Multiperiod Collection of Results

Prof. Serhiy Kozak & James Zhang

January 31, 2025

Contents

1 Setup

All R tensors are *log* returns, not simple and not excess of the one-lag return [1]. The out-of-sample period for all datasets is post January 2005. [2] In our primary specification,

- Rolling window size W is fixed at 120 months.
- Maximum horizon S is fixed at 36 months.
- Maximum lag (lagged characteristics) is 36 months.
- For each dataset and setting, we experiment with $K = 1, 3, 5, 10, 15, 20, 25$, where K represents the rank of PARAFAC and PCA ie. the number of factors. extracted from the tensor or matrix, respectively.
- Default $\gamma = 0$, we do not yet experiment with RP-PCA in the MVE procedure.

For robustness purposes, we present multiperiod results for each of the following 3 datasets:

1. Characteristics Anomalies dataset: 1972-06-01 (ie. $\min(t - L)$) to 2020-08-01 . - 43 characteristic-sorted portfolios (SortedFactors)
2. SCS dataset: 1972-07-31 to 2021-12-31

   - 44 characteristic-based portfolios (UnivariateFactors). All industry and constant factors are now removed.
3. WRDS dataset: 1975-01-31 to 2020-12-31

   - 107 characteristic-based portfolios (UnivariateFactors). I believe this dataset still currently has imputed data points, so we should get a version of this with 0 imputed data points.

For the Characteristic Anomalies dataset, factors are decile-sorted long-short portfolios. For the SCS and WRDS datasets, each stock is weighted by its characteristic signal value and signals are based on cross-sectional ranks, centered and normalized.

[1] Clarify that returns are excess of the risk-free rate, they should be, I know it is for the Characteristics Anomalies dataset

[2] We should standardize this to be 01-31-2005 to 2020-08-01, but in the preliminary, should not be a huge difference across the datasets.

2 Models

2.1 Tensor Factor Model (Approximate)

This is the primary model outlined in the paper. For time T and horizon S, slice the window of three-dimensional data $R \in \mathbb{R}^{W \times L \times N}$, apply PARAFAC to obtain

$$R_{t,i,l} = \sum_{k=1}^{K} \lambda_k F_{t,k} B_{t,k} W_{l,k}$$

or in tensor notation

$$R = \sum_{k=1}^{K} \lambda_k \cdot F \circ B_k \circ W_k \quad \text{for } F \in \mathbb{R}^{T \times K}, B \in \mathbb{R}^{N \times K}, W \in \mathbb{R}^{L \times K} \tag{1}$$

For clarity, let us denote the original returns R simply put as assets or returns and the K latent factors as the basis assets or factors. From here, for an S-period investment, we can compute the vector of multihorizon basis asset means and covariance matrix using the formulas in the paper,

$$\mu^{F,S} = \left(\sum_{l=1}^{S} W_l \right) \cdot \mathbb{E}_t[F_t] \in \mathbb{R}^K \tag{2}$$

$$\Sigma^{F,S} = \left(\sum_{l=1}^{S} W_l W_l^\top \right) \cdot \mathrm{Var}_t(F_t) \in \mathbb{R}^{K \times K} \tag{3}$$

where $*$ represents element-wise multiplication. Importantly, we assume that basis asset returns are uncorrelated over time, which is an empirically reasonable assumption. SDF weights are constructed using the simple Markowitz formula

$$w^S = \left(\Sigma^{F,S}\right)^{-1} \mu^{F,S} = \left(\left(\sum_{l=1}^{S} W_l W_l^\top \right) \cdot \mathrm{Var}_t(F_t) \right)^{-1} \mathrm{diag} \left(\sum_{l=1}^{S} W_k \right) \mathbb{E}_t[F_t] \tag{4}$$

To test our estimators, $\forall\ s \in [1, S]$ ie. for each of the next S months, we solve the regression

$$F_{t+s} \left(\lambda \cdot (W \odot B) \right)^\top = R_{t+s}$$

$$F_{t+s} = \mathrm{flat}(R_{t+s})((\lambda \cdot (W \odot B))^\top)^\dagger \in \mathbb{R}^K \tag{5}$$

Completing this for all S months effectively yields the OOS time series of the basis assets $F' \in \mathbb{R}^{S \times K}$. Then calculate multiperiod returns by combining it with W such that

$$\sum_{s=1}^{S} F_{t+s,k} W_{s,k} \tag{6}$$

and apply the mean-variance weights w_S to compute our returns in the OOS S-month period, which we can then use to compute the Sharpe Ratio.

2.2 Tensor Factor Model (Precise)

This model is also based on our tensor framework; however, it *does not* make the assumption that basis asset returns are uncorrelated over time. For time T and horizon S, slice the same three-dimensional window of data $R \in \mathbb{R}^{W\times L\times N}$ as above. Apply 1 to obtain the time series of basis assets returns, lag components, and cross-sectional loadings. Next, construct a matrix of *overlapping* multiperiod returns that we will denote $FW \in \mathbb{R}^{(W-S+1)\times K}$ such that

$$FW_{t,k} = \sum_{s=1}^{S} F_{t+s-1,k} W_{s,k} \ \forall \ t \in 1 \cdots W-S+1 \text{ and } k \in 1 \dots K \tag{7}$$

where $FW_{t,k}$ represents the S-month return of basis asset k beginning at time t. From here, we can easily compute the vector of observed sample means of basis asset k as

$$\mu^{F,S} = \frac{1}{W-S+1} \sum_{t=1}^{W-S+1} FW_t = \overline{FW_t} \in \mathbb{R}^K \tag{8}$$

and we use the Newey-West Covariance estimator because the matrix FW contains overlapping returns

$$\Sigma^{F,S} = \text{Newey-West}(FW, \text{overlap} = S-1) \in \mathbb{R}^{K\times K} \tag{9}$$

Construct SDF weights w^S using the Markowitz formula. Then, use 5 and 6 to obtain OOS basis asset multiperiod returns and combine with SDF weights to get the OOS return. If the approximate model in 2.1 out-performs this naive model, then we empirically justify our assumption that basis asset returns in separate months are uncorrelated.

2.3 Multihorizon PCA

Given time T and horizon S, consider the tensor $R \in \mathbb{R}^{W\times L\times N}$. Construct a matrix of *overlapping* S-month horizon returns denoted as $R^S \in \mathbb{R}^{(W-S+1)\times N}$ such that $\forall \ t \in 1 \dots W-S+1$ and $i \in 1 \dots N$

$$R^S_{t,i} = \sum_{l=1}^{S} R_{t+l,l,i} = \sum_{l=1}^{S} r_{t+l,i} | C_t \tag{10}$$

or in plain English, the S-month buy-and-hold return of characteristic-based portfolio i conditioned on characteristics from the current month t. Apply PCA to R^S obtain basis assets $\widehat{F}$ and loadings $\widehat{\Lambda}$. Compute basis asset moments

$$\mu^{F,S} = \frac{1}{W-S+1} \sum_{t=1}^{W-S+1} \widehat{F}_t = \overline{\widehat{F}} \in \mathbb{R}^K$$
$$\Sigma^{F,S} = \text{Newey-West}(\widehat{F}, \text{overlap} = S-1) \in \mathbb{R}^{K\times K}$$

OOS factor returns are estimated by a regression of the returns on the estimated loadings

$$\widehat{F}_{OOS} = \overrightarrow{R}_{OOS}\widehat{\Lambda}\left(\widehat{\Lambda}^\top\widehat{\Lambda}\right)^{-1} \in \mathbb{R}^S \tag{11}$$

Construct SDF weights and use 11 to find OOS basis asset values and the model's subsequent OOS returns.

2.4 Multihorizon Model-Free

Once more, given time T and horizon S, obtain the window of data and construct the same matrix of overlapping S-month returns $R^S \in \mathbb{R}^{(W-S+1)\times N}$. Rather than applying PCA to this panel, compute the moments of the factors as is.

$$\begin{aligned}
\mu^{R,S} &= \frac{1}{W-S+1}\sum_{t=1}^{W-S+1} R_t^S = \overline{R^S} \in \mathbb{R}^N \\
\Sigma^{R,S} &= \text{Newey-West}(R^S, \text{overlap} = S-1) \in \mathbb{R}^{N\times N}
\end{aligned}$$

Construct mean-variance weights and combine with the OOS factor values to obtain the S-month return. In theory, as $K \to N$, the Sharpe Ratios from 2.3 should approach this Model-Free method.

3 Further Analysis

3.1 SDF weights as a function of horizon

SDF weights as a function of horizon. Empirically, for a different number of factors, we show w^S as a function of S in a lineplot. We can normalize the l_1 norm of the weights to 1. To study this theoretically is a little bit more challenging. Recall that

$$w^S = \left(\left(\sum_{l=1}^S W_l W_l^\top\right)\cdot \text{Var}_t(F_t)\right)^{-1} \text{diag}\left(\sum_{l=1}^S W_k\right)\mathbb{E}_t[F_t]$$

By Lemma 6.1, there does not exist a closed-form expression for the element-wise product inside the matrix inverse. However, we can take advantage of the additional structure from the outer products of the lag components. By Lemma 6.2,

$$\begin{aligned}
\left(\left(\sum_{l=1}^S W_l W_l^\top\right)\cdot \text{Var}_t(F_t)\right) &= \left(\sum_{l=1}^S W_l W_l^\top \cdot \text{Var}_t(F_t)\right)^{-1} \\
&= \left(\sum_{l=1}^S D_{W_l}\text{Var}_t(F_t)D_{W_l}^\top\right)^{-1}
\end{aligned}$$

Note that the covariance matrix of the basis assets if symmetric, positive semi-definite and so each term in the summation is also summation. Does this help?

3.2 Sharpe Ratio of the SDF as a function of horizon

$$
\begin{aligned}
SR^S &= (\mu^{F,S})^\top w^S \sqrt{\frac{12}{S}} \\
&= (\mu^{F,S})^\top \left(\left(\sum_{l=1}^{S} W_l W_l^\top \right) \cdot \mathrm{Var}_t(F_t) \right)^{-1} \left(\sum_{l=1}^{S} W_k \right) \cdot \mathbb{E}_t[F_t] \sqrt{\frac{12}{S}}
\end{aligned}
$$

4 Collection of Results

List of Figures

List of Tables

4.1 Primary Specification: Full-Sample

4.1.1 Full-Sample Characteristic Anomalies Dataset

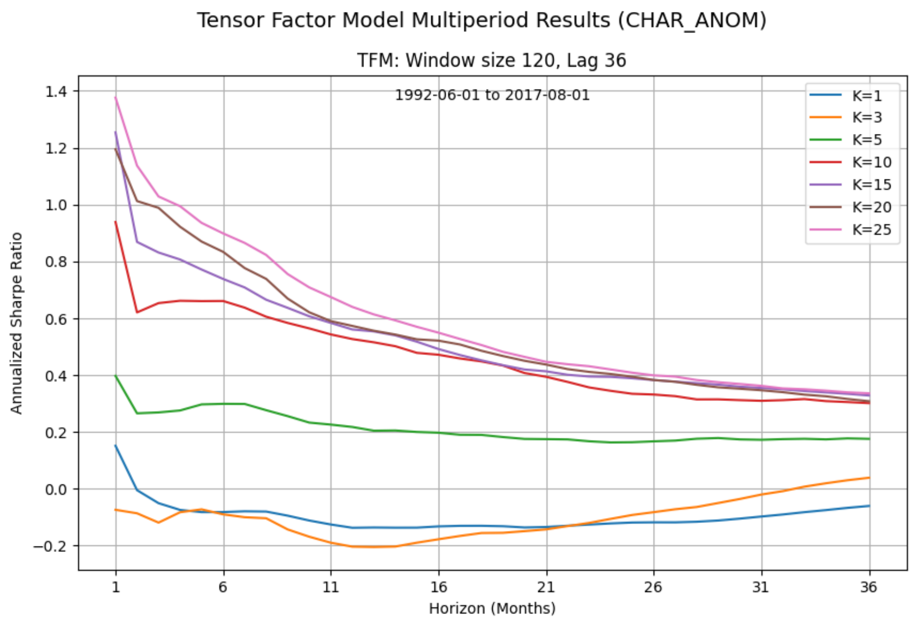

Figure 1: (CA) Tensor Factor Model: Full-Sample Annualized SR

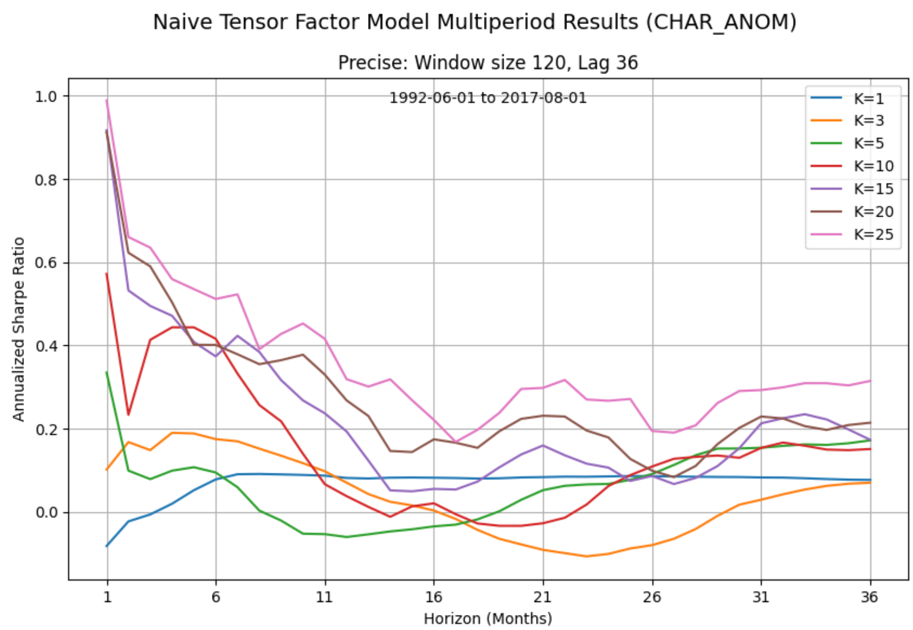

Figure 2: (CA) Precise Tensor Factor Model: Full-Sample Annualized SR

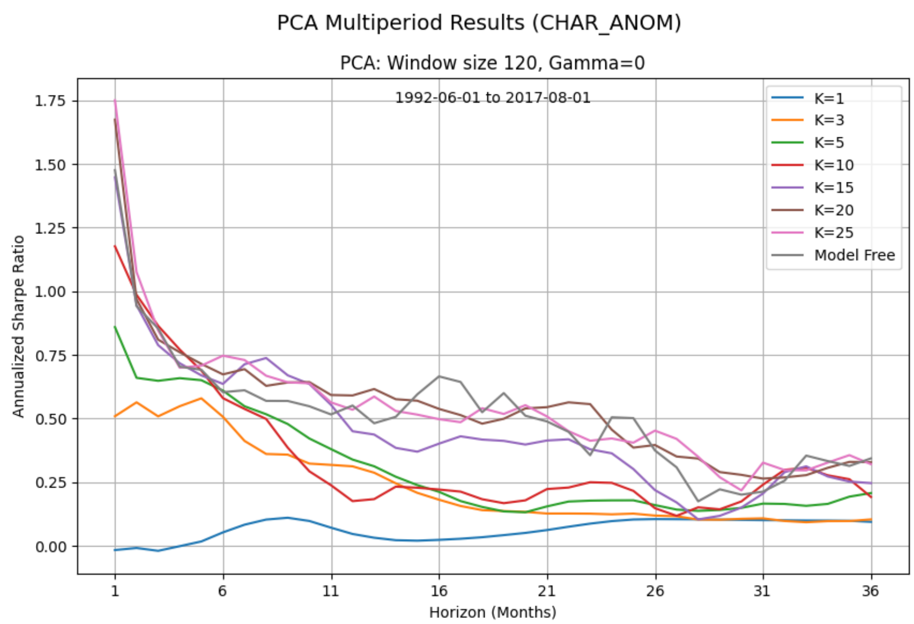

Figure 3: (CA) Multiperiod PCA Model: Full-Sample Annualized SR

maybe a table that's like

model $\implies$ K value (or model-free) $\implies$ sharpes for all 36 horizons

4.1.2 Full-Sample SCS Dataset

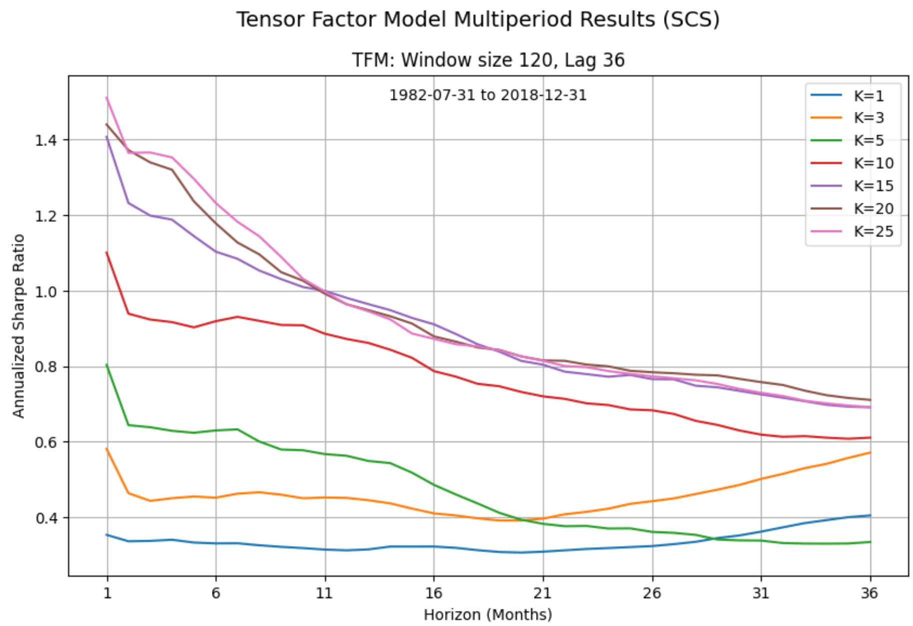

Figure 4: (SCS) Tensor Factor Model: Full-Sample Annualized SR

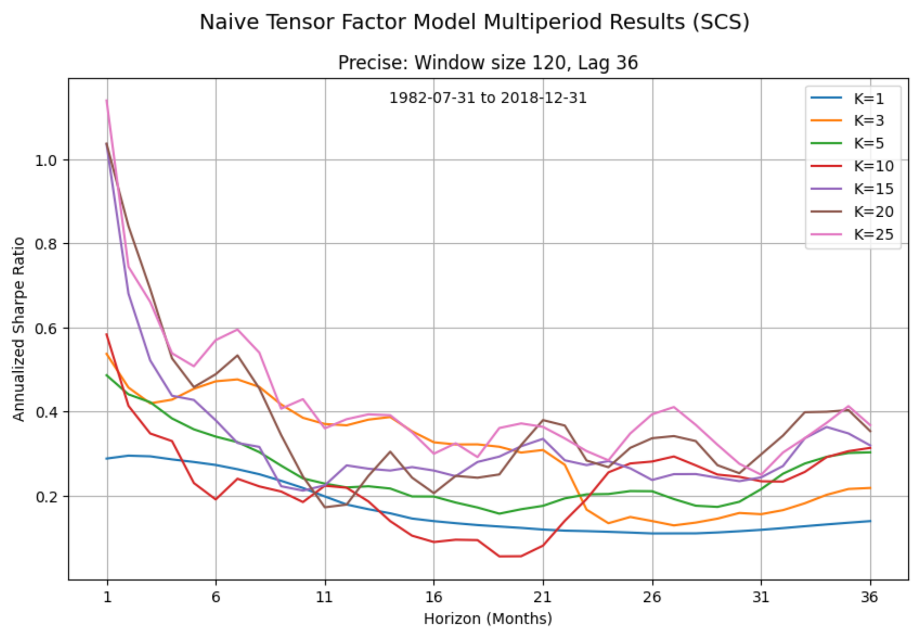

Figure 5: (SCS) Precise Tensor Factor Model: Full-Sample Annualized SR

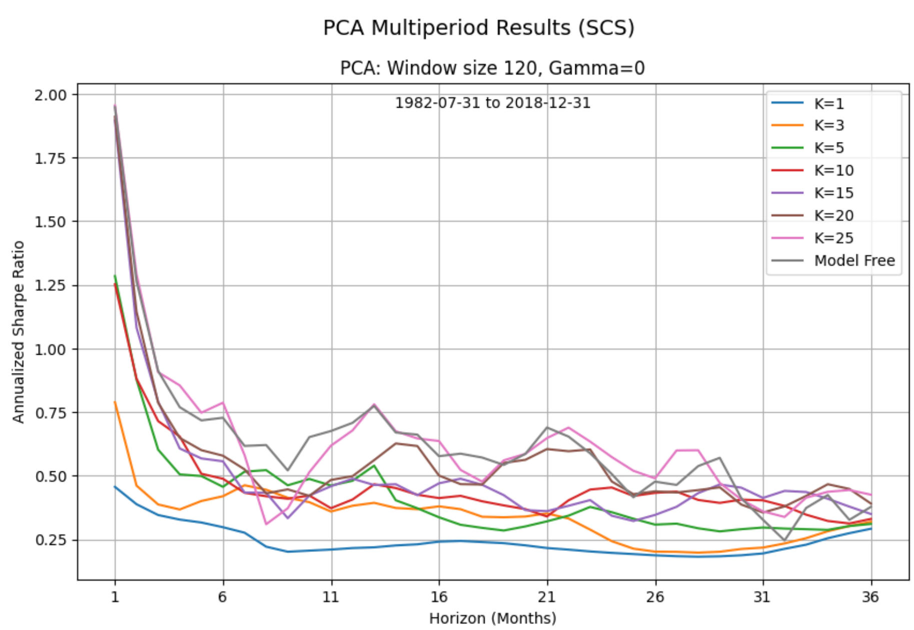

Figure 6: (SCS) Multiperiod PCA Model: Full-Sample Annualized SR

4.1.3 Full-Sample WRDS Dataset

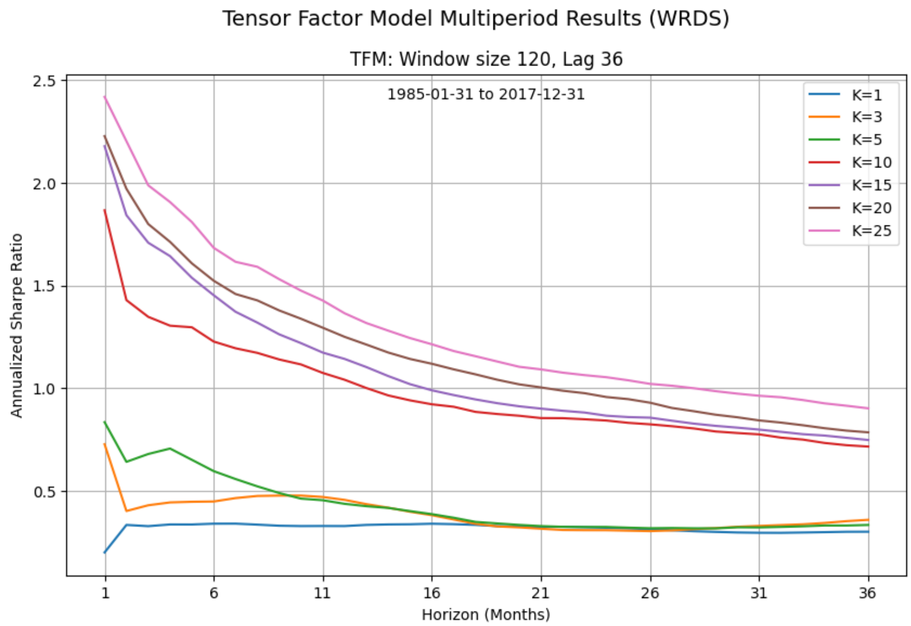

Figure 7: (WRDS) Tensor Factor Model: Full-Sample Annualized SR

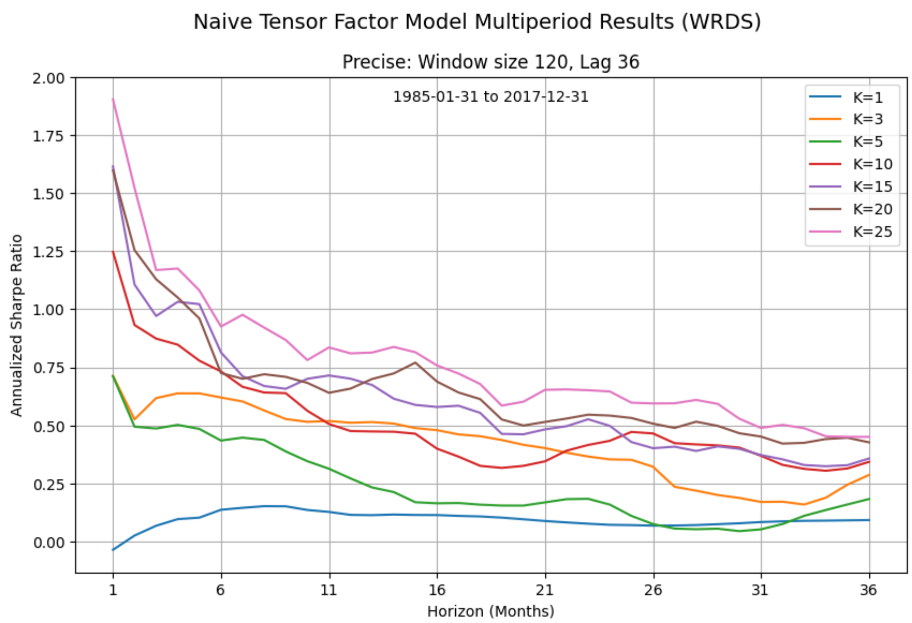

Figure 8: (WRDS) Precise Tensor Factor Model: Full-Sample Annualized SR

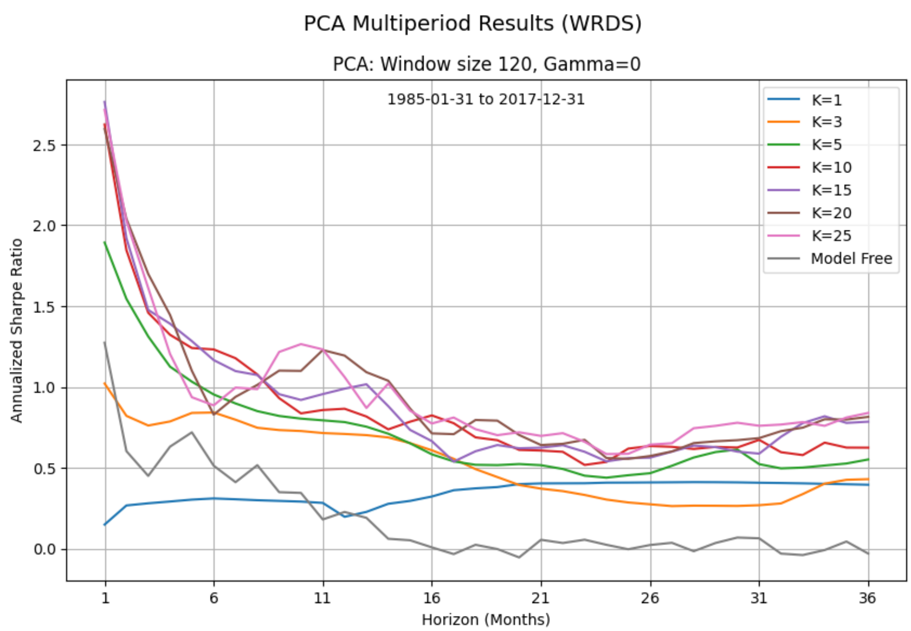

Figure 9: (WRDS) Multiperiod PCA Model: Full-Sample Annualized SR

4.1.4 Full-Sample Fama French 5 with Momentum Dataset

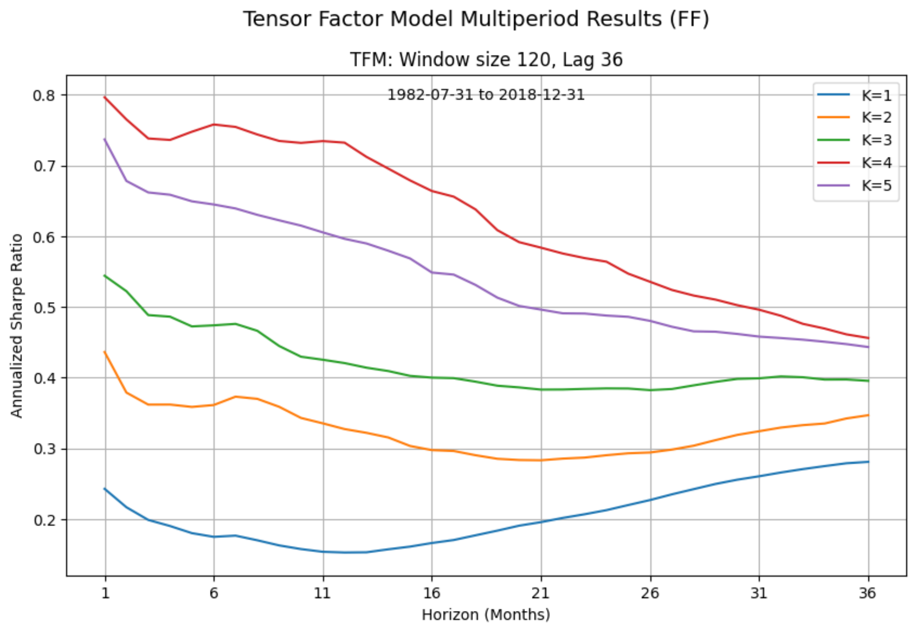

Figure 10: (FF5) Tensor Factor Model: Full-Sample Annualized SR

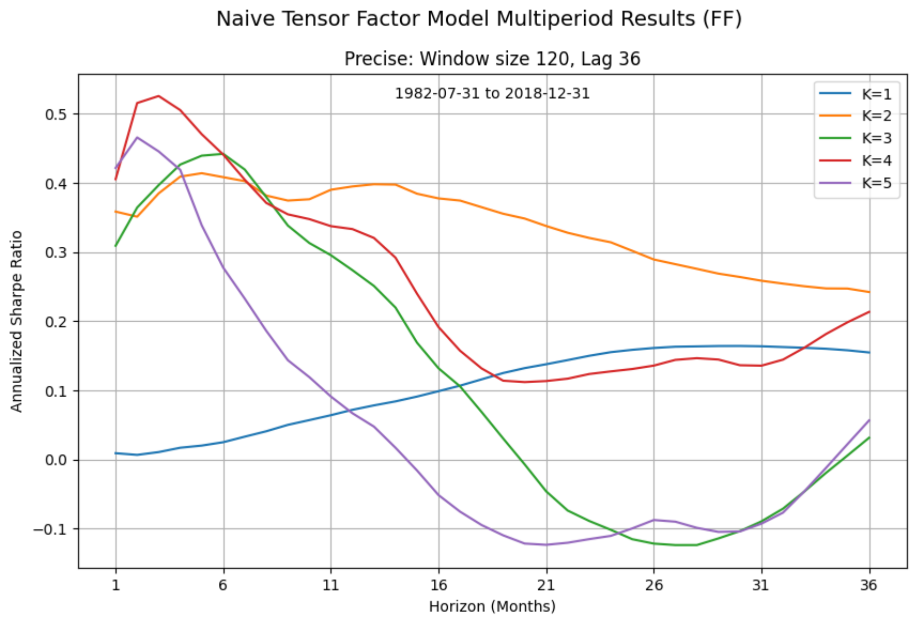

Figure 11: (FF5) Precise Tensor Factor Model: Full-Sample Annualized SR

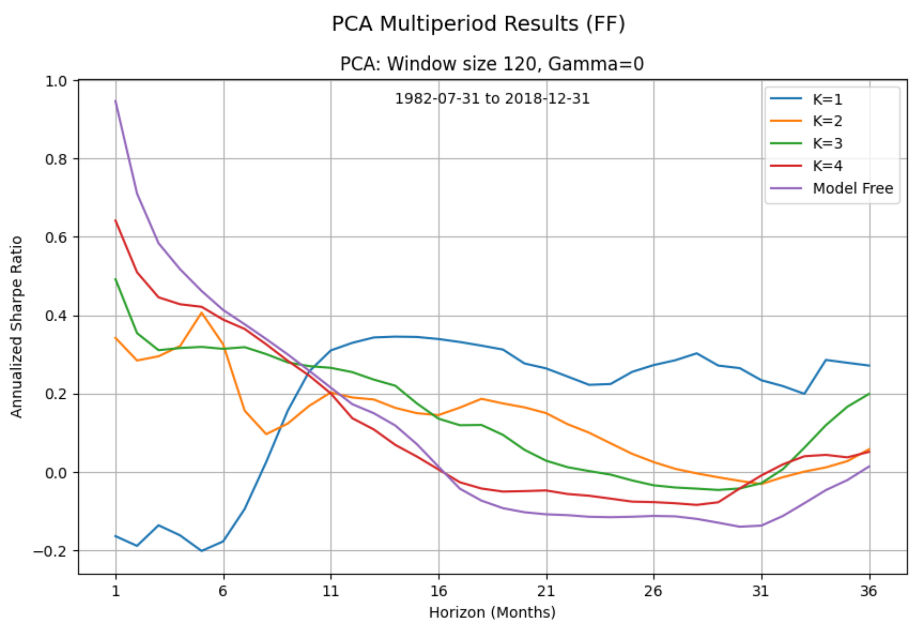

Figure 12: (FF5) Multiperiod PCA Model: Full-Sample Annualized SR

5 Appendix A

5.1 Primary Specification: OOS Results

5.1.1 OOS Characteristic Anomalies Dataset

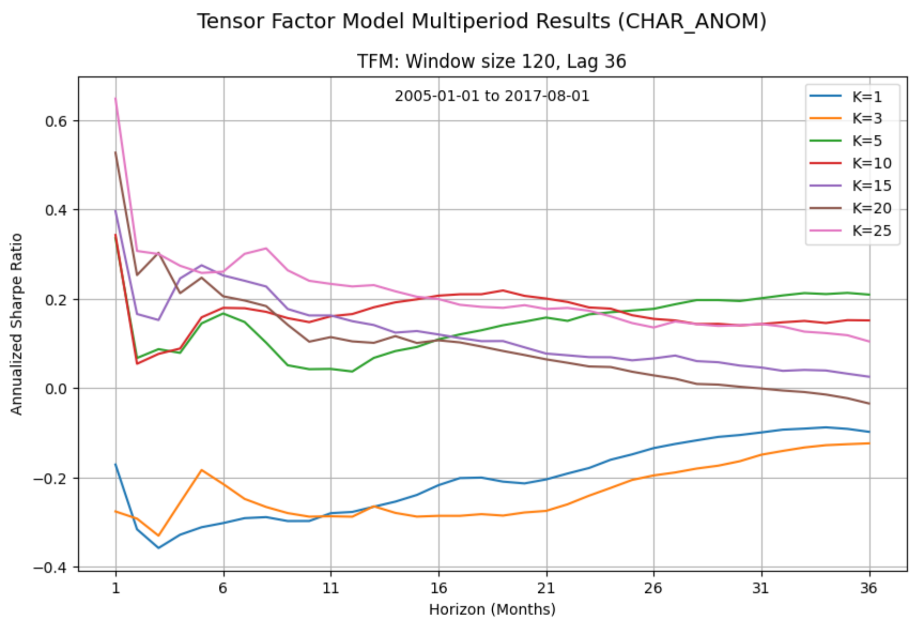

Figure 13: (CA) Tensor Factor Model: OOS Annualized SR

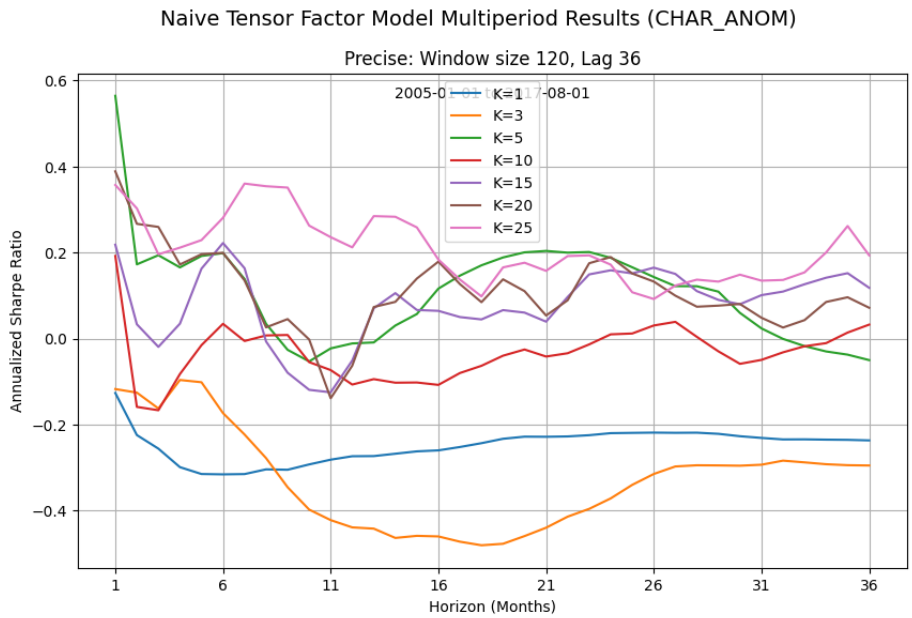

Figure 14: (CA) Precise Tensor Factor Model: OOS Annualized SR

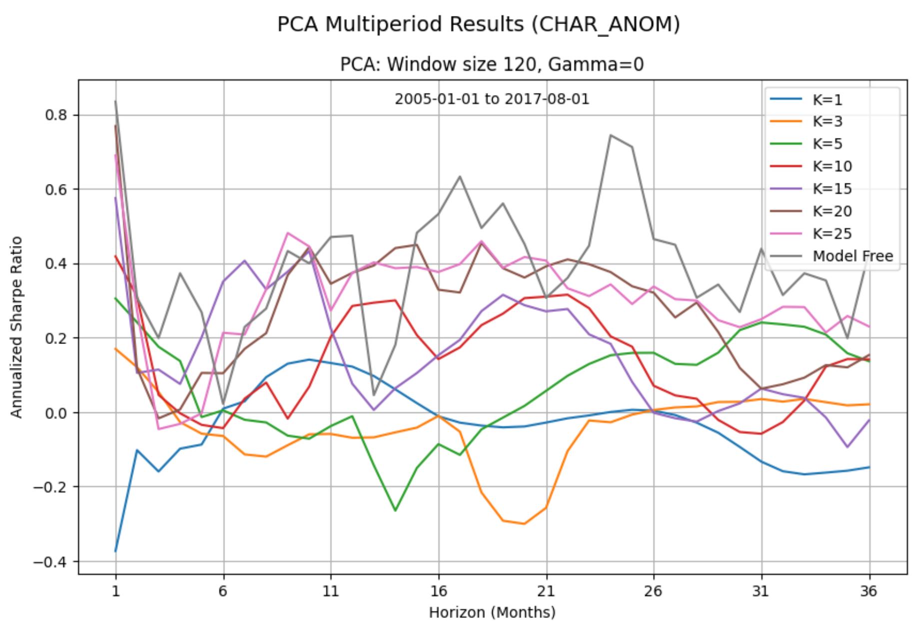

Figure 15: (CA) Multiperiod PCA Model: OOS Annualized SR

5.1.2 OOS SCS Dataset

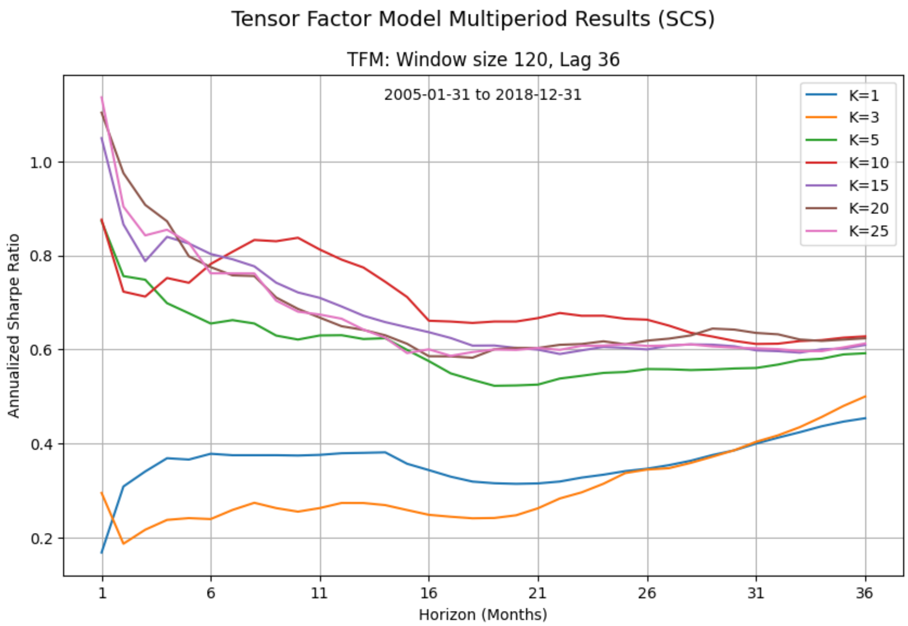

Figure 16: (SCS) Tensor Factor Model: OOS Annualized SR

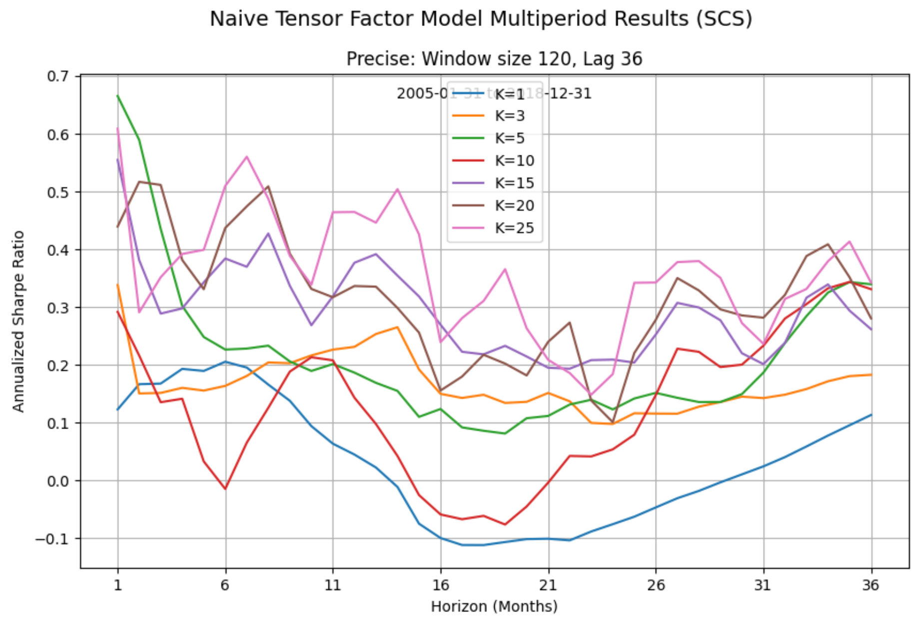

Figure 17: (SCS) Precise Tensor Factor Model: OOS Annualized SR

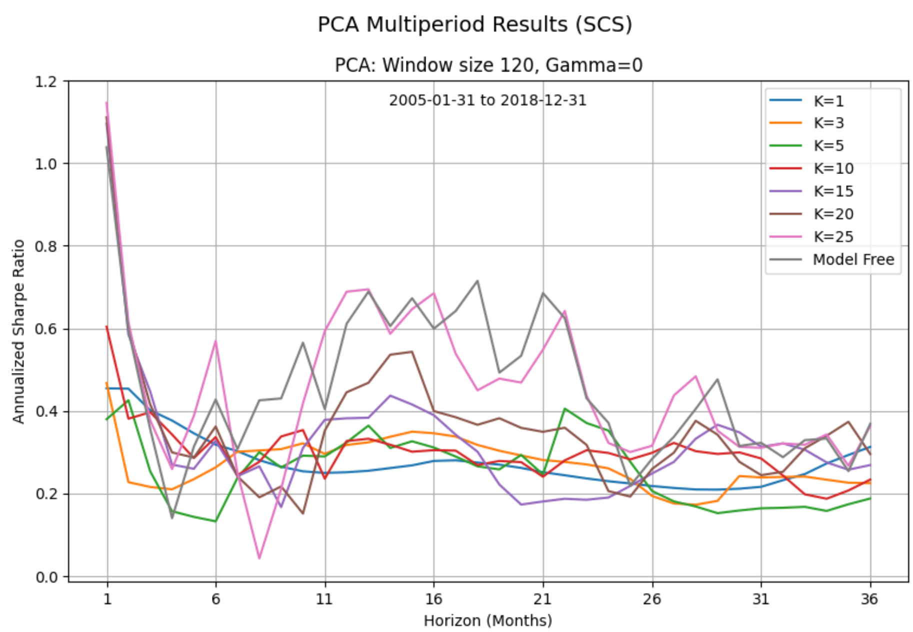

Figure 18: (SCS) Multiperiod PCA Model: OOS Annualized SR

5.1.3 OOS WRDS Dataset

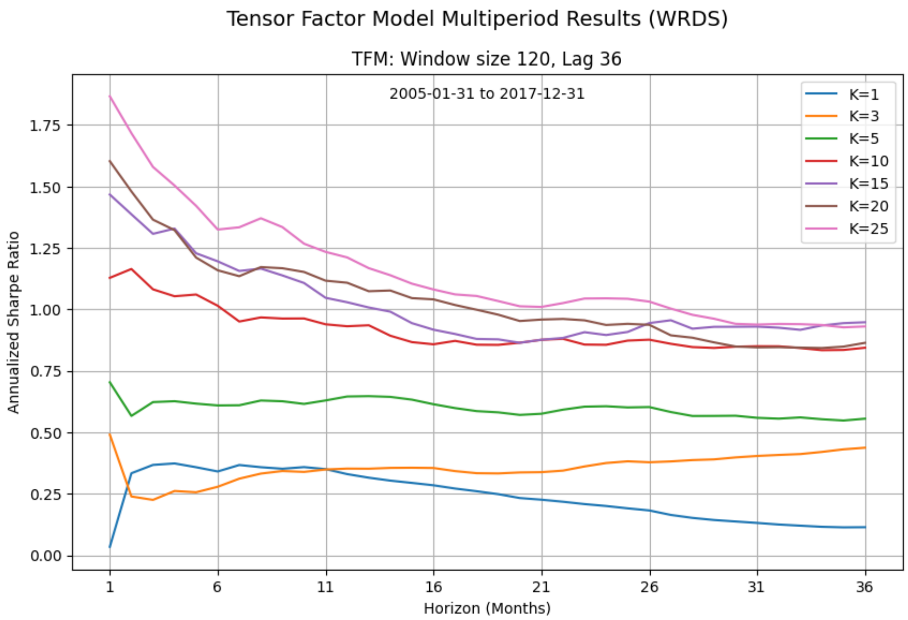

Figure 19: (WRDS) Tensor Factor Model: OOS Annualized SR

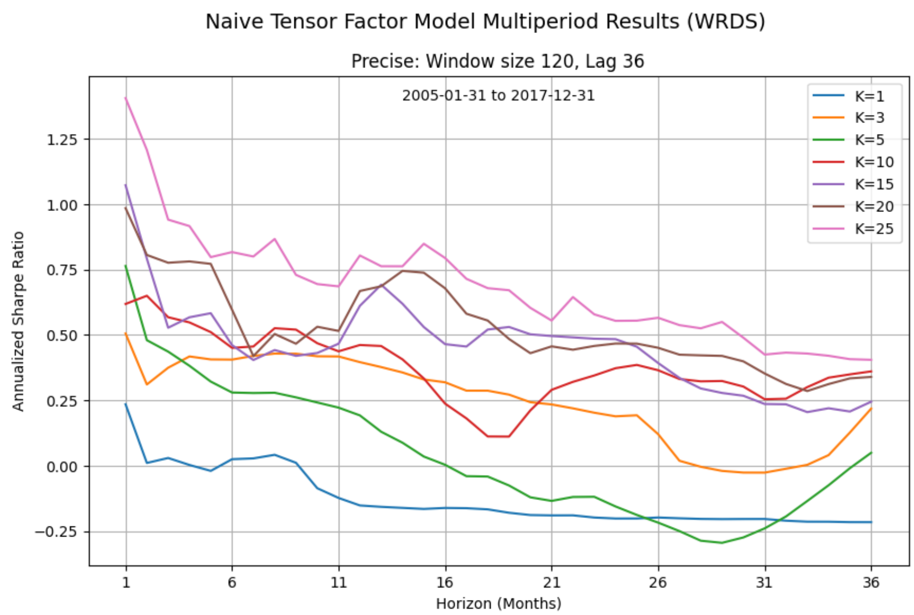

Figure 20: (WRDS) Precise Tensor Factor Model: OOS Annualized SR

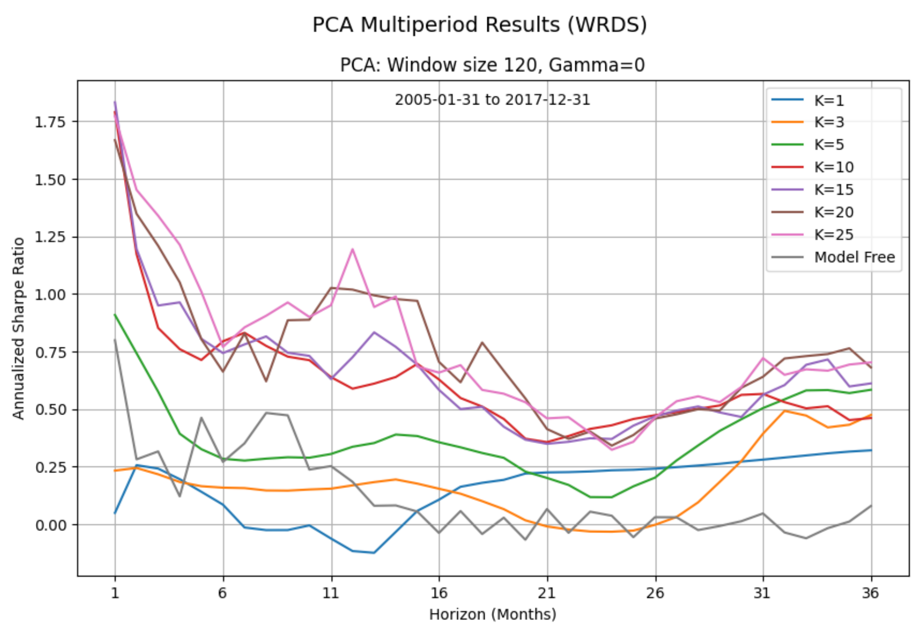

Figure 21: (WRDS) Multiperiod PCA Model: OOS Annualized SR

5.1.4 Full-Sample Fama French 5 with Momentum Dataset

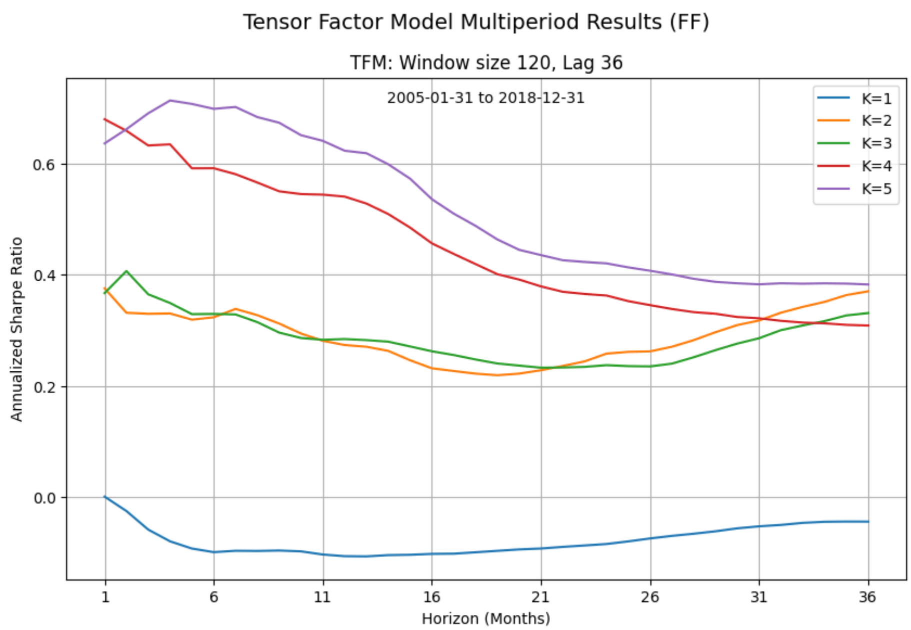

Figure 22: (FF5) Tensor Factor Model: Full-Sample Annualized SR

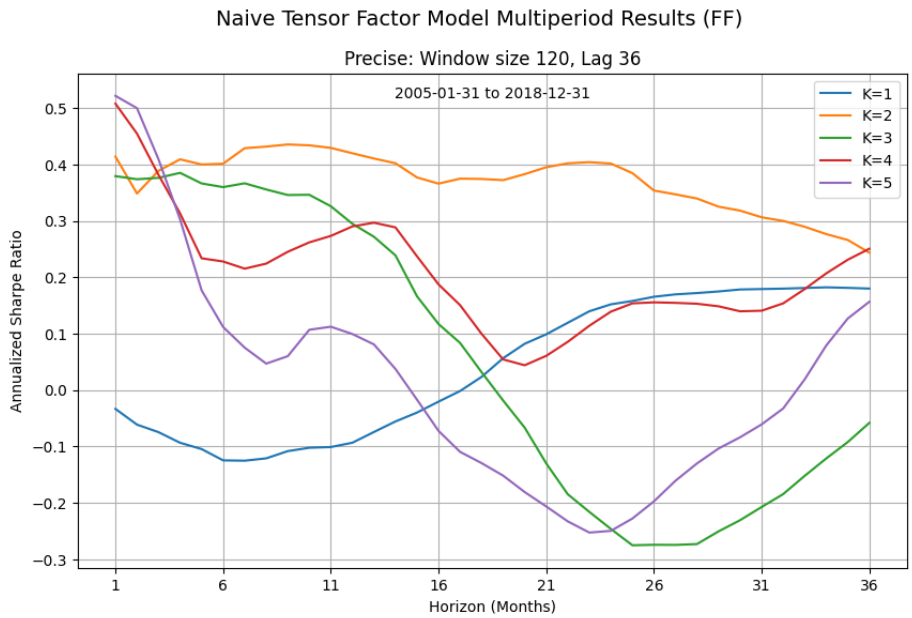

Figure 23: (FF5) Precise Tensor Factor Model: Full-Sample Annualized SR

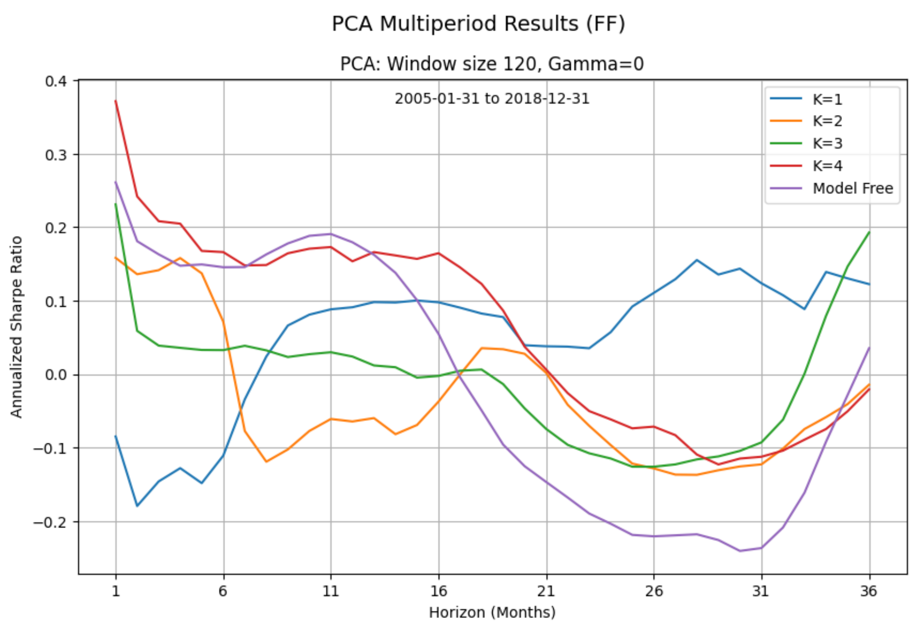

Figure 24: (FF5) Multiperiod PCA Model: Full-Sample Annualized SR

6 Appendix B

Lemma 6.1. *There does not exist a closed-form solution of $(A \cdot B)^{-1}$ where $A, B \in \mathbb{R}^{N \times N}$ and are real and symmetric and $\cdot$ is element-wise multiplication.*

Proof. Consider a counter-example where

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad\qquad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Clearly, A and B are both invertible, but $(A \cdot B)$ is not invertible, and thus we cannot even guarantee existence, much less uniqueness. $\square$

Lemma 6.2. *For two vectors $x, y \in \mathbb{R}^N$ and a square matrix $A \in \mathbb{R}^{N \times N}$, then*

$$(yx^\top) \cdot A = D_y A D_x^\top$$

where $D_y \in \mathbb{R}^{N \times N}$ matrix the elements of y along the diagonal and 0 elsewhere.

Proof. Omitted. Found on Wikipedia page for Hadamard products. $\square$

Note 6.1 (Multihorizon Covariance Matrix)**.** *We want to analyze equation 3.*

Solution. By Lemma 6.2, each term in the summation simplifes such that

$$\left(\left(\sum_{l=1}^{S} (W_l W_l^\top) \right) \cdot \mathrm{Var}(F_t) \right)^{-1} = \left(\sum_{l=1}^{S} D_{W_l} \mathrm{Var}(F_t) D_{W_l} \right)^{-1}$$

Now consider the eigendecomposition of the covariance matrix (which is positive semi-definite) $\mathrm{Var}(F_t) = Q \Lambda Q^\top$ where the columns of Q are the eigenvectors.

$$\left(\sum_{l=1}^{S} D_{W_l} Q \Lambda Q^\top D_{W_l} \right)^{-1}$$

If all of the eigenvalues in Λ are distinct (which should be reasonable despite the lack of orthogonality between the factors in PARAFAC), then Q is orthogonal and by definition, $Q^\top = Q^{-1}$. Thus, $Q^\top D_{W_l} = Q^{-1} D_{W_l}$ maps each of the columns of D_{W_l} onto the eigenspace spanned by the eigenvectors of the covariance matrix.

The mapping of D_{W_l} onto the eigenspace spanned by the covariance matrix's eigenvectors suggests a powerful economic interpretation. This transformation effectively decomposes the investment horizon weights into the same coordinate system where risk factors naturally vary. This means we're aligning our investment horizons with the fundamental modes of variation in the underlying factors.

Let $X_l = Q^T D_{W_l}$, which is now neither diagonal orthogonal, or symmetric. Now we have

$$\left(\sum_{l=1}^{S} X_l^\top \Lambda X_l \right)^{-1}$$

Note that for some l, $X_l^\top \Lambda X_l$ is a positive semi-definite matrix because the eigenvalues in Λ are guaranteed to be nonnegative, and furthermore, the sum of PSD matrices is still PSD. The fact that the sum of PSD matrices remains PSD ensures that the portfolio's risk remains well-defined and manageable. This is crucial in mean-variance optimization, as it guarantees that the covariance matrix used in the optimization process is valid. □

Note 6.2 (PSD Stability)**.** *Solution.* Recall that the covariance matrix is PSD, so eigenvalues are all nonnegative, and Λ is diagonal and therefore symmetric. The sum of PSD matrices is still PSD. Alternatively, just from 3 each element in the outer product is positive so when multiplied by the covariance matrix, one would propose that the result was PSD but this shows it.

We have the following equality

Further exploring, this could be a deadend, is there some nice assumption

$$\begin{aligned}\left(\sum_{l=1}^{S} X_l^\top \Lambda X_l\right)^{(i,j)} &= \sum_{k=1}^{K} \Lambda_k \left(\sum_{l=1}^{S} X_l^{(k,i)} X_l^{(k,j)}\right) \\ &= \sum_{k=1}^{K} \Lambda_k\end{aligned}$$

Each $X_l^{(i,j)} = \sum_{k=1}^{K} Q^{(k,i)} D_{W_l}^{(k,j)}$ □