Machine-Learning the Skill of Mutual Fund Managers

Ron Kaniel, Zihan Lin, Markus Peleger, & Stijn Van Nieuwerburgh

Discussion at the 2023 WFA by
ROHIT ALLENA
University of Houston

June 28, 2023

Summary of the Paper

The paper is forthcoming in JFE. My discussion focuses on

- Analyzing in detail the main insights of the paper
- Posing more questions related to the paper's insights

Main Insights of the Paper:

1. Model abnormal returns $\tilde{R}_{it} = (R_{it} - \sum_k \hat{\beta}_{ik} f_{kt})$, rather than excess returns $(R_{it} - R_f)$ for better detecting the skill of mutual fund managers. Why?
   - Because abnormal returns are less noisier
2. Abnormal returns are modeled using Neural Networks
   - Predictors are fund char, fund-family char, char of stocks that funds hold
   - Fund momentum, fund flow $\implies \tilde{R}_{it}$, but stock-chars $\not\Longrightarrow \tilde{R}_{it}$
3. Long-short portfolios of funds with extreme abnormal returns yield high returns
   - Prediction-weighted long-short portfolios: Overweight funds in the extreme deciles with higher abnormal return forecasts

Discussion on Modeling Abnormal Returns

Why does modeling abnormal returns yield better inferences?

A GLS interpretation:

- Consider $r_{it} = g(z_{it-1}, \theta) + \epsilon_{it}$, where $\epsilon_{it} = \sum_k \beta_{ik} f_{kt} + \eta_i t$
- $\hat{\theta}_{OLS}$ delivers less precise forecasts than $\hat{\theta}_{GLS}$
- $\hat{\theta}_{GLS} \equiv$ implementing OLS on abnormal returns $r_{it} - \sum_k \beta_{ik} f_{kt}$
- Thus, targeting abnormal returns (rather than returns) is useful

Discussion on Modeling Abnormal Returns

Why does modeling abnormal returns yield better inferences?

A GLS interpretation:

- Consider $r_{it} = g(z_{it-1}, \theta) + \epsilon_{it}$, where $\epsilon_{it} = \sum_k \beta_{ik} f_{kt} + \eta_i t$
- $\hat{\theta}_{OLS}$ delivers less precise forecasts than $\hat{\theta}_{GLS}$
- $\hat{\theta}_{GLS} \equiv$ implementing OLS on abnormal returns $r_{it} - \sum_k \beta_{ik} f_{kt}$
- Thus, targeting abnormal returns (rather than returns) is useful

Related Questions:

- The paper sorts funds into deciles based on their abnormal return forecasts, and then it takes long (short) positions on the top (bottom) decile of funds
- If any, this long-short portfolio should deliver large **abnormal returns**
- But the paper documents that it delivers large **returns**. Why?
- Note that ret=abn ret+ factor risk premia ($\beta_{ik} f_{kt}$).
- So, long-short abr port need not deliver large returns
- Does the paper implicitly assumes factor means to be zero? If yes, wouldn't α estimates be biased?

Discussion on Prediction-Weighted Portfolios

The paper uses prediction-weighted long-short strategies (PWS)

- These portfolios are in the spirit of Allena (2020, 2023 RFS *R&R*) which introduced precision-weighted long-short strategies

What are prediction-weighted strategies? Are they universally applicable?

- PWS first sort funds into deciles based on abnormal return forecasts
- Take prediction-weighted long (short) positions on top (bottom) decile funds

Discussion on Prediction-Weighted Portfolios

The paper uses prediction-weighted long-short strategies (PWS)

- These portfolios are in the spirit of Allena (2020, 2023 RFS *R&R*) which introduced precision-weighted long-short strategies

What are prediction-weighted strategies? Are they universally applicable?

- PWS first sort funds into deciles based on abnormal return forecasts
- Take prediction-weighted long (short) positions on top (bottom) decile funds

Do prediction-weighted strategies always outperform? Need not always be

- return forecast $=$ E(ret)$+$ measurement error; forecast $\uparrow \implies E(r)\uparrow$ or *error* $\uparrow$
- Thus, by overweighting funds with high return forecast, one may be overweighting on the noise!
- Thus, prediction-weighted strategies may not always work

Discussion on Prediction-Weighted Portfolios

The paper uses prediction-weighted long-short strategies (PWS)

- These portfolios are in the spirit of Allena (2020, 2023 RFS *R&R*) which introduced precision-weighted long-short strategies

What are prediction-weighted strategies? Are they universally applicable?

- PWS first sort funds into deciles based on abnormal return forecasts
- Take prediction-weighted long (short) positions on top (bottom) decile funds

Do prediction-weighted strategies always outperform? Need not always be

- return forecast = E(ret)+ measurement error; forecast $\uparrow \implies E(r) \uparrow$ or $error \uparrow$
- Thus, by overweighting funds with high return forecast, one may be overweighting on the noise!
- Thus, prediction-weighted strategies may not always work

One way to track noise is to look at **forecast variances**. Put more weights on forecasts with low forecast variances (Allena (2020, 2023))

- Fortunately, for Neural Networks, forecasts and forecast variances are correlated. Thus, prediction-weighted strategies outperform
- But for linear models, prediction-weighted strategies underperform

Discussion on Empirical Findings

Stock chars do not impact abnormal fund returns

Question: Is this result driven by the fact that stock chars are aggregated linearly whereas returns are modeled non-linearly?

- Consider two stocks with $r_{it} = a + bc_{i,t-1}^2 + \epsilon_{it}, \ i = 1, 2$
- Consider a fund that assigns equal weights on both stocks
- Then the expected abnormal return of the fund $E(R_F) = a + b\frac{c_{1,t-1}^2 + c_{2,t-1}^2}{2}$
- However, when the chars are first linearly aggregated to the fund-level, as in the paper, abnormal returns will be modeled using

$$r_{i,t+1} = f\left(\frac{c_{1,t-1} + c_{2,t-1}}{2}\right) + \epsilon_{it}$$

- $f\left(\frac{c_{1,t-1}+c_{2,t-1}}{2}\right) \neq E(R_F)$, and it may result in biased inferences
- Thus, linear aggregation of characteristics may not be suitable for non-linear modeling of fund returns

Discussion on Empirical Findings

Stock chars do not impact abnormal fund returns

Question: Is this result driven by the fact that stock chars are aggregated linearly whereas returns are modeled non-linearly?

- Consider two stocks with $r_{it} = a + bc^2_{i,t-1} + \epsilon_{it}, \ i = 1, 2$
- Consider a fund that assigns equal weights on both stocks
- Then the expected abnormal return of the fund $E(R_F) = a + b\frac{c^2_{1,t-1}+c^2_{2,t-1}}{2}$
- However, when the chars are first linearly aggregated to the fund-level, as in the paper, abnormal returns will be modeled using

$$r_{i,t+1} = f\left(\frac{c_{1,t-1} + c_{2,t-1}}{2}\right) + \epsilon_{it}$$

- $f\left(\frac{c_{1,t-1}+c_{2,t-1}}{2}\right) \neq E(R_F)$, and it may result in biased inferences
- Thus, linear aggregation of characteristics may not be suitable for non-linear modeling of fund returns

Thought for future research: How to aggregate stock-level chars to the fund-level for modeling fund returns non-linearly?

Conclusion

Great paper! Main takeaways are

- Abnormal fund returns are predictable and are persistent
- Modeling abnormal returns is important
- Allowing for a non-linear relation between abnormal fund returns and fund chars is important