Dynamic Portfolio Choice balancing forecasting risk

Ekaterina Kazak[1] Yifan Li[2,3] Ingmar Nolte[3] Sandra Nolte[3]

[1]University of Konstanz, GSDS

[2]University of Manchester

[3]Lancaster University

4th Konstanz - Lancaster Workshop
Konstanz
July 31, 2018

Motivation I

- Estimation risk a huge problem in portfolio allocation
- So far the literature was concentrated on
  1. improving estimation and forecasting of the portfolio "ingredients" (Ledoit and Wolf, 2003, 2004a,b, 2012; Fan et al., 2011; Kourtis et al., 2012; Abadir et al., 2014; Bollerslev et al., 2016a)
  2. new portfolio allocation strategies (Jagannathan and Ma, 2003; Brandt et al., 2009; Brodie et al., 2009; Tu and Zhou, 2011; Fan et al., 2012; DeMiguel et al., 2014; Bodnar et al., 2015; Bollerslev et al., 2016b)

Motivation II

- Problems remaining:
  - high-dimensional covariance forecasts: positive definiteness
  - tuning parameters
  - regularized strategies do not perform "uniformly" well

Motivation II

- Problems remaining:
  - high-dimensional covariance forecasts: positive definiteness
  - tuning parameters
  - regularized strategies do not perform "uniformly" well
- We focus on forecasting the portfolio performance measure directly
  - one-dimensional, very easy to implement
  - using high-frequency measures to construct "realized" performance measures
  - optimizes for a given target (very useful in practice)

Outline

- Introduction
- New Forecasting Approach
- Simulation Evidence
- Empirical Application
- Conclusion

GMVP

- Global Minimum Variance Portfolio : $\min_{\omega_t} \ \omega_t' \Sigma_t \omega_t \quad s.t. \quad \iota' \omega_t = 1$

$$\omega_t^\star = \frac{\Sigma_t^{-1}\iota}{\iota'\Sigma_t^{-1}\iota}.$$

- Minimized portfolio variance:

$$\sigma_{p,t}^2(\omega_t^\star) = \frac{1}{\iota'\Sigma_t^{-1}\iota}.$$

- Σ_t is not observable at time $t-1$
- Common approach: $\Omega_t = \mathrm{E}[\Sigma_t|\mathcal{F}_{t-1}]$ and minimize $\omega_t'\Omega_t\omega_t$ instead
- Plug-in estimator

$$\widehat{\omega}_t = \frac{\Omega_t^{-1}\iota}{\iota'\Omega_t^{-1}\iota}.$$

GMVP

$$\Sigma_t = \Omega_t + \Lambda_t,$$

Λ_t is an $N \times N$ error term, s.t. $\mathrm{E}[\Lambda_t|\mathcal{F}_{t-1}] = \mathbf{0}_{N\times N}$.

- Consider a distance measure $d(\cdot)$

$$d(\omega_t) = \sigma^2_{p,t}(\omega_t) - \sigma^2_{p,t}(\omega^\star_t) \geq 0,$$

- which is a random variable at time $t-1$, but is realized at time t, so that the investor can compute the *ex post* loss

$$\begin{aligned} d_t(\omega_t) &= \omega_t'\Sigma_t\omega_t - \frac{1}{\iota'\Sigma_t^{-1}\iota} \\ &= \omega_t'\Omega_t\omega_t + \omega_t'\Lambda_t\omega_t - \frac{1}{\iota'(\Omega_t+\Lambda_t)^{-1}\iota} \end{aligned}$$

GMVP

- For the plug-in estimator we show that

$$d_t(\widehat{\omega}_t) \approx \frac{\iota'\Omega_t^{-1}\Lambda_t\Omega_t^{-1}\Lambda_t\Omega_t^{-1}\iota}{(\iota'\Omega_t^{-1}\iota)^2} = \widehat{\omega}_t'\Lambda_t\Omega_t^{-1}\Lambda_t\widehat{\omega}_t$$

which is always positive unless $\Lambda_t = \mathbf{0}_{N\times N}$.

GMVP

- For the plug-in estimator we show that

$$d_t(\widehat{\omega}_t) \approx \frac{\iota'\Omega_t^{-1}\Lambda_t\Omega_t^{-1}\Lambda_t\Omega_t^{-1}\iota}{(\iota'\Omega_t^{-1}\iota)^2} = \widehat{\omega}_t'\Lambda_t\Omega_t^{-1}\Lambda_t\widehat{\omega}_t$$

  which is always positive unless $\Lambda_t = \mathbf{0}_{N\times N}$.

- Furthermore,

$$\mathrm{E}[d_t(\widehat{\omega}_t)|\mathcal{F}_{t-1}] \approx tr[\Omega_t^{-1}\mathrm{E}[\widehat{\omega}_t'\Lambda_t\Lambda_t\widehat{\omega}_t|\mathcal{F}_{t-1}]].$$

- Expected relative loss of the plug-in estimator is a function of the second moments of the errors
- The plug-in estimator $\widehat{\omega}_t$ minimizes $\mathrm{E}[d_t(\omega_t)|\mathcal{F}_{t-1}] = \omega_t'\Omega_t\omega_t - \mathrm{E}\left[\frac{1}{\iota'\Sigma_t^{-1}\iota}\middle|\mathcal{F}_{t-1}\right]$

GMVP

Our approach: minimize the square of $d_t(\omega_t)$:

$$\min_{\omega_t} \mathrm{E}[d_t(\omega_t)^2|\mathcal{F}_{t-1}], \quad s.t.\ \iota'\omega_t = 1$$

- By definition

$$E\Big[d_t(\omega_t)^2\Big|\mathcal{F}_{t-1}\Big] = E\Big[\Big(\omega_t'\Sigma_t\omega_t - \frac{1}{\iota'\Sigma_t^{-1}\iota}\Big)^2\Big|\mathcal{F}_{t-1}\Big]$$

- The solution to this comes from matching ω_t:

$$\omega_t'\Omega_t\omega_t \qquad vs. \qquad \mathrm{E}\Big[\frac{1}{\iota'\Sigma_t^{-1}\iota}\Big|\mathcal{F}_{t-1}\Big], \quad s.t.\ \iota'\omega_t = 1$$

GMVP

Our approach: minimize the square of $d_t(\omega_t)$:

$$\min_{\omega_t} \mathrm{E}[d_t(\omega_t)^2 | \mathcal{F}_{t-1}], \quad s.t.\ \iota' \omega_t = 1$$

- Explicitly minimizes the variance of Λ_t

$$E\left[d_t(\omega_t)^2 \middle| \mathcal{F}_{t-1}\right] = V\left[d_t(\omega_t) \middle| \mathcal{F}_{t-1}\right] + E\left[d_t(\omega_t) \middle| \mathcal{F}_{t-1}\right]^2 =$$
$$= \mathrm{V}\left[\omega_t' \Lambda_t \omega_t \middle| \mathcal{F}_{t-1}\right] + \mathrm{E}\left[2\omega_t' \Lambda_t \omega_t C_t \middle| \mathcal{F}_{t-1}\right] + \mathrm{E}\left[C_t^2 \middle| \mathcal{F}_{t-1}\right] + E\left[d_t(\omega_t) \middle| \mathcal{F}_{t-1}\right]^2,$$

where $C_t = \frac{1}{\iota'(\Omega_t + \Lambda_t)^{-1}\iota} - \mathrm{E}\left[\frac{1}{\iota'(\Omega_t + \Lambda_t)^{-1}\iota} \middle| \mathcal{F}_{t-1}\right]$ is a zero-mean random variable which is not a function of ω_t but a function of Λ_t.

General Utility Framework

$$\mathcal{U}_t(\omega_t) = f(\omega_t; r_t, \Sigma_t),$$

$$\omega_t^\star = \arg\max_{\omega_t} \mathcal{U}_t(\omega_t), \quad s.t.\ \iota'\omega_t = 1,$$

$$d_t(\omega_t) = \mathcal{U}_t(\omega_t^\star) - \mathcal{U}_t(\omega_t) \geq 0.$$

- Consider the following form of optimization:

$$\min_{\omega_t} \mathrm{E}[d_t(\omega_t)^p | \mathcal{F}_{t-1}], \quad s.t.\ \iota'\omega_t = 1, \quad p > 0$$

- $p = 1$ case (commonly used in the literature):

$$\mathrm{E}[d_t(\omega_t) | \mathcal{F}_{t-1}] = \mathrm{E}[\mathcal{U}_t(\omega_t^\star) | \mathcal{F}_{t-1}] - \mathrm{E}[\mathcal{U}_t(\omega_t) | \mathcal{F}_{t-1}]$$

General Utility Framework

- $p = 1$ case is equivalent to maximizing $\mathrm{E}[\mathcal{U}_t(\omega_t)|\mathcal{F}_{t-1}]$ and $\widehat{\omega}_t$ is the optimal weight
- Note that $\forall \omega_t$

$$\mathcal{U}_t(\omega_t) = \mathrm{E}[\mathcal{U}_t(\omega_t)|\mathcal{F}_{t-1}] + \varepsilon_t(\omega_t).$$

- The true optimal weight $\omega_t^\star$ maximizes both $\mathrm{E}[\mathcal{U}_t(\omega_t)|\mathcal{F}_{t-1}]$ and $\varepsilon_t(\omega_t)$
- $p = 1$ maximizes **only** $\mathrm{E}[\mathcal{U}_t(\omega_t)|\mathcal{F}_{t-1}]$, therefore $\omega_t^\star = \widehat{\omega}_t$ iff $\varepsilon_t(\omega_t) = 0 \ \forall \omega_t$
- In general $\mathrm{V}[\varepsilon_t(\omega_t)|\mathcal{F}_{t-1}] > 0$ and $\omega_t^\star \neq \widehat{\omega}_t$

We propose to use $p = 2$

General Utility Framework

$$\min_{\omega_t} \mathrm{E}[d_t(\omega_t)^2|\mathcal{F}_{t-1}], \quad s.t.\ \iota'\omega_t = 1,$$

- which can be decomposed similar to the MSE:

$$\mathrm{E}[d_t(\omega_t)^2|\mathcal{F}_{t-1}] = \mathrm{V}[d_t(\omega_t)|\mathcal{F}_{t-1}] + \mathrm{E}[d_t(\omega_t)|\mathcal{F}_{t-1}]^2.$$

- Yields more efficient estimates of the optimal utility compared to the $p = 1$ case.
- Particularly attractive when $\mathrm{V}[d_t(\omega_t)|\mathcal{F}_{t-1}]$ dominates the gain in $\mathrm{E}[d_t(\omega_t)|\mathcal{F}_{t-1}]$
- Advantageous in large dimensional portfolio analysis

Implementation

1. At time t an investor can construct a "realized" utility (portfolio measure of interest) using RC and return realizations
2. The "realized" distance measure $d_t(\omega_t)$ for any ω_t
3. Using standard time series models to forecast $\mathrm{E}[d_t(\omega_t)|\mathcal{F}_{t-1}]$ and $\mathrm{V}[d_t(\omega_t)|\mathcal{F}_{t-1}]$
4. Compute the weight forecast using $p = 2$

$$\min_{\omega_t} \mathrm{E}[d_t(\omega_t)^2|\mathcal{F}_{t-1}], \quad s.t.\ \iota'\omega_t = 1.$$

Finding a good forecasting model for a univariate $d_t(\omega_t)$ is easier than fine-tuning a high dimensional matrix forecast.

Simulation Design

- Bivariate return process with DCC in conditional variance
- Mean return process: zero mean or AR(1)
- Covariance estimators
  - sample covariance
  - Univariate HAR + CCC
- Portfolio strategies:
  - GMVP
  - Efficient (max CE)
- Our approach:
  - Minimize portfolio variance: $U_t(\omega_t^\star)$ - realized portfolio variance of GMVP
  - Maximize certainty equivalent: $U_t(\omega_t^\star)$ - realized CE of Efficient Portfolio
  - Maximize Sharpe Ratio : $U_t(\omega_t^\star)$ - realized SR of Tangency Portfolio

Simulation Results I: zero mean

	Absolute				Relative			
	Variance		CE	SR	Variance		CE	SR
	QLIKE	MSE	MSE	MSE	QLIKE	MSE	MSE	MSE
GMVP + $\widehat{\Sigma}$	0.0468	0.1708	1.4674	1.9432	2.60	18.04	1.20	5.82
GMVP + HAR	0.0197	0.0824	1.3513	1.5147	1.09	8.70	1.10	4.54
EFF + $\widehat{\Sigma}$ + $\hat{\mu}$	0.0508	0.1746	1.4751	1.9409	2.82	18.44	1.20	5.81
EFF + $\widehat{\Sigma}$ + AR(1)	1.1284	0.1847	1.8470	1.9510	62.58	19.51	1.51	5.84
EFF + HAR + $\hat{\mu}$	0.0215	0.0869	1.3551	1.5168	1.19	9.17	1.11	4.54
EFF + HAR+ AR(1)	0.0267	0.0924	1.3641	1.5217	1.48	9.76	1.11	4.56
min VAR	0.0180	0.0095	1.2065	0.6348	*1.00*	*1.00*	0.98	1.90
max CE	0.0209	0.0122	1.2254	0.6059	1.16	1.29	*1.00*	1.81
max SR	0.0445	0.0036	1.2527	0.3339	2.47	0.39	1.02	*1.00*

Average QLIKE and MSE over 1000 Monte Carlo repetitions, $T = 250$ in-sample estimation window length, $H = 1000$ out-of-sample evaluation window, $\gamma = 1$. Persistence for the DCC $\beta = (0.8; 0.9)$, HAR for forecasting conditional variance of $d_t(\omega t)$.

Simulation Results II: AR(1)

	Absolute				Relative			
	Variance		CE	SR	Variance		CE	SR
	QLIKE	MSE	MSE	MSE	QLIKE	MSE	MSE	MSE
GMVP + $\widehat{\Sigma}$	0.0451	0.1702	1.4680	1.9380	3.83	18.13	1.20	5.80
GMVP + HAR	0.0183	0.0822	1.3529	1.5118	1.56	8.75	1.11	4.52
EFF + $\widehat{\Sigma}$ + $\hat{\mu}$	0.0494	0.1737	1.4738	1.9323	4.20	18.50	1.21	5.78
EFF + $\widehat{\Sigma}$ + AR(1)	2.8392	0.1858	2.2310	1.9439	241.42	19.78	1.83	5.81
EFF + HAR + $\hat{\mu}$	0.0248	0.0863	1.3565	1.5111	2.11	9.19	1.11	4.52
EFF + HAR+ AR(1)	0.0637	0.0927	1.3717	1.5159	5.41	9.87	1.12	4.53
min VAR	0.0118	0.0094	1.2106	0.6315	*1.00*	*1.00*	0.99	1.89
max CE	0.0194	0.0123	1.2208	0.6058	1.65	1.31	*1.00*	1.81
max SR	0.0360	0.0036	1.2572	0.3343	3.06	0.38	1.03	*1.00*

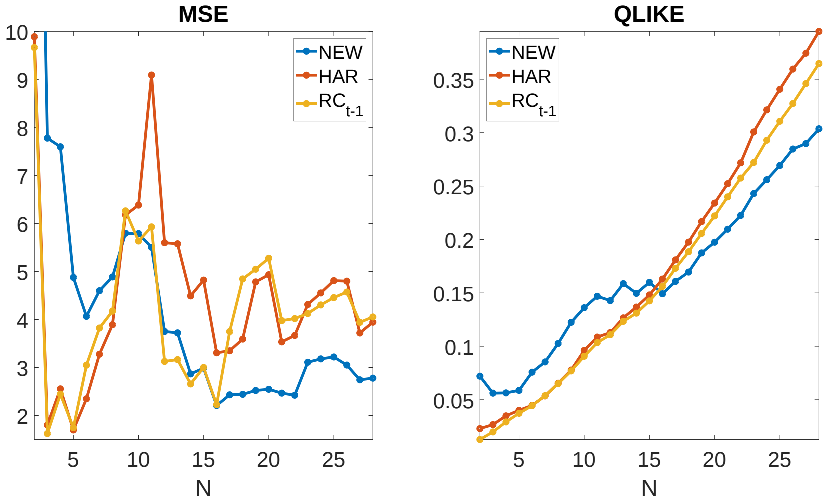
Average QLIKE and MSE over 1000 Monte Carlo repetitions, $T = 250$ in-sample estimation window length, $H = 1000$ out-of-sample evaluation window, $\gamma = 1$. Persistence for the DCC $\beta = (0.8; 0.9)$, AR parameter $\phi = 0.05$, HAR for forecasting conditional variance of $d_t(\omega t)$.

Set-Up

- Realized Covariances based on 5-min returns
- 28 Dow Jones constituents during Jan 2007 - Dec 2015
- Adding assets one by one (starting from $N = 2$)
- Competing strategies:
  1. GMVP + HAR + CCC
  2. GMVP based on RC_{t-1}
  3. Our new approach minimizing portfolio variance
- Evaluation: out-of-sample MSE and QLIKE based on

$$e_t = \widehat{\omega}_t' \Sigma_t \widehat{\omega}_t - {\omega_t^{\star}}' \Sigma_t \omega_t^{\star}$$

Portfolio variance



Estimation window $T = 250$, out-of-sample horizon of $H = 1000$.

Summary

- A new portfolio allocation approach:
  - easy to implement
  - flexible in terms of criterion, dataset, frequency
  - does not suffer from dimensionality problem
  - does not suffer from constraints on positive definiteness
  - makes use of the intra-day data
  - good for variance targeting (and any targeting in general)
- Outlook
  - Theoretical justification (optimality property)
  - Turnover costs

Thank you for your attention!

References I

ABADIR, K. M., W. DISTASO, AND F. ŽIKEŠ (2014): "Design-free estimation of variance matrices," *Journal of Econometrics*, 181, 165–180.

BODNAR, T., S. MAZUR, AND Y. OKHRIN (2015): "Bayesian Estimation of the Global Minimum Variance Portfolio," .

BOLLERSLEV, T., A. J. PATTON, AND R. QUAEDVLIEG (2016a): "Exploiting the Errors: A Simple Approach for Improved Volatility Forecasting," *Journal of Econometrics*, 192, 1–18.

——— (2016b): "Modeling and Forecasting (Un)Reliable Realized Covariances for More Reliable Financial Decisions," Working paper, Duke University.

BRANDT, M. W., P. SANTA-CLARA, AND R. VALKANOV (2009): "Parametric Portfolio Policies: Exploiting Characteristics in the Cross-Section of Equity Returns," *Review of Financial Studies*, 22, 3411–3447.

BRODIE, J., I. DAUBECHIES, C. D. MOL, D. GIANNONE, AND I. LORIS (2009): "Sparse and stable Markowitz portfolios," *Proceedings of the National Academy of Sciences*, 106, 12267–12272, PMID: 19617537.

References II

DeMiguel, V., F. J. Nogales, and R. Uppal (2014): "Stock Return Serial Dependence and Out-of-Sample Portfolio Performance," *Review of Financial Studies*, 27, 1031–1073.

Fan, J., Y. Liao, and M. Mincheva (2011): "High dimensional covariance matrix estimation in approximate factor models," *Annals of statistics*, 39, 3320.

Fan, J., J. Zhang, and K. Yu (2012): "Vast Portfolio Selection With Gross-Exposure Constraints," *Journal of the American Statistical Association*, 107, 592–606.

Jagannathan, R. and T. Ma (2003): "Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps," *The Journal of Finance*, 58, 1651–1684.

Kourtis, A., G. Dotsis, and R. N. Markellos (2012): "Parameter uncertainty in portfolio selection: Shrinking the inverse covariance matrix," *Journal of Banking & Finance*, 36, 2522–2531.

Ledoit, O. and M. Wolf (2003): "Improved estimation of the covariance matrix of stock returns with an application to portfolio selection," *Journal of Empirical Finance*, 10, 603–621.

——— (2004a): "Honey, I Shrunk the Sample Covariance Matrix," *The Journal of Portfolio Management*, 30, 110–119.

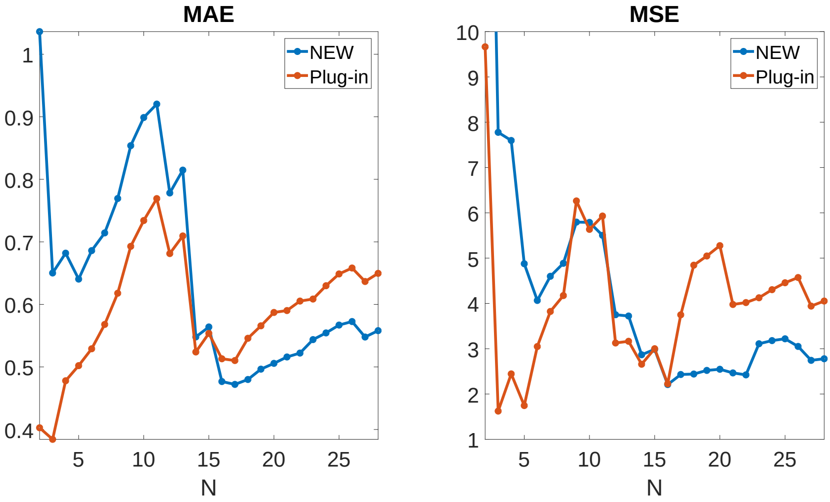
References III

——— (2004b): "A well-conditioned estimator for large-dimensional covariance matrices," *Journal of Multivariate Analysis*, 88, 365–411.

——— (2012): "Nonlinear Shrinkage Estimation of Large-dimensional Covariance Matrices," *The Annals of Statistics*, 40, 1024–1060.

Tu, J. and G. Zhou (2011): "Markowitz meets Talmud: A combination of sophisticated and naive diversification strategies," *Journal of Financial Economics*, 99, 204–215.

Appendix: GMVP



Estimation window $T = 250$, out-of-sample horizon of $H = 1000$.