Time-Varying Hedge Fund Risk Exposure and Skill: A Latent Factor Approach with Big Data

Research Proposal to "Big Data and High-Performance Computing for Financial Economics"

Zhaogang Song* Yinan Su†

April 5, 2019

*The Johns Hopkins Carey Business School, 100 International Drive, Baltimore, MD 21202. E-mail: zsong8@jhu.edu.
†The Johns Hopkins Carey Business School, 100 International Drive, Baltimore, MD 21202. E-mail: ys@jhu.edu.

1 Research Objective

Hedge funds are an important component of the financial system, with their total assets under management increasing from $39 billion in 1990 to more than $2.97 trillion as of the second quarter of 2015 according to the Hedge Fund Research. They have been constantly associated with booms and busts of the U.S. financial markets and economy, notably in the quant crisis of 1998, dot-com bubble era, and 2007-2008 global financial crisis. They also account for a significant fraction of the portfolios of important institutional investors such as endowments, pension funds, and insurance companies. Evaluating the performance of hedge funds is hence an important task not only for the purpose of asset allocation but also for regulation policies.

Evaluating the performance of hedge funds, however, is an challenging task at least. In particular, hedge fund portfolio holdings are highly confidential and rarely available so that methods based on portfolio holdings like those proposed for mutual funds are not applicable. In consequence, a benchmark model is usually needed. For example, a popular benchmark is the 8-factor model based on Fung and Hsieh (1997) and Fung and Hsieh (2001) (FH) that include three equity factors, two bond factors, and three option factors. Yet, the complexity of hedge fund strategies makes benchmarking their returns particularly challenging and raises the serious model mis-specification issues. For example, hedge funds usually take dynamic trading strategies, hold derivative securities, take short positions, and maintain highly leverage positions. Such features lead to several challenges. First, it is difficult to identify the complete set of benchmark factors or their observable proxies. For example, even though the FH model include three option portfolios, these are just for at-the-money options for bonds, currencies, and commodities, so other important asset classes such as credit products and out-of-the-money options that capture the "tail risk" are missing. Second, the dynamic trading strategies can lead to significant variations of risk exposure over time. Third, hedge funds actively manage their strategies, likely generating significant time-varying skills. The last two features raise substantial concerns on the constant-parameter assumptions in standard benchmark models.

2 Research Design

We propose a structured evaluation of hedge fund performance by harnessing the vast amount of data related to fund behavior with cutting-edge econometrics methods. Our econometric framework can handle the three key difficulties with the traditional benchmark models as mentioned above – unobservable risk factors, time-varying risk exposure, and time-varying skill - by adapting the IPCA framework developed in Kelly, Pruitt, and Su (2018) and Kelly, Pruitt, and Su (2019) with modifications tailored to hedge funds. In particular, we adopt the following general conditional factor model with *dynamic* risk exposures (β) and skill (α):

$$r_{i,t+1} = \alpha_{i,t} + \beta_{i,t} F_{t+1} + \epsilon_{i,t+1}$$

where α and β are time-varying conditional moments. The standard methods typically specify a model where α and β are fixed for each fund. However, as a hedge fund changes its size, investment style, or as the aggregate cycle changes over time, the fund's risk loadings and α should change accordingly. Following the spirit of IPCA, we map the conditioning information into the conditional risk loading and α:

$$\alpha_{i,t} = Z_{i,t}\Gamma_\alpha + \mu^\alpha_{i,t}$$
$$\beta_{i,t} = Z_{i,t}\Gamma_\beta + \mu^\beta_{i,t}$$

Here $Z_{i,t}$ represents the fund specific characteristics that are observable by time t, which can potentially be a high-dimensional object including 1) different observable features of a fund that have been proposed to indicate its performances, 2) aggregate series that affect fund risk exposure, and 3) the mixture of the two. Γ_α and Γ_β are the key parameters of the model. They specify the linear mapping from characteristics to α and β respectively. The mappings are allowed to have errors.

Given this specification, a fund's return process is

$$r_{i,t+1} = \underbrace{Z_{i,t}\Gamma_\alpha}_{\alpha_{i,t}} + \underbrace{Z_{i,t}\Gamma_\beta}_{\beta_{i,t}} F_{t+1} + \epsilon_{i,t+1}.$$

This model has several desirable features compared with previous methods. First, it allows for incorporating the vast amount of data into the identification of conditional risk-loading and conditional returns. Many existing methods of hedge fund evaluation rely on sorting hedge funds according to some fund-level (risk or skill) variables, which runs into problems when the number of variables increases. Our framework allows a "kitchen-sink" method to incorporate many variables simultaneously in a unified multivariate environment.

Second, the model is more parsimonious. One reason why the existing literature resorts to a benchmark model with several observable factors is that the number of parameters increases with the increasing sample size of individual funds if standard principal component analysis were used. In contrast, the key parameters Γ_α and Γ_β have fixed dimensions in our framework, which overcomes the problems of unbalanced panel - a known feature of hedge fund return samples.

Third, estimating Γ_α address not only whether hedge funds produce α, but also which fund characteristics are related to generating α in what direction and magnitude. If we construct $Z_{i,t}$ with macro time series, the estimation results can also speak to time-varying risk exposure of funds. Combining Γ_α and Γ_β estimates can distinguish how fund delivers performance - through merely levering up the exposure to systematic risks or from return on top of risk compensations.

To briefly discuss the resources needed, we plan to implement the methodology using a combination of R and Matlab. Large high-performance computing power is a must for implementing such a data-intensive econometric procedure. This is especially important given that our objective of hedge fund performance will focus on a large cross section of individual hedge funds, different from standard asset pricing tests based on portfolios.

References

Fung, W., and Hsieh, D. 2001. The Risk in Hedge Fund Strategies: Theory and Evidence from Trend Followers. *Review of Financial Studies* 14: 313–341.

Fung, W., and Hsieh, D. A. 1997. Empirical Characteristics of Dynamic Trading Strategies: The Case of Hedge Funds. *The Review of Financial Studies* 10 (06): 275–302.

Kelly, B., Pruitt, S., and Su, Y. 2018. Characteristics Are Covariances: A Unified Model of Risk and Return. *Journal of Financial Economics* forthcoming.

Kelly, B., Pruitt, S., and Su, Y. 2019. Instrustrumented Principal Component Analysis. *working paper.*