Why Does Complexity Make Factor Models Fail in Equilibrium?

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

[1] *Indiana University*

[2] *University of Florida*

This Version: September 2024

Abstract

We examine the impact of complexity on asset pricing models and pricing errors in equilibrium when rational, risk-averse agents have imperfect knowledge of the data-generating process. Our model yields three implications as complexity rises: (1) pricing errors of factor models increase, (2) high Sharpe ratio strategies are more likely to exploit estimation error components, and (3) these strategies tend not to span each other. The model rationalizes several asset-pricing puzzles, including the limited explanatory power of factor models, the weak relationship between betas and average returns, and the proliferation of seemingly unrelated anomalies. We provide empirical evidence documenting high complexity in predictive characteristics and covariances and a low correlation between sophisticated strategies.

Asset pricing theory has long sought parsimonious models relating expected returns to risk. The capital asset pricing model (CAPM) provides an elegant one-factor model where assets with higher market betas demand higher expected returns. Multi-factor models extend this reasoning by identifying a small set of underlying risk factors driving returns. However, empirically, these parsimonious models fall short. Instead, recent findings show that complex models have very high Sharpe ratios, but surprisingly, they cannot explain each other's returns. In this paper, we explore how asset pricing models and pricing errors behave in equilibrium when rational and risk-averse agents do not know the data-generating process perfectly and are faced with complexity because of a large number of assets or potential forecasting variables. We find that an increase in complexity makes it more likely for any factor model to fail in equilibrium and for anomalies to proliferate.

We illustrate the impact of complexity with three equations that characterize many asset pricing puzzles in a setup where agents use feasible mean-variance optimization. The first one relates to pricing errors. Let Λ be the covariance matrix of return payoffs and μ the vector with their expected values. Let $w^* = \Lambda^{-1}\mu$ be the optimal weights under the true data-generating process. This portfolio is generally unknown in real-time because expected returns and covariances must be estimated. Let w be any vector of portfolio weights. Let $SR(w)$ denote the Sharpe ratio of the portfolio with weights w. Finally, let α_w be the vector of pricing errors implied by using this portfolio as an asset pricing model $\alpha_w = \mathbb{E}[r] - \beta_w \mathbb{E}[w'r]$. Then, similar to Gibbons, Ross, and Shanken (1989), the first equation states that the squared sum of weighted pricing errors is

$$\alpha_w' \Lambda^{-1} \alpha_w = SR(w^*)^2 - SR(w)^2. \tag{1}$$

We show theoretically that this difference is increasing in complexity. Intuitively, the Sharpe ratio of any feasible portfolio in equilibrium is bounded since it depends on the compensation for

risk. To get some intuition regarding the Sharpe ratio implied by the optimal weights, w^*, the second equation explores w^* when agents have imperfect information. Let $\hat{\mu}$ and $\hat{\Lambda}$ be agents' estimate about expected returns and covariances, γ be agents' risk aversion, ι a vector of ones, and n the number of assets in the economy. Under the model's assumptions, we can decompose the optimal weights as

$$w^* = \underbrace{\frac{\gamma}{n}\iota}_{\gamma\times\text{Market Portfolio}} + \underbrace{\Lambda^{-1}(\mu - \hat{\mu})}_{\text{estimation error in Expected Payoffs}} + \underbrace{\frac{\gamma}{n}\Lambda^{-1}(\hat{\Lambda} - \Lambda)\iota}_{\text{Estimation Error in Covariances}} . \tag{2}$$

Intuitively, there are three parts in the optimal weights. The first corresponds to the market portfolio that agents must hold in equilibrium. With perfect knowledge of the data-generating process and mean-variance agents, prices adjust so that the market portfolio is optimal. Moreover, its implied Sharpe ratio is determined by the total risk in the economy and agents' risk aversion. The second part exploits agents' imperfect information about expected payoffs, and the third part exploits their imperfect information about covariances. The model shows that the Sharpe ratio implied by these estimation error-related parts increases as the economy's complexity increases. Intuitively, if faced with agents that have imperfect information about expected returns or covariances, the optimal strategy is to put positions that exploit these estimation errors. Moreover, a growing complexity implies the estimation becomes harder for agents, and these estimation errors grow. Hence, high Sharpe ratio portfolios must exploit these estimation errors.

The third equation shows how to write the scaled pricing errors of a portfolio y against a benchmark portfolio x. Let σ_y denote the standard deviation of the portfolio y and $\rho_{x,y}$ the correlation between portfolios x and y. Then

$$\frac{\alpha}{\sigma_y} = SR(y) - \rho_{y,x} SR(x). \tag{3}$$

Equation 3 characterizes when portfolios can explain each other. Assume both have positive Sharpe ratios (otherwise, we can create short portfolios with negative Sharpe ratios). If portfolio x has a lower Sharpe ratio, it cannot price portfolio y, albeit the pricing errors decrease if the portfolios are more correlated. More importantly, if the portfolios have little or no correlation with each other, then the magnitude of the Sharpe ratio of the pricing portfolio x plays no role in the amount of pricing errors. In light of Equation 2, we would expect high Sharpe ratio portfolios to be loading on assets where the estimation errors are highest.

We expect high Sharpe ratio portfolios to have little correlation with each other in complex markets because, in a complex market with numerous assets or predictive variables, the space of possible estimation errors—both in expected payoffs and covariances—expands dramatically. Consequently, different high Sharpe ratio strategies are likely to exploit distinct, potentially orthogonal sources of mispricing or risk premia. Different models or agents, each with their own estimation errors, may identify and exploit different aspects of market inefficiencies, and it is likely that these errors are uncorrelated. This framework explains why complex models can simultaneously achieve high Sharpe ratios and fail to explain each other's returns and provides insight into the proliferation of seemingly unrelated anomalies in the asset pricing literature. These anomalies may represent different aspects of the estimation errors.

In summary, the model has three main implications. First, pricing errors should increase as the complexity of the economy grows. We formally show that average square pricing errors grow with the state space or the number of assets in the economy, even without uncertainty about expected returns. Second, high Sharpe ratio strategies are more likely to load on estimation error components in high-complexity environments. Third, because these strategies load on diverse estimation errors and the higher room for anomalies, we expect many high-Sharpe ratio strategies that do not span each other as complexity increases. Hence, the model helps rationalize several asset pricing puzzles.

First, the limited explanatory power of factor models, especially parsimonious ones, can be explained by the high complexity of the economy. Second, because pricing errors increase, a diminished cross-sectional relationship exists between betas and average returns, thus explaining why, counterintuitively, assets with higher exposure to systematic risk do not yield higher returns. For example, higher market exposure is not compensated (the security market line is very flat), and time-series exposure to standard multi-factor models is not rewarded either (Daniel and Titman (1997), Lopez-Lira and Roussanov (2022)).

Third, we explain why many characteristics that predict returns form an 'anomaly zoo.' An ample state space where multiple characteristics have predictive power is generally not anomalous. The anomalies only happen because the expected returns are unrelated to factor exposures. Anomalies are more likely to increase in complex settings because as the number of potential predictive variables increases, both the (total) amount of those with predictability increase, and it is harder for factor models to price the implied portfolios since they are further from being optimal. Finally, while numerous methods have been proposed to predict returns and develop optimal portfolios, these portfolios do not span one another despite their relatively high Sharpe ratios (Baba Yara, Boyer, and Davis 2021). Our model shows that these methods are expected to load on increasingly diverse parts as the complexity of the economy increases.

Further, we provide evidence documenting the high complexity of the economy and thus supporting our model's assumptions. Consistent with the literature on anomalies, we document an extensive proliferation of firm-based characteristics with return predictability thus complicating es.[1] We find them even after the 2000s when return predictability is known to have substantially decreased. Of the 17,517 characteristic-sorted portfolios we examine, 347 (2%) have an absolute t-stat greater than 3 over our entire sample period (1980-2020). 197 have absolute t-stats greater

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

than 3, only considering 2000-2020. The alphas' t-statistics of these characteristic portfolios show a similar pattern exceeding chance levels. Moreover, using the same portfolios, we find evidence for a high dimensionality in the covariance space. About 20% of the overall variance of one-year daily returns is explained by the first ten greatest principal components. For 50% of the variation, we need about 50 factors and 75 factors to explain roughly 75% of the variation. Hence, this high complexity makes it harder to estimate the covariance structure.

Finally, our empirical analysis provides strong support for the model's prediction of low correlations between investment strategies in a high-dimensional world. Examining a diverse set of investment models from the literature, we find an average R^2 of around 7%, indicating substantial divergence in strategy returns. More compellingly, when we analyze the active components of actual systematic hedge fund portfolios using 13F holdings data, we observe an average R^2 lower than 1%. This remarkably low correlation among sophisticated investors' active bets strongly suggests that different funds are exploiting distinct sources of alpha, consistent with our model's implications for a complex, high-dimensional market environment.

Related Literature

Our paper builds on the existing literature modeling economies with high-dimensional properties. The most related work is Da, Nagel, and Xiu (2022), who take a partial equilibrium approach to study the statistical limits of arbitrage. They take the existence of alphas as given and focus on characterizing the attainable Sharpe ratios and optimal arbitrage strategies for investors facing uncertainty about these alphas. While their analysis also involves parameter uncertainty, it does not fully account for the equilibrium implications of these limits.

In a partial equilibrium setting, the Sharpe ratios of the optimal portfolio under the true data-generating process and the feasible portfolios constructed by arbitrageurs are treated as exogenous.

However, these Sharpe ratios are determined endogenously in equilibrium and depend on the interaction between arbitrageurs' trading strategies and the beliefs and preferences of other investors in the economy. Our general equilibrium model, in contrast, explicitly accounts for the endogenous determination of Sharpe ratios, asset prices and pricing errors. The equilibrium shows the compensation for risk does not depend substantially on the economy's complexity, in contrast to the pricing errors.

Martin and Nagel (2022) explore a model with numerous potential predictors and risk-neutral Bayesian investors learning about them. They show that only out-of-sample tests retain their economic meaning. We complement their paper by adding risk-averse agents and imperfect knowledge about the covariance, allowing us to study risk pricing and how strategies relate. Moreover, we show that factor model testing methods are bias against the factors Kelly, Malamud, and Zhou (2021) document that more complex models regularized appropriately do better than simpler models when predicting the market's expected returns.

Kozak et al. (2023) provide significant theoretical results on the conditions under which characteristics-based factor models span the stochastic discount factor and empirical analyses supporting those results. Our paper complements theirs by offering a model explaining asset pricing puzzles through rational, risk-averse investors with imperfect information, formally deriving how pricing errors increase with dimensionality, and confirming empirically the model's novel predictions about rising pricing errors and non-spanning as the number of assets grows.

Our model assumes return predictability based on the extensive literature documenting it. For example, Fama and MacBeth (1973), Jacobs and Levy (1988), Lewellen (2015), Gu, Kelly, and Xiu (2020), and Freyberger, Neuhierl, and Weber (2020a), among many others, provide ample evidence that cross-sectional returns are predictable, and machine learning enhances these predictions.[2]

2. See especially Gu, Kelly, and Xiu (2020) for an excellent survey of standard methods. See Alti and Titman (2019) for a different model that justifies return predictability.

We further contribute to the literature on high-dimensional asset pricing. For example, Kelly, Malamud, and Zhou (2022) constructs long-short portfolios using random features by first applying thousands of random linear combinations and then applying a non-linear transformation, obtaining thousands of new, linearly independent signals. Bryzgalova, Pelger, and Zhu (2020a) uses decision trees to group stocks into portfolios of derived characteristics. Didisheim et al. (2023) show that machine learning factor models improve in model parameterization (complexity). We complement their research by providing an equilibrium framework to understand how the agent's approximation quality of the economy impacts risk premia and abnormal returns.

Our paper extends Andrei, Cujean, and Wilson (2023), who demonstrate how informational differences between investors and empiricists can distort beta estimates and flatten the Security Market Line, helping to explain the failure of the CAPM despite its widespread practical use. We expand on their framework by observing how complexity and incomplete information can generally impact pricing errors and anomalies even without informational differences.

1 Model

We work with economies with an increasing number of assets n. d_t is the n dimensional column vector of dividends in period t. Dividends have a multivariate normal distribution. The n dimensional column vector p_t contains the prices of the assets in period t. Let $P_t = \text{diag}(p_t)$, which means that P_t is the $n \times n$ dimensional matrix with zeros on the off-diagonal and p_t on the diagonal. 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]

The dividends drive the asset payoffs. This allows the covariance matrix of returns to not be

3. The terms share excess returns, dollar excess returns, and excess dollar payoffs are all used in the literature. It is standard to work with returns in this form rather than percentage returns.

a function of endogenous prices in the next period, which makes the model more tractable. The conditional covariance matrix of returns is denoted as $\Lambda_t = \mathbb{V}\text{ar}_t[d_{t+1}] = \mathbb{V}\text{ar}_t[r_{t+1}]$. We denote the (true) expectation of d_t as $a_t = E_t[d_{t+1}]$ and the (true) expectation of r_t as $\mu_t = E_t[r_{t+1}]$. We drop time subscripts when the context is clear.

We first start with the general case. Let $w^* = \Lambda^{-1}\mu$ be the optimal combination of weights. Let w be any vector of portfolio weight. Finally, let α_w be the vector of pricing errors implied by using this portfolio as an asset pricing model:

$$\alpha_w = \mathbb{E}[r] - \beta_w \mathbb{E}[w'r] \tag{4}$$

$$= \mu - \beta_w \mu_w \tag{5}$$

$$= \mu - \frac{I\Lambda w}{w'\Lambda w} w'\mu. \tag{6}$$

Let $SR(w)$ denote the Sharpe ratio of the portfolio with weights w. Then, similar to Gibbons, Ross, and Shanken (1989), we have:

$$\alpha_w'\Lambda^{-1}\alpha_w = (\mu - \frac{\Lambda w w'\mu}{w'\Lambda w})'\Lambda^{-1}(\mu - \frac{\Lambda w w'}{w'\Lambda w}) \tag{7}$$

$$= \mu'\Lambda^{-1}\mu - \frac{(w'\mu)^2}{w'\Lambda w} \tag{8}$$

$$= SR(w^*)^2 - SR(w)^2 \tag{9}$$

The sum of pricing errors squared appropriately scaled generally increases in the difference between the optimal portfolio under the true data-generating process optimal Sharpe ratio and the other portfolio that serves as a factor model.

We will focus on the case where investors are mean-variance optimizers with imperfect infor-

mation and normal shocks for tractability. Moreover, this is the best-case scenario for CAPM to hold. In particular, we will focus on the case of CARA investors with absolute risk aversion γ that have demand of the form:

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

$\hat{\Lambda}$, and $\hat{a}$ are agents' estimates of the covariance matrix and the expected dividends. Agents hold the feasible mean-variance portfolios, which usually differ from the optimal portfolio under the true data-generating process portfolios due to imperfect information about Λ_t and a_t.

In the extreme perfect knowledge case, for example, agents can have $\hat{\Lambda} = \Lambda$ and $\hat{a} = a$, although we will usually relax these assumptions.

We assume there is a unit measure of identical agents in the economy and a supply of $\frac{1}{n}$ per asset.[4] Plugin in the demand function, and solving for the equilibrium leads to clearing prices:

$$p_t = \frac{1}{1 + r_f}\left(\hat{a} - \frac{\gamma}{n}\hat{\Lambda}\iota\right). \tag{10}$$

For notational simplicity, let $\gamma_n = \frac{\gamma}{n}$. The expected excess returns are given by:

$$\mu_t \equiv \mathbb{E}_t[d_{t+1}] - (1 + r_f)p_t = a_t - \hat{a}_t + \gamma_n\hat{\Lambda}\iota. \tag{11}$$

Let $\varepsilon_t = a_t - \hat{a}_t$ denote the estimation error and $s\varepsilon = \varepsilon'\iota$ the sum of the estimation errors. Let $\sigma_m^2 = \iota'\Lambda\iota$ denote the market variance and $\hat{\sigma}_m^2 = \iota'\hat{\Lambda}\iota$ the perceived variance. Notice the market variance does not depend on agents' expectations.

4. This ensures that under perfect information and common variation in the dividends, the maximum Sharpe ratio of the economy neither increases without bound nor decreases to zero as we increase n. With a unit asset supply, $SR(w^*)^2 = \mu'\Lambda^{-1}\mu = \gamma^2\iota'\Lambda\Lambda^{-1}\Lambda\iota = \gamma^2\iota'\Lambda\iota \in O(n^2)$. In the case of uncorrelated assets, $SR(w^*)^2 \in O(n)$.

Hence, the optimal (optimal portfolio under the true data-generating process) weights of the mean-variance portfolio are given by

$$w^* = \Lambda_t^{-1}\mu_t \equiv \mathbb{E}_t[d_{t+1}] - (1+r_f)p_t \tag{12}$$

$$= \Lambda^{-1}(a_t - \hat{a}_t) + \gamma_n \Lambda^{-1}\hat{\Lambda}\iota \tag{13}$$

$$= \Lambda^{-1}(a_t - \hat{a}_t) + \gamma_n \Lambda^{-1}\hat{\Lambda}\iota + \gamma_n \Lambda^{-1}\Lambda\iota - +\gamma_n \Lambda^{-1}\Lambda\iota \tag{14}$$

$$= \Lambda^{-1}\varepsilon_t + \gamma_n \Lambda^{-1}(\hat{\Lambda} - \Lambda)\iota + \gamma_n \iota. \tag{15}$$

In equilibrium, abnormal returns with respect to the market portfolio are given by:

$$\alpha_t = \underbrace{\mathbb{E}_t[r_{t+1}]}_{\text{excess return}} - \underbrace{\frac{\Lambda\iota}{\iota'\Lambda\iota}}_{\beta} \underbrace{\iota'\mathbb{E}_t[r_{t+1}]}_{\text{market excess return}} \tag{16}$$

$$= \left(I - \frac{\Lambda\iota\iota'}{\iota'\Lambda\iota}\right)(\mathbb{E}_t[d_{t+1}] - (1+r_f)p_t) \tag{17}$$

$$= \left(I - \frac{\Lambda\iota\iota'}{\iota'\Lambda\iota}\right)(a - \hat{a} + \gamma_n\hat{\Lambda}\iota). \tag{18}$$

Moreover, the weighted sum of squared pricing errors in the economy is given by

$$\alpha'\Lambda^{-1}\alpha = SR(w^*)^2 - SR(\frac{1}{n}\iota)^2 \tag{19}$$

$$= (\varepsilon + \gamma_n\hat{\Lambda}\iota)'\Lambda^{-1}(\varepsilon + \gamma_n\hat{\Lambda}\iota) - \frac{1}{\sigma_m^2}((\iota'\varepsilon)^2 + 2\gamma_n(\iota'\varepsilon)\hat{\sigma}_m^2 + \frac{\gamma^2}{n^2}(\hat{\sigma}_m^2)^2) \tag{20}$$

$$= \varepsilon'\Lambda^{-1}\varepsilon + 2\gamma_n\iota'\hat{\Lambda}\Lambda^{-1}\varepsilon + \frac{\gamma^2}{n^2}\iota'\hat{\Lambda}\Lambda^{-1}\hat{\Lambda}\iota - \frac{1}{\sigma_m^2}((\iota'\varepsilon)^2 + 2\gamma(\iota'\varepsilon)\hat{\sigma}_m^2 + \frac{\gamma^2}{n^2}(\hat{\sigma}_m^2)^2). \tag{21}$$

Equation 19 completely characterizes the pricing errors given the statistical properties of the

dividends and agents' estimate about them. In what follows, we first consider specific cases when the either the means or covariances are known, and then we impose statistical assumptions to get analytical results.

1.0.1 Known Means and Unknown Covariances

To get some intuition, consider the case of $\varepsilon = 0$, where agents know the true expected value of the fundamental. In that case, equation 19 specializes to:

$$\alpha'\Lambda^{-1}\alpha = \frac{\gamma^2}{n^2}\iota'\hat{\Lambda}\Lambda^{-1}\hat{\Lambda}\iota - \frac{\frac{\gamma^2}{n^2}(\hat{\sigma}_m^2)^2}{\sigma_m^2} \tag{22}$$

$$= \frac{\gamma^2}{n^2}(\iota'\hat{\Lambda}\Lambda^{-1}\hat{\Lambda}\iota - \frac{(\iota'\hat{\Lambda}\iota)^2}{\iota'\Lambda\iota}). \tag{23}$$

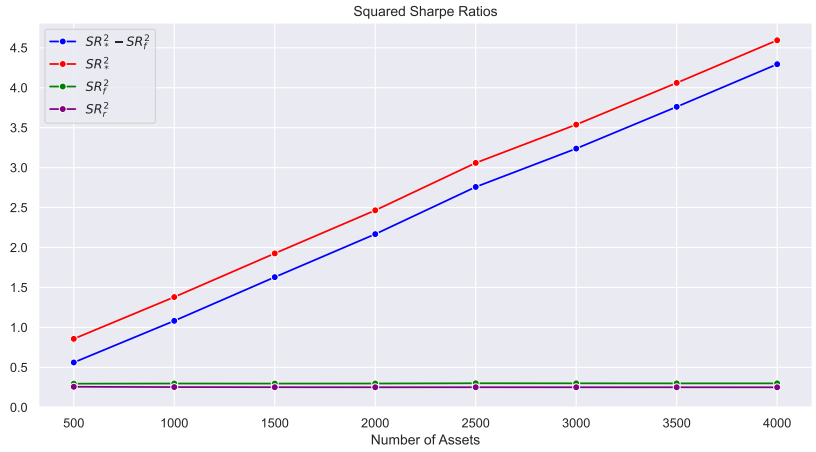
Figure 1 depicts the components of the weighted squared pricing errors for a simulated economy calibrated to target an equilibrium Sharpe ratio of 0.5 when there is perfect information. We assume that asset payoffs are driven by a multi-factor model with 5 factors. The first factor has a dominant influence, 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 that 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 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.

1.0.2 Unknown Means and Known Covariances

If covariances are known, then the squared pricing errors are given by

$$\alpha'\Lambda^{-1}\alpha = \varepsilon'\Lambda^{-1}\varepsilon + 2\gamma_n\iota'\Lambda\Lambda^{-1}\varepsilon + \frac{\gamma^2}{n^2}\iota'\Lambda\Lambda^{-1}\Lambda\iota - \frac{1}{\sigma_m^2}((\iota'\varepsilon)^2 + 2\gamma(\iota'\varepsilon)\sigma_m^2 + \frac{\gamma^2}{n^2}(\sigma_m^2)^2) \tag{24}$$

$$= \varepsilon'\Lambda^{-1}\varepsilon - \frac{1}{\sigma_m^2}(\iota'\varepsilon)^2. \tag{25}$$

Similarly as before, we can see the pricing errors are generally increasing.

1.0.3 Unknown Means Covariances

The last case of interest is considering unknown means and covariances but where the estimation error is approximately zero on average, $\iota'\varepsilon \approx 0$, and a slightly stronger approximation imposing that $\iota' Ab \approx 0$ where A is positive definite. In that case, we have $\iota'\varepsilon \approx 0$ and:

$$\alpha'\Lambda^{-1}\alpha \approx \varepsilon'\Lambda^{-1}\varepsilon + \frac{\gamma^2}{n^2}\iota'\hat{\Lambda}\Lambda^{-1}\hat{\Lambda}\iota - \frac{1}{\sigma_m^2}(\frac{\gamma^2}{n^2}(\hat{\sigma}_m^2)^2) \tag{26}$$

$$= \varepsilon'\Lambda^{-1}\varepsilon - \frac{1}{\sigma_m^2}(\iota'\varepsilon)^2 + \frac{\gamma^2}{n^2}(\iota'\hat{\Lambda}\Lambda^{-1}\hat{\Lambda}\iota - \frac{(\iota'\hat{\Lambda}\iota)^2}{\iota'\Lambda\iota}) \tag{27}$$

$$> \frac{\gamma^2}{n^2}(\iota'\hat{\Lambda}\Lambda^{-1}\hat{\Lambda}\iota - \frac{(\iota'\hat{\Lambda}\iota)^2}{\iota'\Lambda\iota}). \tag{28}$$

Effectively, the pricing errors increase because of the estimation error component ε and the previous unknown covariance component. We now impose more statistical structure in the economy to get analytical results on the bounds of the pricing errors.

1.1 Statistical Assumptions

Assume each asset i has:

$$d_{i,t+1} = b'_{i,t} x_{t+1} + c_{i,t} z_{i,t+1} \tag{29}$$

where x_{t+1} is an s dimensional column vector of state variables, and $b_{i,t}$ is a vector of loadings. We can think of $b_{i,t}$ as firm characteristics such as the firm's size, age, capital structure, productivity, or management quality that can impact its dividend policy when interacted with productivity shocks.[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 thought of as a reduced form specification of the production function, where firm production potentially covaries with each other via common shocks. Moreover, there is a is a firm-specific state variable $z_{i,t+1}$ with loading $c_{i,t}$.

Stacking this last equation, we can write

$$d_{t+1} = B_t X_{t+1} \tag{30}$$

where d_{t+1} is the n dimensional vector of dividends,

5. Alternatively, we can think of x_{t+1} as the coefficients of OLS regressions as in Martin and Nagel (2022).

X_{t+1} is the $s+n$ dimensional column vector defined as

$$x_t = \begin{bmatrix} x_{1,t} \\ x_{2,t} \\ \vdots \\ x_{s,t} \\ z_{1,t} \\ \vdots \\ z_{n,t} \end{bmatrix}, \tag{31}$$

and B_t is the $n \times (s+n)$ matrix defined as

$$B_t = \begin{bmatrix} b'_{1,t} & c_{1t} & \dots & 0 \\ b'_{2,t} & 0 & c_{2t} \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ b'_{n,t} & 0 & 0 & \dots & c_{nt} \end{bmatrix} \tag{32}$$

We assume that investors perfectly observe B_t. Also, assume that conditional on time t information ($\mathcal{F}_t$):

$$X_{t+1}|\mathcal{F}_t \sim N(\pi, V). \tag{33}$$

It follows that

$$\mathbb{E}_t[d_{t+1}] = B_t\pi, \tag{34}$$

and we can also write

$$\Lambda_t = B_t V B'_t. \tag{35}$$

Dividends follow a multivariate normal distribution. Let ι be an n dimensional column vector of ones as before and the supply of each asset to be $\frac{1}{n}$. We first illustrate the baseline case with perfect knowledge.

1.2 Bayesian Learning About Covariance Matrix

Investors use the T observations to learn about X_t. Assume that investors have the following prior about the covariance matrix V:

$$V \sim W_{s+n}^{-1}(V_0, \nu) \quad \text{(Inverse-Wishart Distribution)} \tag{36}$$

This yields a well-known conjugate prior result, and the investor has posterior beliefs:

$$X_{T+1} \sim t_{T+\nu-(s+n)+1}\left(\pi, \frac{1}{T+\nu-(s+n)-1}\left(V_0 + \sum_{t=1}^{T}(X_t-\pi)(X_t-\pi)'\right)\right). \tag{37}$$

Where this notation means indicates a multivariate t distribution. We assume that multivariate t distributions are intractable for investors, so they approximate the posterior using a multivariate normal distribution:

$$X_{T+1} \sim N\left(\pi, \frac{1}{T+\nu-(s+n)-1}\left(V_0 + \sum_{t=1}^{T}(X_t-\pi)(X_t-\pi)'\right)\right) \tag{38}$$

Hence, investors believe that

$$r_{T+1} = B_T X_{T+1} - (1 + r_{f,T}) p_T \sim \tag{39}$$

$$N\left(B_T\pi - (1 + r_{f,T})p_T, \frac{1}{T + \nu - (s + n) - 1} B_T \left(V_0 + \sum_{t=1}^{T} (X_t - \pi)(X_t - \pi)'\right) B_T'\right) \tag{40}$$

Let's define:

$$\mu_T = B_T\pi - (1 + r_{f,T})p_T \tag{41}$$

$$\hat{\Lambda}_T = \frac{1}{T + \nu - (s + n) - 1} B_T \left(V_0 + \sum_{t=1}^{T} (X_t - \pi)(X_t - \pi)'\right) B_T' \tag{42}$$

Theorem 1. *Even if investors have the correct prior, the average pricing errors relative to the feasible optimal portfolio increase in the number of assets and states at a rate of n.*

$$\mathbb{E}[\alpha_t' \Lambda^{-1} \alpha_t] = \frac{\gamma^2}{n^2} \left(\frac{1}{T+1}\right)^2 T(n-1)(s+n)^2 \bar{\varrho} \approx c\frac{n}{T} \tag{43}$$

Proof. See Appendix □

1.3 Uncertainty about Mean and Covariance

Now, we consider Bayesian learners who are uncertain about the mean and the covariance matrix. In particular, we assume they have normal-inverse-Wishart prior:

$$(\pi, V) \sim NIW(\pi_0, \kappa_0, V_0, \nu) \tag{44}$$

where π is the expected value of X_t.

This implies that

$$X_{T+1} \sim t_{\nu_0+T-(s+n)+1}\left(\frac{\kappa_0\pi_0 + T\bar{X}_T}{T+\kappa_0}, \frac{\kappa_0+T+1}{(\kappa_0+T)(\nu_0+T-(s+n)+1)}\left(V_0 + \sum_{t=1}^{T}(X_t-\bar{X}_T)(X_t-\bar{X}_T)' + \frac{\kappa_0 T}{\kappa_0+T}(\bar{X}_T-\pi_0)(\bar{X}_T-\pi_0)'\right)\right) \tag{45}$$

Again, investors use a multivariate normal approximation. In particular, investors believe:

$$X_{T+1} \sim N\left(\frac{\kappa_0\pi_0 + T\bar{X}_T}{T+\kappa_0}, V_T\right) \tag{46}$$

where

$$V_T = \frac{\nu_0+T-(s+n)+1}{\nu_0+T-(s+n)-1}\frac{\kappa_0+T+1}{(\kappa_0+T)(\nu_0+T-(s+n)+1)}\left(V_0 + \sum_{t=1}^{T}(X_t-\bar{X}_T)(X_t-\bar{X}_T)' + \frac{\kappa_0 T}{\kappa_0+T}(\bar{X}_T-\pi_0)(\bar{X}_T-\pi_0)'\right) \tag{47}$$

$$= \frac{\kappa_0+T+1}{(\kappa_0+T)(\nu_0+T-(s+n)-1)}\left(V_0 + \sum_{t=1}^{T}(X_t-\bar{X}_T)(X_t-\bar{X}_T)' + \frac{\kappa_0 T}{\kappa_0+T}(\bar{X}_T-\pi_0)(\bar{X}_T-\pi_0)'\right) \tag{48}$$

This means investors believe that

$$r_{T+1} \sim N\left(\hat{\mu}_T - (1+r_{f,T})p_T, \hat{\Lambda}_T\right) \tag{49}$$

where

$$\hat{\mu}_T = \frac{\kappa_0}{T+\kappa_0}B_T\pi_0 + \frac{T}{T+\kappa_0}B_T\bar{X}_T, \text{ and} \tag{50}$$

$$\hat{\Lambda}_T = \frac{\kappa_0+T+1}{(\kappa_0+T)(\nu_0+T-(s+n)-1)}B_T\left(V_0 + \sum_{t=1}^{T}(X_t-\bar{X}_T)(X_t-\bar{X}_T)' + \frac{\kappa_0 T}{\kappa_0+T}(\bar{X}_T-\pi_0)(\bar{X}_T-\pi_0)'\right)B_T' \tag{51}$$

In this case, the equilibrium price will be

$$p_T = \frac{1}{1 + r_{T,f}} \left(\hat{\mu}_T - \gamma_n \hat{\Lambda}_T \iota \right) \tag{52}$$

The α_t is given as

$$\alpha_t = \underbrace{\mathbb{E}_t[r_{t+1}]}_{\text{excess return}} - \underbrace{\frac{\Lambda\iota}{\iota'\Lambda\iota}}_{\beta} \underbrace{\iota'\mathbb{E}_t[r_{t+1}]}_{\text{market excess return}} = \left(I - \frac{\Lambda\iota\iota'}{\iota'\Lambda\iota} \right) (\mathbb{E}_t[d_{t+1}] - (1 + r_f) p_t) \tag{53}$$

Recall that alpha is:

$$\alpha_T = \left(I - \frac{\Lambda\iota\iota'}{\iota'\Lambda\iota} \right) (\mathbb{E}_T[d_{T+1}] - (1 + r_{T,f}) p_T) \tag{54}$$

$$= \left(I - \frac{\Lambda\iota\iota'}{\iota'\Lambda\iota} \right) (\mu - \hat{\mu}_T + \gamma_n \hat{\Lambda}_T \iota) \tag{55}$$

$$= \left(I - \frac{\Lambda\iota\iota'}{\iota'\Lambda\iota} \right) (\mu - \hat{\mu}_T) + \left(I - \frac{\Lambda\iota\iota'}{\iota'\Lambda\iota} \right) (\gamma_n \hat{\Lambda}_T \iota) \tag{56}$$

So we have:

$$\alpha_T' \Lambda^{-1} \alpha_T = (\mu - \hat{\mu}_T + \gamma_n \hat{\Lambda}_T \iota)' \left(I - \frac{\Lambda\iota\iota'}{\iota'\Lambda\iota} \right)' \Lambda^{-1} \left(I - \frac{\Lambda\iota\iota'}{\iota'\Lambda\iota} \right) (\mu - \hat{\mu}_T + \gamma_n \hat{\Lambda}_T \iota) \tag{57}$$

$$= (\mu - \hat{\mu}_T + \gamma_n \hat{\Lambda}_T \iota)' \left(\Lambda^{-1} - \frac{\iota\iota'}{\iota'\Lambda\iota} \right) (\mu - \hat{\mu}_T + \gamma_n \hat{\Lambda}_T \iota) \tag{58}$$

$$\tag{59}$$

$$= \mu'\Lambda^{-1}\mu + \hat{\mu}_T'\Lambda^{-1}\hat{\mu}_T + \frac{\gamma^2}{n^2}\iota'\hat{\Lambda}_T\Lambda^{-1}\hat{\Lambda}_T\iota - 2\mu'\Lambda^{-1}\hat{\mu}_T + 2\gamma_n\mu'\Lambda^{-1}\hat{\Lambda}_T\iota - 2\gamma_n\iota'\hat{\Lambda}_T\Lambda^{-1}\hat{\mu}_T \tag{60}$$

$$+\frac{1}{\iota'\Lambda\iota}\left(\mu'\iota\iota'\mu + \hat{\mu}_T'\iota\iota'\hat{\mu}_T + \frac{\gamma^2}{n^2}\iota'\hat{\Lambda}_T\iota\iota'\hat{\Lambda}_T\iota - 2\mu'\iota\iota'\hat{\mu}_T + 2\gamma_n\mu'\iota\iota'\hat{\Lambda}_T\iota - 2\gamma_n\iota'\hat{\Lambda}_T\iota\iota'\hat{\mu}_T\right) \tag{61}$$

This can be shown to increase in the number of assets and states similarly than with only unknown covariances or unknown means.

1.4 Bayesian Learning About Means with Heterogenous Beliefs

Now, we consider investors who learn about the mean of X_t, rather than the covariance. The mean is unknown, but the covariance is known. 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. We assume investor type i has the following prior about the mean of returns:

$$\pi \sim N(\pi_i, V_i) \tag{62}$$

We define $\zeta_i = V_i^{-1}\pi_i$. This implies that investors have the following conjugate prior:

$$X_{T+1} \sim N\left(\left(V_i^{-1} + TV^{-1}\right)^{-1}\left(V_i^{-1}\pi_i + V^{-1}\sum_{t=1}^{T} X_t\right), \left(V_i^{-1} + TV^{-1}\right)^{-1} + V\right) \tag{63}$$

This implies the investor believes that returns have the following conditional distribution:

$$r_{T+1} \sim N\left(B_T\left(V_i^{-1} + TV^{-1}\right)^{-1}\left(V_i^{-1}\pi_i + V^{-1}\sum_{t=1}^{T} X_t\right) - (1+r_f)p_t, B_T\left(\left(V_i^{-1} + TV^{-1}\right)^{-1} + V\right)B_T'\right) \tag{64}$$

Then CARA demand for investor type i is given by:

$$x_{i,t} = \frac{1}{\gamma}\left(B_T\left(\left(V_i^{-1} + TV^{-1}\right)^{-1} + V\right)B_T'\right)^{-1}\left(B_T\left(V_i^{-1} + TV^{-1}\right)^{-1}\left(V_i^{-1}\pi_i + V^{-1}\sum_{t=1}^{T} X_t\right) - (1 + r_f)p_t\right) \tag{65}$$

Equilibrium is given by:

$$\iota = \sum_{i=1}^{I} \theta_i x_{i,t}. \tag{66}$$

Again, the left-hand side is the fixed supply, and the right-hand side is the aggregate demand across investor types. Define ζ to be equal to $\sum_{i=1}^{I} \theta_i \zeta_i$.

We consider the following assumptions:

Assumption 1: The prior covariance matrix about the mean is simply proportional to the actual covariance matrix:

$$V_i = \kappa^{-1} V, \tag{67}$$

where κ is a positive constant. This is intuitive: just as the standard errors of the mean are proportional to the covariance matrix, the prior covariance matrix about the mean is proportional to the covariance matrix.

Assumption 2: $B_t = I$, meaning that B_t is just the identity matrix. We just use $s = n$ as the dimensionality of $x_{i,t}$ instead of $s + n$ as it is above, making the problem effectively a lower-dimensional problem.

This means the number of states equals the number of assets ($n = s$), simplifying the problem.

This allows us to introduce the following proposition:

Theorem 2. *Define the beta of investor i type's portfolio relative to investor j's type portfolio. Let*

us consider this beta in the limit, that is define:

$$\beta_{i,j} = \operatorname*{plim}_{\substack{T,S\to\infty \\ S/T=c}} \frac{x'_{i,T}\Lambda x_{j,T}}{x'_{j,T}\Lambda x_{j,T}}. \tag{68}$$

Let a_s be the s^{th} element of $\zeta_i - \zeta_j$ and let b_s be the s^{th} element of $\zeta_j - \zeta$. Define:

$$\omega_S = \frac{1}{s^2}\sum_{s=1}^{s}\sum_{n=1}^{s} a_s\Lambda_{s,n}b_n, \quad \textit{and} \quad \tau_S = \frac{1}{s}\sum_{s=1}^{s}\sum_{n=1}^{s} a_s\Lambda_{s,n}. \tag{69}$$

Then define:

$$\omega = \lim_{s\to\infty}\omega_S, \quad \tau = \lim_{s\to\infty}\tau_S, \quad \textit{and} \quad \sigma^2 = \operatorname*{plim}_{s\to\infty} x'_{j,T}\Lambda x_{j,T}. \tag{70}$$

Then as long as $\gamma_n^{-2}c\omega + \gamma_n^{-1}c\tau \neq 0$, we have

$$\beta_{i,j} = \frac{\gamma_n^{-2}c\omega + \gamma_n^{-1}c\tau + \sigma^2}{\sigma^2} \neq 1. \tag{71}$$

The proof is in the appendix.

This theorem shows that betas do not fully work to price assets in a large dimensional world. In other words, the posteriors of these different investors, if separated across the many states/assets (i.e. $\gamma^{-2}c\omega+\gamma^{-1}c\tau \neq 0$), then the betas will not equal one, even with many time periods. Strikingly, the only difference in these portfolios is the prior. The risk-aversion γ is the same, and it is easy to show that with identical risk-aversion and a small number of assets/states, the $\beta_{i,j}$ term can get arbitrarily close to 1 with large enough T. However, in a large dimensional world, this is not the case. The betas no longer work to price each other, and we have many trading strategies that cannot price each other, despite attempting to form mean-variance optimal portfolios. In short, factor loadings do not capture the return in a large dimensional world.

1.5 Sparse Beliefs

Now consider investors with *sparse beliefs*, which is defined as investors who have the prior:

$$V_0 = SVS + D. \tag{72}$$

S is the identity matrix with some diagonal elements set to zero, and D is a diagonal matrix. The intuition is straightforward: this investor believes only a subset of factors matter, while the rest are ignored (corresponding to the elements of S that are set to zero). Under the investors' priors, the matrix D captures the prior about the remaining risk, which is assumed to be idiosyncratic—thus D is assumed to be diagonal. Assume that prior beliefs about aggregate market volatility are the same, i.e., assume:

$$\iota' B_T (SVS + D) B_T' \iota = \iota' B_T V B_T' \iota \tag{73}$$

Consider an atomistic investor who knows this is the aggregate prior of the market. Furthermore, assume this investor knows which elements of $B_T(SVS + D)B_T'\iota$ are greater than $B_T V B_T'\iota$. In other words, this atomistic investor knows which assets are assessed by the sparse-belief market as having too high risk (relative to the right-prior market with $V_0 = V$), prices that are too low, and returns that are ultimately too high. Let ψ_T be an n dimensional column vector of zeros, but has ones in the corresponding places where $B_T(SVS + D)B_T'\iota$ are greater than $B_T V B_T'\iota$. A naive but simple investment strategy would be to invest with the following portfolio weights:

$$x_T = l(\iota' p_T - \psi_T' p_T)\psi_T - l(\iota - \psi_T)$$

where l is a scalar that controls the leverage. The first term on the right-hand side is the long position in assets with returns that are too high because of the sparse beliefs, while the second term

represents a short position in assets with returns that are too low because of the sparse beliefs. The long and short side of this long-short portfolio are value-weighted.

What is the return on this portfolio? This is

$$x_T' \mathbb{E}_T[r_{T+1}] = x_T'(\mathbb{E}_T[d_{T+1}] - (1 + r_f)p_T) \tag{74}$$

$$= \gamma_n x_T' \hat{\Lambda}_T \iota \tag{75}$$

where the second equation comes from plugging in equation (**??**).

The alpha of this strategy is the scalar:

$$\alpha_{x,T} = \gamma x_T' \hat{\Lambda}_T \iota - \frac{x_T' \Lambda_T \iota}{\iota' \Lambda_T \iota} (\gamma_n \iota' \hat{\Lambda}_T \iota) \tag{76}$$

Because of the assumption that beliefs about market volatility is identical, i.e. $\iota' B_T (SVS + D) B_T' \iota = \iota' B_T V B_T' \iota$, this means that $\iota' \Lambda_T \iota = \iota' \hat{\Lambda}_T \iota$. Then we have:

$$\alpha_{x,T} = \gamma x_T' \hat{\Lambda}_T \iota - \gamma x_T' \Lambda_T \iota \tag{77}$$

$$= \gamma_n \left(l(\iota' p_T - \Psi_T' p_T) \underbrace{\Psi_T'(\hat{\Lambda}_T - \Psi_T' \Lambda_T)}_{>0} - l \underbrace{(\iota - \Psi_T)'(\hat{\Lambda}_T - \Psi_T' \Lambda_T)}_{<0} \right) > 0 \tag{78}$$

Thus, this investment strategy has a positive alpha.

2 Empirical Study

Inspired by the predictions of our theoretical model, we look for those patterns in the data. We are after 1) evidence of the high-dimensionality of stock returns, 2) pricing errors increasing as the

number of states/assets increase, 3) a lower relationship between betas and expected returns as the number of assets/states increase, 4) high Sharpe ratio strategies are less likely to span each other as the number of assets increases.

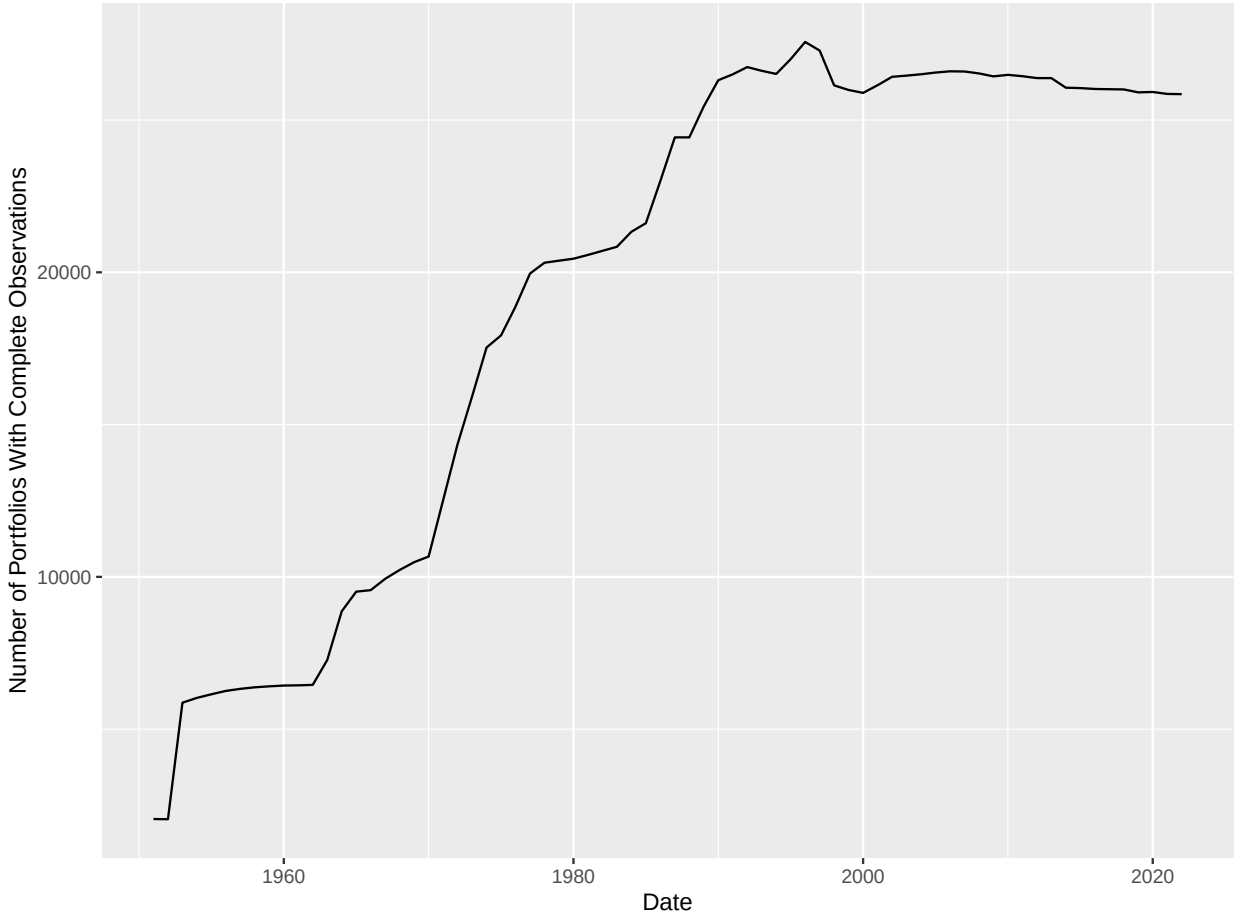
2.1 Data

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



2.2 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

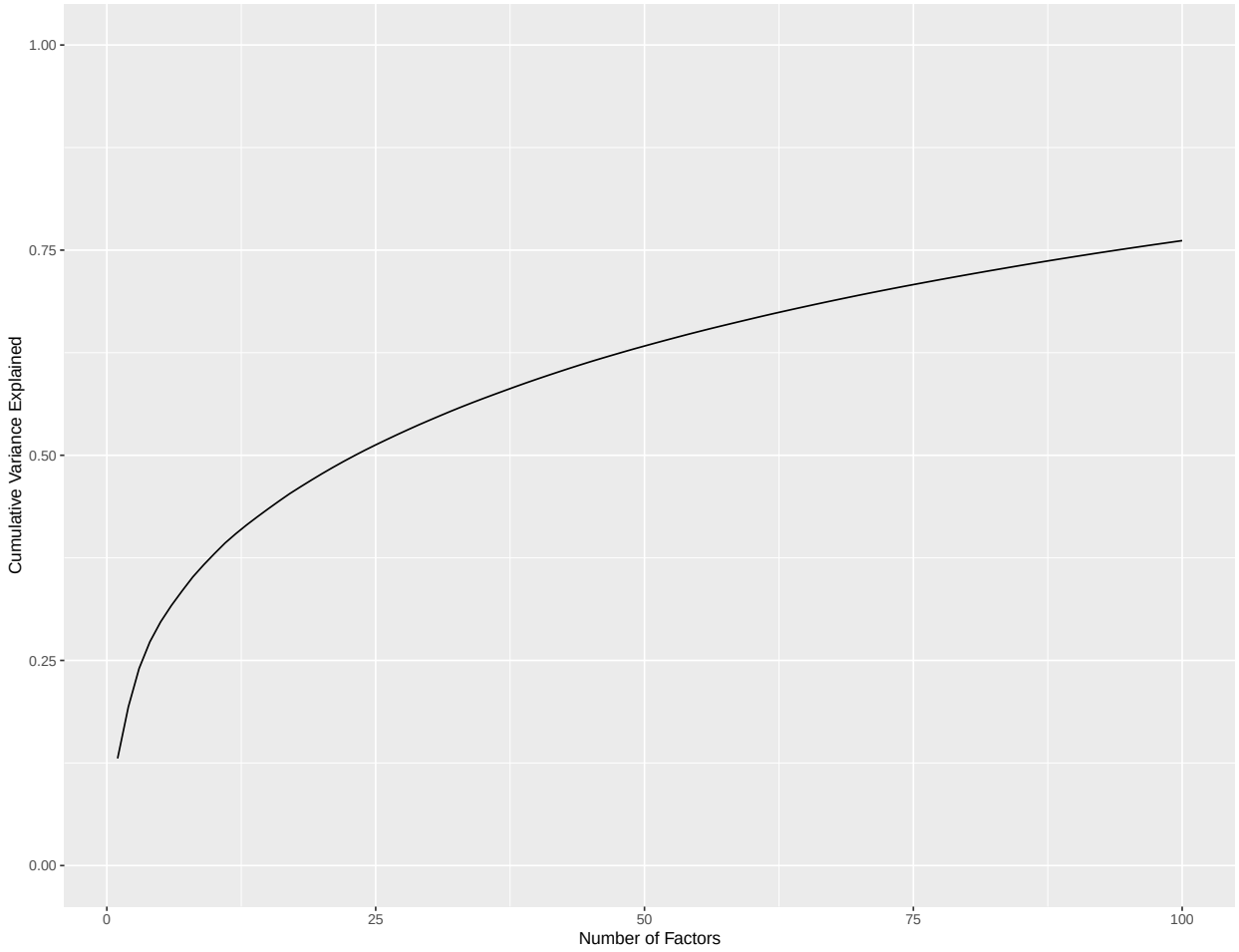
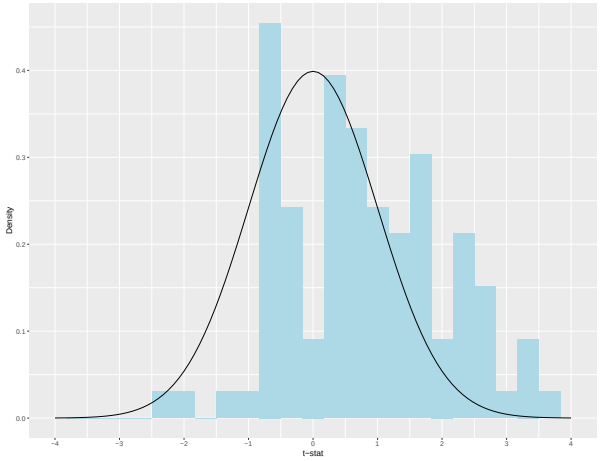
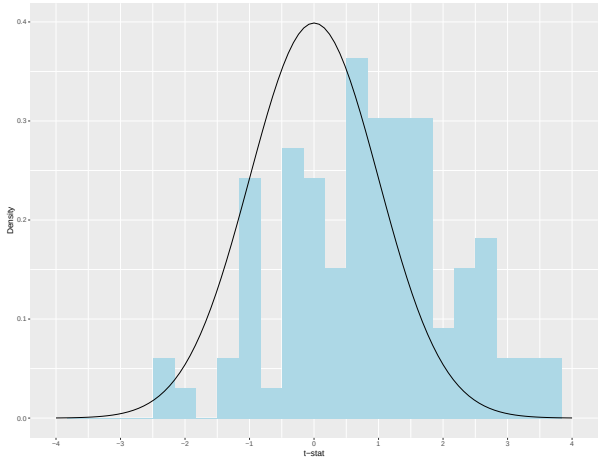


Figure 4: Comparative distributions of Out-of-Sample t-statistics



(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

2.3 Parsimonious Asset Pricing Models

Next, we show that consistent with our model, parsimonious factor models have a more challenging time pricing assets in high-dimension. We show that the average squared pricing errors increase as the number of states increases. Further, betas have less explanatory power as we increase the number of assets/states.

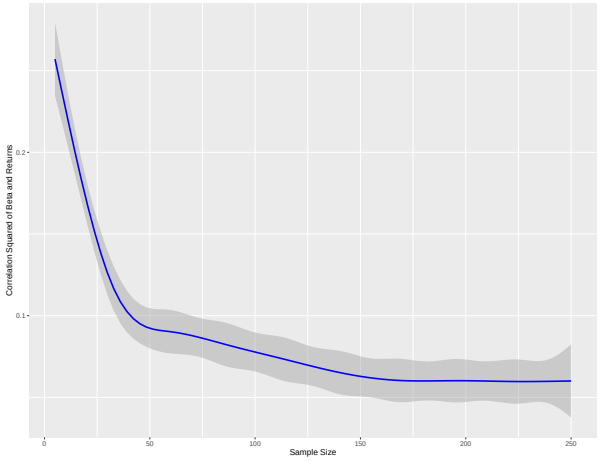
We show this pattern in two distinct ways. First, we consider using K distinct portfolios from the universe of n portfolios. Second, we form the mean-variance combination of these k-assets in the in-sample period. Alternatively, we consider an equal-weighted portfolio of these K assets, mimicking the 'market' portfolio. Third, we run out-of-sample time-series regressions of each K asset on the mean-variance or equal-weighted portfolios. Fourth, we collect the betas and alphas. We then run the procedure 2000 times for each number K and average the results.

Figure 7a shows the average R2 from running cross-sectional regressions of betas and average returns. Betas are calculated using time-series regressions using the entire sample, and average

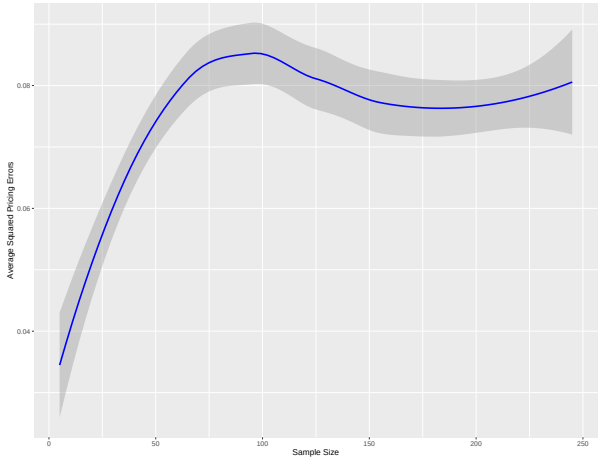
returns are calculated using the time-series average of the returns for each sample.[6] We see a monotonically decreasing pattern as the number of portfolios to select from increases.

Figure 7b explores the out-of-sample relationship between (squared) pricing errors and the out-of-sample returns. It shows the average out-of-sample squared pricing errors of running time-series regressions against the (in-sample) mean-variance portfolio. We calculate the time-series regressions and average returns using the same out-of-sample period. We see a monotonically increasing pattern as the number of portfolios to select from increases.

Figure 5: Feasible Mean-Variance Portfolio Pricing Errors as the Number of Assets Increase



(a) R2 between Betas and OOS Returns

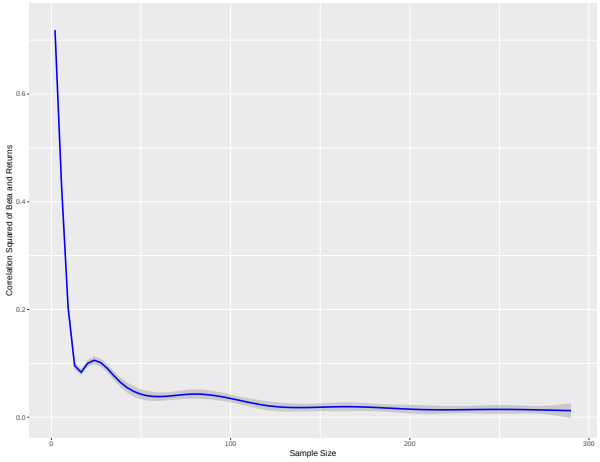


(b) Average Out-of-Sample Squared Pricing Errors Against In-Sample Mean-Variance Portfolio

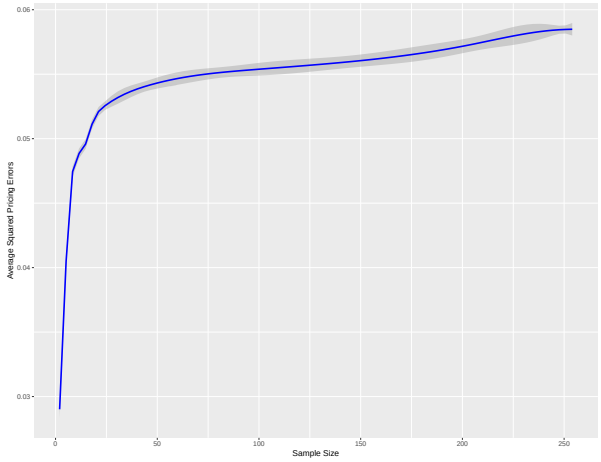
Naturally, it is straightforward to reject CAPM. For example, choosing to test CAPM using a GRS test with the best 100 (in-sample) portfolios during the out-of-sample period completely rejects the model.

6. We use the same sample for average returns and betas because otherwise, the betas may change during the out-of-sample period.

Figure 6: CAPM as the Number of Assets Increase



(a) R2 between Betas and OOS Returns



(b) Average Out-of-Sample Squared Pricing Errors Against In-Sample Mean-Variance Portfolio

2.4 Sharpe ratios that do not span each other

We then show that as the number of assets increases, the best strategies have a harder time spanning each other. As before, we consider using K distinct portfolios from the universe of n portfolios and form the mean-variance combination of these k-assets in the in-sample period and repeat the procedure 2000 times for each K. We then select the 5 portfolios with the highest (in-sample) Sharpe ratio for each K. We use in-sample to mimic the usual way of testing between potential pricing kernels. Only the portfolios with the highest Sharpe ratios are interesting. We then run spanning tests of these 5 portfolios against each other. We then collect the GRS statistics (Gibbons, Ross, and Shanken (1989). Figure 9 depicts the relationship between the average GRS statistic for each k and the number of assets. There is a monotonic pattern. Hence, as the number of potential assets increase, each feasible mean-variance portfolio is further from the ideal, and it becomes easier to find abnormal returns against them.

2.5 Low Correlations Between Strategies (No Spanning)

An important implication from the model is that in a high-dimensional world, portfolios with high Sharpe ratios are likely to have correlations below 1 with different priors. 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 8 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

$$w_{i|M} = w_i - \beta_i w_M. \tag{79}$$

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

$$\text{corr}_{i,j} = \frac{(w_{i|M})'\Sigma(w_{j|M})}{\sqrt{(w_{i|M})'\Sigma(w_{i|M})}\sqrt{(w_{j|M})'\Sigma(w_{j|M})}} \tag{80}$$

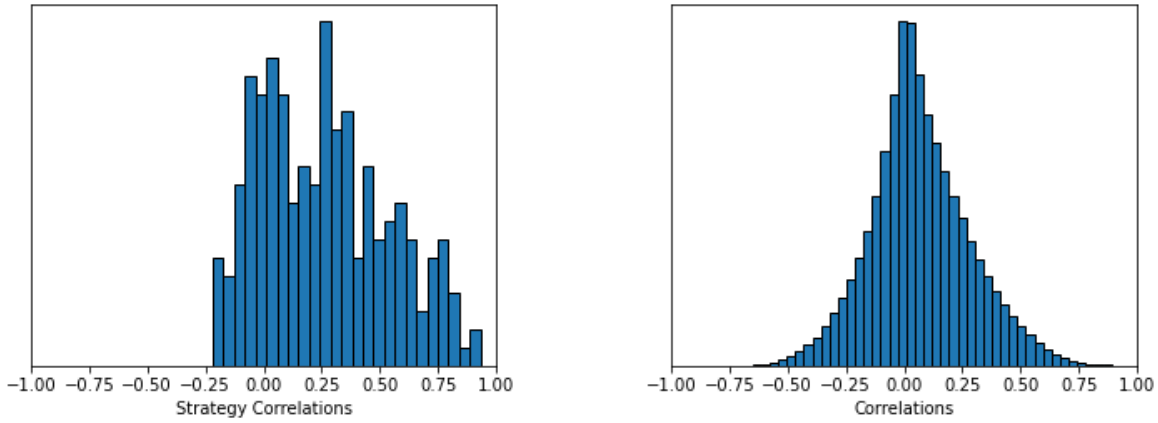
Panel (b) of Figure 8 shows the pair-wise correlations of these strategies. The average pair-wise correlation is about 0.07. This is in many ways similar to Koijen and Yogo 2019, who show that much of demand is not explained by classic characteristics. We go one step further by showing that much of the 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 7: Correlations of Portfolios



(a) Trading Strategy Correlations

(b) Systematic Hedge Fund Cross Correlations

Figure 8: 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

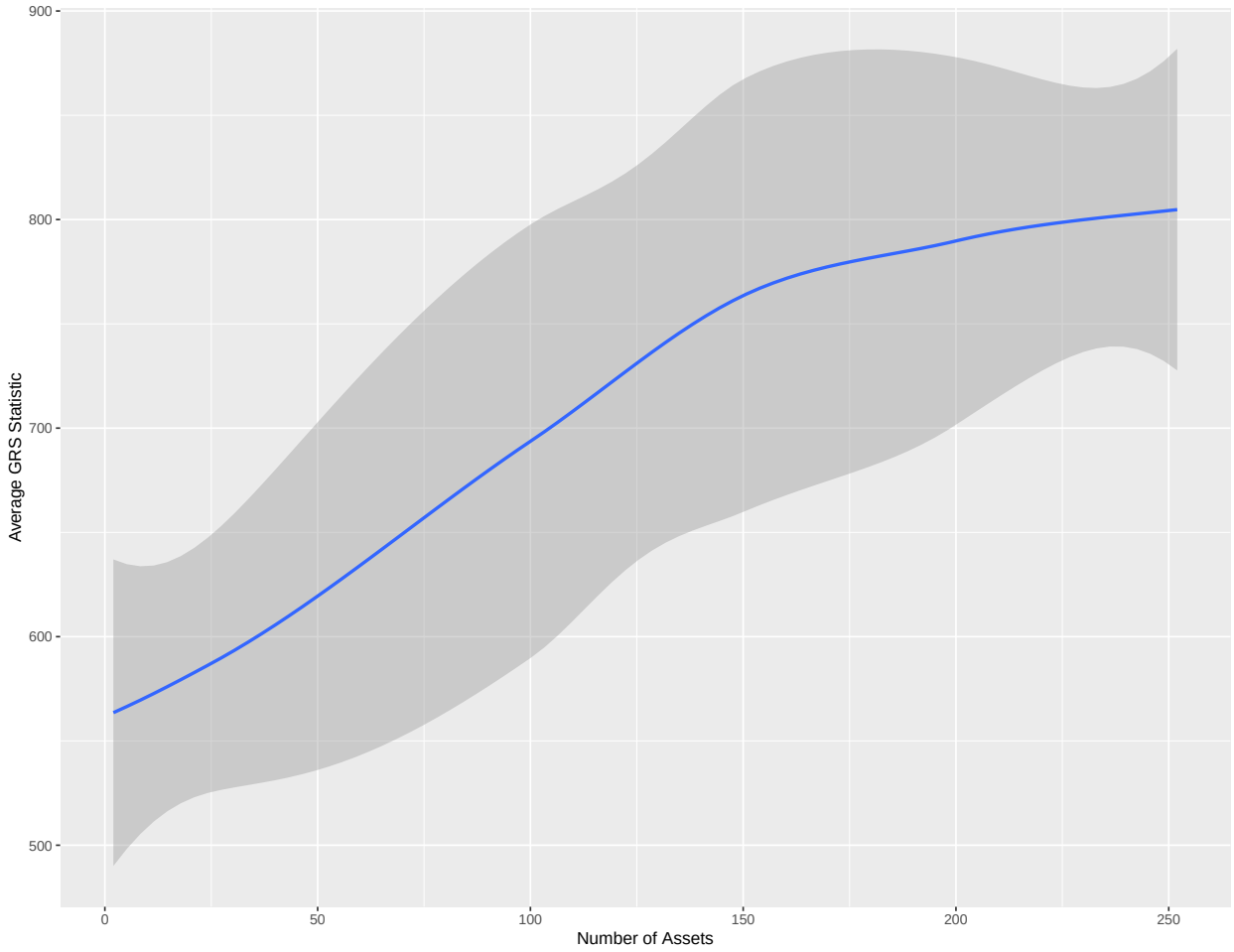
Panel B: Weber Cross Regressions Beta Part 2

	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.

Figure 9: Average GRS Statistic Between the Best Portfolios and Number of Assets



3 Conclusion

Our model provides a comprehensive explanation for several long-standing asset pricing puzzles. We demonstrate that the limitations of feasible optimization in high-dimensional spaces create significant challenges in asset pricing, leading to systematic pricing errors, flatter risk premiums, and the proliferation of return-predicting anomalies. The traditional models, including CAPM and multi-factor models, struggle in high-dimensional settings due to their inability to accurately capture the complex dynamics of asset returns and covariances. This inadequacy becomes more pronounced with the increase in the number of assets and state variables. Our empirical analysis validates these theoretical predictions, showing a clear trend of increased pricing errors and the emergence of numerous high Sharpe ratio strategies that existing models do not fully explain.

Moreover, our findings highlight the importance of considering the high dimensionality of financial markets in asset pricing theory. As the number of relevant characteristics and state variables grows, the need for more sophisticated, complex models becomes evident. These models must account for the growing distance to the optimal portfolio under the true data-generating process mean-variance frontier and the increased difficulty in accurately estimating covariances and expected returns. In doing so, they can better explain the observed phenomena in asset markets, such as the flatness of the security market line and the existence of return-predicting anomalies that elude simpler models.

Our research underscores that parsimonious models cannot fully capture the financial market's complexity. It calls for a paradigm shift in asset pricing towards embracing more intricate, high-dimensional models that can more effectively navigate modern financial markets' vast and nuanced landscape. This approach promises to enhance our understanding of asset pricing dynamics and improve investment strategies, ultimately leading to more efficient and robust financial markets.

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 Ke, Bryan T. Kelly, and Semyon Malamud. 2023. "Complexity in Factor Pricing Models." *SSRN Electronic Journal* (March). https://doi.org/10.2139/SSRN.4388526.

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. 2021. "The Virtue of Complexity in Return Prediction." *SSRN Electronic Journal* (December). https://doi.org/10.2139/SSRN.3984925.

———. 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, Stefan Nagel, Mike Chernov, Kent Daniel, Gene Fama, Stefano Giglio, Lars Hansen, Ian Martin, Andreas Neuhierl, and Markus Pelger. 2023. "When Do Cross-Sectional Asset Pricing Factors Span the Stochastic Discount Factor?" (Cambridge, MA) (May). https://doi.org/10.3386/W31275.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh. 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.

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 CAPM Fails

Proof. Recall that the CAPM alpha is:

$$\alpha_t = \left(I - \frac{\Lambda\iota\iota'}{\iota'\Lambda\iota}\right)(\mathbb{E}_t[d_{t+1}] - (1 + r_f)p_t) \tag{81}$$

$$= \gamma\left(I - \frac{\Lambda\iota\iota'}{\iota'\Lambda\iota}\right)(\hat{\Lambda}\iota) \tag{82}$$

So we need to plug in the price above and solve for the distribution (or at least expected value) of

$$\alpha_t'\Lambda^{-1}\alpha_t \tag{83}$$

So plugging in our price yields:

$$\alpha_t = \left(I - \frac{\Lambda\iota\iota'}{\iota'\Lambda\iota}\right)(\gamma\hat{\Lambda}_T\iota) \tag{84}$$

So we have:

$$\alpha_t'\Lambda^{-1}\alpha_t = \gamma^2\iota'\hat{\Lambda}_T\left(I - \frac{\Lambda\iota\iota'}{\iota'\Lambda\iota}\right)'\Lambda^{-1}\left(I - \frac{\Lambda\iota\iota'}{\iota'\Lambda\iota}\right)\hat{\Lambda}_T\iota \tag{85}$$

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

$$= \gamma^2\left(\iota'\hat{\Lambda}_T\Lambda^{-1}\hat{\Lambda}_T\iota - \frac{(\iota'\hat{\Lambda}_T\iota)^2}{\iota'\Lambda\iota}\right). \tag{87}$$

Simplifying by parts:

$$\mathbb{E}[\iota'\hat{\Lambda}_T\Lambda^{-1}\hat{\Lambda}_T\iota] \tag{88}$$

$$= \left(\frac{1}{T+\nu-NS-1}\right)^2 \mathbb{E}\left[\iota' B_T\left(V_0 + \sum_{t=1}^{T}(X_t-\pi)(X_t-\pi)'\right)B_T'\Lambda^{-1}B_T\left(V_0 + \sum_{t=1}^{T}(X_t-\pi)(X_t-\pi)'\right)B_T'\iota\right] \tag{89}$$

$$= \left(\frac{1}{T+\nu-NS-1}\right)^2 \left(\iota' B_T V_0 B_T'\Lambda^{-1}B_T V_0 B_T'\iota\right. \tag{90}$$

$$+\iota' B_T V_0 B_T'\Lambda^{-1}B_T\mathbb{E}\left[\sum_{t=1}^{T}(X_t-\pi)(X_t-\pi)'\right]B_T'\iota \tag{91}$$

$$+\iota' B_T\mathbb{E}\left[\sum_{t=1}^{T}(X_t-\pi)(X_t-\pi)'\right]B_T'\Lambda^{-1}B_T V_0 B_T'\iota \tag{92}$$

$$\left.+\iota' B_T\mathbb{E}\left[\left(\sum_{t=1}^{T}(X_t-\pi)(X_t-\pi)'\right)B_T'\Lambda^{-1}B_T\left(\sum_{t=1}^{T}(X_t-\pi)(X_t-\pi)'\right)\right]B_T'\iota\right) \tag{93}$$

We can write

$$\mathbb{E}\left[\sum_{t=1}^{T}(X_t-\pi)(X_t-\pi)'\right] = TV \tag{94}$$

We use the iid nature of X_t):

$$\iota' B_T\mathbb{E}\left[\left(\sum_{t=1}^{T}(X_t-\pi)(X_t-\pi)'\right)B_T'\Lambda^{-1}B_T\left(\sum_{t=1}^{T}(X_t-\pi)(X_t-\pi)'\right)\right]B_T'\iota \tag{95}$$

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

$$+2\iota' B_T\mathbb{E}\left[\sum_{t=1}^{T}\sum_{s=1}^{t-1}\left((X_t-\pi)(X_t-\pi)'\right)B_T'\Lambda^{-1}B_T\left((X_s-\pi)(X_s-\pi)'\right)\right]B_T'\iota \tag{97}$$

$$= T\mathbb{E}[\iota' B_T(X_t - \pi)(X_t - \pi)' B_T' \Lambda^{-1} B_T(X_t - \pi)(X_t - \pi)' B_T' \iota] \tag{98}$$

$$+T(T-1)\iota' \Lambda \Lambda^{-1} \Lambda \iota \tag{99}$$

$$= T\mathbb{E}[(\iota' B_T(X_t - \pi))(\Lambda^{-1/2} B_T(X_t - \pi))'(\Lambda^{-1/2} B_T(X_t - \pi))(\iota' B_T(X_t - \pi))'] + T(T-1)\iota' \Lambda \iota \tag{100}$$

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

$$\mathbb{E}[(Ax)(Bx)'(Cx)(Dx)'] = (AVB')(CVD') + (AVC')(BVD') + \text{tr}(BVC')AVD' \tag{101}$$

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

Using this we have

$$T\mathbb{E}[(\iota' B_T(X_t - \pi))(\Lambda^{-1/2} B_T(X_t - \pi))'(\Lambda^{-1/2} B_T(X_t - \pi))(\iota' B_T(X_t - \pi))'] \tag{102}$$

$$= T[2(\iota' B_T V B_T' \Lambda^{-1/2})(\Lambda^{-1/2} B_T V B_T' \iota) + \text{tr}(\Lambda^{-1/2} B_T V B_T' \Lambda^{-1/2})(\iota' B_T V B_T' \iota)] \tag{103}$$

Well, we have

$$\Lambda = B_T V B_T' \tag{104}$$

so

$$\text{tr}(\Lambda^{-1/2} B_T V B_T' \Lambda^{-1/2}) = \text{tr}(\Lambda^{-1}\Lambda) = N \tag{105}$$

hence

$$T\mathbb{E}[(\iota' B_T(X_t - \pi))(\Lambda^{-1/2} B_T(X_t - \pi))'(\Lambda^{-1/2} B_T(X_t - \pi))(\iota' B_T(X_t - \pi))'] = T((2\iota' \Lambda \iota) + N(\iota' \Lambda \iota)) \tag{106}$$

$$= T(N+2)\iota'\Lambda\iota \tag{107}$$

So we have

$$\mathbb{E}[\iota'\hat{\Lambda}_T\Lambda^{-1}\hat{\Lambda}_T\iota] \tag{108}$$

$$= \left(\frac{1}{T+\nu-NS-1}\right)^2 \left(\iota' B_T V_0 B_T' \Lambda^{-1} B_T V_0 B_T' \iota + 2T\iota' B_T V_0 B_T' \iota + T(N+2)\iota'\Lambda\iota + T(T-1)\iota'\Lambda\iota\right) \tag{109}$$

Now onto the next part. We need to compute

$$\mathbb{E}[(\iota'\hat{\Lambda}_T\iota)^2] \tag{110}$$

$$= \mathbb{E}[\iota'\hat{\Lambda}_T\iota\iota'\hat{\Lambda}_T\iota] \tag{111}$$

$$= \left(\frac{1}{T+\nu-NS-1}\right)^2 \mathbb{E}\left[\iota' B_T \left(V_0 + \sum_{t=1}^{T}(X_t-\pi)(X_t-\pi)'\right) B_T'\iota\iota' B_T \left(V_0 + \sum_{t=1}^{T}(X_t-\pi)(X_t-\pi)'\right) B_T'\iota\right] \tag{112}$$

$$= \left(\frac{1}{T+\nu-NS-1}\right)^2 \left(\iota' B_T V_0 B_T'\iota\iota' B_T V_0 B_T'\iota\right. \tag{113}$$

$$+\iota' B_T V_0 B_T'\iota\iota' B_T \mathbb{E}\left[\sum_{t=1}^{T}(X_t-\pi)(X_t-\pi)'\right] B_T'\iota \tag{114}$$

$$+\iota' B_T \mathbb{E}\left[\sum_{t=1}^{T}(X_t-\pi)(X_t-\pi)'\right] B_T'\iota\iota' B_T V_0 B_T'\iota \tag{115}$$

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

$$= \left(\frac{1}{T+\nu-NS-1}\right)^2 \left(\iota' B_T V_0 B_T'\iota\iota' B_T V_0 B_T'\iota\right. \tag{117}$$

$$+2T\iota' B_T V_0 B_T'\iota\iota'\Lambda\iota \tag{118}$$

$$+\iota' B_T \mathbb{E}\left[\left(\sum_{t=1}^{T}(X_t-\pi)(X_t-\pi)'\right) B_T' \iota\iota' B_T \left(\sum_{t=1}^{T}(X_t-\pi)(X_t-\pi)'\right)\right] B_T'\iota\Bigg) \tag{119}$$

Simplifying:

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

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

$$+2\iota' B_T \mathbb{E}\left[\sum_{t=1}^{T}\sum_{s=1}^{t-1}\left((X_t-\pi)(X_t-\pi)'\right) B_T' \iota\iota' B_T \left((X_s-\pi)(X_s-\pi)'\right)\right] B_T'\iota \tag{122}$$

$$= T\mathbb{E}[\iota' B_T (X_t-\pi)(X_t-\pi)' B_T' \iota\iota' B_T (X_t-\pi)(X_t-\pi)' B_T'\iota] \tag{123}$$

$$+T(T-1)\iota'\Lambda\iota\iota'\Lambda\iota \tag{124}$$

$$= T\mathbb{E}[(\iota' B_T (X_t-\pi))(\iota' B_T (X_t-\pi))'(\iota' B_T (X_t-\pi))(\iota' B_T (X_t-\pi))'] + T(T-1)(\iota'\Lambda\iota)^2 \tag{125}$$

Recall the identity:

$$\mathbb{E}[(Ax)(Bx)'(Cx)(Dx)'] = (AVB')(CVD') + (AVC')(BVD') + \mathrm{tr}(BVC')AVD' \tag{126}$$

Using this, we can calculate

$$T\mathbb{E}[(\iota' B_T (X_t-\pi))(\iota' B_T (X_t-\pi))'(\iota' B_T (X_t-\pi))(\iota' B_T (X_t-\pi))'] \tag{127}$$

$$= 3T(\iota'\Lambda\iota)^2 \tag{128}$$

So we can write:

$$\mathbb{E}[(\iota'\hat{\Lambda}_T\iota)^2] \tag{129}$$

$$= \left(\frac{1}{T+\nu-NS-1}\right)^2 \left(\iota' B_T V_0 B_T' \iota \iota' B_T V_0 B_T' \iota + 2T\iota' B_T V_0 B_T' \iota \iota' \Lambda \iota + 3T(\iota' \Lambda \iota)^2 + T(T-1)(\iota' \Lambda \iota)^2\right) \tag{130}$$

So then we have:

$$\mathbb{E}[\alpha_t' \Lambda^{-1} \alpha_t] = \mathbb{E}\left[\gamma^2 \left(\iota' \hat{\Lambda}_T \Lambda^{-1} \hat{\Lambda}_T \iota - \frac{(\iota' \hat{\Lambda}_T \iota)^2}{\iota' \Lambda \iota}\right)\right] \tag{131}$$

$$= \gamma^2 \left(\frac{1}{T+\nu-NS-1}\right)^2 \left(\left(\iota' B_T V_0 B_T' \Lambda^{-1} B_T V_0 B_T' \iota + 2T\iota' B_T V_0 B_T' \iota + T(N+2)\iota' \Lambda \iota + T(T-1)\iota' \Lambda \iota\right)\right. \tag{132}$$

$$\left. -\frac{1}{\iota' \Lambda \iota}\left(\iota' B_T V_0 B_T' \iota \iota' B_T V_0 B_T' \iota + 2T\iota' B_T V_0 B_T' \iota \iota' \Lambda \iota + 3T(\iota' \Lambda \iota)^2 + T(T-1)(\iota' \Lambda \iota)^2\right)\right) \tag{133}$$

$$= \gamma^2 \left(\frac{1}{T+\nu-NS-1}\right)^2 \left(\iota' B_T V_0 B_T' \left(\Lambda^{-1} - \frac{\iota \iota'}{\iota' \Lambda \iota}\right) B_T V_0 B_T' \iota + T(N-1)\iota' \Lambda \iota\right) \tag{134}$$

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

$$\mathbb{E}[\alpha_t' \Lambda^{-1} \alpha_t] = \gamma^2 \left(\frac{1}{T+\nu-NS-1}\right)^2 T(N-1)(NS)^2 \bar{\varrho} \tag{135}$$

where ϱ is

$$\bar{\varrho} = \frac{1}{(NS)^2} \sum_{i=1}^{NS} \sum_{j=1}^{NS} b_{i,T} b_{j,T} V_{i,j} \tag{136}$$

and

$$b_T = B_T \iota \tag{137}$$

We need $\nu > NS + 1$. A weak prior would be $\nu = NS + 2$. Plugging in this prior yields

$$\mathbb{E}[\alpha_t' \Lambda^{-1} \alpha_t] = \gamma^2 \left(\frac{1}{T+1}\right)^2 T(N-1)(NS)^2 \bar{\varrho} \tag{138}$$

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

states.

□

Proof of Theorem 2.

Proof. Given the assumptions, it is easy to see that demand simplifies to

$$x_{i,T} = \frac{1}{\gamma}\frac{1}{1+\kappa+T}(\zeta_i + T\Lambda\bar{x}_T) - \frac{1+r_f}{\gamma}\frac{\kappa+T}{1+\kappa+T}\Lambda^{-1}p_T \tag{139}$$

We can plug this into equilibrium, in order to solve for prices:

$$\iota = \sum_{i=1}^{I}\theta_i\left(\frac{1}{\gamma}\frac{1}{1+\kappa+T}(\zeta_i + T\Lambda\bar{x}_T) - \frac{1+r_f}{\gamma}\frac{\kappa+T}{1+\kappa+T}\Lambda^{-1}p_T\right) \tag{140}$$

We solve for prices and get:

$$p_T = \frac{1}{1+r_f}\frac{1}{\kappa+T}(V\zeta + T\bar{x}_T) - \frac{\gamma}{1+r_f}\frac{1+\kappa_T}{\kappa+T}\Lambda\iota. \tag{141}$$

When we plug this back into demand $x_{i,T}$, we get:

$$x_{i,T} = \frac{1}{\gamma}\frac{1}{1+\kappa+T}(\zeta_i - \zeta) + \iota \tag{142}$$

Define:

$$\beta_{i,j,S,T} = \frac{x'_{i,T}\Lambda x_{j,T}}{x'_{j,T}\Lambda x_{j,T}}. \tag{143}$$

We can plug in demand here:

$$\beta_{i,j,S,T} = \frac{\gamma^{-2}(1+\kappa+T)^{-2}S^2\omega_S + \gamma^{-1}(1+\kappa+T)^{-1}S\tau_S + x'_{j,T}\Lambda x_{j,T}}{x'_{j,T}\Lambda x_{j,T}} \tag{144}$$

Then we have:

$$\beta_{i,j} = \lim_{\substack{T,S\to\infty \\ S/T=c}} \beta_{i,j,S,T} = \frac{\gamma^{-2}c\omega + \gamma^{-1}c\tau + \sigma^2}{\sigma^2}. \tag{145}$$

□