The Empirical Virtue of Complexity in Simple Economic Models

Kerry Back* Alexander Ober† Seth Pruitt‡

September 1, 2025

We show that the empirical virtue of complexity arises in simple economic models. We show that machine-learning estimates, in data simulated from well-known economic frameworks, are superior to those from simpler empirical models. Machine-learning empirical models more accurately estimate the true SDF than sort-based or regression-based methods. Moreover, machine-learning methods more accurately estimate firms' SDF weights, thus providing a superior depiction of the economically important firm heterogeneity that the economic models were designed to explain.

PRELIMINARY AND INCOMPLETE

*Jones Graduate School of Business and School of Social Sciences, Rice University
†Jones Graduate School of Business, Rice University
‡W.P. Carey School of Business, Arizona State University

1. Introduction

Empirical asset pricing seeks to understand the stochastic discount factor (SDF) driving asset prices. A seminal strand of literature flows from the CAPM of Sharpe (1964) and Lintner (1965) through the celebrated Fama and French (2015) multi-factor model, and represents the SDF by the excess returns on a collection of portfolios sorted using a handful empirically-relevant asset characteristics. Meanwhile, an alternative strand flows from Fama and MacBeth (1973) and Rosenberg (1974) through the recent Fama and French (2019), and represents the SDF by cross-sectional regression estimates of excess returns on those characteristics. An important difference between the portfolio-sort and regression-based approaches is that the latter allows a large number of characteristics to be used jointly. Kelly, Pruitt, and Su (2019) combines this insight with machine-learning-inspired dimension reduction to arrive at efficient estimates of a low-dimensional SDF. Recently, Didisheim et al. (2024) argue that the assumption of low dimensionality significantly restricts the SDF estimator—building from the virtue of complexity shown in Kelly, Malamud, and Zhou (2024)—and develop estimates that use both a large number of asset characteristics and an immense number of random transformations of those characteristics. The machine-learning models of Kelly, Pruitt, and Su (2019) and Didisheim et al. (2024) have been shown to significantly outperform simpler portfolio-sort models like Fama and French (2015) in out-of-sample analysis of actual data.

Economists imagine actual data as coming from an unknown data generating process. No economic model claims to completely represent the actual data generating process—the latter is viewed as far too complicated to comprehend and hence "all models are wrong" (Box, 1976). Economic models are tremendous simplifications of reality. As technically involved as their solutions may be, the agents' choices follow from clear economic incentives via well-defined mathematical functions so they "maintain the clear quantitative-parable feature of good economic analysis" (Cochrane, 2011). Relative to the real world, economic models are simple.

We show that the empirical virtue of complexity arises in simple economic models, too. That is, in data simulated from different well-known economic models, we find that

machine-learning estimators (e.g. Kelly et al., 2019; Didisheim et al., 2024) outperform simpler estimators (e.g. Fama and French, 2015, 2019) at revealing the true SDF. Hence, complex empirical models are important tools for understanding even data coming from simple economic models. Furthermore, we show that the estimates can be used to learn about the economics driving different firms' weights in the SDF.[1] These results show that machine-learning estimators can uncover economic structure.

We view the complexity-based methods of Kelly et al. (2024) and Didisheim et al. (2024) as the vanguard of an agenda bringing into financial economics a host of machine-learning techniques that have shown resounding efficacy in non-financial fields. It is common for new lines of scientific inquiry to face skepticism, and this nascent agenda has. As an example, Berk (2023)'s Comment argues that return prediction alone, particularly of the Kelly et al. (2024) type, is unlikely to deliver interpretable economic insights. Our work contributes to building a bridge from successful machine-learning estimates to interpretable economic insights. Complex yet empirically successful models are likely to proliferate in future literature, driven by the surprising existence of "double descent": the finding that complex statistical models provide optimal generalization (out-of-sample performance) for a variety of data-generating processes (Hastie et al., 2022). It will be fruitful for economic theory to learn from insights uncovered by overparameterized models, whose empirical success need not be due to data-snooping or overfitting, and that is what we aim to contribute.

Empirical asset pricing uses heterogeneous asset characteristics as crucial information, and therefore, we require our economic models to produce such characteristics. We use seminal models that explicitly generate heterogeneous firms: Berk, Green, and Naik (1999) and Kogan and Papanikolaou (2014). We use the calibrations of the original papers and simulate synthetic panels of stock returns and characteristics. We then demonstrate that complex models are useful for understanding the linkages between characteristics and returns even in these relatively simple models.

Specifically, we focus on how well the empirical models uncover the true SDF and firms'

[1] We mean the tradable SDF that Hansen and Richard (1987) shows exists. For pricing purposes, this is equivalent to the nontradable SDF that the economic models assume.

(conditional) weights thereon. We find that the machine-learning models of Kelly et al. (2019) and Didisheim et al. (2024) provide a superior depiction of the SDF, measured by both its realized Sharpe Ratio and its Hansen and Jagannathan (1997) Distance, relative to simpler empirical models building on Fama and French (2015) or Fama and French (2019). Then we turn to the related question of how accurately the estimates reveal firms' exposure to the SDF. First, we find that machine-learning methods capture the true, persistent nature of these SDF weights, whereas simpler empirical models imply much less persistence. Second, the machine-learning estimates exhibit a higher cross-sectional correspondence to the true weights than simpler models' estimates. Altogether, these results suggest that machine-learning estimates, though complex, are important tools for learning about economic structures.

Related literature Machine-learning methods in finance have exploded in usage. A partial list of empirical contributions of the past several years include Gu et al. (2020); Kozak et al. (2020); Lettau and Pelger (2020); Freyberger et al. (2020); Bianchi et al. (2021); Bali et al. (2023). Some of this superior empirical performance has been questioned, for instance in Li et al. (2025) which argue that backward-looking signal curation is a significant factor. Meanwhile, machine-learning methods, particularly neural networks, have recently been employed in macroeconomics and finance to solve dynamic models that have not been computationally feasible before (e.g. Barnett et al., 2023; Duarte et al., 2024). Though we solve dynamic models, unlike those papers we do not use machine-learning methods for their solution (it is unnecessary), and instead use the methods to analyze data simulated from the solved models.

Recently, Kelly et al. (2024) explicitly argued for the benefits of complexity in return prediction, defining complexity by a number of parameters that surpasses the number of observations (further elaboration on this idea is found in Kelly and Malamud, 2025), and has generated responding work. Nagel (2025) argues that particular results in Kelly et al. (2024) reflect volatility-managed momentum more than machine-learning performance, and Kelly and Malamud (2025) disagree with that result. Berk (2023)'s Comment also critiques the Kelly et al. (2024) results and questions the ability of complexity to connect

to economic theory—the present paper, in some ways, seeks to address those criticisms.

By studying equilibrium models with a small number of economically meaningful risk factors, theoretical work tries to understand how firms' dynamic decision-making is reconciled with the empirically observed cross-section of returns (Berk et al. (1999); Gomes et al. (2003); Carlson et al. (2004); Zhang (2005); Eisfeldt and Papanikolaou (2013); Kogan and Papanikolaou (2014); Kuehn and Schmid (2014); Ai et al. (2018); Gomes and Schmid (2021)). We show that complex empirical models help to explain the pricing of risk factors in equilibrium models. This result bridges theoretical work, which hopes to explain returns using parsimonious, economically-motivated, factors, and empirical work, which uses statistically-motivated, often highly parameterized, methods to understand patterns in data.

Road map The paper is organized as follows. We concisely discuss the economic framework of Berk, Green, and Naik (1999) and Kogan and Papanikolaou (2014) in Section 2 and how the (tradable) SDF is similarly derived in any model. The empirical models we examine are detailed in Section 3. Section 4 details our results focusing on how well empirical models uncover the true SDF and firm weights thereon. We then conclude. Further model details are given in the appendices.

2. Economic Models

Here we briefly outline salient features of the economic models. We provide more detailed descriptions of the models and their solutions in the appendices. We simulate the models at a monthly frequency to align with the standard frequency of data used in empirical asset-pricing studies.

2.1. Berk, Green, and Naik (1999)

In the model of Berk, Green, and Naik (1999) (henceforth BGN), firms invest optimally given an exogenous pricing kernel and random investment opportunities. The SDF at

date t for pricing cash flows at $t+1$ is

$$m_{t+1} := \mathrm{e}^{-r_t - \frac{1}{2}\sigma_m^2 + \sigma_m \varepsilon_{t+1}} \,. \tag{2.1}$$

The interest rate process is a Vasicek (1979) process:

$$r_{t+1} = r_t + \kappa(\theta - r_t) + \sigma_r \xi_{t+1} \,. \tag{2.2}$$

Here, ε and ξ are sequences of iid standard normals that are contemporaneously correlated with each other.

There are a fixed number of firms. Each firm begins at date 0 with zero capital. Each firm receives a take-it-or-leave-it investment opportunity each period. All projects require the same amount of capital and are fully equity financed. A project that is taken generates operating cash flows each period until it randomly dies. Free cash flow is paid out to shareholders.

The operating cash flow of each project has a time-invariant beta with respect to the SDF process shocks ε and a time-invariant idiosyncratic risk. The betas and idiosyncratic risks are drawn randomly for each firm and date from fixed distributions. A project's net present value (NPV) depends on its beta and on the level of interest rates. Firms accept all positive NPV projects. Because the project arrival processes are the same across firms, all firms have the same value of growth options at any point in time. The value of growth options varies over time, because of variation in the interest rate. The value of assets in place varies across firms at each point in time due to differences in past project quality. The value of assets in place also depends on the interest rate.

The model generates the following data for each firm each period:

- book value of equity = book value of assets
- market value of equity
- net income = operating cash flow
- stock return

From these, we calculate size, book-to-market, ROE, asset growth, and momentum ($t-12$ through $t-2$ returns). BGN show that size, book-to-market, and momentum are correlated with subsequent returns. We calibrate the model following BGN, and simulate 10 panels of data with a period length of one month, as do BGN. We simulate 920-month panels, and we use the first 200 months as burn-in.

2.2. Kogan and Papanikolaou (2014)

In Kogan and Papanikolaou (2014) (hereafter KP), firms likewise invest optimally given an exogenous pricing kernel. The model is set in continuous time. It has two aggregate state variables: a disembodied productivity process x_t that affects the productivity of all capital, and a process z_t that affects only the productivity (equivalently, the cost per productive unit) of newly installed capital. These variables are independent geometric Brownian motions:

$$\begin{aligned}\frac{\mathrm{d}x_t}{x_t} &= \mu_x\,\mathrm{d}t + \sigma_x\,\mathrm{d}B_{xt}\\ \frac{\mathrm{d}z_t}{z_t} &= \mu_z\,\mathrm{d}t + \sigma_z\,\mathrm{d}B_{zt}\,.\end{aligned}$$

There are also firm-specific and project-specific productivity processes that are independent square-root processes. The SDF process is given by

$$\frac{\mathrm{d}M_t}{M_t} = -r\,\mathrm{d}t - \gamma_x\,\mathrm{d}B_{xt} - \gamma_z\,\mathrm{d}B_{zt}\,.$$

The risk-free rate r is constant.

Each firm begins at date 0 with zero capital. Firms acquire potential projects stochastically at a firm-specific rate that switches between high and low states. When presented with a new project at time t, a firm chooses how much capital to install and pays the investment cost, which depends on z. The project then produces cash flows until it randomly dies. Cash flows depend on the capital invested, the economy-wide productivity process x and the firm-specific and project-specific productivity processes.

We implement the continuous-time model in discrete time at a monthly frequency. We use simple Bernoulli approximations to the Poisson arrival and expiration of projects, and we use closed-form transition distributions to discretize the other continuous-time processes that describe the model. The model generates the same list of data as in BGN. We use a monthly version of KP's calibration.[2] We simulate 10 panels of 920-month data, as with BGN.

2.3. SDFs in the Economic Models

BGN and KP specify exogenous SDF processes and use them to compute prices and returns. There is no assumption of market completeness, and the SDFs are nontradable. As is well known, there is always a unique tradable SDF formed by projecting any SDF onto the linear span of the asset returns. In our empirics, we work exclusively with excess returns, as is common, so we do not attempt to price the risk-free asset. We call any random variable that is orthogonal to excess returns an SDF.

We choose a particular SDF following Hansen and Richard (1987) and Britten-Jones (1999). Conditional on information at date t, we project the constant 1 onto the span of the excess stock returns to obtain an excess return $\gamma_t' R_{t+1}$, where R_{t+1} denotes the vector of excess stock returns from t to $t+1$ and γ_t is a vector that depends only on date t information. Hansen and Richard (1987) call $\gamma_t' R_{t+1}$ the random variable that "represents the expectation operator on the space of excess returns" (it is the Riesz representation). The residual $1 - \gamma_t' R_{t+1}$ is orthogonal to excess returns, and we call it the conditional tradable SDF. We compute γ_t and the conditional tradable SDF in each of the BGN and KP models.

The BGN and KP models are dynamic, and the vector γ_t varies randomly over time, depending on date t information. The models are also non-Markovian in the sense that γ_t is history dependent rather than depending only on a set of date-t state variables. As emphasized before, the models are relatively simple, but the dynamic non-Markovian nature complicates an empiricist's estimation of the conditional tradable SDF from past

[2]KP simulate at the weekly frequency and form annual observations.

returns.

3. Empirical Models

In the factor models that we study, all factors are excess returns. We can mimic the construction of Section 2.3 for each factor model to compute β_t such that $1 - \beta_t' F_{t+1}$ is orthogonal to the factors, where F_{t+1} denotes the vector of factor (excess) returns from t to $t+1$. The excess return $\beta_t' F_{t+1}$ has the maximum Sharpe ratio among all portfolios of the factors.

In actual empirical work, we observe the factors, but, of course, we do not know β_t (nor γ_t). We can estimate β_t in-sample by computing the tangency portfolio of factors, but the real test of a model is how well it performs out of sample. This means that we need to estimate β_t using only information available at date t. We follow the usual technique of estimating over rolling samples (as described further below). This produces $\hat{\beta}_t' F_{t+1}$, which we call the estimated mean-variance efficient (MVE) portfolio return, and $1 - \hat{\beta}_t' F_{t+1}$, which we call the estimated factor SDF.

The goal of factor modeling is to minimize pricing errors. This is equivalent to minimizing the Hansen-Jagannathan distance (Hansen and Jagannathan, 1997) and to maximizing the Sharpe ratio (Barillas and Shanken, 2017). So, we evaluate factor models by evaluating how close $\hat{\beta}_t' F_{t+1}$ is to $\gamma_t' R_{t+1}$ either in the L^2 norm or in terms of their Sharpe ratios. Obviously, this comparison depends not just on the alternative sets of factors but also on the method used to estimate the vector β_t for each set of factors.

To estimate β_t, we follow Didisheim et al. (2024) by running the Britten-Jones regression over rolling 360-month samples with the regression being penalized (ridge) regression when the number of factors is large. We also tried ridge regression for the small factor models FFC, FMR, and KPS described below, but it underperformed OLS, and we do not report the results. Next, we describe how the various sets of factors are constructed. We present the details of the ridge regression in Section 3.4.

3.1. FFC and FMR

We replicate the Fama and French (2015) construction of SMB, HML, RMW, and CMA and include UMD as well as the value-weighted market excess return to form the six-factor Fama-French-Carhart (FFC) model.

We also run Fama and MacBeth (1973) regressions on book-to-market, momentum, profitability, and asset growth. We standardize the portfolios implicit in the Fama-MacBeth regressions (Rosenberg, 1974; Fama, 1976) to be 100% long and 100% short and use the portfolio returns in conjunction with the equally weighted market excess return to form what we call the Fama-MacBeth-Rosenberg (FMR) model.[3]

3.2. KPS

We use the same five characteristics to implement the Kelly, Pruitt, and Su (2019) (henceforth KPS) method. Let N denote the number of stocks. The KPS begins by standardizing each characteristic in each cross-section, replacing the raw characteristic values with percentiles and then subtracting 0.5 to get ranks between -0.5 and $+0.5$. Let C_t denote the $N \times 5$ matrix of rank-standardized characteristics at date t.

Let L denote the number of KPS factors. We begin with the $N \times 5$ matrix C_t of rank-standardized characteristics defined above. Each column of this matrix can be interpreted as a characteristic factor portfolio. We augment this matrix with a column of constants. Call the resulting $N \times 6$ matrix Z_t. The returns on the columns of Z_t from t to $t+1$ are $f_{t+1} := Z_t' R_{t+1}$ where R_{t+1} denotes the N-vector of excess stock returns. We estimate IPCA in rolling 360-month windows. The IPCA step at date t uses the factor returns and characteristics over the past 360 months to compute a series of $L \times 6$ matrices $\{A_{t,s}\}_s$ which are functions of the characteristics at date s. The returns of the KPS factors estimated at date t and realized at date $s+1$ are defined to be $F_{t,s+1} = A_{t,s} f_{s+1}$. The six-factor KPS model does not employ dimension reduction and is the same as the FMR

[3] We use the equal weighted market excess return because the FMR regressions weight stocks equally. The equally weighted market excess return is the intercept in the FMR regression when the characteristics are de-meaned in each cross section. We could run weighted FMR regressions to achieve value weighting or something between equal and value weighting, but we do not explore that.

model run on rank-standardized characteristics.[4]

3.3. DKKM

We use the same five characteristics, along with the risk-free rate for the BGN model, to implement the Didisheim, Ke, Kelly, and Malamud (2024) (henceforth DKKM) method. DKKM use random Fourier features to create potentially a very large number of factors.[5] We follow their recipe to form various sets of what we call DKKM factors, ranging from a six factor model to a model with 3,600 factors.

Let M denote the number of DKKM factors. From C_t, we generate an $N \times M$ matrix of random characteristics as follows.[6] Let W denote an $(M/2) \times 5$ matrix whose entries are sampled from the standard normal distribution, and let γ denote a length $M/2$ vector whose entries are sampled uniformly from $\{0.5, 0.6, \ldots, 1\}$. We use the same W and γ for all t. We work with the matrix $C_t(\gamma \odot W)'$, where $\gamma \odot W$ denotes element by element multiplication of γ with each column of W. We compute an $N \times M$ matrix of random characteristics by taking sines and cosines of the elements of $C_t(\gamma \odot W)'$ and stacking the sines and cosines as separate columns. The DKKM factor realizations F_{t+1} from t to $t+1$ are the returns of the portfolios from t to $t+1$ whose weights on each firm's excess return are given by a column of the random characteristic matrix at date t.[7]

3.4. Ridge Regression

The ridge regression is

$$\min_{\beta} \quad \frac{1}{360} \sum_{i=0}^{359} \left(1 - \beta' F_{t-i}\right)^2 + \alpha \beta' \beta \,, \tag{3.1}$$

[4] In our implementation of the FMR model, we do not rank-standardize the characteristics. Instead, we scale the FMR portfolios to be 100% long and 100% short.

[5] Didisheim et al. (2024) discuss complex models in general, and, as they point out, this is only one particular way to generate such a model. See also Kelly and Malamud (2025).

[6] In BGN, we augment C_t with a column equal to the risk-free rate, standardized by subtracting its mean dividing by its standard deviation in the 200 months of burn-in in the simulation. We abuse notation by calling this augmented matrix C_t.

[7] Unlike DKKM, we do not rank-standardize the random characteristics.

where α is the penalty parameter. More penalization is needed when the number of factors is larger. To get a sense for how the penalty should vary with the number of factors, consider doubling the number of factors by simply replicating each factor. Then, to make $\beta'\beta$ small, we will want to split each beta evenly among the duplicate factors in each pair. This reduces the sum of squared betas by 1/2. Therefore, to maintain the same penalization, we should double α. Hence, we set $\alpha = \kappa M$, where M is the number of factors, and we vary κ. If adding more factors is more effective than simply replicating factors, then, for each value of κ, we should see performance improve as M increases. We also look at what DKKM call ridgeless regression, which can be interpreted as the limit of the ridge regression as $\alpha \to 0$ (it is OLS when the number of factors is not larger than the number of time periods).

4. Results

An important takeaway is that the performance of the DKKM model increases with the number of factors, provided the ridge penalty is sufficiently large, up to hundreds of factors, at which point it plateaus. There is no material difference in performance between hundreds of factors and thousands of factors in the simple economic models that we study, but there is a notable difference in performance if we drop below hundreds of factors. Furthermore, for both the BGN and KP models, and on several dimensions, the DKKM and KPS empirical models outperform FFC and FMR. We document these points below.

4.1. Sharpe Ratios and HJ Distances

We begin by analyzing the empirical models' depictions of the MVE portfolio and SDF in the BGN and KP economic models. We report the average, across time and simulations, of the true conditional (unannualized) Sharpe Ratio of the estimated MVE portfolios $\hat{\beta}_t' F_{t+1}$ and the true conditional Hansen and Jagannathan (1997) (HJ) Distances of the estimated SDFs $1 - \hat{\beta}_t' F_{t+1}$ from the true SDF.

Table 1 reports results for the FFC and FMR emprical models. FMR has both a higher mean Sharpe Ratio and a lower mean HJ Distance than FFC in the BGN model. However, the results are reversed in the KP model.

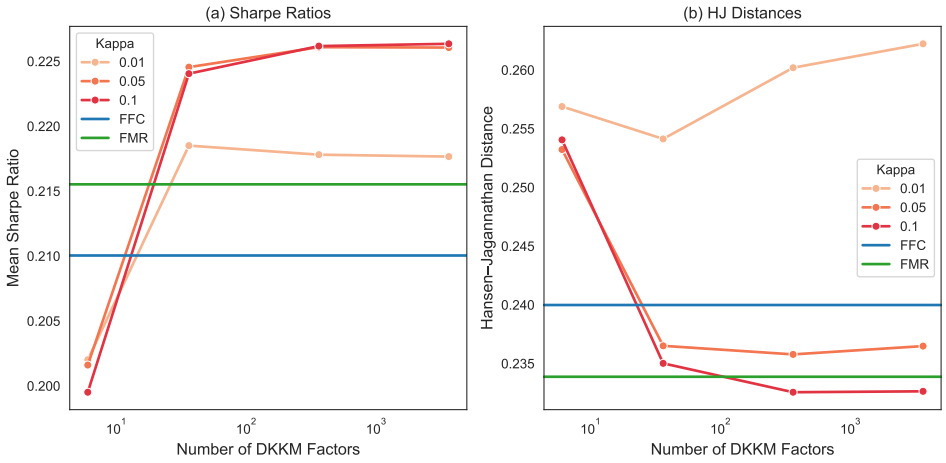
Table 2 reports results for DKKM relative to FMR. Outperformance by DKKM is represented by positive numbers for the Sharpe Ratio and negative numbers for the HJ Distance. We see that the performance of DKKM is increasing in the penalty parameter κ up to $\kappa = 0.1$ after which it declines. For $\kappa \geq 0.05$, performance is increasing in the number of factors up to 360, after which it plateaus. The best performance on either the Sharpe Ratio or the HJ Distance for either the BGN or KP model is achieved at $\kappa = 0.1$ and either 360 or 3,600 factors. These parameter combinations outperform the best performing of the FFC and FMR models on both metrics and in both economic models.

Figure 4.1 presents a visualization of how the performance of DKKM varies with the number of factors and with the penalization parameter κ. The figure shows that the DKKM model outperforms both FFC and FMR with $\kappa = 0.1$ and a sufficiently large number of factors.

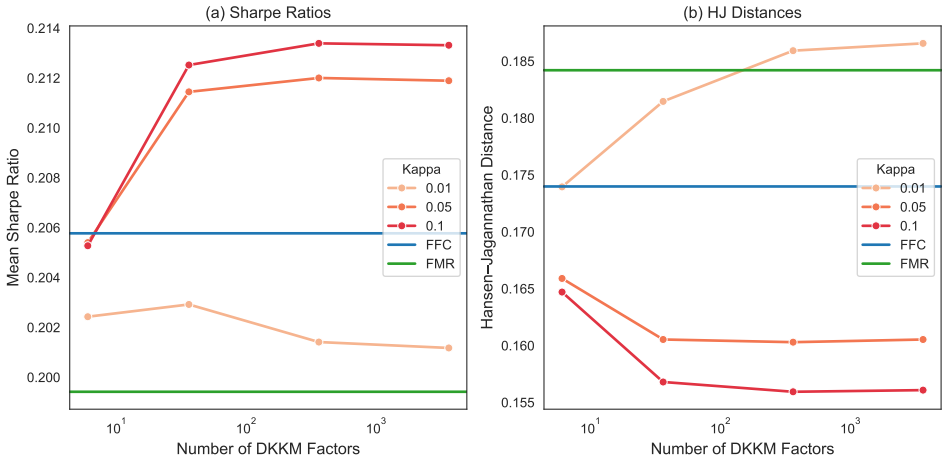
Table 3 shows how the performance of the KPS model varies with the number of factors and how its performance compares to FMR and DKKM. Positive numbers indicate outperformance of KPS for the Sharpe Ratio, and negative numbers indicate outperformance for the HJ distance. It is optimal to use two factors for the BGN economy and three factors for the KP economy. The best performing KPS model outperforms FMR on both metrics. The best-performing KPS model outperforms the best-performing DKKM model in the BGN economy, but the reverse is true in the KP economy.

Figure 4.2 shows how the performance of the KPS model varies with the number of factors and how it compares to the FFC and FMR models. With either two, three, or four factors, the KPS model outperforms both FFC and FMR on both metrics and in both economic models.

The Sharpe ratio of a portfolio is equal to minus its correlation with the SDF multiplied by the standard deviation of the SDF. So, the Sharpe ratios of the portfolios $\hat{\beta}_t' F_t$ are

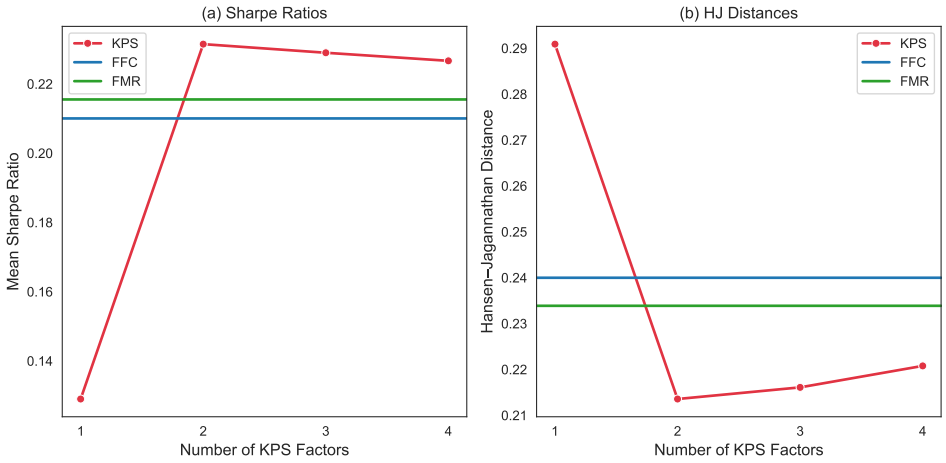


I. Berk, Green, and Naik (1999)

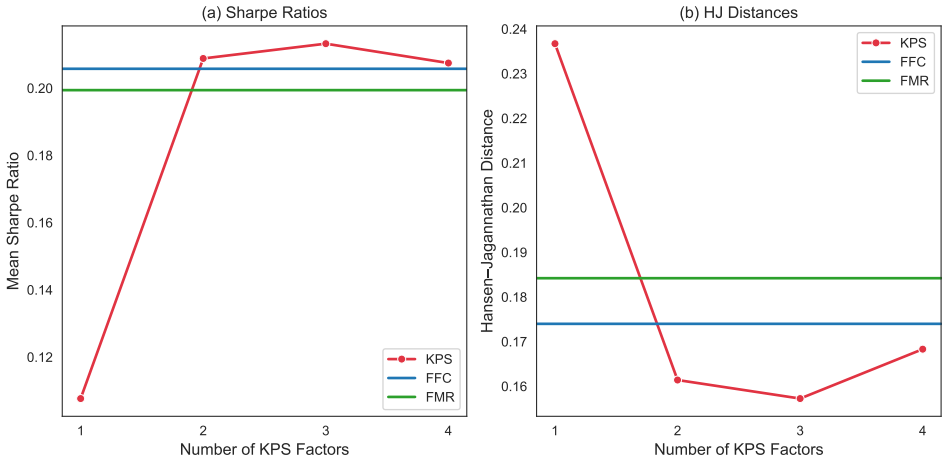


II. Kogan and Papanikolaou (2014)

Figure 4.1: **Performance of DKKM.** 10 panels of data are simulated for the BGN and KP economies. The conditional Sharpe ratio $\hat{\beta}_t'\mathbb{E}_t[F_{t+1}]/\sqrt{\beta_t'\,\mathrm{cov}_t(F_{t+1})\beta_t}$ and $(\hat{\beta}_t'F_{t+1}-\gamma_t'R_{t+1})^2$ are calculated for each out-of-sample month for the DKKM, FFC and FMR models, and the means are computed in the pooled sample of panels. The square roots of the means of $(\hat{\beta}_t'F_{t+1}-\gamma_t'R_{t+1})^2$ are taken for the HJ distance.



I. Berk, Green, and Naik (1999)



II. Kogan and Pananikolaou (2014)

Figure 4.2: **Performance of KPS.** 10 panels of data are simulated for the BGN and KP economies. The conditional Sharpe ratio $\hat{\beta}_t' \mathbb{E}_t[F_{t+1}] / \sqrt{\beta_t' \operatorname{cov}_t(F_{t+1}) \beta_t}$ and $(\hat{\beta}_t' F_{t+1} - \gamma_t' R_{t+1})^2$ are calculated for each out-of-sample month for the KPS, FFC and FMR models, and the means are computed in the pooled sample of panels. The square roots of the means of $(\hat{\beta}_t' F_{t+1} - \gamma_t' R_{t+1})^2$ are taken for the HJ distance.

proportional to the correlations of the estimated SDFs $1-\hat{\beta}'F_t$ with the true SDF. Thus, in both the BGN and KP models, the DKKM and KPS empirical models produce estimated SDFs that are closer to the true SDF in the HJ Distance and have higher correlations with the true SDF than do the FFC and FMR models. In the next subsection, we look at the cross-sectional correlations of the $\hat{\beta}'F_t$ portfolio weights with the true tangency portfolio weights, and we also look at the persistence of weights.

4.2. Portfolio Weights

Another important goal of an empirical model is to estimate assets' exposure to systematic risks. Along this dimension, the machine-learning models of KPS and DKKM are superior to the simpler FFC and FMR models.

To begin with, the true SDF weights in both economic models are quite persistent. We calculate this by fitting a first-order autoregression (AR(1)) separately for each firm and simulation; in Panel I of Table 4 we report the mean, median, and standard deviation of those AR(1) parameter estimates. For what follows, the mean and median are close and we discuss only the former. Panel a.I shows that in BGN the average persistence of weights is 0.963, and Panel b.I shows that in KP the average persistence is 0.931. The standard deviations are quite low in both models (0.021 and 0.028, respectively), which tells us that for the majority of firms their SDF weight is quite persistent.

Before moving onto the empirical models, let us consider what this tells us about economics. In these models, a firm's return is related to aggregate risk through the firm-specific component of its state variables. In BGN, a firm-specific state is its collection of past projects still in operation, along with their risks; and, each project's idiosyncratic productivity state comes from an iid process. In KP, a firm-specific state is its capital stock comprised of past projects still in operation and their productivities; each firm's and each project's idiosyncratic productivity follows modestly persistent processes; and investment opportunities arrive fairly infrequently, on average. In both models, a firm's optimal investment choices lead its firm-specific state to only gradually change over time. The nature of firms' optimal choices, in either model, ensures a slow firm evolution. This

is one force behind firms' persistent SDF weights.

Both KPS and DKKM pick up this important economic feature. In BGN, KPS estimates an average SDF weight persistence of 0.962 and DKKM finds 0.910, both of which are close to the true 0.963. Meanwhile, the simpler FFC and FMR models struggle, with average persistences of 0.612 and 0.592, respectively. The story is similar for the KP model: it's average persistence of 0.931 is close to what KPS (0.907) and DKKM (0.833) deliver, and noticeably higher than what FFC (0.654) and FMR (0.743) would suggest. That is, the simpler empirical models struggle to identify that firms evolve only gradually in their relationship to the SDF.

Furthermore, the standard deviations tell us that simple models allow for more firm heterogeneity in this dimension than the economic models actual have. Recall that for BGN and KP these standard deviations were 0.02–0.03. Across the models, for KPS the standard deviations are both about 0.03 and for DKKM they are 0.06-0.07: this is indicative of the machine-learning models triangulating the economic truth that most firms have very persistent SDF weights. Meanwhile, across the models, for FFC the standard deviations are 0.10–0.18 and for FMR there are 0.18–0.20: this is indicative that the simpler empirical models are finding a type of heterogeneity that does not actually exist.[8]

Now we move to understanding the cross section of firms. This is one of main goals of the economic models we have considered: they are designed to explain why firms differ in their exposure to aggregate risks. In the models, firms are meaningfully heterogeneous in their SDF weights, as a direct function of their firm-specific state. Therefore, a mark of success for an empirical model is its ability to correctly distinguish firms' SDF weights on the basis of observable firm characteristics.

Table 4 Panel II reports cross-sectional correlations. Specifically, for each month and simulation, we calculate the correlation of estimated weights to true weights, and we report the mean, median, and standard deviation of those. A broad-based, first observation

[8]The range of AR(1) estimates also bears this out. The minimum AR(1) parameter in BGN is 0.83: KPS gets 0.62, DKKM gets 0.37, FFC gets -0.02, and FMR gets -0.02. In KP it is 0.75: KPS gets 0.63, DKKM gets 0.56, FFC gets 0.20, and FMR gets -0.05.

is that the average cross-sectional correlations are below one. We view this as reflecting the sampling error that affects all empirical models, and so focus on a comparison of the correlations between estimators.

We see a noteworthy separation between the machine-learning and simpler models. Looking first at Table 4 Panel a.II, we see KPS achieves an average cross-section correlation of 0.467, and DKKM is nearby at 0.441. There is modest variation in this cross-sectional performance, as the standard deviation (across months and simulations) is 0.05 and 0.09, respectively. Meanwhile, the simpler FFC and FMR models achieve average cross-sectional correlations that are 30–40% lower at 0.284 and 0.308, respectively. These are also accompanied by greater variation of the cross-sectional correspondences, with standard deviations of 0.10 and 0.13, respectively.[9]

In Panel b.II using the KP model, the separation is not as large but still present. KPS has an average cross-sectional correlation of 0.313 and DKKM is essentially the same at 0.310. The average correlations for FFC and FMR are about 20–30% lower at 0.265 and 0.232, respectively. The evidence on variability is more mixed across the two groups.

The cross-sectional correlations demonstrate that machine-learning empirical models are superior at capturing the actual economic heterogeneity these models were built to explain.

5. Conclusion

In data simulated from different well-known economic models, we find that machine-learning estimators outperform simpler estimators at revealing the true SDF. Furthermore, we show that the estimates can be used to learn about the economics driving different firms' weights in the SDF. These results show that machine-learning estimators can uncover economic structure better than simpler estimators, even in simple environments reflecting only well-defined economic forces. Thus, machine learned, complex estimates of

[9]Of course, standard deviation is symmetric and higher cross-sectional correlations are good, so if the larger standard deviations of FFC and FMR are driven by positive skew, that would be good for them. This is not the case. The interquartile ranges (across months and simulations) is (0.217, 0.350) for FFC, (0.225, 0.400) for FMR. For KPS it is (0.443,0.5) and for DKKM it is (0.380,0.502).

the SDF in actual data hold the potential to sharpen our understanding of the economic forces driving firm heterogeneity, and hence cross-sectional asset pricing.

Appendices

A. Berk, Green, and Naik (1999)

The SDF at date $t-1$ for pricing cash flows at t is M_t/M_{t-1}, which is given by

$$\log M_t = \log M_{t-1} - r_{t-1} - \frac{1}{2}\sigma_m^2 + \sigma_m \varepsilon_t \,. \tag{A.1}$$

The interest rate process is a Vasicek (1979) process:

$$r_t = r_{t-1} + \kappa(\theta - r_t) + \sigma_r \xi_t \,. \tag{A.2}$$

Here, ε and ξ are sequences of independent standard normals that are contemporaneously correlated with each other, and σ_m, σ_r, κ, and θ are constants. Let ρ_{mr} denote the correlation between ε_t and ξ_t.

Each firm receives a single take-it-or-leave-it investment opportunity each period. The cost of each investment opportunity is the same and is taken as the numeraire. A project that is taken at time s is alive and delivers cash flows at date $s+1$ with probability $1-\pi$ for a constant π. Conditional on being alive at date t for $t > s$, it is likewise alive and delivers cash flows at date $t+1$ with probability $1-\pi$. Conditional on being alive at $t > s$, the cash flow at date t of a project that was taken at date s is given by

$$\log C_{st}^i = \overline{C} - \frac{1}{2}\sigma_{is}^2 + \sigma_{is}\left(\rho_{is}\varepsilon_t + \sqrt{1-\rho_{is}^2}\eta_{st}^i\right) \,, \tag{A.3}$$

where $\overline{C}$ is a constant, σ_{is} and ρ_{is} are observed at date s, and the η_{st}^i are independent standard normals that are independent of the ε and ξ sequences. The correlation between the cash flow and $\Delta \log M$ is ρ_{is}.

Following BGN, set $\beta_{is} = \sigma_m \sigma_{is} \rho_{is}$. Following BGN, β_{is} is drawn independently across time and across firms from a translated exponential distribution with support $(-\infty, \beta^*]$, for a constant β^*, and σ_{is} is drawn from a uniform distribution with endpoints determined

by β_{is}/σ_m to ensure that $\rho_{is} = \beta_{is}/(\sigma_m\sigma_{is})$ is between -1 and $+1$.

The operating cash flow of firm i at date t is

$$\mathrm{OCF}_t^i := \sum_{s<t} C_{st}^i \chi_{st}^i \,,$$

where χ_{st}^i is the indicator for the project that arrived at s being alive at t. The free cash flow is

$$\mathrm{FCF}_t^i = \mathrm{OCF}_t^i - \chi_{tt}^i \,.$$

Here, χ_{tt}^i is the indicator for the project that arrived at t being taken at t. Free cash flow is paid out to shareholders.

The value at date t of a project that arrived at $s < t$, after cash flows at date t are realized, is

$$V_{st}^i = \mathbb{E}_t \left[\sum_{u=t+1}^{\infty} \frac{M_u}{M_t} C_{su}^i \chi_{su} \right]$$

BGN show that, conditional on the project being alive at t,

$$V_{st}^i = \mathrm{e}^{\overline{C} - \beta_{is}} D(r_t)$$

where $D(r)$ denotes the value of a perpetual consol with coupons declining at rate $1 - \pi$. Thus, the value of assets in place at date t for firm i is

$$A_t^i := \mathrm{e}^{\overline{C}} D(r_t) \sum_{s \le t} \chi_{st}^i \mathrm{e}^{-\beta_{is}} \,.$$

The value of growth options is the same for each firm and depends on the interest rate. Denote it by $J^*(r_t)$. BGN show that

$$J^*(r) = \mathrm{e}^{\overline{C}} \sum_{u=1}^{\infty} \sum_{k=1}^{\infty} \int_{-\infty}^{\beta^*} J(r, u, k, r_\beta^*) f(\beta) \, \mathrm{d}\beta \,, \tag{A.4}$$

where $J(r, u, k, r_\beta^*)$ denotes the value at date 0 when the interest rate is r of a European call option maturing at u on a discount bond maturing at $u + k$ with strike price de-

termined by r^*_β, which is the solution of an implicit equation, and where f denotes the translated exponential density function of each β_{is} as described above. Conditional on the strike price, the call option values have closed form solutions.

It follows that the excess return of firm i from t to $t+1$ is

$$\frac{A^i_{t+1} + J^*(r_{t+1}) + \mathrm{FCF}^i_{t+1}}{A^i_t + J^*(r_t)} - \mathrm{e}^{r_t}\,. \tag{A.5}$$

Let R_{t+1} denote the N–vector of excess stock returns from t to $t+1$. As discussed in the text, we can calculate a particular SDF by regressing the constant 1 on the excess returns, which means that the SDF is

$$1 - \mathbb{E}_t[R_{t+1}]'\mathbb{E}_t[R_{t+1}R'_{t+1}]^{-1}R_{t+1}\,.$$

We need to compute the vector

$$\mathbb{E}_t[R_{t+1}R'_{t+1}]^{-1}\mathbb{E}_t[R_{t+1}]\,.$$

Equivalently (in theory, though not numerically), we need to compute the solution x to the system of linear equations $Ax = b$ where $A = \mathbb{E}_t[R_{t+1}R'_{t+1}]$ and $b = \mathbb{E}_t[R_{t+1}]$. BGN calculate $\mathbb{E}_t[R_{t+1}]$. The additional complication is to calculate the expectation of products and squares of excess returns and to solve the system of linear equations.

To calculate the expectation of products and squares of excess returns, we need to evaluate the following integrals:

$$\mathbb{E}_t\left[\int_{-\infty}^{\beta^*}((e^{\overline{C}-\beta}D(r_{t+1})-1)^+)^2 f(\beta)\,\mathrm{d}\beta\right] \tag{A.6}$$

$$\mathbb{E}_t\left[\int_{-\infty}^{\beta^*}\int_{-\infty}^{\beta^*}(e^{\overline{C}-\beta}D(r_{t+1})-1)^+(e^{\overline{C}-\beta'}D(r_{t+1})-1)^+ f(\beta)f(\beta')\,\mathrm{d}\beta\,\mathrm{d}\beta'\right] \tag{A.7}$$

$$\mathbb{E}_t\big[J^*(r_{t+1})^2\big] \tag{A.8}$$

$$\mathbb{E}_t\big[D(r_{t+1})^2\big] \tag{A.9}$$

$$\mathbb{E}_t\big[D(r_{t+1})J^*(r_{t+1})\big] \tag{A.10}$$

$$\mathbb{E}_t\left[\int_{-\infty}^{\beta^*} J^*(r_{t+1})(e^{\overline{C}-\beta}D(r_{t+1})-1)^+ f(\beta)\,\mathrm{d}\beta\right] \tag{A.11}$$

$$\mathbb{E}_t\left[\int_{-\infty}^{\beta^*} D(r_{t+1})(e^{\overline{C}-\beta}D(r_{t+1})-1)^+ f(\beta)\,\mathrm{d}\beta\right] \tag{A.12}$$

$$\mathbb{E}_t\left[J^*(r_{t+1})e^{\gamma\xi_{t+1}}\right] \tag{A.13}$$

$$\mathbb{E}_t\big[D(r_{t+1})e^{\gamma\xi_{t+1}}\big] \tag{A.14}$$

$$\mathbb{E}_t\left[\int_{-\infty}^{\beta^*} (e^{\overline{C}-\beta}D(r_{t+1})-1)^+ e^{\gamma\xi_{t+1}} f(\beta)\,\mathrm{d}\beta\right] \tag{A.15}$$

In (A.13)–(A.15), γ denotes $\sigma_{it}\rho_{it}\rho_{mr}$ and is treated as a constant, because it is known at date t. The conditional expectations are all integrals over the shock ξ_{t+1} to the interest rate at date $t+1$. We compute any inner integrals over β, β' analytically, for given values of $J^*(r_{t+1})$ and $D(r_{t+1})$, and then compute the outer integrals over ξ_{t+1} by Gauss-Hermite quadrature.

B. Kogan and Papanikolaou (2014)

The economy-wide productivity parameters and the SDF process are as described in the text. Firms acquire new projects at firm-specific arrival rates $\lambda_{it} = \lambda_i \times \tilde{\lambda}_{it}$. Here, $\tilde{\lambda}_{it}$ follows a two-state continuous-time Markov chain, and λ_i is a firm-specific constant drawn from a shifted negative log-uniform distribution.

Firm i chooses how much capital to invest in any project j. If project j of firm i arrives at date s and firm i invests K, then firm i incurs an investment cost $z_s^{-1}x_sK$ at date s. Subsequently, the project produces cash flows at rate

$$\epsilon_{it}u_{jt}x_tK^\alpha$$

for $t > s$ until the project expires, which occurs at rate δ. The processes ϵ_i and u_j are independent square-root processes with mean equal to 1. The process ϵ_i is firm specific, while u_j is specific to project j of firm i, and $u_{js} = 1$ if project j arrives at date s.

KP show that the market value of a project that arrived at date s and is alive at date $t \geq s$, conditional on capital K having been invested, is

$$\mathbb{E}_t \left[\int_t^\infty \frac{M_\tau}{M_t} e^{-\delta(\tau - t)} \epsilon_{i\tau} u_{j\tau} x_\tau K^\alpha \, \mathrm{d}\tau \right] = A(\epsilon_{it}, u_{jt}) x_t K^\alpha \,,$$

for a quadratic function A. When choosing how much to invest in a project, a firm trades off the present value of future cash flows with the investment cost. The optimal investment in project j of firm i that arrives at time s is

$$K_{ij} = (\alpha z_s A(\epsilon_{is}, 1))^{\frac{1}{1-\alpha}} \,.$$

As in BGN, the value of the firm is equal to the value of its assets in place plus the value of its growth options. The value of the assets in place at time t is equal to the sum of the present values of live projects:

$$x_t \sum_{j \in \mathcal{J}_{it}} A(\epsilon_{it}, u_{jt}) K_{ij}^\alpha,$$

where $\mathcal{J}_{it}$ is the set of projects that are alive at date t. KP also show that the value of firm i's growth options at time t is equal to

$$z_t^{\frac{\alpha}{1-\alpha}} x_t G(\epsilon_{it}, \lambda_{it}) \,,$$

for a function G that solves an ordinary differential equation. The total market value of the firm is

$$V_t^i = x_t \sum_{j \in \mathcal{J}_{it}} A(\epsilon_{it}, u_{jt}) K_{ij}^\alpha + z_t^{\frac{\alpha}{1-\alpha}} x_t G(\epsilon_{it}, \lambda_{it}) \,.$$

To simulate the model, we discretize it at a monthly frequency. We approximate the Poisson arrival rate and expiration of projects by a simple Bernoulli approximation, and we use closed-form transition distributions to discretize the other continuous-time processes that describe the model. Without loss of generality, take one unit of time to

equal a month. Operating cash flows from date $t-1$ to date t are

$$\mathrm{OCF}_t^i = \sum_{j \in \mathcal{J}_{i,t-1}} \epsilon_{it} u_{jt} x_t K_{ij}^{\alpha} \,,$$

Let χ_{it} denote the indicator function for whether a project arrives to firm i during the period $(t-1,t)$. We assume only one project can arrive during each period, and, if a project arrives, it arrives at date t. Then, free cash flows are

$$\mathrm{FCF}_t^i := \mathrm{OCF}_t^i - \chi_{it} z_t^{-1} x_t (\alpha z_t A(\epsilon_{it}, 1))^{\frac{1}{1-\alpha}} \,.$$

Firm i's return between dates t and $t+1$ is

$$R_{i,t+1} = \frac{V_{t+1}^i + \mathrm{FCF}_{t+1}^i}{V_t^i} \,.$$

As in BGN, we need to compute the vector

$$\mathbb{E}_t[R_{t+1} R_{t+1}']^{-1} \mathbb{E}_t[R_{t+1}] \,,$$

where R_{t+1} denotes the N–vector of excess stock returns from t to $t+1$. Evaluating the first two moments of returns requires computing several expectations numerically:

$$\mathbb{E}_t \left[A(\epsilon_{i,t+1}, 1)^{\frac{1}{1-\alpha}} A(\epsilon_{i,t+1}, u_{s,t+1}) \right] \tag{B.1}$$

$$\mathbb{E}_t \left[A(\epsilon_{i,t+1}, 1)^{\frac{2}{1-\alpha}} \right] \tag{B.2}$$

$$\mathbb{E}_t \left[A(\epsilon_{i,t+1}, 1)^{\frac{1}{1-\alpha}} \right] \tag{B.3}$$

$$E_t \left[G(\epsilon_{i,t+1}, \lambda_{i,t+1}) \right] \tag{B.4}$$

$$\mathbb{E}_t \left[A(\epsilon_{i,t+1}, 1) G(\epsilon_{i,t+1}, \lambda_{i,t+1}) \right] \tag{B.5}$$

$$\mathbb{E}_t \left[G(\epsilon_{i,t+1}, \lambda_{i,t+1})^2 \right] \tag{B.6}$$

$$\mathbb{E}_t \left[G(\epsilon_{i,t+1}, \lambda_{i,t+1}) A(\epsilon_{i,t+1}, 1)^{\frac{1}{1-\alpha}} \right]. \tag{B.7}$$

These expectations integrate over the distribution of $\epsilon_{i,t+1}$ given $\epsilon_{i,t}$, using closed-form expressions for the discrete transition function of a square-root process and standard numerical integration routines in Python. The expectation (B.1) also integrates over the conditional distribution of $u_{s,t+1}$, while the expectations (B.4)-(B.7) also integrate over the conditional distribution of $\lambda_{i,t+1}$.

References

Ai, Hengjie, Mariano Max Croce, Anthony M Diercks, and Kai Li, 2018, News shocks and the productiosn-based term structure of equity returns, The Review of Financial Studies 31, 2423–2467.

Bali, Turan G., David Beckmeyer, Philipp Mörke, and Florian Weigert, 2023, Option return predictability with machine learning and big data, The Review of Financial Studies .

Barillas, Francisco, and Jay Shanken, 2017, Which alpha?, Review of Financial Studies 30, 1316–1338.

Barnett, Michael, William Brock, Lars Peter Hansen, Ruimeng Hu, and Joseph Huang, 2023, A deep learning analysis of climate change, innovation, and uncertainty, Technical report, Arizona State Univeristy, Univesity of Wisconsin, University of Chicago, UCSB, University of Pennsylvania.

Berk, Jonathan, 2023, Comment on 'The virtue of complexity in return prediction', Technical report, Stanford Universty.

Berk, Jonathan B., Richard C. Green, and Vasant Naik, 1999, Optimal investment, growth options, and security returns, Journal of Finance 54, 1553–1607.

Bianchi, Francesco, Markus Büchner, and Andrea Tamoni, 2021, Bond risk premiums with machine learning, The Review of Financial Studies 34, 1046–1089.

Box, George E. P., 1976, Science and statistics, Journal of the American Statistical Association 71, 791–799.

Britten-Jones, Mark, 1999, The sampling error in estimates of mean-variance efficient portfolio weights, Journal of Finance 54, 655–671.

Carlson, Murray, Adlai Fisher, and Ron Giammarino, 2004, Corporate investment and asset price dynamics: Implications for the cross-section of returns, The Journal of Finance 59, 2577–2603.

Cochrane, John H., 2011, Presidential address: Discount rates, Journal of Finance 66, 1047–1108.

Didisheim, Antoine, Shikun Ke, Bryan Kelly, and Semyon Malamud, 2024, APT or "AIPT"? The surprising dominance of large factor models, Research Paper 23-19, Swiss Finance Institute.

Duarte, Jefferson, Jefferson Duarte, and Tiago Silva, 2024, Machine learning for continuous-time finance, The Review of Financial Studies .

Eisfeldt, Andrea, and Dimitris Papanikolaou, 2013, Organization capital and the cross-section of expected returns, The Journal of Finance 68, 1365–1406.

Fama, Eugene F., 1976, Foundations of Finance: Portfolio Decisions and Securities Prices (Basic Books, New York).

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, Journal of Financial Economics 116, 1–22.

Fama, Eugene F, and Kenneth R French, 2019, Comparing cross-section and time-series factor models, The Review of Financial Studies 33, 1891–1926.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, Journal of Political Economy 81, 607–636.

Freyberger, Joachim, Andreas Neuhierl, and Michael Weber, 2020, Dissecting characteristics nonparametrically, The Review of Financial Studies 33, 2321–2377.

Gomes, Joao, Leonid Kogan, and Lu Zhang, 2003, Equilibrium cross section of returns, Journal of Political Economy 111, 693–732.

Gomes, Joao F., and Lukas Schmid, 2021, Equilibrium asset pricing with leverage and default, The Journal of Finance 76, 977–1018.

Gu, Shihao, Bryan Kelly, and Dacheng Xiu, 2020, Empirical asset pricing via machine learning, The Review of Financial Studies 33, 2223–2273.

Hansen, Lars Peter, and Ravi Jagannathan, 1997, Assessing specification errors in stochastic discount factor models, Journal of Finance 52, 557–590.

Hansen, Lars Peter, and Scott F. Richard, 1987, The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models, Econometrica 55, 587–613.

Hastie, Trevor, Andrea Montanari, Saharon Rosset, and Ryan J. Tibshirani, 2022, Surprises in high-dimensional ridgeless least squares interpolation, Annals of Statistics 50, 949–986.

Kelly, Brian T., Seth Pruitt, and Yinan Su, 2019, Characteristics are covariances: A unified model of risk and return, Journal of Financial Economics 134, 501–524.

Kelly, Bryan, Semyon Malamud, and Kangying Zhou, 2024, The virtue of complexity in return prediction, The Journal of Finance 79, 459–503.

Kelly, Bryan T., and Semyon Malamud, 2025, Understanding the virtue of complexity, NBER working paper, National Bureau of Economic Research.

Kogan, Leonid, and Dimitris Papanikolaou, 2014, Growth opportunities, technology shocks, and asset prices, Journal of Finance 69, 675–718.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2020, Shrinking the cross-section, Journal of Financial Economics 135, 271–292.

Kuehn, Lars-Alexander, and Lukas Schmid, 2014, Investment-based corporate bond pricing, The Journal of Finance 69, 2741–2776.

Lettau, Martin, and Markus Pelger, 2020, Factors that fit the time series and cross-section of stock returns, The Review of Financial Studies 33, 2274–2320.

Li, Y., A. Rossi, H. Yan, and L. Zheng, 2025, Machine learning from a 'universe' of signals: The role of feature engineering, Journal of Financial Economics .

Lintner, John, 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, The Review of Economics and Statistics 47, 13–37.

Nagel, Stefan, 2025, Seemingly virtuous complexity in return prediction, NBER Working Paper Series .

Rosenberg, Barr, 1974, Extra-market components of covariance in security returns, Journal of Financial and Quantitative Analysis 9, 263–274.

Sharpe, William F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, The Journal of Finance 19, 425–442.

Vasicek, Oldřich, 1979, An equilibrium characterization of the term structure, Journal of Financial Economics 7, 83–105.

Zhang, Lu, 2005, The value premium, The Journal of Finance 60, 67–103.

	Berk, Green, and Naik (1999)		Kogan and Papanikolaou (2014)	
	Sharpe Ratio	HJ Distance	Sharpe Ratio	HJ Distance
FMR	0.216	0.234	0.199	0.184
FFC	0.210	0.24	0.206	0.174
FMR - FFC	0.005	-0.006	-0.06	0.010

Table 1: **Performance of FMR and FFC Models.** 10 panels of data are generated for the BGN and KP economies. The conditional Sharpe ratio $\hat{\beta}_t'\mathbb{E}_t[F_{t+1}]/\sqrt{\beta_t' \operatorname{cov}_t(F_{t+1})\beta_t}$ and $(\hat{\beta}_t'F_{t+1} - \gamma_t'R_{t+1})^2$ are calculated for each out-of-sample month for the DKKM and FMR models, and the means are computed in the pooled sample of panels. The HJ distance is the square root of the mean of $(\hat{\beta}_t'F_{t+1} - \gamma_t'R_{t+1})^2$.

I. Berk, Green, and Naik (1999)

(a) Sharpe Ratio				
κ	6	36	360	3600
0.000	-0.012	-0.059	-0.212	-0.209
0.001	-0.015	-0.014	-0.021	-0.021
0.010	-0.014	0.003	0.002	0.002
0.050	-0.014	0.009	0.011	0.011
0.100	-0.016	0.009	0.011	0.011
1.000	-0.040	-0.019	-0.015	-0.015
(b) Hansen-Jagannathan Distance				
κ	6	36	360	3600
0.000	0.031	0.210	270.722	7.882
0.001	0.029	0.066	0.089	0.093
0.010	0.023	0.020	0.026	0.028
0.050	0.019	0.003	0.002	0.003
0.100	0.020	0.001	-0.001	-0.001
1.000	0.039	0.030	0.028	0.028

II. Kogan and Pananikolaou (2014)

(a) Sharpe Ratio				
κ	6	36	360	3600
0.000	-0.003	-0.063	-0.172	-0.187
0.001	-0.002	-0.014	-0.022	-0.023
0.010	0.003	0.004	0.002	0.002
0.050	0.006	0.012	0.013	0.012
0.100	0.006	0.013	0.014	0.014
1.000	-0.007	0.004	0.005	0.005
(b) Hansen-Jagannathan Distance				
κ	6	36	360	3600
0.000	0.011	0.182	29.990	3.190
0.001	0.001	0.040	0.060	0.062
0.010	-0.010	-0.003	0.002	0.002
0.050	-0.018	-0.024	-0.024	-0.024
0.100	-0.020	-0.027	-0.028	-0.028
1.000	-0.003	-0.014	-0.016	-0.016

Table 2: **Performance of the DKKM Model Relative to FMR.** 10 panels of data are generated for the BGN and KP economies. The conditional Sharpe ratio $\hat{\beta}_t'\mathbb{E}_t[F_{t+1}]/\sqrt{\beta_t'\,\mathrm{cov}_t(F_{t+1})\beta_t}$ and $(\hat{\beta}_t'F_{t+1}-\gamma_t'R_{t+1})^2$ are calculated for each out-of-sample month for the DKKM and FMR models, and the means are computed in the pooled sample of panels. The table reports the difference in means. The square roots of the means of $(\hat{\beta}_t'F_{t+1}-\gamma_t'R_{t+1})^2$ are taken before differencing for the HJ distance.

I. Berk, Green, and Naik (2014)

(a) Sharpe Ratio				
	1	2	3	4
vs FMR	-0.087	0.016	0.013	0.011
vs DKKM	-0.097	0.005	0.003	0.000
(b) Hansen-Jagannathan Distance				
	1	2	3	4
vs FMR	0.057	-0.020	-0.018	-0.013
vs DKKM	0.058	-0.019	-0.017	-0.012

II. Kogan and Pananikolaou (2014)

(a) Sharpe Ratio				
	1	2	3	4
vs FMR	-0.092	0.009	0.014	0.008
vs DKKM	-0.106	-0.004	-0.000	-0.006
(b) Hansen-Jagannathan Distance				
	1	2	3	4
vs FMR	0.053	-0.023	-0.027	-0.016
vs DKKM	0.081	0.005	0.001	0.012

Table 3: **Performance of the KPS Model Relative to FMR and DKKM.** 10 panels of data are generated for the BGN and KP economies. The conditional Sharpe ratio $\hat{\beta}_t' \mathbb{E}_t[F_{t+1}] / \sqrt{\beta_t' \operatorname{cov}_t(F_{t+1}) \beta_t}$ and $(\hat{y}_{t+1} - z_{t+1})^2$ are calculated for each out-of-sample month for the DKKM, FMR, and KPS models, and the means are computed in the pooled sample of panels. The table reports the difference in means of the KPS model compared to the DKKM model with 3600 factors and $\kappa = 0.1$, and the KPS model compared to the FMR model. The square root of the HJ distance is taken before differencing.

I. Berk et al. (1999)

	Mean	Median	Std. Dev.
I. Serial Correlation			
Truth	0.963	0.968	0.021
FFC	0.612	0.651	0.183
FMR	0.592	0.619	0.179
KPS	0.962	0.968	0.027
DKKM	0.910	0.923	0.055
II. Cross-sectional Correlation			
FFC	0.284	0.284	0.097
FMR	0.308	0.322	0.133
KPS	0.467	0.473	0.050
DKKM	0.441	0.432	0.089

II. Kogan and Papanikolaou (2014)

	Mean	Median	Std. Dev.
I. Serial Correlation			
Truth	0.931	0.935	0.028
FFC	0.654	0.664	0.103
FMR	0.743	0.816	0.202
KPS	0.907	0.910	0.026
DKKM	0.833	0.838	0.059
II. Cross-sectional Correlation			
FFC	0.265	0.277	0.075
FMR	0.232	0.252	0.113
KPS	0.313	0.336	0.101
DKKM	0.310	0.325	0.070

Table 4: **SDF Weight Statistics.** Serial correlations are calculated as the AR(1) parameter in time-series regressions for every firm and simulation; the Mean, Median, and standard devation (Std. Dev.) of those parameters are reported. Cross-sectional correlations are calculated (across all firms) for every month and simulation; the Mean, Median, and Std. Dev. of those parameters are reported. For KPS and DKKM, we use the best-performing models from 4.2 and 4.1: KPS with 2 factors and DKKM with 3600 factors and $\kappa = 0.1$, for BGN in Panel (a); KPS with 3 factors and DKKM with 3600 factors and $\kappa = 0.1$, for KP in Panel (b).