

$$+2\iota' B_t \mathbb{E}\left[\sum_{\tau=1}^{t}\sum_{s=1}^{t-1}\left((X_t - \pi)(X_t - \pi)'\right) B_t' \Lambda^{-1} B_t \left((X_s - \pi)(X_s - \pi)'\right)\right] B_t'\iota \tag{98}$$

$$= t\mathbb{E}[\iota' B_t(X_t - \pi)(X_t - \pi)' B_t' \Lambda^{-1} B_t(X_t - \pi)(X_t - \pi)' B_t' \iota] \tag{99}$$

$$+t(t-1)\iota'\Lambda\Lambda^{-1}\Lambda\iota \tag{100}$$

$$= t\mathbb{E}[(\iota' B_t(X_t - \pi))(\Lambda^{-1/2} B_t(X_t - \pi))'(\Lambda^{-1/2} B_t(X_t - \pi))(\iota' B_t(X_t - \pi))'] + t(t-1)\iota'\Lambda\iota \tag{101}$$

We can use the identity that if x is multivariate normal with mean zero and covariance V, then

$$\mathbb{E}[(Ax)(Bx)'(Cx)(Dx)'] = (AVB')(CVD') + (AVC')(BVD') + \mathrm{tr}(BVC')AVD' \tag{102}$$

where $\mathrm{tr}(\cdot)$ is the trace.

Using this we have

$$t\mathbb{E}[(\iota' B_t(X_t - \pi))(\Lambda^{-1/2} B_t(X_t - \pi))'(\Lambda^{-1/2} B_t(X_t - \pi))(\iota' B_t(X_t - \pi))'] \tag{103}$$

$$= t[2(\iota' B_t V B_t' \Lambda^{-1/2})(\Lambda^{-1/2} B_t V B_t' \iota) + \mathrm{tr}(\Lambda^{-1/2} B_t V B_t' \Lambda^{-1/2})(\iota' B_t V B_t' \iota)] \tag{104}$$

Well, we have

$$\Lambda = B_t V B_t' \tag{105}$$

so

$$\mathrm{tr}(\Lambda^{-1/2} B_t V B_t' \Lambda^{-1/2}) = \mathrm{tr}(\Lambda^{-1}\Lambda) = S + N \tag{106}$$

hence

$$t\mathbb{E}[(\iota' B_t(X_t - \pi))(\Lambda^{-1/2} B_t(X_t - \pi))'(\Lambda^{-1/2} B_t(X_t - \pi))(\iota' B_t(X_t - \pi))'] = t((2\iota'\Lambda\iota) + (S+N)(\iota'\Lambda\iota)) \tag{107}$$

$$= t(S + N + 2)\iota'\Lambda\iota \tag{108}$$

So we have

$$\mathbb{E}[\iota'\hat{\Lambda}_t\Lambda^{-1}\hat{\Lambda}_t\iota] \tag{109}$$

$$= \left(\frac{\nu_0+t}{\nu_0+t-2}\right)^2\left(\frac{1}{t+\nu_0-(S+N)+1}\right)^2\left(\iota'B_tV_0B_t'\Lambda^{-1}B_tV_0B_t'\iota+2t\iota'B_tV_0B_t'\iota+t(S+N+2)\iota'\Lambda\iota+t(t-1)\iota'\Lambda\iota\right) \tag{110}$$

Now onto the next part. We need to compute

$$\mathbb{E}[(\iota'\hat{\Lambda}_t\iota)^2] \tag{111}$$

$$= \mathbb{E}[\iota'\hat{\Lambda}_t\iota\iota'\hat{\Lambda}_t\iota] \tag{112}$$

$$= \left(\frac{\nu_0+t}{\nu_0+t-2}\right)^2\left(\frac{1}{t+\nu_0-(S+N)+1}\right)^2\mathbb{E}\left[\iota'B_t\left(V_0+\sum_{\tau=1}^{t}(X_t-\pi)(X_t-\pi)'\right)B_t'\iota\iota'B_t\left(V_0+\sum_{\tau=1}^{t}(X_t-\pi)(X_t-\pi)'\right)\right. \tag{113}$$

$$= \left(\frac{\nu_0+t}{\nu_0+t-2}\right)^2\left(\frac{1}{t+\nu_0-(S+N)+1}\right)^2\left(\iota'B_tV_0B_t'\iota\iota'B_tV_0B_t'\iota\right. \tag{114}$$

$$+\iota'B_tV_0B_t'\iota\iota'B_t\mathbb{E}\left[\sum_{\tau=1}^{t}(X_t-\pi)(X_t-\pi)'\right]B_t'\iota \tag{115}$$

$$+\iota'B_t\mathbb{E}\left[\sum_{\tau=1}^{t}(X_t-\pi)(X_t-\pi)'\right]B_t'\iota\iota'B_tV_0B_t'\iota \tag{116}$$

$$\left.+\iota'B_t\mathbb{E}\left[\left(\sum_{\tau=1}^{t}(X_t-\pi)(X_t-\pi)'\right)B_t'\iota\iota'B_t\left(\sum_{\tau=1}^{t}(X_t-\pi)(X_t-\pi)'\right)\right]B_t'\iota\right) \tag{117}$$

$$= \left(\frac{\nu_0+t}{\nu_0+t-2}\right)^2\left(\frac{1}{t+\nu_0-(S+N)+1}\right)^2\left(\iota'B_tV_0B_t'\iota\iota'B_tV_0B_t'\iota\right. \tag{118}$$

$$+2t\iota'B_tV_0B_t'\iota\iota'\Lambda\iota \tag{119}$$

$$\left.+\iota'B_t\mathbb{E}\left[\left(\sum_{\tau=1}^{t}(X_t-\pi)(X_t-\pi)'\right)B_t'\iota\iota'B_t\left(\sum_{\tau=1}^{t}(X_t-\pi)(X_t-\pi)'\right)\right]B_t'\iota\right) \tag{120}$$

Simplifying:

$$\iota'B_t\mathbb{E}\left[\left(\sum_{\tau=1}^{t}(X_t-\pi)(X_t-\pi)'\right)B_t'\iota\iota'B_t\left(\sum_{\tau=1}^{t}(X_t-\pi)(X_t-\pi)'\right)\right]B_t'\iota \tag{121}$$

$$= \iota' B_t \mathbb{E}\left[\sum_{\tau=1}^{t} \left((X_t - \pi)(X_t - \pi)'\right) B_t' \iota \iota' B_t \left((X_t - \pi)(X_t - \pi)'\right)\right] B_t' \iota \tag{122}$$

$$+2\iota' B_t \mathbb{E}\left[\sum_{\tau=1}^{t} \sum_{s=1}^{t-1} \left((X_t - \pi)(X_t - \pi)'\right) B_t' \iota \iota' B_t \left((X_s - \pi)(X_s - \pi)'\right)\right] B_t' \iota \tag{123}$$

$$= t\mathbb{E}[\iota' B_t (X_t - \pi)(X_t - \pi)' B_t' \iota \iota' B_t (X_t - \pi)(X_t - \pi)' B_t' \iota] \tag{124}$$

$$+t(t-1)\iota'\Lambda\iota\iota'\Lambda\iota \tag{125}$$

$$= t\mathbb{E}[(\iota' B_t (X_t - \pi))(\iota' B_t (X_t - \pi))'(\iota' B_t (X_t - \pi))(\iota' B_t (X_t - \pi))'] + t(t-1)(\iota'\Lambda\iota)^2 \tag{126}$$

Recall the identity:

$$\mathbb{E}[(Ax)(Bx)'(Cx)(Dx)'] = (AVB')(CVD') + (AVC')(BVD') + \mathrm{tr}(BVC')AVD' \tag{127}$$

Using this, we can calculate

$$t\mathbb{E}[(\iota' B_t (X_t - \pi))(\iota' B_t (X_t - \pi))'(\iota' B_t (X_t - \pi))(\iota' B_t (X_t - \pi))'] \tag{128}$$

$$= 3t(\iota'\Lambda\iota)^2 \tag{129}$$

So we can write:

$$\mathbb{E}[(\iota'\hat{\Lambda}_t\iota)^2] \tag{130}$$

$$= \left(\frac{\nu_0 + t}{\nu_0 + t - 2}\right)^2 \left(\frac{1}{t + \nu_0 - (S+N) + 1}\right)^2 \left(\iota' B_t V_0 B_t' \iota \iota' B_t V_0 B_t' \iota + 2t\iota' B_t V_0 B_t' \iota \iota' \Lambda\iota + 3t(\iota'\Lambda\iota)^2 + t(t-1)(\iota'\Lambda\iota)^2\right) \tag{131}$$

So then we have:

$$\mathbb{E}[\alpha_t'\Lambda^{-1}\alpha_t] = \mathbb{E}\left[\frac{\gamma^2}{N^2}\left(\iota'\hat{\Lambda}_t\Lambda^{-1}\hat{\Lambda}_t\iota - \frac{(\iota'\hat{\Lambda}_t\iota)^2}{\iota'\Lambda\iota}\right)\right] \tag{132}$$

$$= \frac{\gamma^2}{N^2}\left(\frac{\nu_0 + t}{\nu_0 + t - 2}\right)^2 \left(\frac{1}{t + \nu_0 - (S+N) + 1}\right)^2 \left(\left(\iota' B_t V_0 B_t' \Lambda^{-1} B_t V_0 B_t' \iota + 2t\iota' B_t V_0 B_t' \iota + t(S+N+2)\iota'\Lambda\iota + t(t-1)\iota'\Lambda\iota\right)\right) \tag{133}$$

$$-\frac{1}{\iota'\Lambda\iota}\left(\iota' B_t V_0 B_t' \iota \iota' B_t V_0 B_t' \iota + 2t\iota' B_t V_0 B_t' \iota \iota' \Lambda\iota + 3t(\iota'\Lambda\iota)^2 + t(t-1)(\iota'\Lambda\iota)^2\right)\Bigg) \tag{134}$$

$$= \frac{\gamma^2}{N^2}\left(\frac{\nu_0 + t}{\nu_0 + t - 2}\right)^2 \left(\frac{1}{t + \nu_0 - (S+N) + 1}\right)^2 \left(\iota' B_t V_0 B_t' \left(\Lambda^{-1} - \frac{\iota\iota'}{\iota'\Lambda\iota}\right) B_t V_0 B_t' \iota + t(S+N-1)\iota'\Lambda\iota\right) \tag{135}$$

Focusing on the simple case where $V_0 = V$ (right prior). Even in this case, we then have

$$\mathbb{E}[\alpha_t' \Lambda^{-1} \alpha_t] = \frac{\gamma^2}{N^2}\left(\frac{\nu_0 + t}{\nu_0 + t - 2}\right)^2 \left(\frac{1}{t + \nu_0 - (S+N) + 1}\right)^2 t(S+N-1)(S+N)^2 \bar{\varrho} \tag{136}$$

where ϱ is

$$\bar{\varrho} = \frac{1}{((S+N))^2} \sum_{i=1}^{(S+N)} \sum_{j=1}^{(S+N)} b_{i,t} b_{j,t} V_{i,j} \tag{137}$$

and

$$b_t = B_t \iota \tag{138}$$

We need $\nu_0 > (S+N) - 1$. A weak prior would be $\nu_0 = (S+N)$. Plugging in this prior yields

$$\mathbb{E}[\alpha_t' \Lambda^{-1} \alpha_t] = \frac{\gamma^2}{N^2}\left(\frac{(S+N) + t}{(S+N) + t - 2}\right)^2 \left(\frac{1}{t+1}\right)^2 t(S+N-1)(S+N)^2 \bar{\varrho} \tag{139}$$

which shows the average pricing errors relative to CAPM increase in the number of assets and states.

Finally, we establish that pricing errors are higher when agents have incorrect priors about the covariance structure. Let us compare pricing errors when $V_0 = V$ (correct prior) versus $V_0 \neq V$ (incorrect prior).

From our general expression:

$$\mathbb{E}[\alpha_t' \Lambda^{-1} \alpha_t] = \frac{\gamma^2}{N^2}\left(\frac{\nu_0 + t}{\nu_0 + t - 2}\right)^2 \left(\frac{1}{t + \nu_0 - (S+N) + 1}\right)^2 \tag{140}$$

$$\times \left[\iota' B_t V_0 B_t' \left(\Lambda^{-1} - \frac{\iota\iota'}{\iota'\Lambda\iota}\right) B_t V_0 B_t' \iota + t(S+N-1)\iota'\Lambda\iota\right] \tag{141}$$

Define $M = B_t V_0 B_t'$ as the prior-implied covariance matrix and let:

$$A = \iota' M \Lambda^{-1} M \iota - \frac{(\iota' M \iota)^2}{\iota' \Lambda \iota} \tag{142}$$

When $V_0 = V$, we have $M = \Lambda$, so:

$$A = \iota' \Lambda \Lambda^{-1} \Lambda \iota - \frac{(\iota' \Lambda \iota)^2}{\iota' \Lambda \iota} = \iota' \Lambda \iota - \iota' \Lambda \iota = 0 \tag{143}$$

When $V_0 \neq V$, we can show that $A \geq 0$ by interpreting it as a quadratic form. Let $Q = M - \frac{\Lambda \iota \iota' \Lambda}{\iota' \Lambda \iota}$, then $A = \iota' Q \Lambda^{-1} Q \iota$. Since Λ^{-1} is positive definite (as Λ is a covariance matrix), $A \geq 0$ if and only if Q is positive semi-definite.

To show this, note that Q can be interpreted as the residual projection matrix that removes the component in the direction of $\Lambda \iota$. This is a variance-like term that equals zero only when M and Λ are aligned in the ι direction (which happens when $V_0 = V$).

Therefore:

$$\mathbb{E}[\alpha_t' \Lambda^{-1} \alpha_t]\big|_{V_0 \neq V} \geq \mathbb{E}[\alpha_t' \Lambda^{-1} \alpha_t]\big|_{V_0 = V} \tag{144}$$

This confirms our economic intuition: pricing errors are higher when agents have incorrect priors about the covariance structure, with the difference increasing as V_0 deviates further from V.

□

4.4 Unknown Means and Unknown Covariances

Proposition 7. *If investors have perfect information about covariances but only the correct prior about the expectation, the average pricing errors relative to the feasible optimal portfolio increase in the number of assets and states at a rate of* $\frac{N+S}{t}$.

$$\mathbb{E}[\alpha_t' \Lambda_t^{-1} \alpha_t] = \frac{t}{(t+\kappa_0)^2}(N+S) - \frac{t}{(t+\kappa_0)^2} = \frac{t}{(t+\kappa_0)^2}(N+S-1) \sim O(\frac{N+S}{t}) \tag{145}$$

Proof. From the previous derivation, we have that for known covariances, the squared pricing errors are given by:

$$\alpha_t' \Lambda_t^{-1} \alpha_t = \varepsilon' \Lambda_t^{-1} \varepsilon - \frac{1}{\sigma_m^2} (\iota' \varepsilon)^2 \tag{146}$$

where $\varepsilon = a_t - \hat{a}_t = B_t(\pi - \hat{\pi}_t)$ represents the estimation error in the mean.

When investors have the correct prior $\pi_0 = \pi$, the posterior belief about the mean is:

$$\hat{\pi}_t = \frac{\kappa_0 \pi + t \bar{X}_t}{t + \kappa_0} \tag{147}$$

This implies that the estimation error is:

$$\varepsilon = B_t(\pi - \hat{\pi}_t) = B_t \left(\pi - \frac{\kappa_0 \pi + t \bar{X}_t}{t + \kappa_0} \right) = B_t \frac{t}{t + \kappa_0} (\pi - \bar{X}_t) \tag{148}$$

Taking expectations, we need to compute $\mathbb{E}[\alpha_t' \Lambda_t^{-1} \alpha_t]$. We know that $\bar{X}_t \sim N(\pi, \frac{V}{t})$ where V is the true covariance matrix of the state variables. Therefore:

$$\mathbb{E}[\varepsilon' \Lambda_t^{-1} \varepsilon] = \mathbb{E}\left[\left(B_t \frac{t}{t + \kappa_0} (\pi - \bar{X}_t) \right)' \Lambda_t^{-1} \left(B_t \frac{t}{t + \kappa_0} (\pi - \bar{X}_t) \right) \right] \tag{149}$$

$$= \frac{t^2}{(t + \kappa_0)^2} \mathbb{E}\left[(\pi - \bar{X}_t)' B_t' \Lambda_t^{-1} B_t (\pi - \bar{X}_t) \right] \tag{150}$$

Since $\Lambda_t = B_t V B_t'$, we have $B_t' \Lambda_t^{-1} B_t = V^{-1}$. Furthermore, $\mathbb{E}[(\pi - \bar{X}_t)' V^{-1} (\pi - \bar{X}_t)] = \mathbb{E}[\text{trace}(V^{-1}(\pi - \bar{X}_t)(\pi - \bar{X}_t)')] = \text{trace}(V^{-1} \cdot \frac{V}{t}) = \frac{(N+S)}{t}$, where $(N + S)$ is the dimension of the state vector.

Therefore:

$$\mathbb{E}[\varepsilon' \Lambda_t^{-1} \varepsilon] = \frac{t^2}{(t + \kappa_0)^2} \cdot \frac{(N + S)}{t} = \frac{t}{(t + \kappa_0)^2} (N + S) \tag{151}$$

For the second term, we have:

$$\mathbb{E}\left[\frac{1}{\sigma_m^2}(\iota'\varepsilon)^2\right] = \frac{1}{\sigma_m^2}\mathbb{E}\left[\left(\iota' B_t \frac{t}{t+\kappa_0}(\pi - \bar{X}_t)\right)^2\right] \tag{152}$$

$$= \frac{t^2}{(t+\kappa_0)^2} \cdot \frac{1}{\sigma_m^2}\mathbb{E}\left[(\iota' B_t(\pi - \bar{X}_t))^2\right] \tag{153}$$

Note that $\iota' B_t = \sum_i B_{i,t}$ represents the sum of the loadings. When we compute the expectation, we get $\mathbb{E}[(\iota' B_t(\pi - \bar{X}_t))^2] = \iota' B_t \cdot \frac{V}{t} \cdot B_t'\iota = \frac{1}{t}\iota' B_t V B_t'\iota = \frac{\sigma_m^2}{t}$.

Therefore:

$$\mathbb{E}\left[\frac{1}{\sigma_m^2}(\iota'\varepsilon)^2\right] = \frac{t^2}{(t+\kappa_0)^2} \cdot \frac{1}{\sigma_m^2} \cdot \frac{\sigma_m^2}{t} = \frac{t}{(t+\kappa_0)^2} \tag{154}$$

Combining the two terms:

$$\mathbb{E}[\alpha_t'\Lambda_t^{-1}\alpha_t] = \frac{t}{(t+\kappa_0)^2}(N+S) - \frac{t}{(t+\kappa_0)^2} \tag{155}$$

$$= \frac{t}{(t+\kappa_0)^2}(N+S-1) \tag{156}$$

For large t, this scales as $O\left(\frac{N+S}{t}\right)$, confirming that average pricing errors increase linearly with the number of assets and states, and decrease with the number of observations. □

Moreover,if investors have perfect information about covariances but an incorrect prior about the expectation (i.e., $\pi_0 \neq \pi$), the average pricing errors are greater than or equal to those with the correct prior:

$$\mathbb{E}[\alpha_t'\Lambda_t^{-1}\alpha_t] \geq \frac{t}{(t+\kappa_0)^2}(N+S-1) \tag{157}$$

Proof. When investors have an incorrect prior $\pi_0 \neq \pi$, the posterior belief becomes:

$$\hat{\pi}_t = \frac{\kappa_0\pi_0 + t\bar{X}_t}{t+\kappa_0} \tag{158}$$

The estimation error can be written as:

$$\varepsilon = B_t(\pi - \hat{\pi}_t) = B_t\left(\pi - \frac{\kappa_0\pi_0 + t\bar{X}_t}{t+\kappa_0}\right) \tag{159}$$

$$= B_t\frac{t(\pi - \bar{X}_t) + \kappa_0(\pi - \pi_0)}{t+\kappa_0} \tag{160}$$

This error has two components: sampling error $(\pi - \bar{X}_t)$ and prior bias $(\pi - \pi_0)$. Let us denote:

$$\varepsilon_1 = B_t\frac{t}{t+\kappa_0}(\pi - \bar{X}_t) \quad \text{(sampling error)} \tag{161}$$

$$\varepsilon_2 = B_t\frac{\kappa_0}{t+\kappa_0}(\pi - \pi_0) \quad \text{(prior bias)} \tag{162}$$

So $\varepsilon = \varepsilon_1 + \varepsilon_2$. The squared pricing errors are:

$$\alpha_t'\Lambda_t^{-1}\alpha_t = \varepsilon'\Lambda_t^{-1}\varepsilon - \frac{1}{\sigma_m^2}(\iota'\varepsilon)^2 \tag{163}$$

$$= (\varepsilon_1 + \varepsilon_2)'\Lambda_t^{-1}(\varepsilon_1 + \varepsilon_2) - \frac{1}{\sigma_m^2}(\iota'(\varepsilon_1 + \varepsilon_2))^2 \tag{164}$$

Expanding:

$$\alpha_t'\Lambda_t^{-1}\alpha_t = \varepsilon_1'\Lambda_t^{-1}\varepsilon_1 + 2\varepsilon_1'\Lambda_t^{-1}\varepsilon_2 + \varepsilon_2'\Lambda_t^{-1}\varepsilon_2 \tag{165}$$

$$- \frac{1}{\sigma_m^2}[(\iota'\varepsilon_1)^2 + 2(\iota'\varepsilon_1)(\iota'\varepsilon_2) + (\iota'\varepsilon_2)^2] \tag{166}$$

Taking expectations, note that $\mathbb{E}[\varepsilon_1] = 0$ since $\mathbb{E}[\bar{X}_t] = \pi$, but $\mathbb{E}[\varepsilon_2] = B_t\frac{\kappa_0}{t+\kappa_0}(\pi - \pi_0) \neq 0$.

Additionally, ε_1 and ε_2 are independent because ε_1 depends on random sampling while ε_2 is deterministic given the prior. Therefore:

$$\mathbb{E}[\varepsilon_1' \Lambda_t^{-1} \varepsilon_2] = \mathbb{E}[\varepsilon_1'] \Lambda_t^{-1} \varepsilon_2 = 0 \tag{167}$$

$$\mathbb{E}[(\iota' \varepsilon_1)(\iota' \varepsilon_2)] = \mathbb{E}[\iota' \varepsilon_1](\iota' \varepsilon_2) = 0 \tag{168}$$

This gives us:

$$\mathbb{E}[\alpha_t' \Lambda_t^{-1} \alpha_t] = \mathbb{E}[\varepsilon_1' \Lambda_t^{-1} \varepsilon_1] + \mathbb{E}[\varepsilon_2' \Lambda_t^{-1} \varepsilon_2] \tag{169}$$

$$- \frac{1}{\sigma_m^2} (\mathbb{E}[(\iota' \varepsilon_1)^2] + \mathbb{E}[(\iota' \varepsilon_2)^2]) \tag{170}$$

We've already shown that:

$$\mathbb{E}[\varepsilon_1' \Lambda_t^{-1} \varepsilon_1] - \frac{1}{\sigma_m^2} \mathbb{E}[(\iota' \varepsilon_1)^2] = \frac{t}{(t + \kappa_0)^2} (N + S - 1) \tag{171}$$

For the prior bias term:

$$\mathbb{E}[\varepsilon_2' \Lambda_t^{-1} \varepsilon_2] = \left(B_t \frac{\kappa_0}{t + \kappa_0} (\pi - \pi_0) \right)' \Lambda_t^{-1} \left(B_t \frac{\kappa_0}{t + \kappa_0} (\pi - \pi_0) \right) \tag{172}$$

$$= \frac{\kappa_0^2}{(t + \kappa_0)^2} (\pi - \pi_0)' B_t' \Lambda_t^{-1} B_t (\pi - \pi_0) \tag{173}$$

$$= \frac{\kappa_0^2}{(t + \kappa_0)^2} (\pi - \pi_0)' V^{-1} (\pi - \pi_0) \tag{174}$$

Similarly:

$$\mathbb{E}[(\iota' \varepsilon_2)^2] = \left(\iota' B_t \frac{\kappa_0}{t + \kappa_0} (\pi - \pi_0) \right)^2 \tag{175}$$

$$= \frac{\kappa_0^2}{(t + \kappa_0)^2} (\iota' B_t (\pi - \pi_0))^2 \tag{176}$$

By the Cauchy-Schwarz inequality, we have:

$$(\iota' B_t(\pi - \pi_0))^2 \leq (\iota' B_t V^{1/2})(V^{-1/2}(\pi - \pi_0))^2 = \sigma_m^2 \cdot (\pi - \pi_0)' V^{-1}(\pi - \pi_0) \tag{177}$$

Therefore:

$$\mathbb{E}[\varepsilon_2' \Lambda_t^{-1} \varepsilon_2] - \frac{1}{\sigma_m^2} \mathbb{E}[(\iota' \varepsilon_2)^2] \geq 0 \tag{178}$$

This proves that the expected pricing errors with an incorrect prior are greater than or equal to those with the correct prior:

$$\mathbb{E}[\alpha_t' \Lambda_t^{-1} \alpha_t] \geq \frac{t}{(t + \kappa_0)^2}(N + S - 1) \tag{179}$$

Equality holds if and only if $\pi_0 = \pi$ or if $(\pi - \pi_0)$ is perfectly aligned with the market portfolio direction. □

4.5 Unknown Means and Unknown Covariances

Proposition (Bounding Expected Pricing Errors)**.** *Consider the setting of unknown means and/or unknown covariances . Suppose that agents use (possibly incorrect) priors but observe t realizations of dividends* $\{d_\tau\}_{\tau=1}^t$ *before making period-t decisions, and that under the* true *data-generating process,* $\Lambda_t = \mathbb{V}ar_t[d_{t+1}]$ *and* $\alpha_t = (a_t - \hat{a}_t) + \gamma_N(\hat{\Lambda}_t - \Lambda_t)\,\iota$ *as before. Let* $\alpha_t' \Lambda_t^{-1} \alpha_t$ *denote the weighted sum of squared pricing errors (equivalently, the difference in squared Sharpe ratios).*

Then, under correct priors (or sufficiently informative priors) for the mean and covariance, one can show that

$$\mathbb{E}\Big[\alpha_t' \Lambda_t^{-1} \alpha_t\Big] \;\gtrsim\; c\,\frac{N+S}{t},$$

for some positive constant $c > 0$ *that depends on the baseline parameterization (e.g. risk aversion* γ*, dimensions* S, N*, prior degrees of freedom, etc.). More generally, even when priors are not fully*

correct, the following lower bound holds in expectation:

$$\mathbb{E}\big[\alpha_t' \Lambda_t^{-1} \alpha_t\big] \;\geq\; \kappa_1 \, \frac{N+S-1}{t+\kappa_2}, \tag{180}$$

for suitable constants κ_1, κ_2 (both depending on the prior strength and distribution of dividends). Hence, the expected magnitude of the pricing errors grows proportionally to the effective dimensionality $(N+S)$ of the problem, for a fixed number of observations t.

Proof. We analyze the general case where agents must estimate both the mean vector a_t and the covariance matrix Λ_t. The proof proceeds as follows.

Step 1: Decomposition of the pricing error term. From Lemma 1, we have the exact decomposition:

$$\alpha_t' \Lambda_t^{-1} \alpha_t = \underbrace{(a_t - \hat{a}_t)' \Lambda_t^{-1} (a_t - \hat{a}_t)}_{\text{Term (A): Mean estimation errors}} + \underbrace{2\gamma_N (a_t - \hat{a}_t)' \Lambda_t^{-1} (\hat{\Lambda}_t - \Lambda_t)\iota}_{\text{Term (B): Interaction term}} + \underbrace{(\gamma_N)^2 \iota' (\hat{\Lambda}_t - \Lambda_t) \Lambda_t^{-1} (\hat{\Lambda}_t - \Lambda_t)\iota}_{\text{Term (C): Covariance estimation errors}}.$$

Step 2: Analysis of Bayesian estimation errors. Under the normal-inverse-Wishart (NIW) prior framework, after observing t dividend realizations $\{d_\tau\}_{\tau=1}^t$, the posterior distributions for the mean and covariance matrix induce estimation errors with the following properties:

- For the mean estimation error, we have:

$$\mathbb{E}\big[(a_t - \hat{a}_t)' \Lambda_t^{-1} (a_t - \hat{a}_t)\big] \geq c_1 \frac{N+S}{t+\nu_0},$$

  where $c_1 > 0$ is a constant and ν_0 represents the prior degrees of freedom.

- For the covariance estimation error, standard results from multivariate statistics yield:

$$\mathbb{E}\big[\|(\hat{\Lambda}_t - \Lambda_t)\|_F^2\big] \geq c_2 \frac{N(N+1)}{t+\nu_0},$$

  where $\|\cdot\|_F$ denotes the Frobenius norm and $c_2 > 0$ is a constant.

Step 3: Bounding each term in the decomposition.

Term (A): By the properties of positive definite matrices and the bound established in Step 2:

$$\mathbb{E}[\text{Term (A)}] = \mathbb{E}\big[(a_t - \hat{a}_t)'\Lambda_t^{-1}(a_t - \hat{a}_t)\big] \geq c_1 \frac{N+S}{t+\nu_0}.$$

Term (C): Using the bound on covariance estimation error and the fact that $\gamma_N = \gamma/N$:

$$\begin{aligned}\mathbb{E}[\text{Term (C)}] &= \mathbb{E}\big[(\gamma_N)^2 \iota'(\hat{\Lambda}_t - \Lambda_t)\Lambda_t^{-1}(\hat{\Lambda}_t - \Lambda_t)\iota\big] \\ &\geq c_3 \frac{\gamma^2}{N^2} \cdot \frac{N(N+S)}{t+\nu_0} \\ &= c_3\gamma^2 \frac{N+S}{N(t+\nu_0)},\end{aligned}$$

where $c_3 > 0$ is another constant.

Term (B): For the interaction term, applying the Cauchy-Schwarz inequality and using our bounds from Terms (A) and (C):

$$\begin{aligned}|\mathbb{E}[\text{Term (B)}]| &= |2\gamma_N \mathbb{E}\big[(a_t - \hat{a}_t)'\Lambda_t^{-1}(\hat{\Lambda}_t - \Lambda_t)\iota\big]| \\ &\leq 2\gamma_N \sqrt{\mathbb{E}[\text{Term (A)}] \cdot \mathbb{E}[\text{Term (C)}]/\gamma_N^2} \\ &\leq \sqrt{c_1 c_3}\frac{2\gamma}{N}\sqrt{\frac{N+S}{t+\nu_0} \cdot \frac{N+S}{N(t+\nu_0)}} \\ &= 2\sqrt{c_1 c_3}\gamma \frac{N+S}{(t+\nu_0)\sqrt{N}}.\end{aligned}$$

Step 4: Combining the bounds. Combining all three terms (noting that Term (B) could be positive or negative), and using the fact that Term (A) and Term (C) are always positive, we have:

$$\begin{aligned}\mathbb{E}[\alpha_t'\Lambda_t^{-1}\alpha_t] &\geq \min\{c_1, c_3\gamma^2/N\}\frac{N+S}{t+\nu_0} - |\mathbb{E}[\text{Term (B)}]| \\ &\geq \min\{c_1, c_3\gamma^2/N\}\frac{N+S}{t+\nu_0} - 2\sqrt{c_1 c_3}\gamma \frac{N+S}{(t+\nu_0)\sqrt{N}}.\end{aligned}$$

For sufficiently large N or small γ, the positive terms dominate, yielding:

$$\mathbb{E}[\alpha_t' \Lambda_t^{-1} \alpha_t] \geq \kappa_1 \frac{N + S - 1}{t + \kappa_2},$$

where $\kappa_1, \kappa_2 > 0$ are constants that depend on the risk aversion parameter γ, the dimensions N and S, and the strength of the prior.

This lower bound establishes that expected pricing errors scale proportionally with the problem dimensionality $(N + S)$ relative to the number of observations t, highlighting the fundamental challenge of high-dimensional parameter estimation in asset pricing. □