THE OHIO STATE UNIVERSITY
FISHER COLLEGE OF BUSINESS

The Factor Model Failure Puzzle

Fahiz Baba-Yara, Brian Boyer and Carter Davis

Discussant: **Andrei S. Gonçalves**

2024 MFA

Outline

This Paper in a Nutshell

- Factor Models: $\mathbb{E}[r] = \beta \cdot \mathbb{E}[r^*]$, with $r^* =$ max SR portfolio
- First factor model was the CAPM ($r^* = r_m$)
- Then we started to add other factors:
- This paper: the "Factor Model Failure Puzzle"

This Paper in a Nutshell

- Factor Models: $\mathbb{E}[r] = \beta \cdot \mathbb{E}[r^*]$, with $r^* =$ max SR portfolio
- First factor model was the CAPM ($r^* = r_m$)
- Then we started to add other factors:
- This paper: the "Factor Model Failure Puzzle"

This Paper in a Nutshell

- Factor Models: $\mathbb{E}[r] = \beta \cdot \mathbb{E}[r^*]$, with $r^* =$ max SR portfolio
- First factor model was the CAPM ($r^* = r_m$)
- Then we started to add other factors:
- This paper: the "Factor Model Failure Puzzle"

This Paper in a Nutshell

- Factor Models: $\mathbb{E}[r] = \beta \cdot \mathbb{E}[r^*]$, with $r^* =$ max SR portfolio
- First factor model was the CAPM ($r^* = r_m$)
- Then we started to add other factors:
  - $r^* = b_m \cdot r_m + b_{\text{SMB}} \cdot r_{\text{SMB}} + b_{\text{HML}} \cdot r_{\text{HML}}$
  - $r^* = b_m \cdot r_m + b_{\text{SMB}} \cdot r_{\text{SMB}} + b_{\text{HML}} \cdot r_{\text{HML}} + b_{\text{CMA}} \cdot r_{\text{CMA}} + b_{\text{RMW}} \cdot r_{\text{RMW}}$
  - $r^* = b_m \cdot r_m + b_{\text{Size}} \cdot r_{\text{Size}} + b_{\text{I/A}} \cdot r_{\text{I/A}} + b_{\text{ROE}} \cdot r_{\text{ROE}}$

    $\vdots$
  - r^* built using Machine Learning techniques
- This paper: the "Factor Model Failure Puzzle"

This Paper in a Nutshell

- Factor Models: $\mathbb{E}[r] = \beta \cdot \mathbb{E}[r^*]$, with $r^* =$ max SR portfolio
- First factor model was the CAPM ($r^* = r_m$)
- Then we started to add other factors:
    - $r^* = b_m \cdot r_m + b_{\text{SMB}} \cdot r_{\text{SMB}} + b_{\text{HML}} \cdot r_{\text{HML}}$
    - $r^* = b_m \cdot r_m + b_{\text{SMB}} \cdot r_{\text{SMB}} + b_{\text{HML}} \cdot r_{\text{HML}} + b_{\text{CMA}} \cdot r_{\text{CMA}} + b_{\text{RMW}} \cdot r_{\text{RMW}}$
    - $r^* = b_m \cdot r_m + b_{\text{Size}} \cdot r_{\text{Size}} + b_{\text{I/A}} \cdot r_{\text{I/A}} + b_{\text{ROE}} \cdot r_{\text{ROE}}$

      $\vdots$
    - r^* built using Machine Learning techniques
- This paper: the "Factor Model Failure Puzzle"

This Paper in a Nutshell

- Factor Models: $\mathbb{E}[r] = \beta \cdot \mathbb{E}[r^*]$, with r^* = max SR portfolio
- First factor model was the CAPM ($r^* = r_m$)
- Then we started to add other factors:
  - $r^* = b_m \cdot r_m + b_{\mathrm{SMB}} \cdot r_{\mathrm{SMB}} + b_{\mathrm{HML}} \cdot r_{\mathrm{HML}}$
  - $r^* = b_m \cdot r_m + b_{\mathrm{SMB}} \cdot r_{\mathrm{SMB}} + b_{\mathrm{HML}} \cdot r_{\mathrm{HML}} + b_{\mathrm{CMA}} \cdot r_{\mathrm{CMA}} + b_{\mathrm{RMW}} \cdot r_{\mathrm{RMW}}$
  - $r^* = b_m \cdot r_m + b_{\mathrm{Size}} \cdot r_{\mathrm{Size}} + b_{\mathrm{I/A}} \cdot r_{\mathrm{I/A}} + b_{\mathrm{ROE}} \cdot r_{\mathrm{ROE}}$

    $\vdots$
  - r^* built using Machine Learning techniques
- This paper: the "Factor Model Failure Puzzle"

This Paper in a Nutshell

- Factor Models: $\mathbb{E}[r] = \beta \cdot \mathbb{E}[r^*]$, with $r^* =$ max SR portfolio
- First factor model was the CAPM ($r^* = r_m$)
- Then we started to add other factors:
  - $r^* = b_m \cdot r_m + b_{\mathrm{SMB}} \cdot r_{\mathrm{SMB}} + b_{\mathrm{HML}} \cdot r_{\mathrm{HML}}$
  - $r^* = b_m \cdot r_m + b_{\mathrm{SMB}} \cdot r_{\mathrm{SMB}} + b_{\mathrm{HML}} \cdot r_{\mathrm{HML}} + b_{\mathrm{CMA}} \cdot r_{\mathrm{CMA}} + b_{\mathrm{RMW}} \cdot r_{\mathrm{RMW}}$
  - $r^* = b_m \cdot r_m + b_{\mathrm{Size}} \cdot r_{\mathrm{Size}} + b_{\mathrm{I/A}} \cdot r_{\mathrm{I/A}} + b_{\mathrm{ROE}} \cdot r_{\mathrm{ROE}}$

    $\vdots$
  - r^* built using Machine Learning techniques
- This paper: the "Factor Model Failure Puzzle"

This Paper in a Nutshell

- Factor Models: $\mathbb{E}[r] = \beta \cdot \mathbb{E}[r^*]$, with $r^* =$ max SR portfolio
- First factor model was the CAPM ($r^* = r_m$)
- Then we started to add other factors:
  - $r^* = b_m \cdot r_m + b_{\text{SMB}} \cdot r_{\text{SMB}} + b_{\text{HML}} \cdot r_{\text{HML}}$
  - $r^* = b_m \cdot r_m + b_{\text{SMB}} \cdot r_{\text{SMB}} + b_{\text{HML}} \cdot r_{\text{HML}} + b_{\text{CMA}} \cdot r_{\text{CMA}} + b_{\text{RMW}} \cdot r_{\text{RMW}}$
  - $r^* = b_m \cdot r_m + b_{\text{Size}} \cdot r_{\text{Size}} + b_{\text{I/A}} \cdot r_{\text{I/A}} + b_{\text{ROE}} \cdot r_{\text{ROE}}$

    $\vdots$
  - r^* built using Machine Learning techniques
- This paper: the "Factor Model Failure Puzzle"

This Paper in a Nutshell

- Factor Models: $\mathbb{E}[r] = \beta \cdot \mathbb{E}[r^*]$, with r^* = max SR portfolio
- First factor model was the CAPM ($r^* = r_m$)
- Then we started to add other factors:
  - $r^* = b_m \cdot r_m + b_{\text{SMB}} \cdot r_{\text{SMB}} + b_{\text{HML}} \cdot r_{\text{HML}}$
  - $r^* = b_m \cdot r_m + b_{\text{SMB}} \cdot r_{\text{SMB}} + b_{\text{HML}} \cdot r_{\text{HML}} + b_{\text{CMA}} \cdot r_{\text{CMA}} + b_{\text{RMW}} \cdot r_{\text{RMW}}$
  - $r^* = b_m \cdot r_m + b_{\text{Size}} \cdot r_{\text{Size}} + b_{\text{I/A}} \cdot r_{\text{I/A}} + b_{\text{ROE}} \cdot r_{\text{ROE}}$

    $\vdots$
  - r^* built using Machine Learning techniques
- This paper: the "Factor Model Failure Puzzle"
  - Expectation: $\uparrow T \;\Rightarrow\;$ better r^* estimates $\;\Rightarrow\;$ r^* converge
  - Finding: no r^* prices well the r^* from other models OOS
  - Explanation: large T but also many $\mathbb{E}[r]$ predictors

This Paper in a Nutshell

- Factor Models: $\mathbb{E}[r] = \beta \cdot \mathbb{E}[r^*]$, with $r^* =$ max SR portfolio
- First factor model was the CAPM ($r^* = r_m$)
- Then we started to add other factors:
  - $r^* = b_m \cdot r_m + b_{\text{SMB}} \cdot r_{\text{SMB}} + b_{\text{HML}} \cdot r_{\text{HML}}$
  - $r^* = b_m \cdot r_m + b_{\text{SMB}} \cdot r_{\text{SMB}} + b_{\text{HML}} \cdot r_{\text{HML}} + b_{\text{CMA}} \cdot r_{\text{CMA}} + b_{\text{RMW}} \cdot r_{\text{RMW}}$
  - $r^* = b_m \cdot r_m + b_{\text{Size}} \cdot r_{\text{Size}} + b_{\text{I/A}} \cdot r_{\text{I/A}} + b_{\text{ROE}} \cdot r_{\text{ROE}}$

    $\vdots$
  - r^* built using Machine Learning techniques
- This paper: the "Factor Model Failure Puzzle"
  - Expectation: $\uparrow T \Rightarrow$ better r^* estimates $\Rightarrow$ r^* converge
  - Finding: no r^* prices well the r^* from other models OOS
  - Explanation: large T but also many $\mathbb{E}[r]$ predictors

This Paper in a Nutshell

- Factor Models: $\mathbb{E}[r] = \beta \cdot \mathbb{E}[r^*]$, with r^* = max SR portfolio
- First factor model was the CAPM ($r^* = r_m$)
- Then we started to add other factors:
  - $r^* = b_m \cdot r_m + b_{\mathrm{SMB}} \cdot r_{\mathrm{SMB}} + b_{\mathrm{HML}} \cdot r_{\mathrm{HML}}$
  - $r^* = b_m \cdot r_m + b_{\mathrm{SMB}} \cdot r_{\mathrm{SMB}} + b_{\mathrm{HML}} \cdot r_{\mathrm{HML}} + b_{\mathrm{CMA}} \cdot r_{\mathrm{CMA}} + b_{\mathrm{RMW}} \cdot r_{\mathrm{RMW}}$
  - $r^* = b_m \cdot r_m + b_{\mathrm{Size}} \cdot r_{\mathrm{Size}} + b_{\mathrm{I/A}} \cdot r_{\mathrm{I/A}} + b_{\mathrm{ROE}} \cdot r_{\mathrm{ROE}}$

    $\vdots$
  - r^* built using Machine Learning techniques
- This paper: the "Factor Model Failure Puzzle"
  - Expectation: $\uparrow T \;\Rightarrow\;$ better r^* estimates $\;\Rightarrow\;$ r^* converge
  - Finding: no r^* prices well the r^* from other models OOS
  - Explanation: large T but also many $\mathbb{E}[r]$ predictors

This Paper in a Nutshell

- Factor Models: $\mathbb{E}[r] = \beta \cdot \mathbb{E}[r^*]$, with r^* = max SR portfolio
- First factor model was the CAPM ($r^* = r_m$)
- Then we started to add other factors:
  - $r^* = b_m \cdot r_m + b_{\text{SMB}} \cdot r_{\text{SMB}} + b_{\text{HML}} \cdot r_{\text{HML}}$
  - $r^* = b_m \cdot r_m + b_{\text{SMB}} \cdot r_{\text{SMB}} + b_{\text{HML}} \cdot r_{\text{HML}} + b_{\text{CMA}} \cdot r_{\text{CMA}} + b_{\text{RMW}} \cdot r_{\text{RMW}}$
  - $r^* = b_m \cdot r_m + b_{\text{Size}} \cdot r_{\text{Size}} + b_{\text{I/A}} \cdot r_{\text{I/A}} + b_{\text{ROE}} \cdot r_{\text{ROE}}$

    $\vdots$
  - r^* built using Machine Learning techniques
- This paper: the "Factor Model Failure Puzzle"
  - Expectation: $\uparrow T \Rightarrow$ better r^* estimates $\Rightarrow$ r^* converge
  - Finding: no r^* prices well the r^* from other models OOS
  - Explanation: large T but also many $\mathbb{E}[r]$ predictors

Models for r^* Estimation

Table 1. Zoo of Asset Pricing Models

Initialism	Factor Method $f(Z_t)$	MVE Collapse Method b
BPZ_F	Bryzgalova et al. (2020) forest	Bryzgalova et al. (2020)
BPZ_L	linear characteristic-weighted portfolios	Bryzgalova et al. (2020)
BSV	linear characteristic-weighted portfolios	Brandt et al. (2009)
DGU	linear characteristic-weighted portfolios	DeMiguel et al. (2007)
DMRS	Daniel, Mota, Rottke, and Santos (2020)	Bryzgalova et al. (2020)
D_F	Davis (2021) random forest	$b = 1$
D_{NN}	Davis (2021) neural network	$b = 1$
FF3	Fama and French (1993)	Bryzgalova et al. (2020)
FF6	Fama and French (2015) with Carhart (1997) momentum	Bryzgalova et al. (2020)
GKX	Gu et al. (2021)	Bryzgalova et al. (2020)
HXZ	Hou et al. (2014)	Bryzgalova et al. (2020)
KNS	linear characteristic-weighted portfolios	Kozak et al. (2020)
KPS	Kelly et al. (2019)	Bryzgalova et al. (2020)
SL	Sharpe (1964b) and Lintner (1965) CAPM	$b = 1$
SY	Stambaugh and Yuan (2016)	Bryzgalova et al. (2020)

Correlations Between r^* from Different Models

Table 3. Dimson Adjusted Factor-MVE Portfolio Correlations

	(1) SL	(2) FF_3	(3) GKX	(4) BPZ_F	(5) DGU	(6) FF_6	(7) D_F	(8) SY	(9) DMRS	(10) HXZ	(11) D_{NN}	(12) BPZ_L	(13) KNS	(14) KPS	(15) BSV
SL	1.														
FF3	0.764	1.													
GKX	0.758	0.71	1.												
BPZ_F	0.219	0.222	0.163	1.											
DGU	0.439	0.734	0.446	0.593	1.										
FF6	-0.023	0.492	0.153	-0.02	-0.052	1.									
D_F	1.069	0.932	0.657	0.391	0.346	0.017	1.								
SY	-0.018	0.34	0.165	-0.021	-0.041	0.635	0.021	1.							
DMRS	0.186	0.438	0.21	0.054	0.043	0.553	0.21	0.293	1.						
HXZ	-0.042	0.227	0.115	-0.089	-0.126	0.648	-0.05	0.745	0.38	1.					
D_{NN}	0.457	0.741	0.463	0.67	0.968	0.011	0.951	-0.127	0.241	-0.153	1.				
BPZ_L	0.515	0.861	0.483	0.552	0.896	0.126	0.871	0.098	0.226	-0.007	0.814	1.			
KNS	0.51	0.847	0.481	0.53	0.857	0.144	0.852	0.118	0.243	0.017	0.798	0.978	1.		
KPS	0.468	0.426	0.338	0.425	0.435	-0.042	0.618	-0.103	0.083	-0.054	0.561	0.551	0.572	1.	
BSV	0.403	0.589	0.365	0.39	0.564	0.159	0.624	0.232	0.219	0.154	0.613	0.749	0.796	0.626	1.

Alpha from $r^*_{Row} = \alpha + \beta \cdot r^*_{Column} + \epsilon$

Table 5. Out-of-Sample Unconditional Dimson Alphas

	α^*_{ij}														
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
	Pricing Portfolio j														
	SL	FF_3	GKX	BPZ_F	DGU	FF_6	D_F	SY	DMRS	HXZ	D_{NN}	BPZ_L	KNS	KPS	BSV
Test Asset i															
SL		0.009	0.011	0.028**	0.043***	0.028**	-0.010	0.022*	0.024*	0.018	0.055***	0.042***	0.040***	0.013	0.020
FF_3	0.017*		0.017	0.034**	0.036***	0.009	-0.008	0.009	0.018	0.010	0.033**	0.006	0.001	0.021	-0.009
GKX	0.027*	0.022		0.042**	0.045***	0.034*	-0.001	0.022	0.033*	0.023	0.037**	0.020	0.013	0.010	-0.007
BPZ_F	0.056***	0.054***	0.055***		0.047***	0.065***	0.037***	0.065***	0.065***	0.068***	0.025***	0.023***	0.021**	0.011	0.018*
DGU	0.052***	0.036**	0.043**	0.029		0.079***	-0.009	0.081***	0.064***	0.098***	-0.070***	-0.089***	-0.100***	-0.082***	-0.112***
FF_6	0.088***	0.067***	0.080***	0.089***	0.091***		0.086***	0.017	0.027**	-0.007	0.082***	0.063***	0.060***	0.111***	0.057***
D_F	0.053***	0.048***	0.053***	0.062***	0.064***	0.085***		0.079***	0.074***	0.085***	0.031**	0.017	0.009	-0.018	-0.016
SY	0.102***	0.087***	0.093***	0.102***	0.104***	0.045***	0.099***		0.067***	0.006	0.106***	0.083***	0.080***	0.130***	0.068***
DMRS	0.105***	0.093***	0.100***	0.107***	0.108***	0.062***	0.092***	0.081***		0.061***	0.086***	0.078***	0.073***	0.091***	0.064***
HXZ	0.134***	0.124***	0.127***	0.139***	0.141***	0.076***	0.137***	0.058***	0.091***		0.145***	0.129***	0.127***	0.161***	0.116***
D_{NN}	0.162***	0.146***	0.153***	0.135***	0.113***	0.176***	0.094***	0.190***	0.150***	0.197***		0.001	-0.016*	-0.038**	-0.057***
BPZ_L	0.159***	0.139***	0.150***	0.140***	0.116***	0.164***	0.099***	0.165***	0.150***	0.176***	0.031***		-0.016***	0.004	-0.055***
KNS	0.173***	0.154***	0.164***	0.155***	0.132***	0.176***	0.115***	0.177***	0.162***	0.187***	0.048***	0.017***		0.018	-0.046***
KPS	0.218***	0.215***	0.215***	0.205***	0.203***	0.236***	0.178***	0.243***	0.223***	0.239***	0.133***	0.136***	0.124***		0.079***
BSV	0.236***	0.224***	0.230***	0.224***	0.211***	0.234***	0.194***	0.225***	0.224***	0.228***	0.140***	0.117***	0.098***	0.103***	
Span	0	2	2	1	0	1	4	3	1	5	0	4	3	7	4

Alpha from $r^*_{Row} = \alpha + \beta \cdot r^*_{Column} + \epsilon$

Table 5. Out-of-Sample Unconditional Dimson Alphas

	α^*_{ij}														
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8) *Pricing Portfolio j*	(9)	(10)	(11)	(12)	(13)	(14)	(15)
	SL	FF_3	GKX	BPZ_F	DGU	FF_6	D_F	SY	DMRS	HXZ	D_{NN}	BPZ_L	KNS	KPS	BSV
Test Asset i															
SL		0.009	0.011	0.028**	0.043***	0.028**	-0.010	0.022*	0.024*	0.018	0.055***	0.042***	0.040***	0.013	0.020
FF_3	0.017*		0.017	0.034**	0.036***	0.009	-0.008	0.009	0.018	0.010	0.033**	0.006	0.001	0.021	-0.009
GKX	0.027*	0.022		0.042**	0.045***	0.034*	-0.001	0.022	0.033*	0.023	0.037**	0.020	0.013	0.010	-0.007
BPZ_F	0.056***	0.054***	0.055***		0.047***	0.065***	0.037***	0.065***	0.065***	0.068***	0.025***	0.023***	0.021**	0.011	0.018*
DGU	0.052***	0.036**	0.043**	0.029		0.079***	-0.009	0.081***	0.064***	0.098***	-0.070***	-0.089***	-0.100***	-0.082***	-0.112***
FF_6	0.088***	0.067***	0.080***	0.089***	0.091***		0.086***	0.017	0.027**	-0.007	0.082***	0.063***	0.060***	0.111***	0.057***
D_F	0.053***	0.048***	0.053***	0.062***	0.064***	0.085***		0.079***	0.074***	0.085***	0.031**	0.017	0.009	-0.018	-0.016
SY	0.102***	0.087***	0.093***	0.102***	0.104***	0.045***	0.099***		0.067***	0.006	0.106***	0.083***	0.080***	0.130***	0.068***
DMRS	0.105***	0.093***	0.100***	0.107***	0.108***	0.062***	0.092***	0.081***		0.061***	0.086***	0.078***	0.073***	0.091***	0.064***
HXZ	0.134***	0.124***	0.127***	0.139***	0.141***	0.076***	0.137***	0.058***	0.091***		0.145***	0.129***	0.127***	0.161***	0.116***
D_{NN}	0.162***	0.146***	0.153***	0.135***	0.113***	0.176***	0.094***	0.190***	0.150***	0.197***		0.001	-0.016*	-0.038**	-0.057***
BPZ_L	0.159***	0.139***	0.150***	0.140***	0.116***	0.164***	0.099***	0.165***	0.150***	0.176***	0.031***		-0.016***	0.004	-0.055***
KNS	0.173***	0.154***	0.164***	0.155***	0.132***	0.176***	0.115***	0.177***	0.162***	0.187***	0.048***	0.017***		0.018	-0.046***
KPS	0.218***	0.215***	0.215***	0.205***	0.203***	0.236***	0.178***	0.243***	0.223***	0.239***	0.133***	0.136***	0.124***		0.079***
BSV	0.236***	0.224***	0.230***	0.224***	0.211***	0.234***	0.194***	0.225***	0.224***	0.228***	0.140***	0.117***	0.098***	0.103***	
Span	0	2	2	1	0	1	4	3	1	5	0	4	3	7	4

Alpha from $r^*_{Row} = \alpha + \beta \cdot r^*_{Column} + \epsilon$

Table 8. Out-of-Sample Unconditional Dimson Alphas for Meta Portfolios

	(1) meta_{BSV}	(2) meta_{DGU}	(3) meta_{BPZ}
SL	0.004	-0.004	0.035**
FF_3	0.036**	-0.039***	0.037**
GKX	0.032*	-0.020	0.046**
BPZ_F	0.061***	0.020**	0.041***
DGU	0.079***	-0.047**	0.014
FF_6	0.098***	-0.024	0.048***
D_F	0.065***	-0.018	0.060***
SY	0.115***	0.006	0.105***
DMRS	0.092***	0.011	0.058***
HXZ	0.128***	0.053***	0.119***
D_NN	0.178***	-0.016	0.028
BPZ_L	0.181***	-0.014	0.052**
KNS	0.192***	-0.003	0.059**
KPS	0.185***	0.080***	0.030*
BSV	0.242***	0.072***	0.108***
Span	1	9	2

Alpha from $r^*_{Row} = \alpha + \beta \cdot r^*_{Column} + \epsilon$

Table 8. Out-of-Sample Unconditional Dimson Alphas for Meta Portfolios

	(1) meta_{BSV}	(2) meta_{DGU}	(3) meta_{BPZ}
SL	0.004	-0.004	0.035**
FF_3	0.036**	-0.039***	0.037**
GKX	0.032*	-0.020	0.046**
BPZ_F	0.061***	0.020**	0.041***
DGU	0.079***	-0.047**	0.014
FF_6	0.098***	-0.024	0.048***
D_F	0.065***	-0.018	0.060***
SY	0.115***	0.006	0.105***
DMRS	0.092***	0.011	0.058***
HXZ	0.128***	0.053***	0.119***
D_NN	0.178***	-0.016	0.028
BPZ_L	0.181***	-0.014	0.052**
KNS	0.192***	-0.003	0.059**
KPS	0.185***	0.080***	0.030*
BSV	0.242***	0.072***	0.108***
Span	1	9	2

Theory: Empirical Results Follow if Large $P = N_{Predictors}$

- Traditional World: $T \to \infty$ and P fixed
- "Big Data World": $T \to \infty$ and $P \to \infty$ (with $c = P/T$)

Theory: Empirical Results Follow if Large $P = N_{Predictors}$

- Traditional World: $T \to \infty$ and P fixed
- "Big Data World": $T \to \infty$ and $P \to \infty$ (with $c = P/T$)

Theory: Empirical Results Follow if Large $P = N_{Predictors}$

- Traditional World: $T \to \infty$ and P fixed
- "Big Data World": $T \to \infty$ and $P \to \infty$ (with $c = P/T$)
  - Observable predictors but unknown r^* parameters:
  - Observable r^* parameters but noisy predictors:

Theory: Empirical Results Follow if Large $P = N_{Predictors}$

- Traditional World: $T \to \infty$ and P fixed
- "Big Data World": $T \to \infty$ and $P \to \infty$ (with $c = P/T$)
    - Observable predictors but unknown r^* parameters:

$$\underset{P,T\to\infty}{Plim}\ \mathbb{V}ar[\hat{r}^*] = \mathbb{V}ar[r^*] + \frac{c}{\gamma^2}$$

    - Observable r^* parameters but noisy predictors:

Theory: Empirical Results Follow if Large $P = N_{Predictors}$

- Traditional World: $T \to \infty$ and P fixed

- "Big Data World": $T \to \infty$ and $P \to \infty$ (with $c = P/T$)

  - Observable predictors but unknown r^* parameters:

$$\underset{P,T\to\infty}{Plim}\ \mathbb{V}ar[\widehat{r^*}] = \mathbb{V}ar[r^*] + \frac{c}{\gamma^2}$$

  - Observable r^* parameters but noisy predictors:

Theory: Empirical Results Follow if Large $P = N_{Predictors}$

- Traditional World: $T \to \infty$ and P fixed
- "Big Data World": $T \to \infty$ and $P \to \infty$ (with $c = P/T$)
  - Observable predictors but unknown r^* parameters:

$$\underset{P,T\to\infty}{Plim} \; \mathbb{V}ar[\widehat{r^*}] = \mathbb{V}ar[r^*] + \frac{c}{\gamma^2}$$

  - Observable r^* parameters but noisy predictors:

$$\underset{P,T\to\infty}{Plim} \; \mathbb{V}ar[\widehat{r^*}] = \mathbb{V}ar[r^*] + N \cdot c \cdot \varphi_t$$

Theory: Empirical Results Follow if Large $P = N_{Predictors}$

- Traditional World: $T \to \infty$ and P fixed
- "Big Data World": $T \to \infty$ and $P \to \infty$ (with $c = P/T$)
  - Observable predictors but unknown r^* parameters:

$$\underset{P,T \to \infty}{Plim} \mathbb{V}ar[\widehat{r^*}] = \mathbb{V}ar[r^*] + \frac{c}{\gamma^2}$$

  - Observable r^* parameters but noisy predictors:

$$\underset{P,T \to \infty}{Plim} \mathbb{V}ar[\widehat{r^*}] = \mathbb{V}ar[r^*] + N \cdot c \cdot \varphi_t$$

Outline

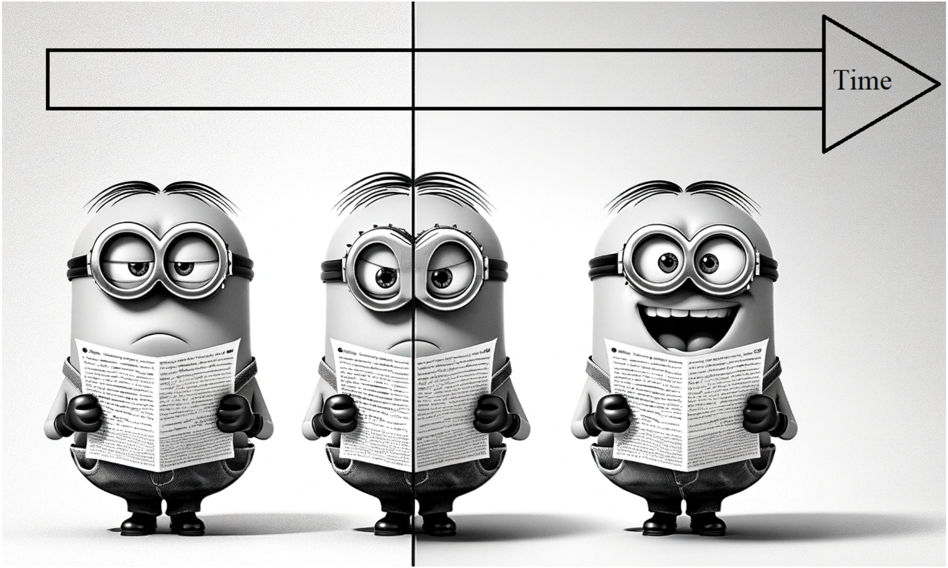
The Paper

My Comments

Final Remarks

My Reaction

My Reaction



1) The Theory Cannot Explain the Meta r^* Results

- There are two steps to build r^*

- Underlying argument for the theory in the paper:

- But this theory cannot explain the meta r^* results

1) The Theory Cannot Explain the Meta r^* Results

- There are two steps to build r^*

    1. Convert predictors into factors (f)

    2. Collapse factors into $r^* = b^{'} f$

- Underlying argument for the theory in the paper:

- But this theory cannot explain the meta r^* results

1) The Theory Cannot Explain the Meta r^* Results

- There are two steps to build r^*
    1. Convert predictors into factors (f)
    2. Collapse factors into $r^* = b^{'} f$
- Underlying argument for the theory in the paper:
- But this theory cannot explain the meta r^* results

1) The Theory Cannot Explain the Meta r^* Results

- There are two steps to build r^*
  1. Convert predictors into factors (f)
  2. Collapse factors into $r^* = b^{'} f$
- Underlying argument for the theory in the paper:
- But this theory cannot explain the meta r^* results

1) The Theory Cannot Explain the Meta r^* Results

- There are two steps to build r^*
  1. Convert predictors into factors (f)
  2. Collapse factors into $r^* = b^{'} f$
- Underlying argument for the theory in the paper:
  - Step 2 is fine
  - But in a "Big Data World" Step 1 fails even with large T
- But this theory cannot explain the meta r^* results

1) The Theory Cannot Explain the Meta r^* Results

- There are two steps to build r^*
  1. Convert predictors into factors (f)
  2. Collapse factors into $r^* = b^{'} f$
- Underlying argument for the theory in the paper:
  - Step 2 is fine
  - But in a "Big Data World" Step 1 fails even with large T
- But this theory cannot explain the meta r^* results

1) The Theory Cannot Explain the Meta r^* Results

- There are two steps to build r^*
  1. Convert predictors into factors (f)
  2. Collapse factors into $r^* = b' f$
- Underlying argument for the theory in the paper:
  - Step 2 is fine
  - But in a "Big Data World" Step 1 fails even with large T
- But this theory cannot explain the meta r^* results

1) The Theory Cannot Explain the Meta r^* Results

- There are two steps to build r^*
  1. Convert predictors into factors (f)
  2. Collapse factors into $r^* = b^{'} f$
- Underlying argument for the theory in the paper:
  - Step 2 is fine
  - But in a “Big Data World” Step 1 fails even with large T
- But this theory cannot explain the meta r^* results
  - The authors take the r^* from each factor model (15 on total)
  - They then construct $r^*_{meta} = b^{'} r^*$
  - And show that r^*_{meta} cannot price the underlying r^* OOS
  - This analysis does not require Step 1 (or predictors)
  - A “Big Data World” theory is silent about this meta r^* result

1) The Theory Cannot Explain the Meta r^* Results

- There are two steps to build r^*
  1. Convert predictors into factors (f)
  2. Collapse factors into $r^* = b^{'} f$
- Underlying argument for the theory in the paper:
  - Step 2 is fine
  - But in a "Big Data World" Step 1 fails even with large T
- But this theory cannot explain the meta r^* results
  - The authors take the r^* from each factor model (15 on total)
  - They then construct $r^*_{meta} = b^{'} r^*$
  - And show that r^*_{meta} cannot price the underlying r^* OOS
  - This analysis does not require Step 1 (or predictors)
  - A "Big Data World" theory is silent about this meta r^* result

1) The Theory Cannot Explain the Meta r^* Results

- There are two steps to build r^*
  1. Convert predictors into factors (f)
  2. Collapse factors into $r^* = b^{'} f$
- Underlying argument for the theory in the paper:
  - Step 2 is fine
  - But in a "Big Data World" Step 1 fails even with large T
- But this theory cannot explain the meta r^* results
  - The authors take the r^* from each factor model (15 on total)
  - They then construct $r^*_{meta} = b^{'} r^*$
  - And show that r^*_{meta} cannot price the underlying r^* OOS
  - This analysis does not require Step 1 (or predictors)
  - A "Big Data World" theory is silent about this meta r^* result

1) The Theory Cannot Explain the Meta r^* Results

- There are two steps to build r^*
  1. Convert predictors into factors (f)
  2. Collapse factors into $r^* = b^{'} f$
- Underlying argument for the theory in the paper:
  - Step 2 is fine
  - But in a "Big Data World" Step 1 fails even with large T
- But this theory cannot explain the meta r^* results
  - The authors take the r^* from each factor model (15 on total)
  - They then construct $r^*_{meta} = b^{'} r^*$
  - And show that r^*_{meta} cannot price the underlying r^* OOS
  - This analysis does not require Step 1 (or predictors)
  - A "Big Data World" theory is silent about this meta r^* result

1) The Theory Cannot Explain the Meta r^* Results

- There are two steps to build r^*
  1. Convert predictors into factors (f)
  2. Collapse factors into $r^* = b^{'} f$
- Underlying argument for the theory in the paper:
  - Step 2 is fine
  - But in a "Big Data World" Step 1 fails even with large T
- But this theory cannot explain the meta r^* results
  - The authors take the r^* from each factor model (15 on total)
  - They then construct $r^*_{meta} = b^{'} r^*$
  - And show that r^*_{meta} cannot price the underlying r^* OOS
  - This analysis does not require Step 1 (or predictors)
  - A "Big Data World" theory is silent about this meta r^* result

1) The Theory Cannot Explain the Meta r^* Results

- There are two steps to build r^*
  1. Convert predictors into factors (f)
  2. Collapse factors into $r^* = b^{'} f$
- Underlying argument for the theory in the paper:
  - Step 2 is fine
  - But in a "Big Data World" Step 1 fails even with large T
- But this theory cannot explain the meta r^* results
  - The authors take the r^* from each factor model (15 on total)
  - They then construct $r^*_{meta} = b^{'} r^*$
  - And show that r^*_{meta} cannot price the underlying r^* OOS
  - This analysis does not require Step 1 (or predictors)
  - A "Big Data World" theory is silent about this meta r^* result

2) I Would Like to See Simulations

- It seems the theory does not fully explain the empirics
- I think it would be useful to explore simulations:

2) I Would Like to See Simulations

- It seems the theory does not fully explain the empirics
- I think it would be useful to explore simulations:

2) I Would Like to See Simulations

- It seems the theory does not fully explain the empirics
- I think it would be useful to explore simulations:
  - Simulate returns with a known SDF
  - You can make realistic choices for N, T, and P
  - Apply the r^* methods you explored to the simulations
  - Do you still observe the "Factor Model Failure Puzzle"?
  - Do you observe it even with r^*_{meta}?
  - Try to explore why (or why not)

2) I Would Like to See Simulations

- It seems the theory does not fully explain the empirics
- I think it would be useful to explore simulations:
  - Simulate returns with a known SDF
  - You can make realistic choices for N, T, and P
  - Apply the r^* methods you explored to the simulations
  - Do you still observe the "Factor Model Failure Puzzle"?
  - Do you observe it even with r^*_{meta}?
  - Try to explore why (or why not)

2) I Would Like to See Simulations

- It seems the theory does not fully explain the empirics
- I think it would be useful to explore simulations:
  - Simulate returns with a known SDF
  - You can make realistic choices for N, T, and P
  - Apply the r^* methods you explored to the simulations
  - Do you still observe the "Factor Model Failure Puzzle"?
  - Do you observe it even with r^*_{meta}?
  - Try to explore why (or why not)

2) I Would Like to See Simulations

- It seems the theory does not fully explain the empirics
- I think it would be useful to explore simulations:
  - Simulate returns with a known SDF
  - You can make realistic choices for N, T, and P
  - Apply the r^* methods you explored to the simulations
  - Do you still observe the "Factor Model Failure Puzzle"?
  - Do you observe it even with r^*_{meta}?
  - Try to explore why (or why not)

2) I Would Like to See Simulations

- It seems the theory does not fully explain the empirics
- I think it would be useful to explore simulations:
  - Simulate returns with a known SDF
  - You can make realistic choices for N, T, and P
  - Apply the r^* methods you explored to the simulations
  - Do you still observe the "Factor Model Failure Puzzle"?
  - Do you observe it even with r^*_{meta}?
  - Try to explore why (or why not)

2) I Would Like to See Simulations

- It seems the theory does not fully explain the empirics
- I think it would be useful to explore simulations:
  - Simulate returns with a known SDF
  - You can make realistic choices for N, T, and P
  - Apply the r^* methods you explored to the simulations
  - Do you still observe the "Factor Model Failure Puzzle"?
  - Do you observe it even with r^*_{meta}?
  - Try to explore why (or why not)

2) I Would Like to See Simulations

- It seems the theory does not fully explain the empirics
- I think it would be useful to explore simulations:
  - Simulate returns with a known SDF
  - You can make realistic choices for N, T, and P
  - Apply the r^* methods you explored to the simulations
  - Do you still observe the "Factor Model Failure Puzzle"?
  - Do you observe it even with r^*_{meta}?
  - Try to explore why (or why not)

3) My Conclusion from the Paper is Different from Yours

- Two purposes for building a factor model:

- (1) $\neq$ (2) if $\mathbb{E}[r]$ also have non-risk sources (likely the case)
- The authors focus on (1) and conclude:

- I would instead conclude:

3) My Conclusion from the Paper is Different from Yours

- Two purposes for building a factor model:
  - (1) Build r^* to price all assets
  - (2) Identify factors that compensates for fundamental risks
- (1) $\neq$ (2) if $\mathbb{E}[r]$ also have non-risk sources (likely the case)
- The authors focus on (1) and conclude:
- I would instead conclude:

3) My Conclusion from the Paper is Different from Yours

- Two purposes for building a factor model:
  (1) Build r^* to price all assets
  (2) Identify factors that compensates for fundamental risks
- (1) $\neq$ (2) if $\mathbb{E}[r]$ also have non-risk sources (likely the case)
- The authors focus on (1) and conclude:
- I would instead conclude:

3) My Conclusion from the Paper is Different from Yours

- Two purposes for building a factor model:
  (1) Build r^* to price all assets
  (2) Identify factors that compensates for fundamental risks
- (1) $\neq$ (2) if $\mathbb{E}[r]$ also have non-risk sources (likely the case)
- The authors focus on (1) and conclude:
- I would instead conclude:

3) My Conclusion from the Paper is Different from Yours

- Two purposes for building a factor model:
  - (1) Build r^* to price all assets
  - (2) Identify factors that compensates for fundamental risks
- (1) $\neq$ (2) if $\mathbb{E}[r]$ also have non-risk sources (likely the case)
- The authors focus on (1) and conclude:
- I would instead conclude:

3) My Conclusion from the Paper is Different from Yours

- Two purposes for building a factor model:
  - (1) Build r^* to price all assets
  - (2) Identify factors that compensates for fundamental risks
- (1) $\neq$ (2) if $\mathbb{E}[r]$ also have non-risk sources (likely the case)
- The authors focus on (1) and conclude:

  > "We advocate for more research to understand the impact of measurement error in factor models and the relation between factor characteristics and the number of predictors needed to explain average returns."

- I would instead conclude:

3) My Conclusion from the Paper is Different from Yours

- Two purposes for building a factor model:
  (1) Build r^* to price all assets
  (2) Identify factors that compensates for fundamental risks
- (1) $\neq$ (2) if $\mathbb{E}[r]$ also have non-risk sources (likely the case)
- The authors focus on (1) and conclude:

  > "We advocate for more research to understand the impact of measurement error in factor models and the relation between factor characteristics and the number of predictors needed to explain average returns."

- I would instead conclude:

3) My Conclusion from the Paper is Different from Yours

- Two purposes for building a factor model:
  (1) Build r^* to price all assets
  (2) Identify factors that compensates for fundamental risks
- (1) $\neq$ (2) if $\mathbb{E}[r]$ also have non-risk sources (likely the case)
- The authors focus on (1) and conclude:

  "We advocate for more research to understand the impact of measurement error in factor models and the relation between factor characteristics and the number of predictors needed to explain average returns."
- I would instead conclude:

  "We cannot Identify r^* given the typical T and P. So, we are better off as a profession if we focus on (2) when building factor models" Chabi-Yo, Gonçalves, Loudis (2023)

3) My Conclusion from the Paper is Different from Yours

- Two purposes for building a factor model:
  (1) Build r^* to price all assets
  (2) Identify factors that compensates for fundamental risks
- (1) $\neq$ (2) if $\mathbb{E}[r]$ also have non-risk sources (likely the case)
- The authors focus on (1) and conclude:

  > "We advocate for more research to understand the impact of measurement error in factor models and the relation between factor characteristics and the number of predictors needed to explain average returns."

- I would instead conclude:

  > "We cannot Identify r^* given the typical T and P. So, we are better off as a profession if we focus on (2) when building factor models" Chabi-Yo, Gonçalves, Loudis (2023)

Some Minor Comments

1) Section 2 has to be shorter (mostly notation + section 2.4)

2) Some alphas are very large (e.g., $\alpha > 20\%$)

3) Given (2), you should consider trading costs
(Jensen, Kelly, Malamud, Pedersen (2022))

4) If machine learning fails due to the multi-step process of converging predictors into weights, then Reinforcement Learning could help (Cong, Tang, Wang, Zhang (2022))

Outline

The Paper

My Comments

Final Remarks

Final Remarks

- Nice paper (and extremely careful implementation)

- It would be useful to:

- Good luck!

Final Remarks

- Nice paper (and extremely careful implementation)
  - Finding: no r^* prices well the r^* from other models OOS
  - Explanation: large T but also many $\mathbb{E}[r]$ predictors
- It would be useful to:
- Good luck!

Final Remarks

- Nice paper (and extremely careful implementation)
  - Finding: no r^* prices well the r^* from other models OOS
  - Explanation: large T but also many $\mathbb{E}[r]$ predictors
- It would be useful to:
- Good luck!

Final Remarks

- Nice paper (and extremely careful implementation)
  - Finding: no r^* prices well the r^* from other models OOS
  - Explanation: large T but also many $\mathbb{E}[r]$ predictors
- It would be useful to:
- Good luck!

Final Remarks

- Nice paper (and extremely careful implementation)
  - Finding: no r^* prices well the r^* from other models OOS
  - Explanation: large T but also many $\mathbb{E}[r]$ predictors
- It would be useful to:
  - Provide a theory that can also explain the r^*_{meta} results
  - Add simulations to better understand the empirical results
  - Reconsider the conclusion: should factor models focus on risk?
- Good luck!

Final Remarks

- Nice paper (and extremely careful implementation)
  - Finding: no r^* prices well the r^* from other models OOS
  - Explanation: large T but also many $\mathbb{E}[r]$ predictors
- It would be useful to:
  - Provide a theory that can also explain the r^*_{meta} results
  - Add simulations to better understand the empirical results
  - Reconsider the conclusion: should factor models focus on risk?
- Good luck!

Final Remarks

- Nice paper (and extremely careful implementation)
  - Finding: no r^* prices well the r^* from other models OOS
  - Explanation: large T but also many $\mathbb{E}[r]$ predictors
- It would be useful to:
  - Provide a theory that can also explain the r^*_{meta} results
  - Add simulations to better understand the empirical results
  - Reconsider the conclusion: should factor models focus on risk?
- Good luck!

Final Remarks

- Nice paper (and extremely careful implementation)
  - Finding: no r^* prices well the r^* from other models OOS
  - Explanation: large T but also many $\mathbb{E}[r]$ predictors
- It would be useful to:
  - Provide a theory that can also explain the r^*_{meta} results
  - Add simulations to better understand the empirical results
  - Reconsider the conclusion: should factor models focus on risk?
- Good luck!

Final Remarks

- Nice paper (and extremely careful implementation)
  - Finding: no r^* prices well the r^* from other models OOS
  - Explanation: large T but also many $\mathbb{E}[r]$ predictors
- It would be useful to:
  - Provide a theory that can also explain the r^*_{meta} results
  - Add simulations to better understand the empirical results
  - Reconsider the conclusion: should factor models focus on risk?
- Good luck!