# Robust Forecasting by Regularization

Dobrislav Dobrev, Federal Reserve Board of Governors
Ernst Schaumburg, Federal Reserve Bank of New York

$7^{\text{th}}$ Workshop on Forecasting Techniques

ECB
May 4-5, 2012

- Forecasting of multivariate outcomes with many regressors
  - Forecast vector of $m$ outcomes $Y_{t+h}$
  - Large number $n$ of potential predictors $X_t$
  - $T$ observations

$$\underset{T\times m}{\mathbf{Y}} = \underset{T\times n}{\mathbf{X}} \underset{n\times m}{\Theta} + \underset{T\times m}{\mathbf{e}}$$

- Features of many finance and macro applications:
  - Large panels, $n, T \to \infty$, with $n \approx T$
  - Ill conditioned design matrix $\mathbf{X}$
  - Strong correlation structure in outcomes Y
- Goal: Extract from X a parsimonious set of predictors for all Y

1. Very large number of potential regressors $\Rightarrow$ in-sample over-fitting!
    - Variable selection problem $\Rightarrow$ Curse of dimensionality $(\sim 2^n)$
    - Factor selection problem $\Rightarrow$ Lack of Interpretability
2. Nearly collinear regressors $\Rightarrow$ Shrinkage estimate of $S_{XX}$
    - Bayesian shrinkage estimation
    - Penalized Estimation / Regularization
3. Strong factor structure in outcomes $\Rightarrow$ Shrinkage estimate of “$S_{XY}$”
    - Reduced Rank Regression

- Combine the two types of shrinkage in new estimator

*Regularized Reduced Rank Regression*

  - a.k.a. Penalized Reduced Rank Regression
  - Reduced Rank $\Rightarrow$ Exploit $S_{XY}$ info to determine forecasting factors
  - Penalty scheme $\Leftrightarrow$ Regularization Scheme ($\Leftrightarrow$ Bayesian prior on $\Theta$) $\Rightarrow$ Overcome ill-conditioned $S_{XX}$

- Two key “shrinkage” parameters:
  - Degree of regularization (“$\rho$”)
  - Number of forecasting factors (“$k$”)

- Empirical illustrations - Out-of-Sample Forecasting:
  - Quarterly Macro Series (n=143, T=195)
  - Monthly Bond Excess Returns (n=15, T=468)

$$\mathbf{Y} \quad = \quad \mathbf{X}\Theta + \mathbf{e}$$

- Properties of the linear system determined by the SVD of $\mathbf{X}$:

$$\underset{T\times n}{\mathbf{X}} = \underset{T\times n}{U}\,\underset{n\times n}{\Sigma}\,\underset{n\times n}{V'} = \sum_{i=1}^{n} \sigma_i u_i v_i'$$

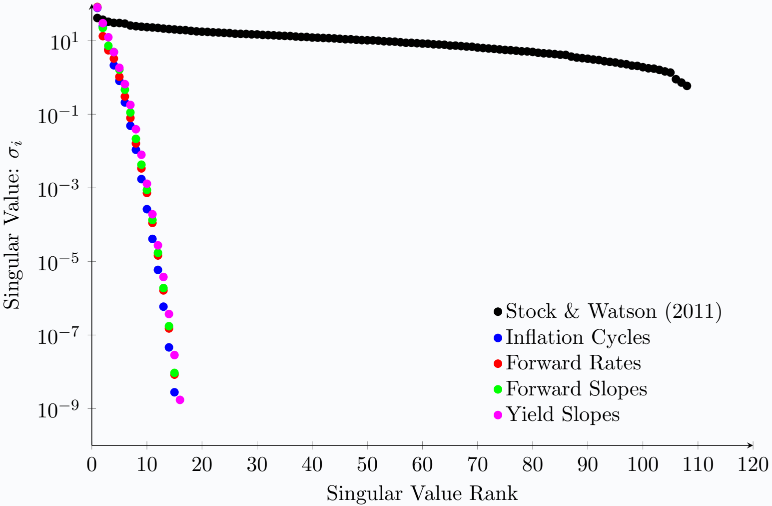
- The OLS estimate is

$$\hat{\Theta}_{OLS} = \sum_{i=0}^{n} v_i \frac{u_i'\mathbf{Y}}{\sigma_i} = \Theta_0 + \sum_{i=0}^{n} v_i \frac{u_i'\mathbf{e}}{\sigma_i}$$

- When $\mathbf{X}$ is ill-conditioned (i.e. $\sigma_{max}/\sigma_{min}$ large), $\hat{\Theta}_{OLS}$ is unstable

$$E[\mathbf{e}'\mathbf{e}] = \kappa^2 I_m \Rightarrow E\|\hat{\Theta}_{OLS} - \Theta_0\|^2 = \kappa^2\, tr\{(\mathbf{X}'\mathbf{X})^{-1}\} = \kappa^2 \sum_{i=1}^{n} \sigma_i^{-2}$$

- In many applications, the design matrix $\mathbf{X}$ is ill-conditioned



- Regularization introduces filter factors: $0 \leq f_i \leq 1$

$$\tilde{\Theta} \;=\; \sum_{i=1}^{n} f_i v_i \left(\frac{u_i'\mathbf{Y}}{\sigma_i}\right), \quad \|\tilde{\Theta}\|^2 = \sum_{i=1}^{n} f_i^2 \left(\frac{u_i'\mathbf{Y}}{\sigma_i}\right)^2 \leq \|\hat{\Theta}_{OLS}\|$$

such that $f_i/\sigma_i \approx 0$ for large $i$ (smallest singular values)

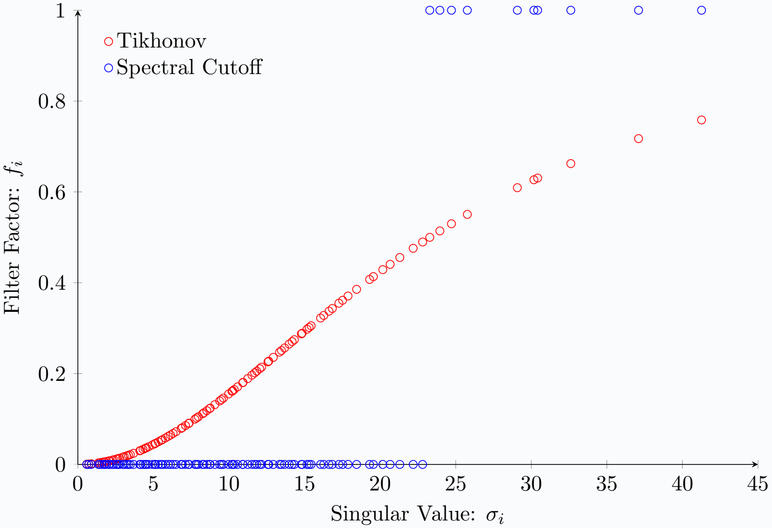
- Let $\tilde{\Theta}_\infty \equiv p\lim_{T\to\infty} \tilde{\Theta} \neq \Theta_0 \Rightarrow$ Bias-Variance Trade-off

$$\underbrace{E\|\Theta_0 - \tilde{\Theta}\|}_{\substack{\text{root mean squared} \\ \text{shrinkage estimation error}}} \leq \underbrace{\|(\Theta_0 - \tilde{\Theta}_\infty\|}_{\text{bias due to regularization}} + \underbrace{E\|(\tilde{\Theta}_\infty - \tilde{\Theta}\|}_{\substack{\text{dampened variance} \\ \text{due to regularization}}}$$

- Common shrinkage estimators are given by a set of filter factors:
  - Tikhonov (Ridge Regression): $f_i = \frac{\sigma_i^2}{\sigma_i^2 + \rho^2}$
  - Spectral Truncation (Princ. Comp. Regr.): $f_i = \mathbf{1}_{\{\sigma_i \geq \rho\}}$ or $f_i = \mathbf{1}_{\{i \leq r\}}$

# Robust Forecasting via Regularization

## Regularizing $S_{XX}$ – The Stock & Watson (2011) Macro Series



- Reduced rank regression imposes $\text{Rank}(\Theta) = k < \min(n, m)$
  - Reparametrize $\Theta = AB$ where $A \in \mathbb{R}^{n \times k}, B \in \mathbb{R}^{k \times m}$

$$\min_{\{A,B\}} \|(\mathbf{Y} - \mathbf{X}AB)W^{\frac{1}{2}}\|^2$$

- Interpretation: $k$ factors $A'X$ jointly forecast $Y$
- Optimal $A^*$ minimizes conditional variance $\text{Var}[W^{1/2}Y|A'X]$:
  - Solution is $k$ principal eigenvectors of the Generalized Eigenvalue problem:

$$|S_{XY}WS_{YX} - \lambda S_{XX}| = 0$$

  - *When $S_{XX}$ is ill-conditioned, the solution is highly unstable*
  - Note: In PCR, $A$ is chosen to minimize $\text{Var}[X|A'X]$

- Idea: Penalize the Reduced Rank Regression objective

$$\min_{\{A,B\}} \|(\mathbf{Y} - \mathbf{X}AB)W^{\frac{1}{2}}\|^2 + \rho^2 \|R(AB)W^{\frac{1}{2}}\|^2$$

  - $\rho$ is a scalar regularization parameter
  - $R$ is an optional matrix parameter
  - $R \neq I$ differentially penalizes certain directions in parameter space

  Solution is given by the Generalized Eigenvalue problem:

$$|S_{XY}WS_{YX} - \lambda(S_{XX} + \rho^2 R'R)| = 0$$

  *Note:* Only $S_{XX}$ is regularized while $\mathbf{X}$ enters unaltered in $S_{XY}$

- Each choice of $(\rho, R)$ leads to a different regularization scheme
- In this paper: *Tikhonov* and *Spectral Truncation*.

**Tikhonov Regularization:**

$$\min_{\{A,B\}} \|(\mathbf{Y} - \mathbf{X}AB)W^{\frac{1}{2}}\|^2 + \rho^2 \|(AB)W^{\frac{1}{2}}\|^2$$

- Let $\rho > 0$ and $R = I_n$

  The optimal A is given by the k principal eigenvectors of:

$$|S_{XY}WS_{YX} - \lambda(S_{XX} + \rho^2 I_n)| = 0$$

- Penalty analogous to multivariate Ridge Regression
  - The penalty is on the norm of $AB$ *not* $A$ and $B$ separately
  - *To our knowledge, the formulation of the corresponding Bayesian prior on the subspace of reduced rank matrices is unknown*

## Spectral Truncation Regularization:

$$\min_{\{A,B\}} \|(\mathbf{Y} - \mathbf{X}AB)W^{\frac{1}{2}}\|^2 + \rho^2\|R(AB)W^{\frac{1}{2}}\|^2$$

- Mimic PCR by putting infinite penalty on the appropriate subspace:
  - Let $R = [V_{r+1}, \ldots, V_n]'$ (the last $n - r$ right singular vectors of $\mathbf{X}$)
  - Let $\rho \to \infty$
- The optimal $A$ is given by the $k$ principal eigenvectors of:

$$\begin{aligned} & |S^{\dagger}_{XX} S_{XY} W S_{YX} - \lambda I_N| \\ S^{\dagger}_{XX} &= V\, diag(\sigma_1^{-2}, \ldots, \sigma_r^{-2}, 0, \ldots, 0)\, V' \end{aligned}$$

- Corresponding (improper) Bayesian prior sets Pr$\{Span(V_1, \ldots, V_r)\} = 1$

- When regularizing $S_{XX}$ we need to decide on a threshold for distinguishing between large and small singular values
- Key choice variables: Regularization parameter $\rho$ (Tikhonov) or $r$ (Spectral Truncation)

- We consider both data driven and ex-ante fixed choices
- Spectral Truncation
  - Fixed choice of $r$ – **RRRRk-PCr**
    - no dependence of empirical distribution of singular values
  - Data driven – **RRRRk-SMP**
    - Threshold based on random matrix theory

- Tikhonov
  - Data driven – **RRRRk-TMP**
    - Threshold based on random matrix theory

**Stylized "signal+noise" Model:**

- Assume a $r$ dimensional factor structure for X

$$\mathbf{X} = \mathbf{F}\Lambda + \mathbf{E}, \quad \Lambda \in \mathbb{R}^{r \times n}$$

- Assume that $\mathbf{F}$ contains the $k \leq r$ relevant forecasting factors for Y
- If we can identify $r$ then we can attempt to "kill" or down-weight the directions in parameter space corresponding to the noise

$$S_{XX} = \underbrace{\Lambda'\Lambda}_{\text{Signal}} + \underbrace{\Omega_n}_{\text{Noise}}$$

## $\mathbf{H_0}$ : No factor structure

- Assume noise i.i.d. zero mean and variance 1 with finite $4^{th}$ moment
- Assume $n, T \to \infty$ with $n/T \to \gamma \in (0;1)$

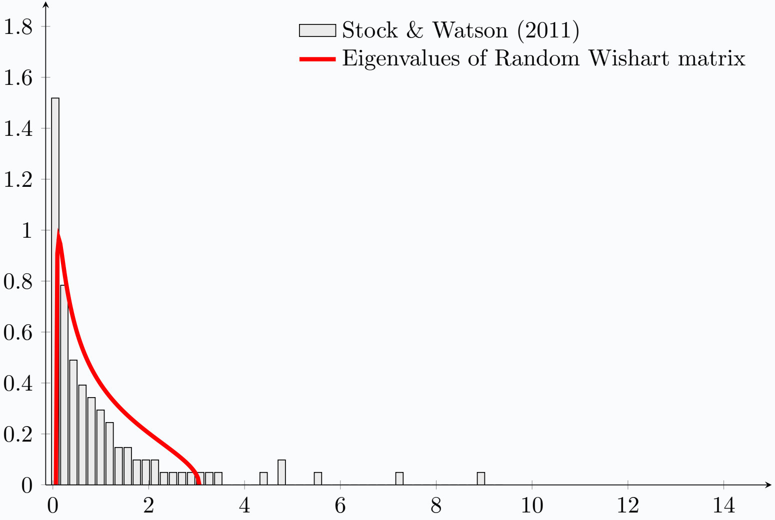
### Theorem (Marcenko & Pastur 1967)

*The limiting distribution of the spectrum of $\Omega_n = \frac{1}{T}\mathbf{E}'\mathbf{E}$ is given by the measure $\mu_\Omega$ with support on the interval $[(1-\sqrt{\gamma})^2; (1+\sqrt{\gamma})^2]$ and density:*

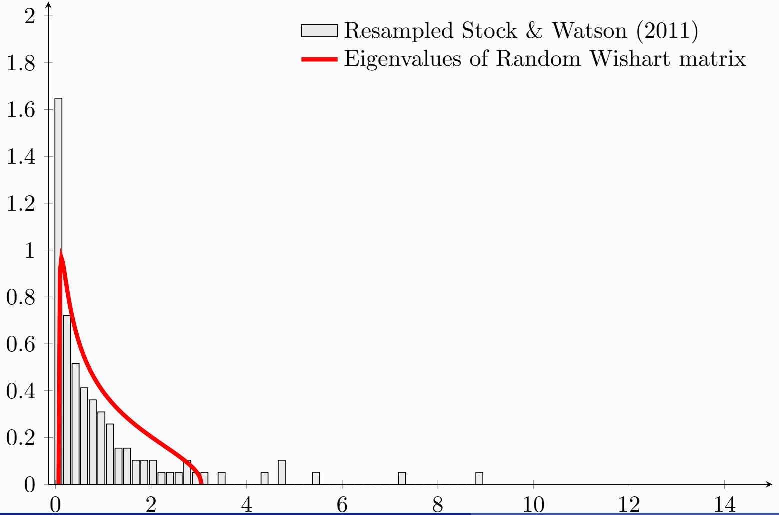
$$d\mu_\Omega(x) = \frac{\sqrt{(x-b_-)(b_+-x)}}{2\pi\gamma x}\,dx \text{ , where } b_\pm = (1 \pm \sqrt{\gamma})^2 \tag{1}$$

- Any eigenvalue smaller than $(1+\sqrt{\gamma})^2$ is “small” (asymptotically) and should be down-weighted

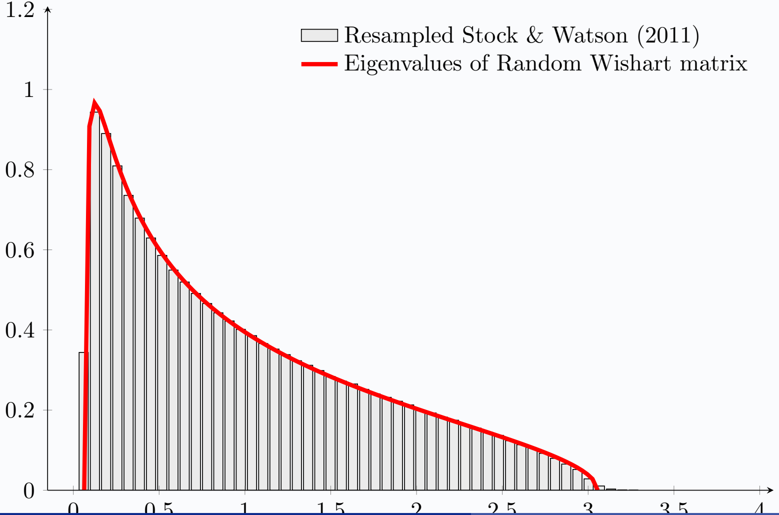
- Empirical spectrum of $S_{XX}$ for the Stock & Watson (2011) data



- AR(12) filtering of $\mathbf{X}$ breaks time-series dependence but spectrum unchanged



- Resampling each series independently breaks cross-sectional dependence and restores MP theory!



# Empirical applications
## Model taxonomy

| # Regressor Components | # Forecasting Factors 1 | 2 | 3 | 4 | 5 | ... | m |
|---|---|---|---|---|---|---|---|
| **Panel A: Fixed Number of Regressor Components** | | | | | | | |
| **1** | **PCR-1** | | | | | | RR-PC1<br>PLS-1<br>3PRF-1 |
| **2** | RRRR1-PC2 | **PCR-2** | | | | | RR-PC2<br>PLS-2<br>3PRF-2 |
| **3** | RRRR1-PC3 | RRRR2-PC3 | **PCR-3** | | | | RR-PC3<br>PLS-3<br>3PRF-3 |
| **4** | RRRR1-PC4 | RRRR2-PC4 | RRRR3-PC4 | **PCR-4** | | | RR-PC4<br>PLS-4<br>3PRF-4 |
| **5** | RRRR1-PC5 | RRRR2-PC5 | RRRR3-PC5 | RRRR4-PC5 | **PCR-5** | | RR-PC5<br>PLS-5<br>3PRF-5 |
| **...** | ... | ... | ... | ... | ... | **...** | ... |
| **n** | RRRR1-PCn | RRRR2-PCn | RRRR3-PCn | RRRR4-PCn | RRRR5-PCn | ... | **OLS** |
| **Panel B: Data Driven Number of Regressor Components** | | | | | | | |
| **MP MAX Spectral** | RRRR1-SMP | RRRR2-SMP | RRRR3-SMP | RRRR4-SMP | RRRR5-SMP | ... | RR-SMP |
| **MP MAX Tikhonov** | RRRR1-TMP | RRRR2-TMP | RRRR3-TMP | RRRR4-TMP | RRRR5-TMP | ... | RR-TMP |

- Replication of the setup in Stock & Watson (2011):
  - 35 aggregates and 108 disaggregates for a total of 143 series
  - 195 quarterly observations from 1960:Q2 through 2008:Q4
  - Rolling out-of-sample forecasts for window size 100 quarters
  - Five-factor benchmark obtained via PCR
- Main findings:
  - RR-SMP and RRRR5-SMP are notable competitors to PCR-5
  - MP-implied cutoff has a median of 5 and varies from 3 to 8
  - MP rationalizes PCR-5 benchmark and little room for improvement
  - Mixed results for all other methods across different series
  - Trade-off between better left and worse right tail of RMSE distribution

- Bond returns largely driven by 1 or 2 common forecasting factors:
  - using forwards by Cochrane & Piazzesi (2005)
  - using maturity "cycles" by Cieslak & Povala (2011)
- Ideal setting for applying our RRRR models:
  - Multiple outcomes explained by few common factors
  - Factors constructed from ill-conditioned set of regressors
- Our application:
  - GSW-interpolated yield curve data
  - 14 excess returns of bonds with maturity from 2 to 15 years
  - 468 *monthly* observations from 1972 to 2010
  - Rolling *out-of-sample* forecasts for window size 120 months
  - Single-factor benchmark extracted from forwards or cycles

- Five forecasting exercises imposing common (1 or 2) factors:
  1. Based on forwards
  2. Based on forward slopes
  3. Based on cycles
  4. Based on yield slopes
  5. Based on yield slopes and cycles
- Report out-of-sample $R^2$ vis-a-vis rolling average benchmark for each forecasted series

# Forecasting bond excess returns

## Forwards

| Out-of-sample $R^2$ | Bond Excess Returns | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Models** | $rx^{(2)}$ | $rx^{(3)}$ | $rx^{(4)}$ | $rx^{(5)}$ | $rx^{(6)}$ | $rx^{(7)}$ | $rx^{(8)}$ | $rx^{(9)}$ | $rx^{(10)}$ | $rx^{(11)}$ | $rx^{(12)}$ | $rx^{(13)}$ | $rx^{(14)}$ | $rx^{(15)}$ |
| Panel A: Naïve benchmark models | | | | | | | | | | | | | | |
| Rolling Average | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Random Walk | -0.514 | -0.585 | -0.640 | -0.686 | -0.728 | -0.765 | -0.797 | -0.824 | -0.846 | -0.862 | -0.874 | -0.881 | -0.883 | -0.882 |
| Panel B: Models with 1 forecasting factor | | | | | | | | | | | | | | |
| PCR-1 | -0.003 | -0.010 | -0.013 | -0.014 | -0.014 | -0.013 | -0.012 | -0.011 | -0.010 | -0.009 | -0.007 | -0.006 | -0.006 | -0.005 |
| RRRR1-PC2 | -0.020 | -0.024 | -0.027 | -0.029 | -0.030 | -0.031 | -0.031 | -0.031 | -0.032 | -0.032 | -0.032 | -0.032 | -0.032 | -0.032 |
| RRRR1-PC3 | -0.001 | -0.002 | -0.001 | 0.000 | 0.000 | -0.001 | -0.003 | -0.006 | -0.009 | -0.012 | -0.014 | -0.017 | -0.018 | -0.020 |
| RRRR1-PC4 | -0.080 | -0.066 | -0.058 | -0.053 | -0.050 | -0.048 | -0.046 | -0.045 | -0.044 | -0.043 | -0.042 | -0.042 | -0.041 | -0.040 |
| RRRR1-PC5 | -0.097 | -0.084 | -0.077 | -0.073 | -0.070 | -0.067 | -0.064 | -0.062 | -0.059 | -0.056 | -0.053 | -0.049 | -0.047 | -0.044 |
| RRRR1-SMP | -0.011 | -0.020 | -0.024 | -0.026 | -0.026 | -0.026 | -0.025 | -0.023 | -0.021 | -0.020 | -0.018 | -0.017 | -0.016 | -0.014 |
| RRRR1-TMP | 0.001 | -0.005 | -0.008 | -0.009 | -0.010 | -0.010 | -0.010 | -0.010 | -0.010 | -0.010 | -0.010 | -0.010 | -0.010 | -0.010 |
| Panel C: Models with 2 forecasting factors | | | | | | | | | | | | | | |
| PCR-2 | -0.026 | -0.034 | -0.037 | -0.038 | -0.036 | -0.034 | -0.031 | -0.028 | -0.026 | -0.023 | -0.021 | -0.019 | -0.017 | -0.016 |
| RRRR2-PC3 | -0.011 | -0.010 | -0.006 | -0.002 | 0.000 | -0.001 | -0.004 | -0.008 | -0.012 | -0.017 | -0.020 | -0.023 | -0.026 | -0.027 |
| RRRR2-PC4 | -0.066 | -0.053 | -0.045 | -0.041 | -0.041 | -0.043 | -0.048 | -0.054 | -0.060 | -0.066 | -0.070 | -0.074 | -0.076 | -0.077 |
| RRRR2-PC5 | -0.083 | -0.065 | -0.054 | -0.050 | -0.051 | -0.055 | -0.062 | -0.070 | -0.078 | -0.085 | -0.091 | -0.095 | -0.098 | -0.099 |
| RRRR2-SMP | -0.027 | -0.035 | -0.039 | -0.039 | -0.038 | -0.035 | -0.033 | -0.030 | -0.027 | -0.024 | -0.022 | -0.020 | -0.019 | -0.017 |
| RRRR2-TMP | -0.001 | -0.004 | -0.001 | 0.004 | 0.007 | 0.009 | 0.009 | 0.008 | 0.006 | 0.004 | 0.002 | 0.000 | -0.001 | -0.002 |
| Panel D: Models with 14 forecasting factors | | | | | | | | | | | | | | |
| RR-SMP | -0.013 | -0.023 | -0.027 | -0.028 | -0.028 | -0.027 | -0.025 | -0.023 | -0.021 | -0.019 | -0.017 | -0.015 | -0.014 | -0.013 |
| RR-TMP | 0.012 | 0.003 | -0.006 | -0.011 | -0.012 | -0.012 | -0.012 | -0.012 | -0.012 | -0.013 | -0.015 | -0.015 | -0.011 | -0.010 |
| PLS-1 | -0.002 | -0.006 | -0.019 | -0.020 | -0.019 | -0.017 | -0.011 | -0.011 | -0.010 | -0.010 | -0.010 | -0.010 | -0.005 | -0.005 |
| PLS-2 | -0.051 | -0.055 | -0.056 | -0.056 | -0.054 | -0.052 | -0.052 | -0.051 | -0.051 | -0.046 | -0.043 | -0.043 | -0.041 | -0.029 |
| PLS-3 | -0.086 | -0.058 | -0.032 | -0.018 | -0.012 | -0.017 | -0.017 | -0.018 | -0.017 | -0.026 | -0.038 | -0.046 | -0.051 | -0.057 |
| PLS-4 | -0.102 | -0.068 | -0.049 | -0.045 | -0.044 | -0.047 | -0.055 | -0.063 | -0.073 | -0.077 | -0.080 | -0.084 | -0.085 | -0.080 |
| PLS-5 | -0.128 | -0.093 | -0.081 | -0.073 | -0.066 | -0.067 | -0.072 | -0.082 | -0.092 | -0.098 | -0.096 | -0.100 | -0.102 | -0.102 |
| 3PRF-1 | -0.054 | -0.043 | -0.033 | -0.026 | -0.020 | -0.019 | -0.019 | -0.020 | -0.020 | -0.020 | -0.020 | -0.022 | -0.021 | -0.011 |
| 3PRF-2 | -0.088 | -0.044 | -0.014 | 0.001 | -0.001 | 0.004 | 0.012 | 0.014 | 0.009 | 0.001 | -0.007 | -0.014 | -0.024 | -0.032 |
| 3PRF-3 | -0.121 | -0.082 | -0.054 | -0.040 | -0.039 | -0.037 | -0.038 | -0.043 | -0.049 | -0.053 | -0.056 | -0.059 | -0.058 | -0.059 |
| 3PRF-4 | -0.118 | -0.076 | -0.056 | -0.054 | -0.053 | -0.049 | -0.055 | -0.065 | -0.067 | -0.067 | -0.067 | -0.066 | -0.064 | -0.063 |
| 3PRF-5 | -0.160 | -0.124 | -0.102 | -0.089 | -0.084 | -0.081 | -0.084 | -0.085 | -0.087 | -0.086 | -0.080 | -0.081 | -0.082 | -0.082 |
| OLS | -0.567 | -0.465 | -0.413 | -0.387 | -0.377 | -0.378 | -0.385 | -0.395 | -0.404 | -0.412 | -0.417 | -0.421 | -0.424 | -0.427 |

# Forecasting bond excess returns

## Forward slopes

| Out-of-sample $R^2$ | Bond Excess Returns | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Models** | $rx^{(2)}$ | $rx^{(3)}$ | $rx^{(4)}$ | $rx^{(5)}$ | $rx^{(6)}$ | $rx^{(7)}$ | $rx^{(8)}$ | $rx^{(9)}$ | $rx^{(10)}$ | $rx^{(11)}$ | $rx^{(12)}$ | $rx^{(13)}$ | $rx^{(14)}$ | $rx^{(15)}$ |
| Panel A: Naïve benchmark models | | | | | | | | | | | | | | |
| Rolling Average | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Random Walk | -0.514 | -0.585 | -0.640 | -0.686 | -0.728 | -0.765 | -0.797 | -0.824 | -0.846 | -0.862 | -0.874 | -0.881 | -0.883 | -0.882 |
| Panel B: Models with 1 forecasting factor | | | | | | | | | | | | | | |
| PCR-1 | 0.005 | 0.006 | 0.008 | 0.010 | 0.012 | 0.014 | 0.015 | 0.016 | 0.017 | 0.018 | 0.018 | 0.019 | 0.019 | 0.019 |
| RRRR1-PC2 | 0.046 | 0.035 | 0.029 | 0.026 | 0.023 | 0.021 | 0.019 | 0.017 | 0.014 | 0.012 | 0.009 | 0.007 | 0.004 | 0.002 |
| RRRR1-PC3 | 0.069 | 0.058 | 0.054 | 0.052 | 0.050 | 0.047 | 0.044 | 0.041 | 0.037 | 0.033 | 0.029 | 0.026 | 0.022 | 0.019 |
| RRRR1-PC4 | -0.017 | -0.009 | 0.002 | 0.010 | 0.017 | 0.021 | 0.023 | 0.024 | 0.023 | 0.023 | 0.021 | 0.020 | 0.018 | 0.017 |
| RRRR1-PC5 | -0.047 | -0.041 | -0.031 | -0.022 | -0.015 | -0.010 | -0.007 | -0.004 | -0.003 | -0.002 | -0.002 | -0.002 | -0.002 | -0.003 |
| RRRR1-SMP | 0.006 | 0.006 | 0.008 | 0.011 | 0.013 | 0.014 | 0.016 | 0.017 | 0.018 | 0.018 | 0.019 | 0.019 | 0.020 | 0.020 |
| RRRR1-TMP | 0.044 | 0.039 | 0.037 | 0.037 | 0.037 | 0.037 | 0.037 | 0.037 | 0.037 | 0.037 | 0.036 | 0.035 | 0.035 | 0.034 |
| Panel C: Models with 2 forecasting factors | | | | | | | | | | | | | | |
| PCR-2 | 0.014 | 0.011 | 0.011 | 0.013 | 0.015 | 0.017 | 0.018 | 0.018 | 0.018 | 0.017 | 0.017 | 0.016 | 0.015 | 0.015 |
| RRRR2-PC3 | 0.059 | 0.048 | 0.044 | 0.042 | 0.041 | 0.040 | 0.039 | 0.038 | 0.037 | 0.037 | 0.037 | 0.037 | 0.037 | 0.038 |
| RRRR2-PC4 | -0.007 | -0.001 | 0.008 | 0.015 | 0.019 | 0.020 | 0.019 | 0.016 | 0.013 | 0.009 | 0.006 | 0.003 | 0.000 | -0.003 |
| RRRR2-PC5 | -0.041 | -0.030 | -0.017 | -0.010 | -0.007 | -0.008 | -0.011 | -0.015 | -0.020 | -0.024 | -0.028 | -0.032 | -0.035 | -0.037 |
| RRRR2-SMP | 0.016 | 0.012 | 0.012 | 0.014 | 0.016 | 0.018 | 0.019 | 0.019 | 0.019 | 0.019 | 0.018 | 0.017 | 0.016 | 0.016 |
| RRRR2-TMP | 0.044 | 0.039 | 0.037 | 0.037 | 0.037 | 0.037 | 0.037 | 0.037 | 0.037 | 0.037 | 0.036 | 0.035 | 0.035 | 0.034 |
| Panel D: Models with 14 forecasting factors | | | | | | | | | | | | | | |
| RR-SMP | 0.006 | 0.006 | 0.008 | 0.011 | 0.013 | 0.014 | 0.016 | 0.017 | 0.018 | 0.018 | 0.019 | 0.019 | 0.020 | 0.020 |
| RR-TMP | 0.075 | 0.065 | 0.060 | 0.055 | 0.049 | 0.044 | 0.040 | 0.036 | 0.033 | 0.032 | 0.030 | 0.029 | 0.029 | 0.029 |
| PLS-1 | 0.021 | 0.019 | 0.033 | 0.018 | 0.020 | 0.019 | 0.020 | 0.020 | 0.020 | 0.020 | 0.020 | 0.020 | 0.020 | 0.020 |
| PLS-2 | 0.024 | 0.025 | 0.026 | 0.027 | 0.027 | 0.027 | 0.028 | 0.028 | 0.028 | 0.029 | 0.030 | 0.031 | 0.033 | 0.035 |
| PLS-3 | 0.016 | 0.021 | 0.026 | 0.029 | 0.029 | 0.028 | 0.027 | 0.026 | 0.027 | 0.028 | 0.030 | 0.032 | 0.033 | 0.032 |
| PLS-4 | -0.054 | -0.024 | -0.003 | 0.008 | 0.013 | 0.014 | 0.012 | 0.000 | -0.005 | -0.012 | -0.020 | -0.024 | -0.024 | -0.018 |
| PLS-5 | -0.089 | -0.046 | -0.026 | -0.025 | -0.030 | -0.032 | -0.036 | -0.028 | -0.028 | -0.030 | -0.032 | -0.038 | -0.041 | -0.031 |
| 3PRF-1 | -0.020 | -0.019 | -0.014 | -0.013 | -0.013 | -0.013 | -0.012 | -0.013 | -0.016 | -0.019 | -0.022 | -0.024 | -0.023 | -0.023 |
| 3PRF-2 | -0.046 | -0.034 | -0.022 | -0.016 | -0.011 | -0.011 | -0.012 | -0.014 | -0.016 | -0.015 | -0.013 | -0.010 | -0.008 | -0.008 |
| 3PRF-3 | -0.098 | -0.063 | -0.038 | -0.025 | -0.018 | -0.020 | -0.019 | -0.026 | -0.031 | -0.037 | -0.043 | -0.050 | -0.055 | -0.056 |
| 3PRF-4 | -0.110 | -0.072 | -0.049 | -0.029 | -0.029 | -0.031 | -0.033 | -0.047 | -0.061 | -0.065 | -0.071 | -0.074 | -0.080 | -0.075 |
| 3PRF-5 | -0.123 | -0.089 | -0.067 | -0.055 | -0.048 | -0.050 | -0.057 | -0.064 | -0.071 | -0.076 | -0.087 | -0.084 | -0.082 | -0.081 |
| OLS | -0.510 | -0.397 | -0.335 | -0.301 | -0.287 | -0.284 | -0.289 | -0.298 | -0.306 | -0.314 | -0.319 | -0.321 | -0.323 | -0.323 |

# Forecasting bond excess returns

## Cycles of Cieslak & Povala (2011)

| Out-of-sample $R^2$ | Bond Excess Returns | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Models** | $rx^{(2)}$ | $rx^{(3)}$ | $rx^{(4)}$ | $rx^{(5)}$ | $rx^{(6)}$ | $rx^{(7)}$ | $rx^{(8)}$ | $rx^{(9)}$ | $rx^{(10)}$ | $rx^{(11)}$ | $rx^{(12)}$ | $rx^{(13)}$ | $rx^{(14)}$ | $rx^{(15)}$ |
| Panel A: Naïve benchmark models | | | | | | | | | | | | | | |
| Rolling Average | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Random Walk | -0.514 | -0.585 | -0.640 | -0.686 | -0.728 | -0.765 | -0.797 | -0.824 | -0.846 | -0.862 | -0.874 | -0.881 | -0.883 | -0.882 |
| Panel B: Models with 1 forecasting factor | | | | | | | | | | | | | | |
| PCR-1 | 0.015 | 0.019 | 0.023 | 0.026 | 0.028 | 0.030 | 0.031 | 0.032 | 0.033 | 0.034 | 0.034 | 0.035 | 0.036 | 0.037 |
| RRRR1-PC2 | 0.021 | 0.026 | 0.030 | 0.034 | 0.038 | 0.041 | 0.044 | 0.047 | 0.050 | 0.052 | 0.053 | 0.054 | 0.055 | 0.056 |
| RRRR1-PC3 | 0.039 | 0.032 | 0.031 | 0.033 | 0.035 | 0.038 | 0.041 | 0.044 | 0.047 | 0.049 | 0.051 | 0.052 | 0.053 | 0.055 |
| RRRR1-PC4 | 0.037 | 0.031 | 0.032 | 0.035 | 0.040 | 0.044 | 0.048 | 0.051 | 0.054 | 0.056 | 0.057 | 0.058 | 0.059 | 0.060 |
| RRRR1-PC5 | -0.048 | -0.037 | -0.025 | -0.014 | -0.005 | 0.004 | 0.011 | 0.018 | 0.023 | 0.027 | 0.030 | 0.032 | 0.034 | 0.035 |
| RRRR1-SMP | 0.018 | 0.024 | 0.030 | 0.035 | 0.040 | 0.043 | 0.047 | 0.050 | 0.052 | 0.055 | 0.057 | 0.058 | 0.060 | 0.061 |
| RRRR1-TMP | 0.011 | 0.019 | 0.026 | 0.031 | 0.036 | 0.040 | 0.044 | 0.047 | 0.050 | 0.052 | 0.054 | 0.056 | 0.057 | 0.059 |
| OLS with 1 cycle | 0.037 | 0.041 | 0.046 | 0.050 | 0.054 | 0.057 | 0.060 | 0.062 | 0.064 | 0.065 | 0.066 | 0.067 | 0.068 | 0.069 |
| Panel C: Models with 2 forecasting factors | | | | | | | | | | | | | | |
| PCR-2 | 0.011 | 0.017 | 0.024 | 0.030 | 0.035 | 0.040 | 0.044 | 0.047 | 0.050 | 0.052 | 0.054 | 0.055 | 0.056 | 0.057 |
| RRRR2-PC3 | 0.021 | 0.023 | 0.028 | 0.033 | 0.038 | 0.042 | 0.045 | 0.047 | 0.048 | 0.048 | 0.048 | 0.048 | 0.047 | 0.047 |
| RRRR2-PC4 | 0.034 | 0.029 | 0.030 | 0.033 | 0.038 | 0.042 | 0.046 | 0.050 | 0.052 | 0.055 | 0.056 | 0.058 | 0.059 | 0.060 |
| RRRR2-PC5 | -0.051 | -0.039 | -0.025 | -0.013 | -0.003 | 0.004 | 0.009 | 0.012 | 0.014 | 0.016 | 0.016 | 0.017 | 0.017 | 0.018 |
| RRRR2-SMP | 0.006 | 0.014 | 0.022 | 0.029 | 0.034 | 0.039 | 0.044 | 0.047 | 0.050 | 0.052 | 0.054 | 0.055 | 0.056 | 0.056 |
| RRRR2-TMP | 0.019 | 0.027 | 0.034 | 0.040 | 0.044 | 0.048 | 0.051 | 0.053 | 0.055 | 0.057 | 0.059 | 0.060 | 0.061 | 0.062 |
| OLS with 2 cycles | 0.017 | 0.024 | 0.031 | 0.037 | 0.042 | 0.046 | 0.050 | 0.053 | 0.055 | 0.057 | 0.058 | 0.060 | 0.060 | 0.061 |
| Panel D: Models with 14 forecasting factors | | | | | | | | | | | | | | |
| RR-SMP | 0.019 | 0.025 | 0.030 | 0.035 | 0.040 | 0.043 | 0.047 | 0.049 | 0.052 | 0.054 | 0.056 | 0.058 | 0.060 | 0.061 |
| RR-TMP | 0.010 | 0.018 | 0.025 | 0.031 | 0.036 | 0.040 | 0.044 | 0.047 | 0.050 | 0.052 | 0.054 | 0.055 | 0.057 | 0.058 |
| PLS-1 | 0.014 | 0.020 | 0.025 | 0.029 | 0.032 | 0.035 | 0.037 | 0.039 | 0.040 | 0.042 | 0.043 | 0.045 | 0.046 | 0.047 |
| PLS-2 | 0.022 | 0.024 | 0.029 | 0.033 | 0.037 | 0.041 | 0.045 | 0.048 | 0.050 | 0.051 | 0.053 | 0.056 | 0.057 | 0.057 |
| PLS-3 | 0.019 | 0.024 | 0.035 | 0.042 | 0.044 | 0.045 | 0.045 | 0.045 | 0.047 | 0.052 | 0.056 | 0.060 | 0.061 | 0.061 |
| PLS-4 | -0.002 | 0.012 | 0.023 | 0.030 | 0.037 | 0.044 | 0.052 | 0.055 | 0.058 | 0.061 | 0.062 | 0.067 | 0.062 | 0.057 |
| PLS-5 | -0.095 | -0.052 | -0.027 | -0.010 | -0.014 | -0.007 | -0.001 | 0.000 | 0.001 | 0.004 | 0.007 | 0.013 | 0.019 | 0.026 |
| 3PRF-1 | 0.059 | 0.053 | 0.048 | 0.045 | 0.045 | 0.044 | 0.044 | 0.044 | 0.043 | 0.043 | 0.042 | 0.042 | 0.041 | 0.041 |
| 3PRF-2 | 0.040 | 0.035 | 0.038 | 0.046 | 0.048 | 0.051 | 0.054 | 0.054 | 0.056 | 0.057 | 0.058 | 0.059 | 0.059 | 0.060 |
| 3PRF-3 | 0.020 | 0.026 | 0.035 | 0.040 | 0.047 | 0.053 | 0.055 | 0.059 | 0.062 | 0.063 | 0.064 | 0.062 | 0.064 | 0.068 |
| 3PRF-4 | -0.073 | -0.046 | -0.023 | -0.014 | -0.007 | -0.001 | 0.003 | 0.005 | 0.007 | 0.009 | 0.010 | 0.012 | 0.015 | 0.019 |
| 3PRF-5 | -0.109 | -0.065 | -0.046 | -0.010 | -0.001 | 0.001 | -0.001 | -0.003 | -0.003 | -0.013 | -0.004 | 0.000 | 0.005 | 0.011 |
| OLS with all cycles | -0.542 | -0.430 | -0.365 | -0.328 | -0.308 | -0.300 | -0.297 | -0.298 | -0.299 | -0.298 | -0.296 | -0.293 | -0.289 | -0.284 |

# Forecasting bond excess returns
## Yield slopes

| Out-of-sample $R^2$ | Bond Excess Returns | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Models** | $rx^{(2)}$ | $rx^{(3)}$ | $rx^{(4)}$ | $rx^{(5)}$ | $rx^{(6)}$ | $rx^{(7)}$ | $rx^{(8)}$ | $rx^{(9)}$ | $rx^{(10)}$ | $rx^{(11)}$ | $rx^{(12)}$ | $rx^{(13)}$ | $rx^{(14)}$ | $rx^{(15)}$ |
| Panel A: Naïve benchmark models | | | | | | | | | | | | | | |
| Rolling Average | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Random Walk | -0.514 | -0.585 | -0.640 | -0.686 | -0.728 | -0.765 | -0.797 | -0.824 | -0.846 | -0.862 | -0.874 | -0.881 | -0.883 | -0.882 |
| Panel B: Models with 1 forecasting factor | | | | | | | | | | | | | | |
| PCR-1 | 0.019 | 0.019 | 0.020 | 0.022 | 0.024 | 0.025 | 0.026 | 0.027 | 0.027 | 0.027 | 0.028 | 0.028 | 0.028 | 0.028 |
| RRRR1-PC2 | 0.012 | 0.010 | 0.010 | 0.012 | 0.013 | 0.015 | 0.016 | 0.017 | 0.019 | 0.020 | 0.020 | 0.021 | 0.022 | 0.022 |
| RRRR1-PC3 | 0.053 | 0.039 | 0.034 | 0.032 | 0.030 | 0.030 | 0.029 | 0.029 | 0.029 | 0.029 | 0.029 | 0.030 | 0.031 | 0.032 |
| RRRR1-PC4 | 0.030 | 0.016 | 0.012 | 0.012 | 0.013 | 0.015 | 0.017 | 0.019 | 0.021 | 0.023 | 0.025 | 0.027 | 0.029 | 0.032 |
| RRRR1-PC5 | 0.059 | 0.051 | 0.053 | 0.056 | 0.059 | 0.060 | 0.061 | 0.060 | 0.059 | 0.058 | 0.056 | 0.055 | 0.054 | 0.053 |
| RRRR1-SMP | 0.019 | 0.019 | 0.019 | 0.021 | 0.022 | 0.023 | 0.023 | 0.024 | 0.024 | 0.024 | 0.025 | 0.025 | 0.025 | 0.025 |
| RRRR1-TMP | 0.045 | 0.038 | 0.036 | 0.036 | 0.036 | 0.037 | 0.037 | 0.038 | 0.038 | 0.038 | 0.039 | 0.039 | 0.039 | 0.040 |
| Panel C: Models with 2 forecasting factors | | | | | | | | | | | | | | |
| PCR-2 | 0.009 | 0.007 | 0.009 | 0.011 | 0.013 | 0.015 | 0.017 | 0.017 | 0.018 | 0.018 | 0.018 | 0.018 | 0.018 | 0.018 |
| RRRR2-PC3 | 0.030 | 0.024 | 0.024 | 0.025 | 0.027 | 0.029 | 0.030 | 0.031 | 0.031 | 0.032 | 0.033 | 0.033 | 0.034 | 0.035 |
| RRRR2-PC4 | 0.034 | 0.019 | 0.014 | 0.013 | 0.014 | 0.015 | 0.016 | 0.017 | 0.018 | 0.019 | 0.021 | 0.023 | 0.025 | 0.027 |
| RRRR2-PC5 | 0.071 | 0.063 | 0.063 | 0.064 | 0.063 | 0.060 | 0.056 | 0.050 | 0.045 | 0.040 | 0.036 | 0.034 | 0.032 | 0.031 |
| RRRR2-SMP | 0.012 | 0.010 | 0.011 | 0.014 | 0.016 | 0.017 | 0.018 | 0.019 | 0.020 | 0.020 | 0.020 | 0.020 | 0.019 | 0.019 |
| RRRR2-TMP | 0.039 | 0.030 | 0.028 | 0.028 | 0.029 | 0.030 | 0.031 | 0.032 | 0.033 | 0.034 | 0.035 | 0.035 | 0.036 | 0.037 |
| Panel D: Models with 14 forecasting factors | | | | | | | | | | | | | | |
| RR-SMP | 0.018 | 0.016 | 0.017 | 0.019 | 0.020 | 0.022 | 0.023 | 0.024 | 0.024 | 0.025 | 0.025 | 0.026 | 0.026 | 0.026 |
| RR-TMP | 0.074 | 0.058 | 0.050 | 0.045 | 0.041 | 0.039 | 0.037 | 0.036 | 0.036 | 0.035 | 0.035 | 0.035 | 0.035 | 0.036 |
| PLS-1 | 0.054 | 0.032 | 0.028 | 0.027 | 0.027 | 0.027 | 0.028 | 0.029 | 0.029 | 0.030 | 0.030 | 0.030 | 0.031 | 0.031 |
| PLS-2 | 0.005 | 0.002 | 0.004 | 0.014 | 0.021 | 0.014 | 0.026 | 0.029 | 0.029 | 0.028 | 0.026 | 0.026 | 0.025 | 0.024 |
| PLS-3 | 0.031 | 0.026 | 0.026 | 0.025 | 0.025 | 0.015 | 0.011 | 0.012 | 0.013 | 0.015 | 0.018 | 0.018 | 0.019 | 0.021 |
| PLS-4 | 0.006 | 0.001 | 0.002 | 0.009 | 0.019 | 0.026 | 0.028 | 0.051 | 0.034 | 0.032 | 0.033 | 0.034 | 0.038 | 0.037 |
| PLS-5 | -0.025 | 0.002 | 0.027 | 0.042 | 0.047 | 0.044 | 0.036 | 0.027 | 0.017 | 0.005 | -0.002 | 0.004 | 0.015 | 0.026 |
| 3PRF-1 | -0.032 | -0.001 | 0.010 | 0.019 | 0.024 | 0.022 | 0.017 | 0.012 | 0.007 | 0.003 | -0.001 | -0.002 | -0.002 | -0.002 |
| 3PRF-2 | 0.023 | 0.021 | 0.022 | 0.024 | 0.021 | 0.017 | 0.018 | 0.017 | 0.017 | 0.016 | 0.016 | 0.019 | 0.021 | 0.023 |
| 3PRF-3 | 0.006 | 0.001 | 0.002 | 0.009 | 0.014 | 0.015 | 0.014 | 0.012 | 0.010 | 0.010 | 0.010 | 0.012 | 0.015 | 0.021 |
| 3PRF-4 | -0.025 | -0.002 | 0.015 | 0.030 | 0.038 | 0.035 | 0.026 | 0.015 | 0.003 | -0.003 | 0.001 | 0.011 | 0.017 | 0.022 |
| 3PRF-5 | -0.047 | -0.016 | 0.001 | 0.008 | 0.011 | 0.007 | 0.001 | -0.009 | -0.018 | -0.026 | -0.025 | -0.022 | -0.020 | -0.016 |
| OLS | -0.549 | -0.424 | -0.361 | -0.330 | -0.318 | -0.316 | -0.320 | -0.327 | -0.333 | -0.337 | -0.340 | -0.340 | -0.338 | -0.336 |

# Forecasting bond excess returns

## Yield slopes and cycles

| Out-of-sample $R^2$ | Bond Excess Returns | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Models** | $rx^{(2)}$ | $rx^{(3)}$ | $rx^{(4)}$ | $rx^{(5)}$ | $rx^{(6)}$ | $rx^{(7)}$ | $rx^{(8)}$ | $rx^{(9)}$ | $rx^{(10)}$ | $rx^{(11)}$ | $rx^{(12)}$ | $rx^{(13)}$ | $rx^{(14)}$ | $rx^{(15)}$ |
| Panel A: Naïve benchmark models | | | | | | | | | | | | | | |
| Rolling Average | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Random Walk | -0.514 | -0.585 | -0.640 | -0.686 | -0.728 | -0.765 | -0.797 | -0.824 | -0.846 | -0.862 | -0.874 | -0.881 | -0.883 | -0.882 |
| Panel B: Models with 1 forecasting factor | | | | | | | | | | | | | | |
| PCR-1 | -0.019 | -0.011 | -0.006 | -0.002 | 0.000 | 0.002 | 0.003 | 0.003 | 0.004 | 0.004 | 0.004 | 0.004 | 0.004 | 0.005 |
| RRRR1-PC2 | 0.034 | 0.036 | 0.040 | 0.043 | 0.047 | 0.051 | 0.054 | 0.056 | 0.059 | 0.061 | 0.062 | 0.063 | 0.065 | 0.065 |
| RRRR1-PC3 | 0.012 | 0.017 | 0.024 | 0.030 | 0.035 | 0.041 | 0.045 | 0.050 | 0.053 | 0.057 | 0.059 | 0.062 | 0.064 | 0.065 |
| RRRR1-PC4 | 0.027 | 0.021 | 0.022 | 0.025 | 0.029 | 0.034 | 0.038 | 0.042 | 0.045 | 0.048 | 0.051 | 0.053 | 0.056 | 0.058 |
| RRRR1-PC5 | 0.057 | 0.046 | 0.045 | 0.048 | 0.053 | 0.059 | 0.065 | 0.071 | 0.077 | 0.082 | 0.087 | 0.091 | 0.095 | 0.099 |
| RRRR1-SMP | 0.034 | 0.036 | 0.040 | 0.043 | 0.047 | 0.051 | 0.054 | 0.056 | 0.059 | 0.061 | 0.062 | 0.063 | 0.065 | 0.065 |
| RRRR1-TMP | 0.027 | 0.029 | 0.034 | 0.039 | 0.044 | 0.049 | 0.054 | 0.057 | 0.061 | 0.064 | 0.066 | 0.068 | 0.070 | 0.071 |
| Panel C: Models with 2 forecasting factors | | | | | | | | | | | | | | |
| PCR-2 | 0.023 | 0.028 | 0.034 | 0.040 | 0.045 | 0.049 | 0.053 | 0.056 | 0.058 | 0.060 | 0.062 | 0.063 | 0.064 | 0.065 |
| RRRR2-PC3 | 0.017 | 0.022 | 0.028 | 0.034 | 0.039 | 0.043 | 0.047 | 0.050 | 0.052 | 0.054 | 0.055 | 0.056 | 0.057 | 0.058 |
| RRRR2-PC4 | 0.004 | 0.008 | 0.017 | 0.025 | 0.031 | 0.037 | 0.041 | 0.044 | 0.046 | 0.048 | 0.049 | 0.050 | 0.051 | 0.051 |
| RRRR2-PC5 | 0.033 | 0.030 | 0.035 | 0.042 | 0.050 | 0.059 | 0.066 | 0.073 | 0.079 | 0.084 | 0.089 | 0.093 | 0.096 | 0.099 |
| RRRR2-SMP | 0.020 | 0.026 | 0.033 | 0.040 | 0.045 | 0.050 | 0.054 | 0.057 | 0.059 | 0.061 | 0.063 | 0.065 | 0.066 | 0.067 |
| RRRR2-TMP | -0.012 | 0.003 | 0.016 | 0.028 | 0.038 | 0.045 | 0.051 | 0.056 | 0.059 | 0.062 | 0.064 | 0.065 | 0.066 | 0.067 |
| Panel D: Models with 14 forecasting factors | | | | | | | | | | | | | | |
| RR-SMP | 0.020 | 0.026 | 0.033 | 0.040 | 0.045 | 0.050 | 0.054 | 0.057 | 0.059 | 0.061 | 0.063 | 0.065 | 0.066 | 0.067 |
| RR-TMP | 0.015 | 0.023 | 0.032 | 0.039 | 0.045 | 0.050 | 0.054 | 0.058 | 0.060 | 0.063 | 0.064 | 0.066 | 0.067 | 0.068 |
| PLS-1 | -0.003 | 0.011 | 0.023 | 0.033 | 0.042 | 0.048 | 0.053 | 0.057 | 0.060 | 0.062 | 0.064 | 0.066 | 0.067 | 0.068 |
| PLS-2 | 0.019 | 0.027 | 0.034 | 0.044 | 0.041 | 0.046 | 0.052 | 0.054 | 0.052 | 0.055 | 0.058 | 0.060 | 0.063 | 0.065 |
| PLS-3 | 0.011 | 0.016 | 0.025 | 0.032 | 0.039 | 0.044 | 0.047 | 0.048 | 0.053 | 0.055 | 0.058 | 0.062 | 0.061 | 0.063 |
| PLS-4 | -0.008 | -0.006 | 0.000 | 0.012 | 0.022 | 0.026 | 0.035 | 0.035 | 0.041 | 0.052 | 0.057 | 0.059 | 0.062 | 0.066 |
| PLS-5 | 0.019 | 0.023 | 0.027 | 0.037 | 0.035 | 0.039 | 0.046 | 0.051 | 0.054 | 0.056 | 0.059 | 0.061 | 0.067 | 0.069 |
| 3PRF-1 | -0.028 | -0.013 | -0.002 | 0.013 | 0.024 | 0.018 | 0.014 | 0.009 | 0.001 | -0.002 | -0.003 | -0.002 | -0.001 | 0.001 |
| 3PRF-2 | 0.062 | 0.063 | 0.066 | 0.067 | 0.064 | 0.060 | 0.061 | 0.061 | 0.057 | 0.056 | 0.056 | 0.053 | 0.052 | 0.051 |
| 3PRF-3 | 0.050 | 0.036 | 0.037 | 0.046 | 0.055 | 0.058 | 0.062 | 0.065 | 0.066 | 0.068 | 0.068 | 0.069 | 0.069 | 0.069 |
| 3PRF-4 | 0.024 | 0.033 | 0.037 | 0.044 | 0.051 | 0.055 | 0.057 | 0.063 | 0.066 | 0.072 | 0.080 | 0.083 | 0.084 | 0.086 |
| 3PRF-5 | 0.019 | 0.031 | 0.042 | 0.048 | 0.052 | 0.050 | 0.053 | 0.056 | 0.055 | 0.051 | 0.055 | 0.057 | 0.061 | 0.066 |
| OLS | -0.597 | -0.478 | -0.414 | -0.381 | -0.364 | -0.358 | -0.357 | -0.357 | -0.358 | -0.357 | -0.354 | -0.349 | -0.343 | -0.336 |

- RRRR models are always among the best performing ones
- Yield curve slopes have almost the same predictive power as cycles
- Less predictive power of forwards slopes & forwards in particular
- Combining yield curve slopes and cycles as predictors almost doubles the out-of-sample predictive power
- Strong case for using RRRR models for forecasting bond excess returns by combining multiple predictor sets

# Conclusions and Future Research

- No one method is uniformly best across data sets and sub-samples
- RRRR is a robust method for jointly forecasting multiple outcomes in situations with correlated outcomes and many or nearly collinear predictors
- Can often deliver more parsimonious forecasting models than competing methods when predicting multiple outcomes
- It is possible to almost double the predictability of 1-month bond excess returns by combining yield slopes and cycles
- More work in progress on data-driven RRRR methods and other regularization schemes

A key shortcoming so far: Lack of factor interpretability!

- Consider imposing $f$ *linear* constraints: $P'A = 0$
- Solution is $A = P^{\perp}a$, where $a \in \mathbb{R}^{(n-f)\times k}$ containins the $k$ principal eigenvectors of the generalized eigenvalue problem

$$|P^{\perp\prime}S_{XY}S_{YX}P^{\perp} - \lambda\, P^{\perp\prime}(S_{XX} + \rho^2 R'R)P^{\perp}| = 0$$

- Constraints on the individual columns of $A = [A_1, \ldots, A_k]$ can be imposed recursively:

$$P_1'A_1 = 0, P_2'A_2 = 0, \ldots, P_k'A_k = 0$$

so to get the $j+1^{st}$ factor, orthogonal to factors $i_1$, $i_2$, ..., $i_g$, set

$$P = [\underbrace{S_{XX}A_{i_1}}_{n\times 1}, \ldots, \underbrace{S_{XX}A_{i_g}}_{n\times 1}, \underbrace{P_{j+1}}_{n\times f_{j+1}}]$$

## Example: Imposing Interpretability

- Regressors classified into 4 *non-exclusive* groups
- Goal: Find principal factor attributable to a single group.

| Variable | Memberships |
|---|---|
| 1 | $G_1, G_2$ |
| 2 | $G_2, G_3, G_4$ |
| 3 | $G_1, G_4$ |
| 4 | $G_3, G_4$ |
| 5 | $G_1, G_3, G_4$ |
| $\vdots$ | $\vdots$ |
| N | $G_3$ |

$$\Rightarrow P' = P'_{\{G_1^\perp\}} = \underbrace{\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & 0 & \cdots & 0 \\ & \vdots & & & \vdots & & \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}}_{\text{Select } G_1\text{-factor}}$$

- Solve GEV problem for each $P \in \{P_{\{G_1^\perp\}}, \ldots, P_{\{G_4^\perp\}}\}$.
  - The principal factor yields the largest eigenvalue
  - Additional factors extracted iteratively imposing within-group orthog.

# Forecasting macroeconomic series

- Two forecasting exercises:
  1. All 143 SW series without imposing common (few) factors
  2. The 35 SW aggregates with imposing common (few) factors
- Report distribution of relative RMSE ratio vis-a-vis PCR-5 benchmark across all forecasted series
- Substantial downward/upward deviations from a ratio equal to 1 indicates better/worse performance than PCR-5

# Forecasting macroeconomic series
## All 143 SW series

| Relative RMSE to PCR-5 | Percentiles | | | | | Empirical Distribution | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Models | 5 | 25 | 50 | 75 | 95 | <0.90 | 0.90-0.97 | 0.97-1.03 | 1.03-1.10 | >1.10 |
| Panel A: Naïve benchmark models | | | | | | | | | | |
| AR-4 | 0.918 | 0.979 | 1.007 | 1.041 | 1.144 | 0.014 | 0.189 | 0.490 | 0.182 | 0.126 |
| PCR-50 | 0.968 | 1.061 | 1.110 | 1.179 | 1.281 | 0.007 | 0.056 | 0.091 | 0.273 | 0.573 |
| Panel B: PCR models | | | | | | | | | | |
| PCR-1 | 0.929 | 0.975 | 0.995 | 1.034 | 1.114 | 0.035 | 0.189 | 0.517 | 0.175 | 0.084 |
| PCR-2 | 0.930 | 0.975 | 0.993 | 1.010 | 1.057 | 0.014 | 0.189 | 0.664 | 0.133 | 0.000 |
| PCR-3 | 0.954 | 0.982 | 0.992 | 1.008 | 1.029 | 0.000 | 0.126 | 0.832 | 0.042 | 0.000 |
| PCR-4 | 0.981 | 0.990 | 0.999 | 1.008 | 1.027 | 0.000 | 0.035 | 0.916 | 0.049 | 0.000 |
| PCR-5 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.000 | 0.000 | 1.000 | 0.000 | 0.000 |
| PCR-6 | 0.976 | 0.993 | 1.002 | 1.009 | 1.020 | 0.000 | 0.042 | 0.937 | 0.021 | 0.000 |
| PCR-7 | 0.973 | 0.995 | 1.005 | 1.017 | 1.042 | 0.000 | 0.021 | 0.846 | 0.133 | 0.000 |
| Panel C: RR models | | | | | | | | | | |
| RR-SMP | 0.977 | 0.990 | 0.996 | 1.003 | 1.013 | 0.000 | 0.028 | 0.965 | 0.007 | 0.000 |
| RR-TMP | 0.975 | 1.026 | 1.069 | 1.111 | 1.187 | 0.000 | 0.042 | 0.252 | 0.413 | 0.294 |
| Panel D: PLS models | | | | | | | | | | |
| PLS-1 | 0.950 | 0.987 | 1.009 | 1.035 | 1.087 | 0.000 | 0.133 | 0.594 | 0.224 | 0.049 |
| PLS-2 | 0.976 | 1.038 | 1.082 | 1.130 | 1.271 | 0.000 | 0.021 | 0.196 | 0.406 | 0.378 |
| PLS-3 | 1.019 | 1.088 | 1.153 | 1.234 | 1.422 | 0.000 | 0.000 | 0.063 | 0.217 | 0.720 |
| PLS-4 | 1.046 | 1.143 | 1.228 | 1.324 | 1.609 | 0.000 | 0.000 | 0.028 | 0.098 | 0.874 |
| PLS-5 | 1.086 | 1.207 | 1.301 | 1.428 | 1.733 | 0.000 | 0.000 | 0.007 | 0.063 | 0.930 |
| PLS-6 | 1.123 | 1.261 | 1.363 | 1.519 | 1.841 | 0.000 | 0.000 | 0.000 | 0.035 | 0.965 |
| PLS-7 | 1.130 | 1.309 | 1.420 | 1.606 | 1.906 | 0.000 | 0.000 | 0.000 | 0.007 | 0.993 |
| Panel E: 3PRF models | | | | | | | | | | |
| 3PRF-1 | 0.947 | 0.980 | 1.002 | 1.026 | 1.081 | 0.000 | 0.147 | 0.629 | 0.203 | 0.021 |
| 3PRF-2 | 0.979 | 1.020 | 1.060 | 1.103 | 1.239 | 0.000 | 0.035 | 0.273 | 0.427 | 0.266 |
| 3PRF-3 | 1.010 | 1.080 | 1.144 | 1.229 | 1.424 | 0.000 | 0.007 | 0.084 | 0.231 | 0.678 |
| 3PRF-4 | 1.035 | 1.135 | 1.225 | 1.323 | 1.601 | 0.000 | 0.000 | 0.049 | 0.091 | 0.860 |
| 3PRF-5 | 1.070 | 1.198 | 1.302 | 1.426 | 1.726 | 0.000 | 0.000 | 0.007 | 0.063 | 0.930 |
| 3PRF-6 | 1.126 | 1.258 | 1.368 | 1.514 | 1.853 | 0.000 | 0.000 | 0.000 | 0.042 | 0.958 |
| 3PRF-7 | 1.140 | 1.307 | 1.420 | 1.585 | 1.914 | 0.000 | 0.000 | 0.007 | 0.021 | 0.972 |

# Forecasting macroeconomic series

## The 35 SW aggregates

| Relative RMSE to PCR-5 | Percentiles | | | | | Empirical Distribution | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Models** | **5** | **25** | **50** | **75** | **95** | **<0.90** | **0.90-0.97** | **0.97-1.03** | **1.03-1.10** | **>1.10** |
| Panel A: Models with 1 forecasting factor | | | | | | | | | | |
| PCR-1 | 0.590 | 0.951 | 1.000 | 1.036 | 1.145 | 0.086 | 0.229 | 0.371 | 0.171 | 0.143 |
| RRRR1-PC2 | 0.561 | 0.935 | 0.999 | 1.013 | 1.087 | 0.086 | 0.200 | 0.543 | 0.143 | 0.029 |
| RRRR1-PC4 | 0.546 | 0.939 | 1.003 | 1.019 | 1.189 | 0.086 | 0.200 | 0.514 | 0.114 | 0.086 |
| RRRR1-PC6 | 0.535 | 0.950 | 1.003 | 1.026 | 1.201 | 0.057 | 0.229 | 0.486 | 0.143 | 0.086 |
| RRRR1-PC8 | 0.526 | 0.953 | 1.005 | 1.038 | 1.231 | 0.057 | 0.257 | 0.400 | 0.171 | 0.114 |
| RRRR1-PC10 | 0.523 | 0.940 | 0.997 | 1.023 | 1.188 | 0.114 | 0.200 | 0.514 | 0.086 | 0.086 |
| RRRR1-SMP | 0.534 | 0.939 | 1.004 | 1.028 | 1.241 | 0.086 | 0.229 | 0.457 | 0.143 | 0.086 |
| RRRR1-TMP | 0.534 | 0.954 | 1.001 | 1.033 | 1.165 | 0.114 | 0.171 | 0.429 | 0.143 | 0.143 |
| Panel B: Models with 3 forecasting factors | | | | | | | | | | |
| PCR-3 | 0.681 | 0.972 | 0.988 | 1.006 | 1.033 | 0.057 | 0.143 | 0.743 | 0.057 | 0.000 |
| RRRR3-PC4 | 0.495 | 0.971 | 0.987 | 1.006 | 1.031 | 0.057 | 0.171 | 0.714 | 0.057 | 0.000 |
| RRRR3-PC6 | 0.500 | 0.990 | 0.998 | 1.026 | 1.082 | 0.086 | 0.057 | 0.629 | 0.229 | 0.000 |
| RRRR3-PC8 | 0.482 | 0.982 | 0.996 | 1.054 | 1.131 | 0.086 | 0.029 | 0.600 | 0.171 | 0.114 |
| RRRR3-PC10 | 0.436 | 0.983 | 1.000 | 1.026 | 1.116 | 0.086 | 0.057 | 0.686 | 0.086 | 0.086 |
| RRRR3-SMP | 0.467 | 0.987 | 0.995 | 1.012 | 1.037 | 0.086 | 0.057 | 0.771 | 0.086 | 0.000 |
| RRRR3-TMP | 0.445 | 0.993 | 1.042 | 1.096 | 1.148 | 0.057 | 0.114 | 0.314 | 0.314 | 0.200 |
| Panel C: Models with 5 forecasting factors | | | | | | | | | | |
| PCR-5 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.000 | 0.000 | 1.000 | 0.000 | 0.000 |
| RRRR5-PC6 | 0.476 | 0.983 | 0.993 | 1.003 | 1.022 | 0.057 | 0.114 | 0.829 | 0.000 | 0.000 |
| RRRR5-PC8 | 0.469 | 0.991 | 0.998 | 1.020 | 1.050 | 0.057 | 0.086 | 0.686 | 0.171 | 0.000 |
| RRRR5-PC10 | 0.407 | 0.965 | 0.988 | 1.006 | 1.035 | 0.057 | 0.257 | 0.600 | 0.086 | 0.000 |
| RRRR5-SMP | 0.470 | 0.989 | 0.997 | 1.000 | 1.013 | 0.057 | 0.029 | 0.914 | 0.000 | 0.000 |
| RRRR5-TMP | 0.380 | 0.997 | 1.045 | 1.127 | 1.180 | 0.057 | 0.000 | 0.400 | 0.286 | 0.257 |
| Panel D: Models with 7 forecasting factors | | | | | | | | | | |
| PCR-7 | 0.969 | 0.996 | 1.004 | 1.029 | 1.332 | 0.000 | 0.057 | 0.714 | 0.171 | 0.057 |
| RRRR7-PC8 | 0.470 | 0.987 | 1.002 | 1.026 | 1.037 | 0.057 | 0.029 | 0.714 | 0.200 | 0.000 |
| RRRR7-PC10 | 0.407 | 0.964 | 0.992 | 1.010 | 1.035 | 0.057 | 0.200 | 0.686 | 0.057 | 0.000 |
| RRRR7-SMP | 0.473 | 0.985 | 1.000 | 1.012 | 1.037 | 0.057 | 0.086 | 0.743 | 0.114 | 0.000 |
| RRRR7-TMP | 0.376 | 1.003 | 1.045 | 1.125 | 1.172 | 0.057 | 0.000 | 0.400 | 0.257 | 0.286 |
| Panel E: Models with 35 forecasting factors | | | | | | | | | | |
| RR-SMP | 0.467 | 0.981 | 0.995 | 1.007 | 1.017 | 0.057 | 0.057 | 0.886 | 0.000 | 0.000 |
| RR-TMP | 0.330 | 0.983 | 1.052 | 1.113 | 1.170 | 0.057 | 0.114 | 0.257 | 0.314 | 0.257 |
| PLS-1 | 0.472 | 0.965 | 0.995 | 1.016 | 1.046 | 0.114 | 0.171 | 0.571 | 0.143 | 0.000 |
| PLS-2 | 0.351 | 0.986 | 1.036 | 1.086 | 1.170 | 0.057 | 0.029 | 0.343 | 0.343 | 0.229 |
| PLS-3 | 0.291 | 1.052 | 1.153 | 1.277 | 1.375 | 0.057 | 0.029 | 0.114 | 0.086 | 0.714 |
| PLS-5 | 0.230 | 1.243 | 1.354 | 1.524 | 1.777 | 0.057 | 0.000 | 0.029 | 0.057 | 0.857 |
| 3PRF-1 | 0.455 | 0.965 | 1.007 | 1.036 | 1.092 | 0.086 | 0.171 | 0.457 | 0.286 | 0.000 |
| 3PRF-2 | 0.360 | 1.016 | 1.074 | 1.114 | 1.203 | 0.057 | 0.029 | 0.200 | 0.371 | 0.343 |
| 3PRF-3 | 0.309 | 1.061 | 1.170 | 1.297 | 1.411 | 0.057 | 0.029 | 0.057 | 0.171 | 0.686 |
| 3PRF-5 | 0.245 | 1.241 | 1.369 | 1.545 | 1.790 | 0.057 | 0.000 | 0.029 | 0.057 | 0.857 |