FACTOR TIMING

Valentin Haddad[1], Serhiy Kozak[2], Shri Santosh[3]

April 2019 – Q Group

[1]UCLA and NBER
[2]University of Michigan
[3]University of Maryland

Introduction

Factor Timing

- *When* to invest in *which* assets?

Factor Timing

- *When* to invest in *which* assets?
    - What is the *conditional stochastic discount factor*?

Factor Timing

- *When* to invest in *which* assets?
    - What is the *conditional stochastic discount factor*?
- Market timing $\Leftrightarrow$ Aggregate risk premium varies over time
    - Predictive regressions: Shiller 1981, Campbell and Shiller 1988, Fama and French 1988

Factor Timing

- *When* to invest in *which* assets?
  - What is the *conditional stochastic discount factor*?
- Market timing $\Leftrightarrow$ Aggregate risk premium varies over time
  - Predictive regressions: Shiller 1981, Campbell and Shiller 1988, Fama and French 1988
- Factor investing $\Leftrightarrow$ Different compensation for various sources of risk
  - Factor models: Fama and French 1993

Factor Timing

- *When* to invest in *which* assets?
  - What is the *conditional stochastic discount factor*?
- Market timing $\Leftrightarrow$ Aggregate risk premium varies over time
  - Predictive regressions: Shiller 1981, Campbell and Shiller 1988, Fama and French 1988
- Factor investing $\Leftrightarrow$ Different compensation for various sources of risk
  - Factor models: Fama and French 1993
- **Factor timing** $\Leftrightarrow$ Time-varying risk premium for various sources of risk

What We Do

- Hard problem! N factors and K predictors $\rightarrow N \times K$ regression coefficients. Potentially severe *overfitting*

What We Do

- Hard problem! N factors and K predictors $\rightarrow N \times K$ regression coefficients. Potentially severe *overfitting*

- **Economic restriction**: absence of near-arbitrage
  - Too many sources of predictability $\rightarrow$ implausible trading profits
  - Summarize factors using a few Principal Components, forecast them individually

What We Find

- Significant and robust benefits to factor timing uncovered using economic restriction

What We Find

- Significant and robust benefits to factor timing uncovered using economic restriction

- Pure timing portfolio collects a Sharpe ratio of 0.85
    - Mostly from timing non-market factors
    - Equivalently, SDF more volatile than we thought

What We Find

- Significant and robust benefits to factor timing uncovered using economic restriction

- Pure timing portfolio collects a Sharpe ratio of 0.85
    - Mostly from timing non-market factors
    - Equivalently, SDF more volatile than we thought

- Conditional and unconditional optimal portfolios
    - Can have very different composition
    - Often weakly correlated

Methodology

Methodology

Assumption
There are (usually) no near-arbitrage opportunities: (average) max SR not too high.

Methodology

Assumption
There are (usually) no near-arbitrage opportunities: (average) max SR not too high.

- Ross 1976, Cochrane and Saá-Requejo 2000, Kozak et al. 2017

- robust across many economic models

SDF Variance

- Max *conditional* squared SR, S_t^2. On average:

$$\mathbb{E}\left(S_t^2\right) = \sum_i \frac{\mathbb{E}\left[Z_{i,t+1}\right]^2}{\lambda_i} + \sum_i \left(\frac{R_i^2}{1-R_i^2}\right)$$

SDF Variance

- Max *conditional* squared SR, S_t^2. On average:

$$\mathbb{E}\left(S_t^2\right) = \sum_i \frac{\mathbb{E}\left[Z_{i,t+1}\right]^2}{\lambda_i} + \sum_i \left(\frac{R_i^2}{1 - R_i^2}\right)$$

- Bounded $\mathbb{E}\left(S_t^2\right)$ implies no individual R^2 can be large
- Many R^2 cannot be even moderately large

SDF Variance

- Max *conditional* squared SR, S_t^2. On average:

$$\mathbb{E}\left(S_t^2\right) = \sum_i \frac{\mathbb{E}\left[Z_{i,t+1}\right]^2}{\lambda_i} + \sum_i \left(\frac{R_i^2}{1 - R_i^2}\right)$$

- Bounded $\mathbb{E}\left(S_t^2\right)$ implies no individual R^2 can be large
- Many R^2 cannot be even moderately large

- Most total pred. must come from few large PCs. Predict those directly

Results:
Predicting Stock Returns

Anomaly Factors

Size	Investment	Long-term Reversals
Value	Inv/Cap	Value (M)
Profitability	Investment Growth	Net Issuance (M)
Value-Profitability	Sales Growth	SUE
F-score	Leverage	Return on Book Equity
Debt Issuance	Return on Assets (A)	Return on Market Equity
Share Repurchases	Return on Equity (A)	Return on Assets
Net Issuance (A)	Sales/Price	Short-term Reversals
Accruals	Growth in LTNOA	Idiosyncratic Volatility
Asset Growth	Momentum (6m)	Beta Arbitrage
Asset Turnover	Industry Momentum	Seasonality
Gross Margins	Value-Momentum	Industry Rel. Reversals
D/P	Value-Prof-Momentum	Industry. Rel. Rev. (LV)
E/P	Short Interest	Industry Momentum-Rev
CF/P	Momentum (12m)	Composite Issuance
Net Operating Assets	Momentum-Reversals	Stock Price

Percentage of Variance Explained by Anomaly PCs

Table 1 – Percentage of variance explained by anomaly PCs

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
% var. explained	19.2	16.9	10.7	5.7	4.8	3.7	3.3	3.2	2.3	2.0
Cumulative	19.2	36.1	46.8	52.6	57.4	61.1	64.4	67.6	69.9	72.0

- Focus on forecasting two largest PCs
- Forecast returns using log B/M ratios

Forecasting PCs of Anomalies

Table 2 – Predicting PC returns of all anomalies with BE/ME ratios

$$Z_{i,t:t+1} = a_i + b_i \times bm_{i,t} + \varepsilon_{i,t:t+1}$$

	MKT	PC1	PC2
Own *bm*	0.027	0.024	0.027
	(1.64)	(3.40)	(3.42)
R^2	0.044	0.240	0.209
OOS R^2	0.136	0.245	0.208
Wald test *p*-value	0.259	0.003	0.003

- Anomaly returns highly predictable, more so than aggregate market

Forecasting PCs of Anomalies

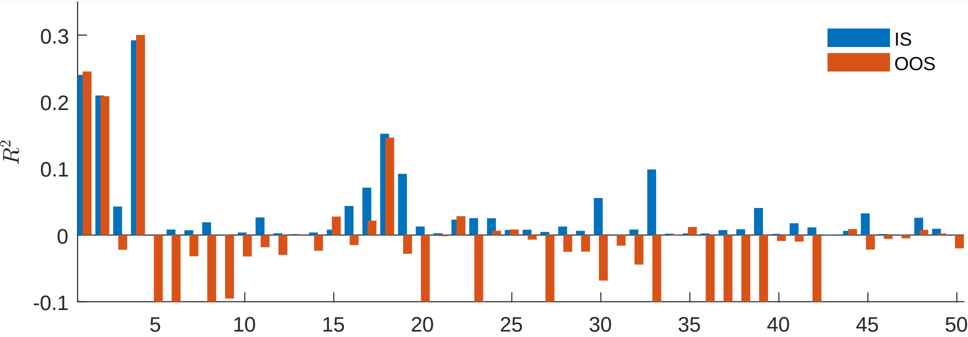
Table 3 – Predicting PC returns of all anomalies with BE/ME ratios

$$Z_{i,t:t+1} = a_i + \left[\begin{array}{ccc} bm_{mkt,t} & bm_{1,t} & bm_{2,t} \end{array} \right] \times b_i + \varepsilon_{i,t:t+1}$$

	MKT	PC1	PC2
bm_{mkt}	0.025	-0.043	0.017
	(1.39)	(2.71)	(1.05)
bm_1	-0.003	0.030	-0.002
	(0.42)	(4.92)	(0.24)
bm_2	-0.014	0.009	0.027
	(1.79)	(1.18)	(3.49)
R^2	0.097	0.374	0.226
OOS R^2	-0.334	0.235	0.217
Wald test p-value	0.214	0.000	0.009

Forecasting All PCs With Own BE/ME

$$Z_{i,t:t+1} = a_i + b_i \times bm_{i,t} + \varepsilon_{i,t:t+1}$$



- Small PCs not predictable: weak in-sample and appear spurious out-of-sample
- Validates economic restriction

Predicting Individual Factors

Are Individual Anomalies Predictable?

- Map PC predictability back to individual anomalies

Are Individual Anomalies Predictable?

- Map PC predictability back to individual anomalies

- Median IS and OOS R^2 of 12% and 9%

Are Individual Anomalies Predictable?

- Map PC predictability back to individual anomalies
- Median IS and OOS R^2 of 12% and 9%
- *Cannot* obtain if predictability concentrated in small PCs

Alternative Approaches

	Our method (restr.)	3 predictors (unrestr.)	Own B/M	Own char. spread	Pooled regression
	In-sample R^2, %				
Mean	11.44	18.25	8.27	3.48	4.22
Median	11.79	16.70	4.44	1.67	4.44
Std. Dev.	10.33	10.43	9.30	4.46	8.65
	Out-of-sample R^2, %				
Mean	8.70	-17.97	0.81	-62.06	2.88
Median	9.29	-6.73	1.22	-0.23	4.48
Std. Dev.	13.94	41.07	19.91	428.73	13.93

Alternative Approaches

	Our method (restr.)	3 predictors (unrestr.)	Own B/M	Own char. spread	Pooled regression
	In-sample R^2, %				
Mean	11.44	18.25	8.27	3.48	4.22
Median	11.79	16.70	4.44	1.67	4.44
Std. Dev.	10.33	10.43	9.30	4.46	8.65
	Out-of-sample R^2, %				
Mean	8.70	-17.97	0.81	-62.06	2.88
Median	9.29	-6.73	1.22	-0.23	4.48
Std. Dev.	13.94	41.07	19.91	428.73	13.93

$$F_{i,t:t+1} = a_i + \begin{bmatrix} bm_{mkt,t} & bm_{1,t} & bm_{2,t} \end{bmatrix} \times b_i + \varepsilon_{i,t:t+1}$$

Alternative Approaches

	Our method (restr.)	3 predictors (unrestr.)	Own B/M	Own char. spread	Pooled regression
			In-sample R^2, %		
Mean	11.44	18.25	8.27	3.48	4.22
Median	11.79	16.70	4.44	1.67	4.44
Std. Dev.	10.33	10.43	9.30	4.46	8.65
			Out-of-sample R^2, %		
Mean	8.70	-17.97	0.81	-62.06	2.88
Median	9.29	-6.73	1.22	-0.23	4.48
Std. Dev.	13.94	41.07	19.91	428.73	13.93

$$F_{i,t:t+1} = a_i + bm_{i,t} \times b_i + \varepsilon_{i,t:t+1}$$

Alternative Approaches

	Our method (restr.)	3 predictors (unrestr.)	Own B/M	Own char. spread	Pooled regression
			In-sample R^2, %		
Mean	11.44	18.25	8.27	3.48	4.22
Median	11.79	16.70	4.44	1.67	4.44
Std. Dev.	10.33	10.43	9.30	4.46	8.65
			Out-of-sample R^2, %		
Mean	8.70	-17.97	0.81	-62.06	2.88
Median	9.29	-6.73	1.22	-0.23	4.48
Std. Dev.	13.94	41.07	19.91	428.73	13.93

$$F_{i,t:t+1} = a_i + X_{i,t} \times b_i + \varepsilon_{i,t:t+1}$$

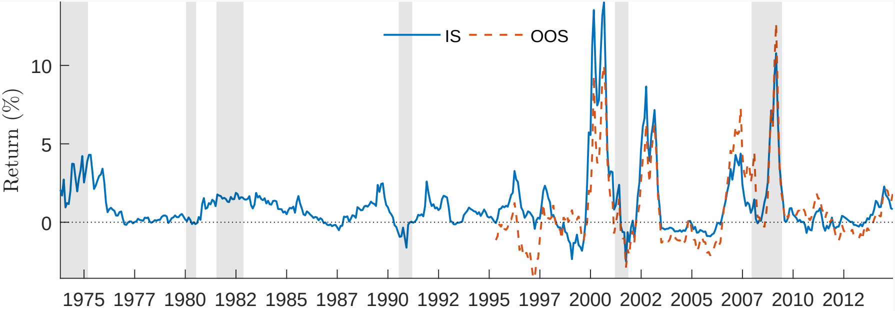
Alternative Approaches

	Our method (restr.)	3 predictors (unrestr.)	Own B/M	Own char. spread	Pooled regression
	In-sample R^2, %				
Mean	11.44	18.25	8.27	3.48	4.22
Median	11.79	16.70	4.44	1.67	4.44
Std. Dev.	10.33	10.43	9.30	4.46	8.65
	Out-of-sample R^2, %				
Mean	8.70	-17.97	0.81	-62.06	2.88
Median	9.29	-6.73	1.22	-0.23	4.48
Std. Dev.	13.94	41.07	19.91	428.73	13.93

$$F_{i,t:t+1} = a_i + bm_{i,t} \times \mathbf{b} + \varepsilon_{i,t:t+1}$$

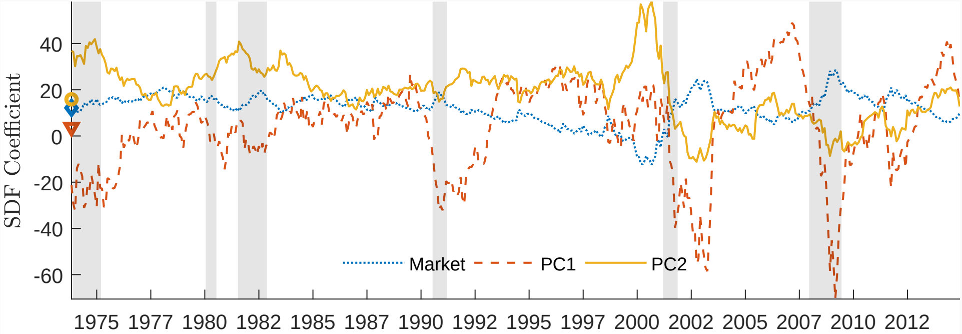
Optimal Factor Timing

Pure Timing Portfolio Returns



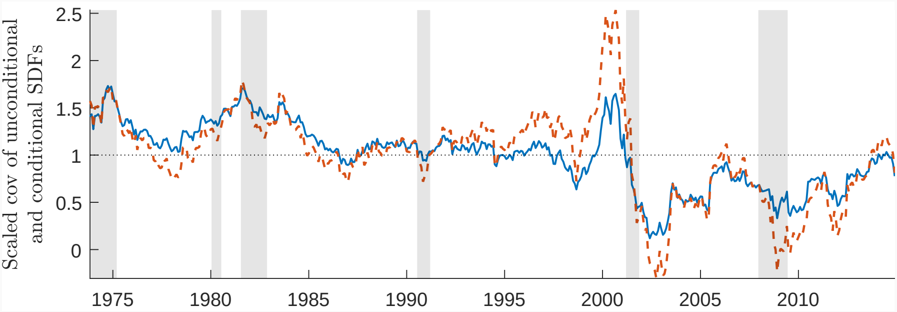
- Zero average weight on all factors. 0.85 average Sharpe ratio

Portfolio Weights



- PC1 not important unconditionally (0.07 Sharpe ratio)
- But high predictability means often very important conditionally

Covariance of Uncond. and Conditional Optimal Portfolios



- $\approx \beta$ of conditional portfolio wrt unconditional portfolio

Other Asset Classes

Data and Approach

- Use the same logic in two other asset classes: Treasury Bonds and Foreign Exchange
- Two factors: average return and long-short portfolios weighted by maturity and interest rate differential
- Use log valuation ratios: forward spreads

Foreign Exchange

Table 4 – Predicting PC returns of Foreign Exchange rates

	Dollar-Carry	Relative-Carry
FS1	0.52	-0.89
	(1.24)	(2.76)
FS2	0.07	1.30
	(0.11)	(2.48)
R^2	0.04	0.19
OOS R^2	-0.06	0.05
Wald test *p*-value	0.67	0.00

Bonds

Table 5 – Predicting PC Returns with Forward Spreads & CFNAI

	LevelR			SlopeR		
FS1	0.31	0.28	-	0.53	0.65	-
	(1.72)	(1.56)		(2.65)	(4.04)	
FS2	-1.76	-1.91	-	0.12	0.83	-
	(2.85)	(2.93)		(0.18)	(1.44)	
FS3	-3.66	-2.98	-	3.96	0.76	-
	(1.97)	(1.48)		(1.95)	(0.43)	
GRO	-	-0.50	-0.55	-	2.34	1.78
		(0.70)	(0.82)		(3.60)	(2.51)
R^2	0.22	0.23	0.01	0.22	0.37	0.13
OOS R^2	-1.74	-1.97	-0.01	0.11	0.18	-0.03
Wald test p-value	0.00	0.01	0.71	0.03	0.00	0.04

Conclusions

Conclusions

- Factor timing *seemingly* more difficult than market timing due to huge dimensionality
- Economic restriction on implied SDF yields tractability and robustness: only large PCs of returns are predictable
- Across asset classes: large PCs of index-neutral returns highly predictable, at least as much as index
- Conditional and unconditional SDFs often very different
  - Risk prices/premia do NOT all move together $\rightarrow$ strongly changing nature of risks investors care about

Appendix

First PC and Individual Regressions

- Often $\{R_{i,t+1}\}_{i \in I}$ has a strong common component F_{t+1}
- When this component is predictable by a variable X_t, does this imply that the individual returns are predictable by X_t?

First PC and Individual Regressions

- Often $\{R_{i,t+1}\}_{i \in I}$ has a strong common component F_{t+1}
- When this component is predictable by a variable X_t, does this imply that the individual returns are predictable by X_t?

$$\mathbb{E}_i\left[\mathrm{R}^2_{X,i}\right] = \mathrm{R}^2_1 \mathrm{R}^2_X + \mathrm{var}_i\left(\mathrm{R}^2_{X,i}\right), \tag{1}$$

$$\mathrm{R}^2_X \geq \frac{\left(1 - \mathrm{R}^2_{\mathrm{max}}\right)\mathbb{E}_i\left[\mathrm{R}^2_{X,i}\right] + \mathbb{E}_i\left[\mathrm{R}^2_{X,i}\right]^2}{\mathrm{R}^2_1}, \tag{2}$$

- where:
  - R^2_1 is the R^2 of an asset on the factor
  - R^2_X is the R^2 of the predictive regression of the factor
  - $\mathrm{R}^2_{X,i}$ is the R^2 of the predictive regression of an asset
  - $\mathrm{R}^2_{\mathrm{max}}$ is the max $\mathrm{R}^2_{X,i}$ from any individual asset forecasting reg.

First PC and Individual Regressions

- Often $\{R_{i,t+1}\}_{i \in I}$ has a strong common component F_{t+1}
- When this component is predictable by a variable X_t, does this imply that the individual returns are predictable by X_t?

$$\mathbb{E}_i \left[\mathrm{R}^2_{X,i}\right] = \mathrm{R}^2_1 \mathrm{R}^2_X + \mathrm{var}_i \left(\mathrm{R}^2_{X,i}\right), \tag{1}$$

$$\mathrm{R}^2_X \geq \frac{\left(1 - \mathrm{R}^2_{\max}\right) \mathbb{E}_i \left[\mathrm{R}^2_{X,i}\right] + \mathbb{E}_i \left[\mathrm{R}^2_{X,i}\right]^2}{\mathrm{R}^2_1}, \tag{2}$$

- where:
  - R^2_1 is the R^2 of an asset on the factor
  - R^2_X is the R^2 of the predictive regression of the factor
  - $\mathrm{R}^2_{X,i}$ is the R^2 of the predictive regression of an asset
  - $\mathrm{R}^2_{\max}$ is the max $\mathrm{R}^2_{X,i}$ from any individual asset forecasting reg.
- Example:
  - level factor explains $\approx 90\%$ of the variation in individual returns; can be predicted with an $R^2 \approx 25\%$
  - Thus $\mathbb{E}_i \left[\mathrm{R}^2_{X,i}\right] \geq 22.5\%$ by (1); $\mathrm{R}^2_X \geq 18\%$ by (2)

Eigenvector Loadings of Long-Short Anomaly Portfolios