# Discussion of *Arbitrage Portfolios*

## Paper by Soohun Kim, Robert Korajczyk and Andreas Neuhierl

Guillaume Roussellet[1]

[1] *McGill University*

**AFFI conference**

**December 2019**

- Start with a standard linear asset pricing model:

$$R_{i,t} = \alpha_i + \boldsymbol{\beta}_i' \mathbf{f}_t + \varepsilon_{i,t}\,. \tag{1}$$

- $\alpha_i$ and $\beta_i$ are random variable/vector distributed according to some unknown cross-sectional distribution.
- After "purging" from systematic exposure, can we detect and exploit the cross-sectional differences in $\alpha$?

**Question**

→ Can we use observable characteristics of stocks to exploit potential arbitrages more efficiently?

# Novelty: Projected PCA (Fan et. al. 2016)

- Formulate the model as:

$$R_{i,t} = \underbrace{\boldsymbol{\theta}' \boldsymbol{X}_i + \Gamma_{\alpha_i}}_{=\alpha_i} + \underbrace{\left(\boldsymbol{\Theta}' \boldsymbol{X}_i + \boldsymbol{\Gamma}_{\beta_i}\right)'}_{=\beta_i'} \mathbf{f}_t + \varepsilon_{i,t} \,. \tag{2}$$

- Key conditions are $\mathbb{E}(\Gamma_{\alpha_i} | \boldsymbol{X}) = 0$ and $\mathbb{E}(\boldsymbol{\Gamma}_{\beta_i} | \boldsymbol{X}) = \mathbf{0}$, so the $\Gamma$'s are *residuals* from cross-sectional regressions.
- Imposing identification conditions, factors and cross-sectional dispersion functions can be estimated with asymptotics requiring only $N \to +\infty$, not $T$.

## General Comments (1/2)

### Question #1

$\rightarrow$ What is the exact relationship with the PPCA of Fan et. al. (2016)?

- The blue term in Equation 2 is a natural extension from the 2016 paper.
  - Is it trivial?
  - Are all asymptotic properties preserved?
  - Do the authors need the same assumptions?
- If $\Theta = \mathbf{0}$ for instance, it seems to me that I can estimate the model with regression then PCA (Frisch-Waugh)
- Theoretical justification for $\alpha(\boldsymbol{X}_i)$? (endogeneity of $\beta$ seems more natural than for $\alpha$).
- One justification of PPCA is that $\Gamma_{\beta_i}$ can be more precisely estimated conditioning by $\boldsymbol{X}_i$ if the true model features $\boldsymbol{\Theta} \neq \mathbf{0}$. Same here?

# General Comments (1/2)

## Question #1

$\rightarrow$ What is the exact relationship with the PPCA of Fan et. al. (2016)?

- The blue term in Equation 2 is a natural extension from the 2016 paper.
  - Is it trivial?
  - Are all asymptotic properties preserved?
  - Do the authors need the same assumptions?
- If $\Theta = \mathbf{0}$ for instance, it seems to me that I can estimate the model with regression then PCA (Frisch-Waugh)
- Theoretical justification for $\alpha(\boldsymbol{X}_i)$? (endogeneity of $\beta$ seems more natural than for $\alpha$).
- One justification of PPCA is that $\Gamma_{\beta_i}$ can be more precisely estimated conditioning by $\boldsymbol{X}_i$ if the true model features $\Theta \neq \mathbf{0}$. Same here?

## General Comments (1/2)

### Question #1

→ What is the exact relationship with the PPCA of Fan et. al. (2016)?

- The blue term in Equation 2 is a natural extension from the 2016 paper.
  - Is it trivial?
  - Are all asymptotic properties preserved?
  - Do the authors need the same assumptions?
- If $\Theta = \mathbf{0}$ for instance, it seems to me that I can estimate the model with regression then PCA (Frisch-Waugh)
- Theoretical justification for $\alpha(\boldsymbol{X}_i)$? (endogeneity of $\beta$ seems more natural than for $\alpha$).
- One justification of PPCA is that $\Gamma_{\beta_i}$ can be more precisely estimated conditioning by $\boldsymbol{X}_i$ if the true model features $\Theta \neq \mathbf{0}$. Same here?

# General Comments (1/2)

## Question #1

$\rightarrow$ What is the exact relationship with the PPCA of Fan et. al. (2016)?

- The blue term in Equation 2 is a natural extension from the 2016 paper.
  - Is it trivial?
  - Are all asymptotic properties preserved?
  - Do the authors need the same assumptions?
- If $\Theta = \mathbf{0}$ for instance, it seems to me that I can estimate the model with regression then PCA (Frisch-Waugh)
- Theoretical justification for $\alpha(\boldsymbol{X}_i)$? (endogeneity of $\beta$ seems more natural than for $\alpha$).
- One justification of PPCA is that $\boldsymbol{\Gamma}_{\beta_i}$ can be more precisely estimated conditioning by $\boldsymbol{X}_i$ if the true model features $\boldsymbol{\Theta} \neq \mathbf{0}$. Same here?

## General Comments (2/2)

### Asymptotics and time dependency

→ $T$ is not that small and $\boldsymbol{X}_i$ is more like $\boldsymbol{X}_{i,t}$. How can I understand the inference conducted?

- "*Our procedure will tend to perform better in situations where characteristics are relatively stable over our short estimation period*"
  - How can I interpret rolling-window estimation then? Stability over rolling intervals? Same values of factors $\hat{f}_t$ over rolling intervals?
  - Is this the case empirically? $\boldsymbol{X}_{i,t}$ stable over a year?
  - If you assume time variation of relationships, wouldn't it make sense to model it?

$$R_{i,t} = \underbrace{\boldsymbol{\theta}_{t-1}{}' \boldsymbol{X}_{i,t-1} + \Gamma_{\alpha_i}}_{=\alpha_{i,t-1}} + \underbrace{\left(\boldsymbol{\Theta}'_{t-1}\boldsymbol{X}_{i,t-1} + \boldsymbol{\Gamma}_{\beta_i}\right)'}_{=\beta'_{i,t-1}} \mathbf{f}_t + \varepsilon_{i,t}\,. \qquad (3)$$

- Raises the question of stationarity, over estimation sample.
- No reference to Gagliardini, Ossola, Scaillet (2015, 2019)?

# General Comments (2/2) C'td

- Finite $T = 624$, large $N = 2,458$, $L = 61$ characteristics,
- **Extremely large** number of estimated parameters overall,
- Rolling window estimation (time correlation),
- $\boldsymbol{\theta}$ estimated through noisy 2-step OLS,

**Comment**

I need either asymptotic theory or a bootstrap procedure for inference.

- Multiple testing problem for $\boldsymbol{\theta}$ and $\boldsymbol{\Theta}$, and significance of arbitrage over benchmarks.