# Performance of Factor Models in a Simple Economy

Kerry Back* Alexander Ober†

February 18, 2025

We simulate multiple panels of firm characteristics and stock returns from the Berk, Green, and Naik (1999) equilibrium model. The characteristics identified in the model are the four Fama and French (2015) characteristics plus momentum. We evaluate the performance of the Fama-French-Carhart model, the Fama-MacBeth-Rosenberg model, and the random Fourier features model proposed by Didisheim et al. (2024). We find that the last model outperforms the other two, and performance is increasing in the number of Fourier features up to at least several hundred.

PRELIMINARY AND INCOMPLETE – PLEASE DO NOT CITE

*Jones Graduate School of Business and School of Social Sciences, Rice University

†Jones Graduate School of Business, Rice University

The goal of this project is to perform Monte Carlo analysis of factor construction methodologies. We examine classical methodologies (Fama and French, 2015; Fama and MacBeth, 1973; Rosenberg, 1974) and a very recent proposal: random Fourier features (Didisheim et al., 2024, , hereafter DKKM). We compare methods based on Hansen-Jagannathan distances and Sharpe ratios. We assess statistical significance by generating a sample of independent panels. We are able to avoid some of the sampling error inherent in empirical evaluations because we can compute the true theoretical conditional stochastic discount factor and conditional moments at each date in each panel.

A prime difficulty in assessing factor models via Monte Carlo is that one must choose a data generating process, and the choice of a process may dictate the outcome. To avoid biasing the outcome through selection of the data generating process, we use an off-the-shelf equilibrium model. Berk, Green, and Naik (1999), hereafter BGN, develop a rational pricing model in which firm characteristics such as size and book-to-market have explanatory power for returns. We simulate their model. In the model, it is possible to calculate the four characteristics used in the five-factor Fama and French (2015) model and also momentum. We evaluate factor construction methods based on those five characteristics.

DKKM consider a very large number of factors, prompting the term "complexity" in their title. However, as the authors make clear, and as is well understood, the ultimate goal is to derive a single factor model, the single factor being an estimate of the stochastic discount factor (SDF). The essence of the DKKM methodology is to generate many random factor portfolios and then use penalized regression to form a single factor of the form $\hat{\beta}_t' f_{t+1}$, where $f_{t+1}$ denotes the (perhaps very large) vector of generated factor returns from $t$ to $t+1$, and $\hat{\beta}_t$ is the vector of regression coefficients at $t$. This process should in principle allow the data more freedom to speak regarding what the SDF is than if we start with a small number of factors as is commonly done. The question we address in this paper is whether this principle has any effect in the relatively simple BGN economy.

Our findings are that the DKKM methodology outperforms the classical models in the BGN economy. Furthermore, performance is increasing in the number of factors up to several hundred factors, at which point it plateaus. Remarkably, despite the fact that the DKKM method is based on a large number of *randomly constructed* factors, the standard deviations across simulated panels of sample Sharpe ratios and Hansen-Jagannathan distances of the DKKM model are less than those of the classical models. In more complex economies, and especially when more characteristics are available to study, it seems likely that, as DKKM argue, it is beneficial to generate thousands or even hundreds of thousands of factors before attempting to consolidate them into an estimate of the SDF.

The DKKM methodology is an extension of the portfolio construction methodology of Brandt, Santa-Clara, and Valkanov (2009), hereafter BSV. BSV de-mean characteristics in each cross-section, so they can be used as portfolio weights in a long-short portfolio, and consider portfolios as linear combinations of characteristics. They recommend using the portfolio of this type that maximizes the past sample mean of a utility function. DKKM generate many additional characteristics as sines or cosines of random linear combinations of the original characteristics. Like BSV, they de-mean these generated characteristics in each cross section, and they consider portfolios that are linear combinations of the characteristics. They select the portfolio of this type that maximizes the past sample mean of the quadratic utility function $-(1-r)^2$, where $r$ is the portfolio excess return, except that they impose $L^2$ penalization in the maximization to avoid overfitting. They use this portfolio as an estimate of a mean-variance frontier portfolio and use itto estimate the stochastic discount factor.

We describe the BGN model in the next section. Section 2 describes how we assess factor models. Section 3 describes the factor models we study. Section 4 presents our results. Section 5 concludes.

# 1 Berk-Green-Naik Model

In the BGN model, firms invest optimally given an exogenous pricing kernel and random investment opportunities. The stochastic discount factor (SDF) at date $t$ for pricing cash flows at $t+1$ is

$$m_{t+1} := \mathrm{e}^{-r_t - \frac{1}{2}\sigma_m^2 + \sigma_m \varepsilon_{t+1}} . \tag{1.1}$$

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

$$r_{t+1} = r_t + \kappa(\mu - r_t) + \sigma_r \eta_{t+1} . \tag{1.2}$$

Here, $\varepsilon$ and $\eta$ are independent sequences of iid standard normals.

There are a fixed number of firms. Each firm begins at date 0 with zero capital. Each firm receives an investment opportunity each period. The opportunities expire if not taken in the period in which they arrive. All projects require the same amount of capital $I$ 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 $\varepsilon$ 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 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 it with a period length of one month, as do BGN. Like BGN, we discard the first 200 months to allow the economy to reach a steady state. The following figures present a single simulated panel of 1,000 firms for 720 months (after discarding the first 200). The exact data shown in these figures is not important, but we provide them to illustrate general features of the model.

Figure 1.1 shows a path of the interest rate process. Figures 1.2–1.5 provide information regarding the cross-section of firms at four distinct dates. Figure 1.2 shows the number of active projects across firms. The book equity of a firm equals its number of active projects multiplied by the cost of each project, so Figure 1.2 also provides information about the dispersion of book equity across firms. Figure 1.3 shows the distribution of market equity across firms. Aggregate market equity is relatively high at month 400 and relatively low at month 900 due to differences in the level and history of the interest rate. The level affects both the value of assets in place and the value of growth options, and the history affects the value of assets in place due to the effect of the interest rate on project choice. Figure 1.4 shows the distribution across firms of four firm characteristics: book-to-market, momentum, profitability, and asset growth. Figure 1.5 shows the distribution of returns. As in the actual data, the cross-sectional distribution of returns is leptokurtic and positively skewed.

To compute theoretical conditional moments in the BGN model, we need the list of all current projects for every firm – the number of projects and each project's beta and idiosyncratic risk, and we need to know the current interest rate. Past project decisions depend on past interest rates as well as project betas, so the economy is path dependent. To put it

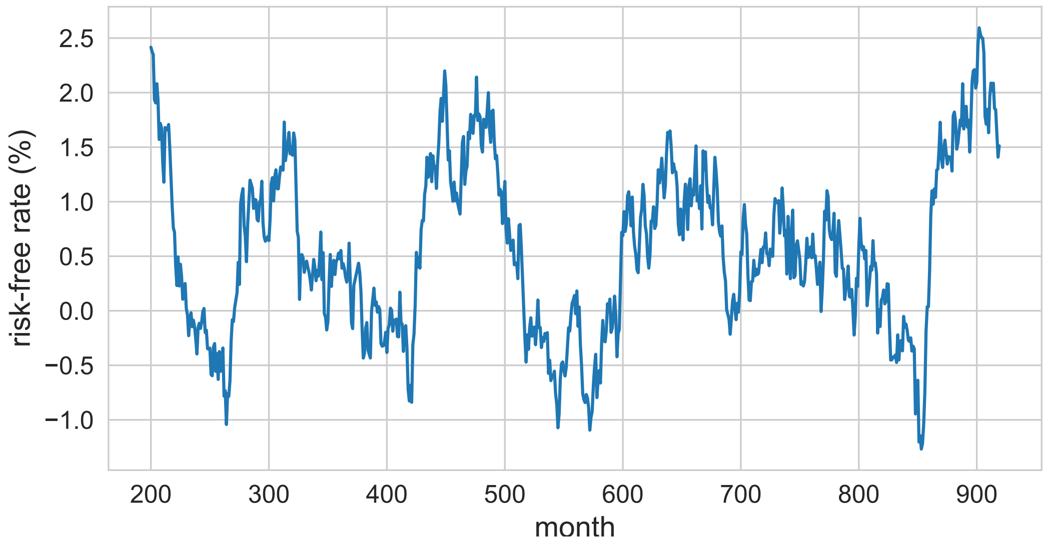


Figure 1.1: A path of the Vasicek interest rate process with the BGN calibration.

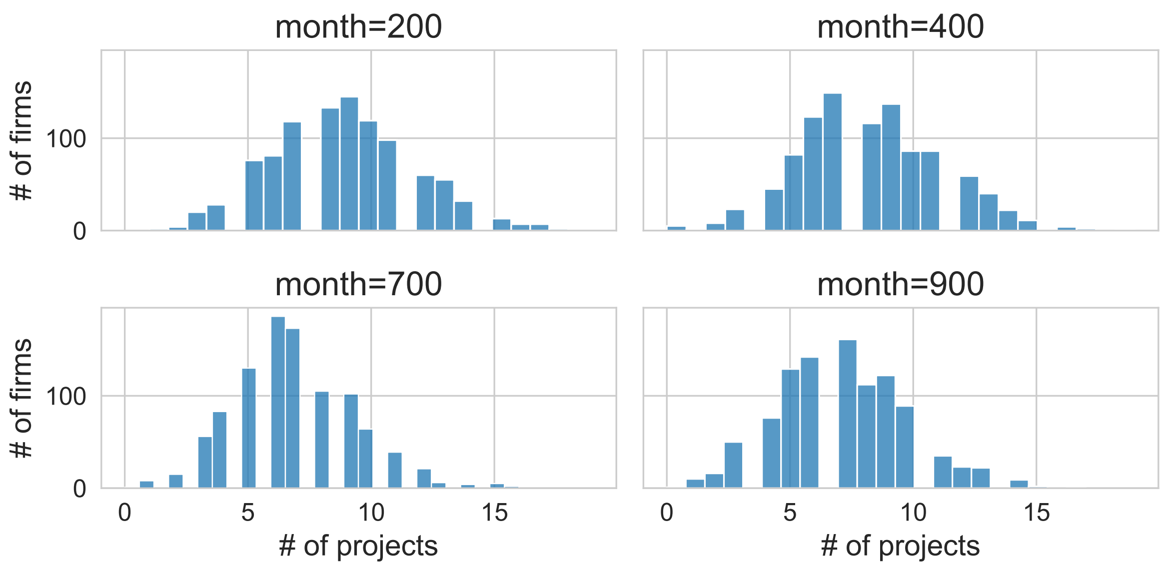


Figure 1.2

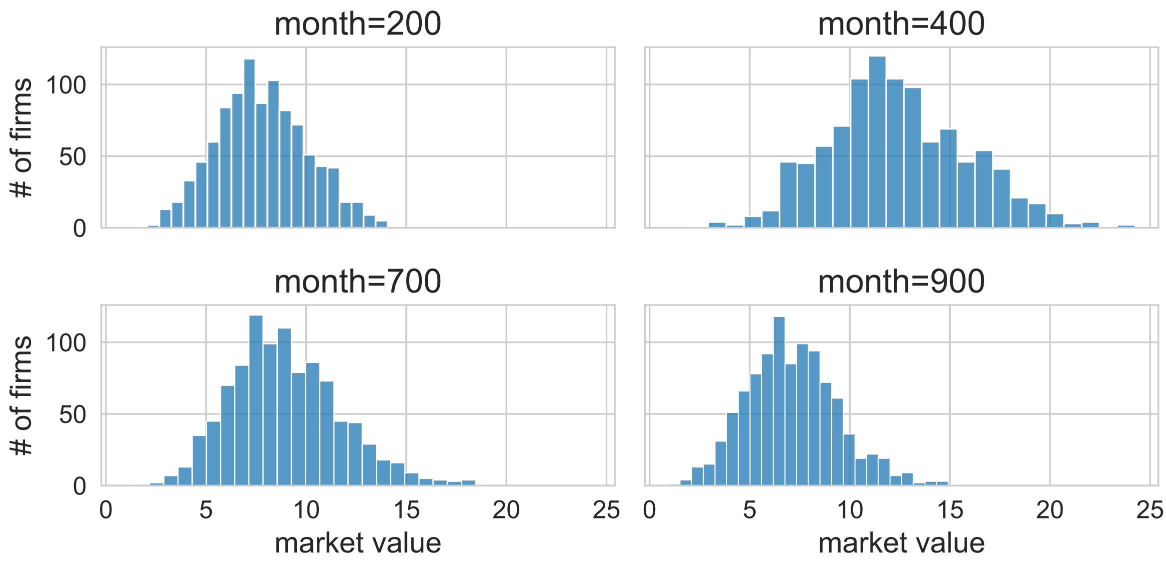


Figure 1.3

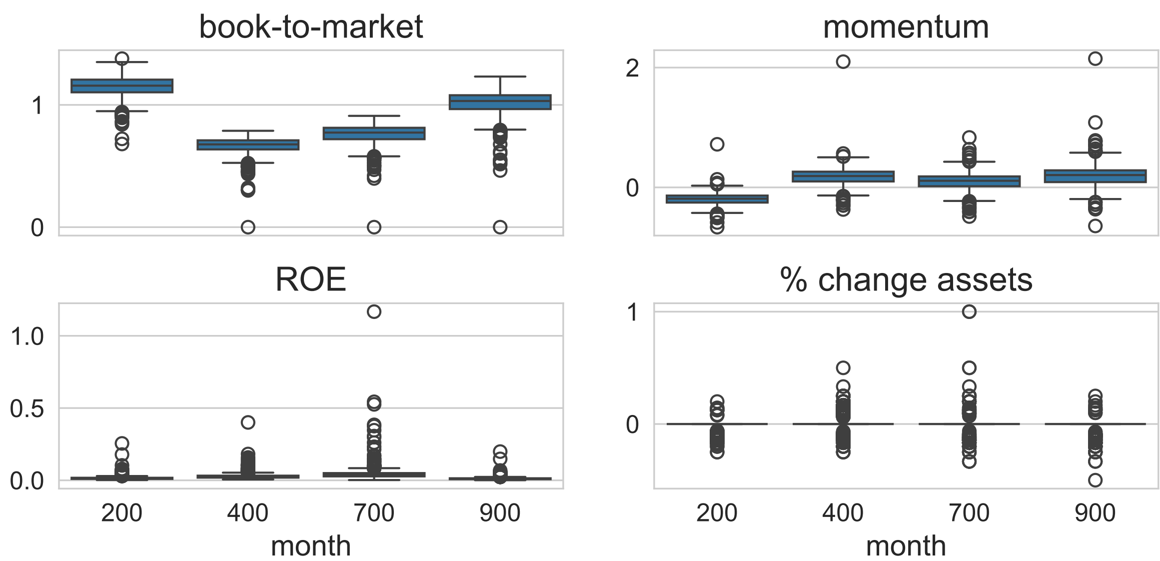


Figure 1.4

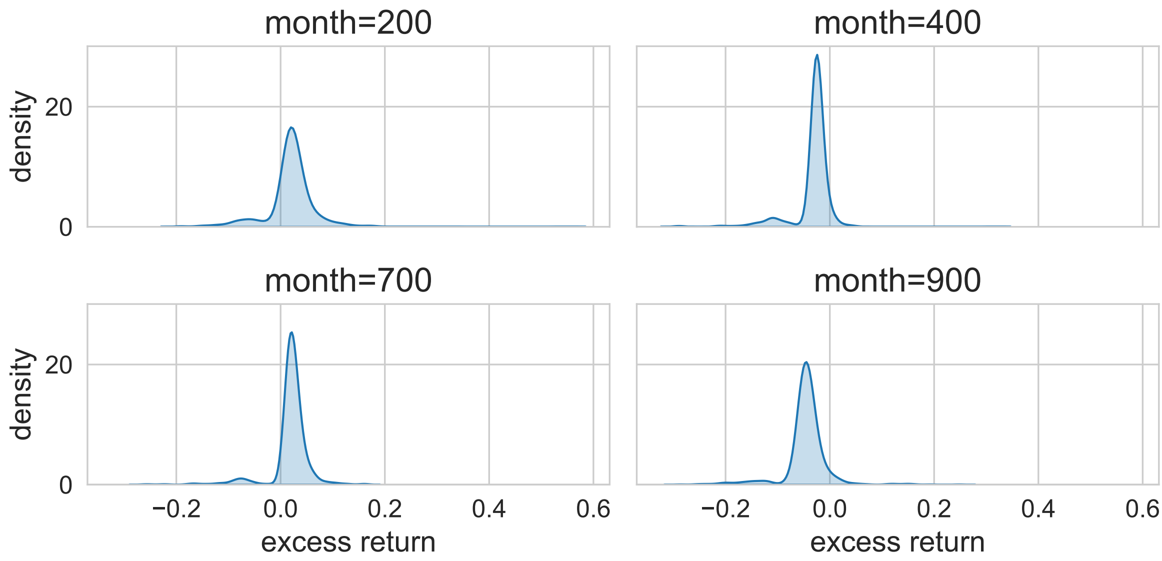


Figure 1.5

another way, the state space of the economy is very large. Nevertheless, we can compute the theoretical conditional moments. We compute the true conditional SDF each period and the true conditional Sharpe ratios of factor portfolios.

## 2 Assessing Factor Models

We follow DKKM closely in our evaluation of models. Hansen and Richard (1987) show that the efficient part of the mean-variance frontier is the set of returns $r_{f,t} + bz_{t+1}$ for $b \geq 0$, where $r_{f,t}$ denotes the risk-free rate from $t$ to $t+1$, and $z_{t+1}$ is the projection of the constant 1 on the space of excess returns from $t$ to $t+1$. The residual $1-z_{t+1}$ in the projection is orthogonal to excess returns. The unique conditional SDF in the span of the asset returns is

$$\frac{1 - z_{t+1}}{(1 + r_{f,t})\mathsf{E}_t[(1 - z_{t+1})^2]} . \tag{2.1}$$

In the factor models that we study, all factors are excess returns. There is a similar representation of the mean-variance frontier spanned by each set of factors and the risk-free asset. Given a set of factors, let $y_{t+1}$ denote the projection of the constant 1 on the set of factor portfolio returns from $t$ to $t+1$, so $\{r_{f,t} + by_{t+1} \mid b \geq 0\}$ is the efficient part of the frontier spanned by the risk-free asset and the factor returns. The unique conditional SDF in the span of the factors and the risk-free asset for pricing the factors and the risk-free asset is

$$\frac{1 - y_{t+1}}{(1 + r_{f,t})\mathsf{E}_t[(1 - y_{t+1})^2]} \,. \tag{2.2}$$

In each factor model and at each date $t$, we estimate $y_{t+1}$ by regressing the constant 1 on the factors (without an intercept and possibly with penalization) using the previous 360 months of returns. Britten-Jones (1999) uses this type of regression (without penalization) to compute the mean-variance frontier. DKKM use ridge regression to mitigate overfitting and to allow even more factors than time periods in the regression. Denoting the vector of regression coefficients by $\hat{\beta}_t$ and the factor returns from $t$ to $t+1$ by $f_{t+1}$, we compute $\hat{y}_{t+1} = \hat{\beta}_t' f_{t+1}$.

Motivated by Hansen and Jagannathan (1997), we calculate the realization of $(\hat{y}_{t+1} - z_{t+1})^2$ each period as a measure of how far the estimated SDF is from the SDF and we compare mean values across models. The mean of $(\hat{y}_{t+1} - z_{t+1})^2$ is a measure of how accurately the factor model prices assets and also a measure of how close the factor model comes to spanning the mean-variance frontier. As a second measure of how well the factors span the frontier, we compute the theoretical Sharpe ratio of $\hat{y}_{t+1}$ each period and compare mean values across models.

# 3 Models

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.[1]

We use the same five characteristics to implement the DKKM method. DKKM use random Fourier features to create potentially a very large number of factors. We follow their recipe to form various sets of what we call DKKM factors, ranging from a six factor model to a model with 36,000 factors.

The DKKM method 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 $5 \times n$ matrix of standardized characteristics at date $t$, where $n$ is the number of assets. From $C_t$, we generate an $n_f \times n$ matrix of random characteristics as follows. Let $W$ denote a $\frac{n_f}{2} \times 5$ matrix whose entries are sampled from the standard normal distribution, and let $\gamma$ denote a length $\frac{n_f}{2}$ vector whose entries are sampled uniformly from $\{0.5, 0.6, \ldots, 1\}$. We use the same $W$ and $\gamma$ for all $t$. Set $A_t = \gamma \odot WC_t$, where $\gamma \odot W$ denotes element by element multiplication of $\gamma$ with each column of $W$. We compute an $n_f \times n$ matrix of random characteristics from the $\frac{n_f}{2} \times n$ matrix $A_t$ by taking sines and cosines of the elements of $A_t$ and stacking the sines and cosines as separate rows. We then rank standardize the rows of this matrix, replacing the raw characteristic values with percentiles and then subtracting 0.5 to get ranks between $-0.5$ and $+0.5$. Each row of this matrix can be interpreted as a long-short portfolio. The returns of the portfolios from $t$ to $t+1$ are the DKKM factor realizations $f_{t+1}$ from $t$ to $t+1$.

We use ridge regression to form the estimate $\hat{\beta}_t' f_{t+1}$. We vary the penalty parameter in the ridge regression to create multiple estimates. To mitigate the effect of randomness in the draws of the random Fourier features, we

[1] 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.

follow DKKM by generating 20 samples of $W$ and $\gamma$, and, for each value of $\alpha$, we average the 20 estimates $\hat{\beta}_t' f_{t+1}$ to produce our final estimate $\hat{y}_{t+1}$ for that value of $\alpha$.

The ridge regression is

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

where $\alpha$ is the penalty parameter. More penalization is needed when the number of factors $n_f$ 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 $\alpha$. Hence, we set $\alpha = \kappa n_f$ and vary $\kappa$. If adding more factors is more effective than simply replicating factors, then, for each value of $\kappa$, we should see performance improve as the number of factors 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). We explored ridge regression to form the estimates $\hat{\beta}_t' f_{t+1}$ for the FFC and FMR factors, but it always underperformed OLS, so in the next section we only report the OLS results for FFC and FMR.

# 4 Results

As discussed before, we compute $(\hat{y}_{t+1} - z_{t+1})^2$ and the theoretical conditional Sharpe ratios $\mathsf{E}_t[\hat{y}_{t+1}]/\text{stdev}_t(\hat{y}_{t+1})$. We do this for each of the 360 "out of sample" months in each of 500 simulated panels. We compute sample means in each panel and compare the sample means across panels.

### 4.1 Sharpe Ratios

The mean Sharpe ratio for the FFC model is 21.1% per month. The mean Sharpe ratio for the FMR model is slightly higher at 21.4% per month. Treating each panel mean as a single observation and conducting a paired $t$-test for the 500 panel means, the difference between the FMR and FMC Sharpe ratios is statistically significant, with a $p$-value of 0.003. We will use the higher Sharpe ratios of the FMR model as a benchmark for evaluating the DKKM model.

Figure 4.1 shows the mean Sharpe ratios for the DKKM model. With sufficient penalization, performance improves up to 360 factors. Table 1 reports $t$-statistics for DKKM Sharpe ratios versus FMR Sharpe ratios. We compute mean conditional Sharpe ratios in each panel and then compare models across panels. With sufficient penalization, and a sufficient number of factors, DKKM outperforms FMR and the outperformance is statistically significant. The $t$ statistics plateau at 360 factors.

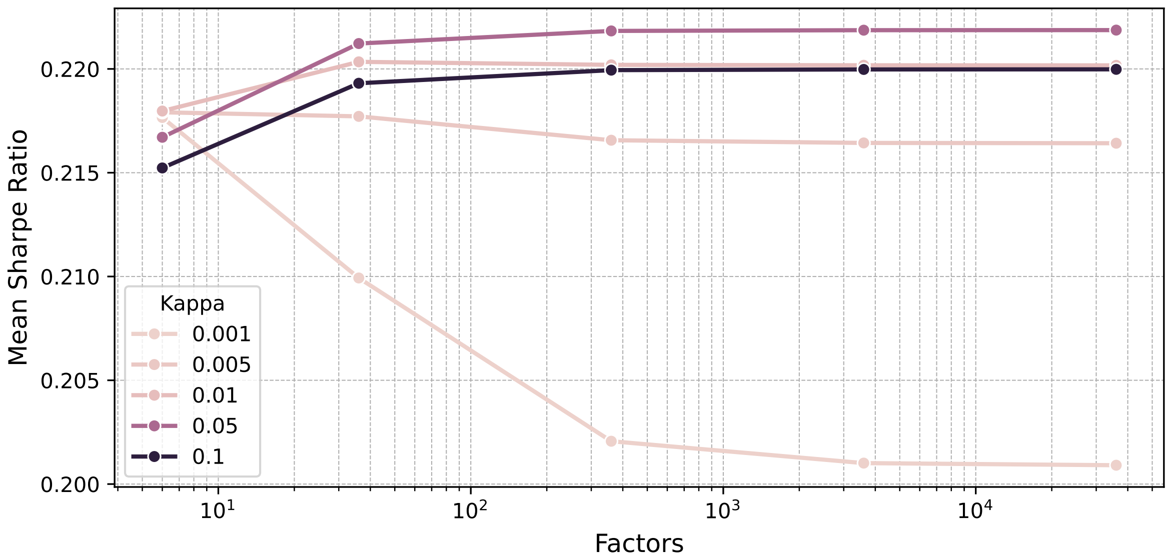


Figure 4.1: The mean (across months and panels) conditional Sharpe ratio is shown for various numbers of factors and various penalization parameters $\kappa$. Ridgeless regression underperforms and is omitted for reasons of scale.

| | Number of Factors | | | | |
|---|---|---|---|---|---|
| $\kappa$ | 6 | 36 | 360 | 3600 | 36000 |
| 0 | 4.88 | -32.33 | -163.09 | -153.82 | -153.51 |
| 0.001 | 5.00 | -5.51 | -14.99 | -16.19 | -16.31 |
| 0.005 | 5.25 | 5.20 | 3.52 | 3.34 | 3.32 |
| 0.01 | 5.24 | 8.99 | 8.78 | 8.76 | 8.76 |
| 0.05 | 3.30 | 9.85 | 10.82 | 10.90 | 10.90 |
| 0.1 | 1.41 | 6.89 | 7.83 | 7.90 | 7.91 |

Table 1: ***t*-statistics for sample Sharpe ratios of DKKM versus FMR.** 500 panels of 720 months are generated for the BGN model. Conditional Sharpe ratios are computed for each of the 360 out-of-sample months and the mean value is computed in each panel for each model. The table reports $t$-statistics for 500 panel means of the DKKM model compared to the FMR model. $\kappa = 0$ is ridgeless regression. A positive sign means that the DKKM model outperforms the FMR model.

## 4.2 Hansen-Jagannathan Distances

We also compare models using the realized values $(\hat{y}_{t+1} - z_{t+1})^2$. The mean values, across months and panels, for the FFC and FMR models are 0.0589 and 0.0603, respectively. The difference is statistically insignificant, but, because the FFC model slightly outperforms the FMR model on this dimension, we compare the DKKM model to the FFC model in what follows.

Figure 4.2 shows the means for the DKKM model. The means are estimates of the unconditional HJ distances. With sufficient penalization, performance improves up to 360 factors. Table 2 reports $t$-statistics for the 500 panel means of the DKKM models versus the FFC model. As with Sharpe ratios, DKKM outperforms on HJ distances for sufficent penalization and a sufficient number of factors. Also as with Sharpe ratios, the $t$ statistics plateau at 360 factors.

Figure 4.3 shows the distribution across panels of mean Sharpe ratios and mean values of $(\hat{y}_{t+1} - z_{t+1})^2$ of the DKKM model with $\kappa = 0.05$ and 360 factors and the FMR model. The DKKM model outperforms on both dimensions in a statistically significant fashion. An interesting property of the data that is illustrated in the figure is that the dispersion of values

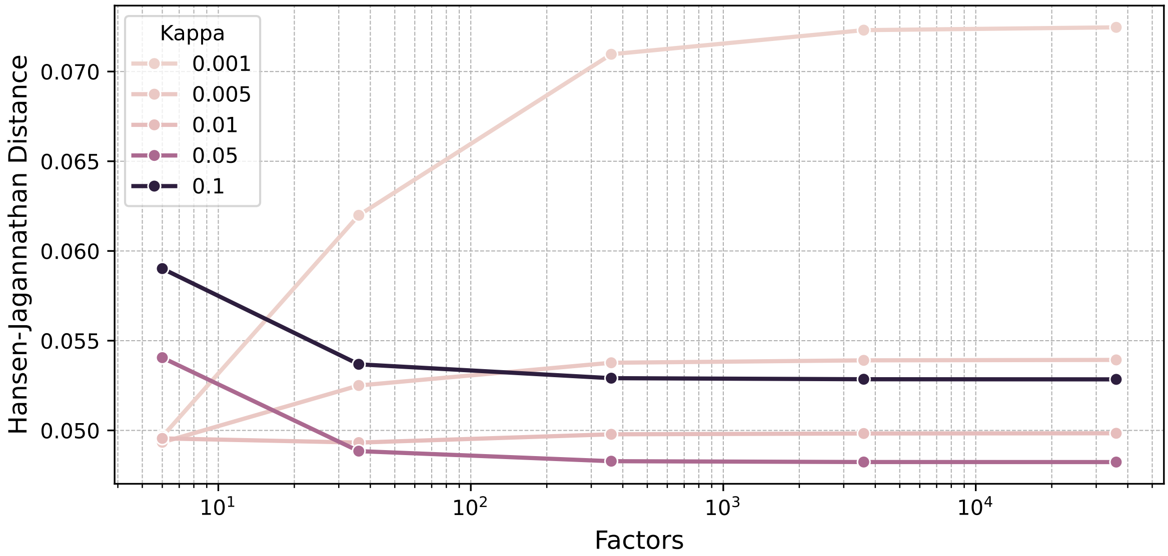


Figure 4.2: The mean (across months and panels) value of $(\hat{y}_{t+1} - z_{t+1})^2$ is shown for various numbers of factors and various penalization parameters $\kappa$. Ridgeless regression underperforms and is omitted for reasons of scale.

for the DKKM model is less than that of the FMR model. This is despite the fact that the DKKM model has 360 randomly generated factors. This highlights the fact that building a composite factor from a large number of factors can better allow the data to speak regarding what the single factor should be.

| $\kappa$ | Number of Factors | | | | |
|---|---|---|---|---|---|
| | 6 | 36 | 360 | 3600 | 36000 |
| 0 | -10.00 | 7.04 | 26.15 | 35.16 | 35.60 |
| 0.001 | -10.12 | 2.14 | 6.35 | 6.77 | 6.88 |
| 0.005 | -9.59 | -6.20 | -4.46 | -4.27 | -4.25 |
| 0.01 | -8.57 | -10.67 | -9.75 | -9.59 | -9.58 |
| 0.05 | -3.31 | -8.19 | -8.89 | -8.96 | -8.97 |
| 0.1 | 0.08 | -3.44 | -4.01 | -4.06 | -4.07 |

Table 2: ***t*-statistics for sample Hansen-Jagannathan distances of DKKM versus FFC.** 500 panels of 720 months are generated for the BGN model. The squared difference between $\hat{y}_{t+1}$ and $z_{t+1}$ is computed in each of the 360 out-of-sample months and the mean value is computed in each panel for each model. The table reports $t$-statistics for the 500 panel means of DKKM model compared to the FFC model. $\kappa = 0$ is ridgeless regression. A negative sign means that the DKKM model outperforms the FFC model.

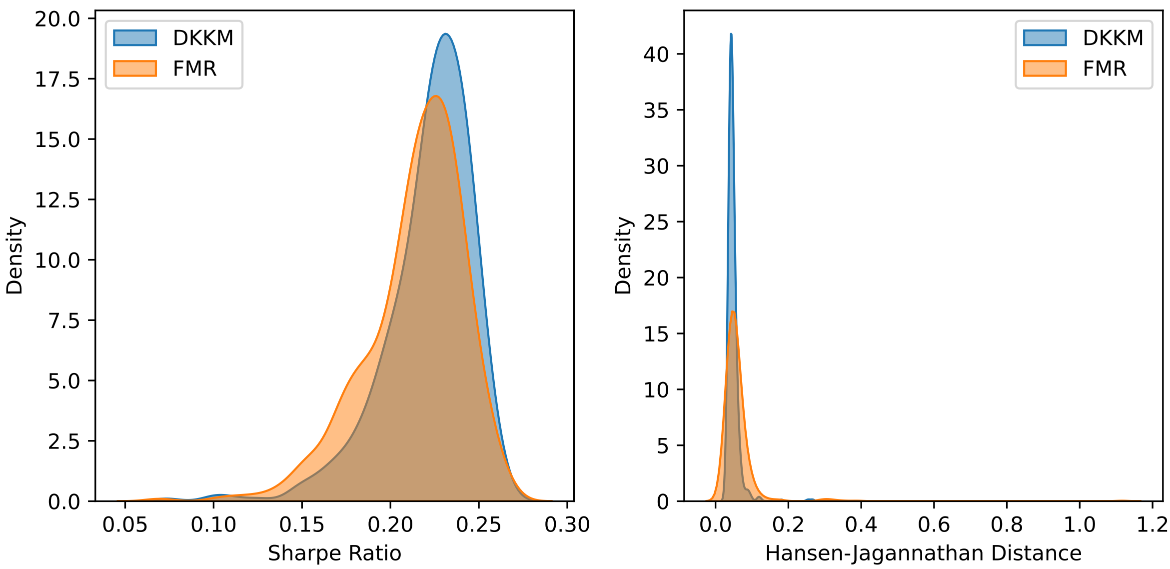


Figure 4.3: The distributions across panels of mean Sharpe ratios and mean values of $(\hat{y}_{t+1} - z_{t+1})^2$ are shown for the DKKM model with $\kappa = 0.05$ and 360 factors and the FMR model.

## 5 Conclusion

The dynamics of the BGN economy does not have a simple state-variable representation, but it is still a fairly simple economy with only two macro shocks each period. Despite the simplicity of the environment, the DKKM "complexity" method outperforms classical factor models. Subsequent research should analyze the DKKM method and other recent proposals for factor construction in more complex simulated economies.

## References


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

Brandt, Michael W., Pedro Santa-Clara, and Rossen Valkanov, 2009, Parametric portfolio policies: Exploiting characteristics in the cross section of equity returns, Review of Financial Studies 22, 3411–3447.

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

Didisheim, Antoine, Shikun Ke, Bryan Kelly, and Semyon Malamud, 2024, Complexity in factor pricing models, Research Paper 23-19, Swiss Finance Institute.

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

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.

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

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