Semiparametric Portfolio Policies

João F. Caldeira[2], André A. P. Santos[1,2], Hudson Torrent[3]

[1]UC3M-Santander Big Data Institute
Universidad Carlos III de Madrid

[2]Departamento de Economia
Universidade Federal de Santa Catarina

[3]Departamento de Estatística
Universidade Federal do Rio Grande do Sul

May 2021

Introduction

- High dimensional portfolio optimization is one the central topics in financial economics, multivariate statistics and machine learning.
- The traditional way requires the modelization of the first two moments of financial asset returns, e.g. Markowitz mean-variance method:

$$w = f(\hat{\mu}, \hat{\Sigma}).$$

- This method is known to fail when the cross-section dimension is large:
  - Estimation error in $\hat{\mu}$ and $\hat{\Sigma}$ inherited to portfolio weights;
  - When $N > T$, $\hat{\Sigma}$ is singular;
  - Poor out-of-sample performance.

Introduction

- Tentative remedies:
  - Shrinkage principle of Stein (1956) and Jorion (1986);
  - Optimal combinations of different portfolio policies (Kan and Zhou, 2007; Tu and Zhou, 2011);
  - Using simple investment rules (DeMiguel et al., 2009);

Introduction

- Tentative remedies:
    - Shrinkage principle of Stein (1956) and Jorion (1986);
    - Optimal combinations of different portfolio policies (Kan and Zhou, 2007; Tu and Zhou, 2011);
    - Using simple investment rules (DeMiguel et al., 2009);
    - Model portfolio weights directly as a function of firm characteristics (Brandt, Santa-Clara, and Valkanov, 2009) - Parametric portfolio policy.

Parametric portfolio policy

Brandt, Santa-Clara, and Valkanov (2009)

- The policy consists in parameterizing the portfolio weights directly as a linear function of asset-specific characteristics - yielding "parametric portfolios".
- Substantial reduction in dimensionality.
    - For a problem with N stocks, the traditional Markowitz approach requires modeling N first and $(N^2 + N)/2$ second moments of stock returns.
    - The parametric policy involves modeling only N portfolio weights.

Parametric portfolio policy

Brandt, Santa-Clara, and Valkanov (2009)

- The portfolios are obtained by adding to a given benchmark portfolio a linear combination of long-short portfolios:

$$w_t(\theta) = w_{b_t} + \left(x_{1,t}\theta_1 + x_{2,t}\theta_2 + \ldots + x_{K,t}\theta_K\right)/N_t, \qquad (1)$$

  where $x_{1,t}, \ldots, x_{K,t}$ are asset-specific characteristics cross-sectionally standardized so that they have zero mean and unit standard deviation.

Parametric portfolio policy

Brandt, Santa-Clara, and Valkanov (2009)

- A large number of asset-specific characteristics have been considered for that purpose such as book-to-market, size and momentum.
    - Fama and French (1993, 2015)
    - Feng, Giglio, and Xiu (2020)
    - Hou, Xue, and Zhang (2020)
    - DeMiguel, Martin-Utrera, Nogales, and Uppal (2020)
    - Chen and Zimmermann (2020)
    - ...

Parametric portfolio policy

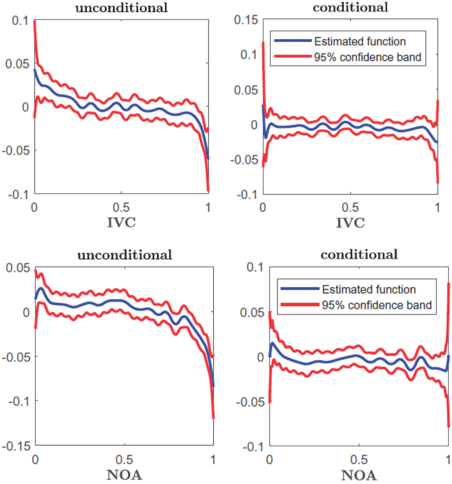
Brandt, Santa-Clara, and Valkanov (2009)

- The linear portfolio policy assumes that the sensitivity of portfolio weights to changes in firm characteristics is *constant* across all percentiles of the cross-section distribution of firm characteristic values.

- Is this assumption supported by the data?
  - Freyberger, Neuhierl, and Weber (2020) and Gu et al. (2021) find that the cross-section relation between characteristic values and asset returns is nonlinear.
  - Apparently, the data refutes the linear relation.

Parametric portfolio policies

Is the linearity assumption supported by the data?



Freyberger, J., Neuhierl, A., & Weber, M. (2020). Dissecting characteristics nonparametrically. *The Review of Financial Studies*, 33(5).

Semiparametric portfolio policies

- Can we relax the linearity restriction without compromising the advantages of the parametric policy?
- We put forward a flexible semiparametric formulation of the linear parametric policy based on the notion of penalized splines (Ruppert, Wand, and Carroll, 2003).

Semiparametric portfolio policies

Formulation

Portfolio weights are specified as an additive model of nonlinear functions of asset-specific characteristics:

$$w_t(\theta) = w_{b_t} + \left(f_1(x_{1,t}, \theta_1) + f_2(x_{2,t}, \theta_2) + \ldots + f_K(x_{K,t}, \theta_K)\right)/N_t,$$

Semiparametric portfolio policies

Formulation

Portfolio weights are specified as an additive model of nonlinear functions of asset-specific characteristics:

$$w_t(\theta) = w_{b_t} + \left(f_1(x_{1,t}, \theta_1) + f_2(x_{2,t}, \theta_2) + \ldots + f_K(x_{K,t}, \theta_K)\right) / N_t,$$

where

$$f_k(x_{k,t}, \theta_k) = \theta_{0,k} x_{k,t} + \theta_{1,k}(x_{k,t} - z_{1,k,t})_+ + \cdots + \theta_{J_k,k}(x_{k,t} - z_{J_k,k,t})_+,$$

$$(x_{k,t} - z_{1,k,t})_+ = \begin{cases} x_{k,t} - z_{1,k,t} & \text{if positive} \\ 0 & \text{otherwise} \end{cases}$$

Semiparametric portfolio policies

Formulation

Portfolio weights are specified as an additive model of nonlinear functions of asset-specific characteristics:

$$w_t(\theta) = w_{b_t} + \left(f_1(x_{1,t}, \theta_1) + f_2(x_{2,t}, \theta_2) + \ldots + f_K(x_{K,t}, \theta_K)\right) / N_t,$$

where

$$f_k(x_{k,t}, \theta_k) = \underbrace{\theta_{0,k} x_{k,t}}_{\text{Leading element}} + \overbrace{\theta_{1,k}(x_{k,t} - z_{1,k,t})_+ + \cdots + \theta_{J_k,k}(x_{k,t} - z_{J_k,k,t})_+}^{\text{Non-leading elements}},$$

$$(x_{k,t} - z_{1,k,t})_+ = \begin{cases} x_{k,t} - z_{1,k,t} & \text{if positive} \\ 0 & \text{otherwise} \end{cases}$$

Semiparametric portfolio policies

Splines

- The term $z_{j,k,t}$ in $(x_{k,t} - z_{j,k,t})_+$ is the jth *knot* of the kth characteristic.
- The knots are therefore the endpoint intervals of the function $(x_{k,t} - z_{j,k,t})_+$.
  - Often set to the percentiles of the distribution of $x_{k,t}$ in empirical applications.

Semiparametric portfolio policies

Splines

- A key aspect is to determine the number of knots (Ruppert, 2002).
    - More knots can better capture nonlinear functions, at the cost of computational complexity.

Semiparametric portfolio policies

Splines

- A key aspect is to determine the number of knots (Ruppert, 2002).
  - More knots can better capture nonlinear functions, at the cost of computational complexity.
  - Rule of thumb proposed in Ruppert, Wand, and Carroll (2003) is $J_k = \min((T-1)N_t, 35)$.

Semiparametric portfolio policies

Splines

- A key aspect is to determine the number of knots (Ruppert, 2002).
  - More knots can better capture nonlinear functions, at the cost of computational complexity.
  - Rule of thumb proposed in Ruppert, Wand, and Carroll (2003) is $J_k = \min((T-1)N_t, 35)$.
  - We set $J_k = 9$ such that each interval corresponds to a decile of the cross-section distribution of firm characteristics.

Semiparametric portfolio policies

Splines

- A key aspect is to determine the number of knots (Ruppert, 2002).
    - More knots can better capture nonlinear functions, at the cost of computational complexity.
    - Rule of thumb proposed in Ruppert, Wand, and Carroll (2003) is $J_k = \min((T-1)N_t, 35)$.
    - We set $J_k = 9$ such that each interval corresponds to a decile of the cross-section distribution of firm characteristics.
    - Cross-validation?

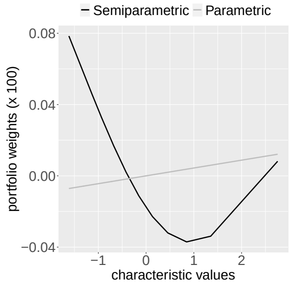
Semiparametric portfolio policies

Advantages

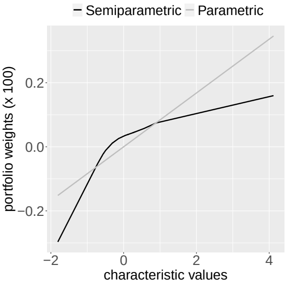
- Changes in portfolio weights can vary across percentiles of the cross-section distribution of each characteristic in a nonlinear fashion;
- When the coefficients associated to the nonleading terms are equal to zero, we recover the linear parametric portfolio specification.

Semiparametric portfolio weights

Example



(a) market equity



(b) return momentum

Semiparametric portfolio policies

Matrix notation

Let $X_t \in \mathbb{R}^{N_t \times (J+1)K}$ be a matrix containing the leading and nonleading terms of the spline basis:

$$X_{k,t} = \begin{pmatrix} x_{k,1} & (x_{k,1} - z_{1,k,1})_+ & (x_{k,1} - z_{2,k,1})_+ & \cdots & (x_{k,1} - z_{J_k,k,1})_+ \\ x_{k,2} & (x_{k,2} - z_{1,k,2})_+ & (x_{k,2} - z_{2,k,2})_+ & \cdots & (x_{k,2} - z_{J_k,k,2})_+ \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{k,N_t} & (x_{k,N_t} - z_{1,k,N_t})_+ & (x_{k,N_t} - z_{2,k,N_t})_+ & \cdots & (x_{k,N_t} - z_{J_k,k,N_t})_+ \end{pmatrix}.$$

Semiparametric portfolio policies

Matrix notation

We can compactly write the semiparametric portfolio weights as

$$w_t(\theta) = w_{b_t} + X_t\theta/N_t,$$

where $\theta^\intercal = (\theta_1^\intercal, \theta_2^\intercal, \cdots, \theta_K^\intercal) \in \mathbb{R}^{(J+1)K}$ is the parameter vector.

Semiparametric portfolio policies

Matrix notation

We can compactly write the semiparametric portfolio weights as

$$w_t(\theta) = w_{b_t} + X_t \theta / N_t,$$

where $\theta^\intercal = (\theta_1^\intercal, \theta_2^\intercal, \cdots, \theta_K^\intercal) \in \mathbb{R}^{(J+1)K}$ is the parameter vector.

The return of the semiparametric portfolio at time $t+1$ is

$$\begin{aligned} r_{p,t+1}(\theta) &= w_{b,t}^\intercal r_{t+1} + \theta^\intercal X_t^\intercal r_{t+1} / N_t \\ &= r_{b,t+1} + \theta^\intercal r_{c,t+1}. \end{aligned}$$

Semiparametric portfolio policies

Matrix notation

We can compactly write the semiparametric portfolio weights as

$$w_t(\theta) = w_{b_t} + X_t \theta / N_t,$$

where $\theta^\intercal = (\theta_1^\intercal, \theta_2^\intercal, \cdots, \theta_K^\intercal) \in \mathbb{R}^{(J+1)K}$ is the parameter vector.

The return of the semiparametric portfolio at time $t+1$ is

$$\begin{aligned} r_{p,t+1}(\theta) &= w_{b,t}^\intercal r_{t+1} + \theta^\intercal X_t^\intercal r_{t+1} / N_t \\ &= \boxed{r_{b,t+1}} + \theta^\intercal \boxed{r_{c,t+1}} \end{aligned}$$

Benchmark portfolio return

Characteristics-return vector

Semiparametric portfolio policies

Mean-variance formulation

We assume an investor that optimizes the mean-variance utility:

$$\min_{\theta} \ \frac{\gamma}{2} \mathrm{var}\left[r_{p,t+1}(\theta)\right] - E\left[r_{p,t+1}(\theta)\right],$$

Semiparametric portfolio policies

Mean-variance formulation

We assume an investor that optimizes the mean-variance utility:

$$\min_{\theta} \; \frac{\gamma}{2}\mathrm{var}\left[r_{p,t+1}(\theta)\right] - E\left[r_{p,t+1}(\theta)\right],$$

DeMiguel et al. (2020) shows that the above problem is equivalent to the quadratic optimization problem:

$$\min_{\theta} \; (\gamma/2)\theta^{\intercal}\hat{\Sigma}_c\theta + \gamma\theta^{\intercal}\hat{\sigma}_{bc} - \theta^{\intercal}\hat{\mu}_c, \tag{2}$$

Semiparametric portfolio policies

Mean-variance formulation

We assume an investor that optimizes the mean-variance utility:

$$\min_{\theta} \ \frac{\gamma}{2}\text{var}\left[r_{p,t+1}(\theta)\right] - E\left[r_{p,t+1}(\theta)\right],$$

DeMiguel et al. (2020) shows that the above problem is equivalent to the quadratic optimization problem:

$$\min_{\theta} \ (\gamma/2)\theta^{\intercal}\boxed{\hat{\Sigma}_c}\theta + \gamma\theta^{\intercal}\boxed{\hat{\sigma}_{bc}} + \theta^{\intercal}\boxed{\hat{\mu}_c}$$

Covaríance of charac. returns ($K \times K$)

Covariance between bench. and charact. returns ($K \times 1$)

Charac. mean return ($K \times 1$)

Semiparametric portfolio policies

Drawbacks of the mean-variance formulation

Two problems with this formulation:

- Depending on the number of knots considered, the problem might be excessively parameterized;
- The solution can be excessively rough depending on the values of the estimated coefficients.

Semiparametric portfolio policies

Drawbacks of the mean-variance formulation

Two problems with this formulation:

- Depending on the number of knots considered, the problem might be excessively parameterized;
- The solution can be excessively rough depending on the values of the estimated coefficients.

We consider a regularization scheme to circumvent both problems.

Semiparametric portfolio policies

Penalized mean-variance formulation

$$\min_{\theta}\ (\gamma/2)\theta^{\intercal}\hat{\Sigma}_{c}\theta+\gamma\theta^{\intercal}\hat{\sigma}_{bc}-\theta^{\intercal}\hat{\mu}_{c},$$

$$\text{subject to}\ \sum_{j=1}^{J_k}\theta_{j,k}^{2}\leq c_{k},\ \ k=1,2,\ldots,K.$$

For sufficiently small values for the penalty coefficients c_k the boils down to the traditional linear parametric portfolio.

Semiparametric portfolio policies

Penalized mean-variance formulation

Using a Lagrange multiplier approach:

$$\min_{\theta} \quad (\gamma/2)\theta^{\intercal}\hat{\Sigma}_c\theta + \gamma\theta^{\intercal}\hat{\sigma}_{bc} - \theta^{\intercal}\hat{\mu}_c + \underbrace{(1/2)\theta^{\intercal}\Omega\theta}_{\text{Penalty}},$$

Semiparametric portfolio policies

Penalized mean-variance formulation

Using a Lagrange multiplier approach:

$$\min_{\theta} \; (\gamma/2)\theta^{\intercal}\hat{\Sigma}_c\theta + \gamma\theta^{\intercal}\hat{\sigma}_{bc} - \theta^{\intercal}\hat{\mu}_c + \underbrace{(1/2)\theta^{\intercal}\Omega\theta}_{\text{Penalty}},$$

where Ω is a diagonal matrix containing the penalty parameters.

Semiparametric portfolio policies

Penalized mean-variance formulation

Using a Lagrange multiplier approach:

$$\min_{\theta} \ (\gamma/2)\theta^\intercal \hat{\Sigma}_c \theta + \gamma \theta^\intercal \hat{\sigma}_{bc} - \theta^\intercal \hat{\mu}_c + \underbrace{(1/2)\theta^\intercal \Omega \theta}_{\text{Penalty}},$$

where Ω is a diagonal matrix containing the penalty parameters.

The selection of penalty values takes into account the in-sample transaction costs measured by the portfolio turnover.

Penalized mean-variance semiparametric portfolios

Analytical results

Proposition

The solution to the semiparametric portfolio problem is given by

$$\hat{\theta} = \left(\gamma\hat{\Sigma}_c + \Omega\right)^{-1}\left(\hat{\mu}_c - \gamma\hat{\sigma}_{bc}\right).$$

The solution can be viewed as a *ridge* estimator: The matrix Ω is responsible to shrink the nonlinear function, given by the splines representation, towards a more parsimonious model.

Penalized mean-variance semiparametric portfolios

Analytical results

Proposition

For a given risk aversion parameter γ *and a penalty matrix* Ω, *the optimal parameter vector* θ^* *for the penalied mean-variance semiparametric portfolio problem is equal to the penalized spline estimate of the slope vector in the following penalized regression model:*

$$r_{b,t} = \alpha - \beta^\intercal r_{c,t} + \epsilon_t$$

subject to the constraints that

$$\beta^\intercal \mu_c = (\theta^*)^\intercal \mu_c \qquad \textit{and} \qquad \beta^\intercal \Omega \beta = (\theta^*)^\intercal \Omega \theta^*.$$

Empirical application

Empirical application

- We collect from CRSP monthly data on 16,855 US companies traded on NYSE, NYMEX and NASDAQ.
- Sample period: 1970:11 until 2018:12.
- We follow Brandt, Santa-Clara, and Valkanov (2009) and compute for each company 3 characteristics:
  - Size;
  - Book-to-market;
  - Momentum (cumulative returns in the last 12 months excluding the last month).
- Portfolio policies considered:
  - Mean-variance parametric portfolio;
  - Mean-variance semiparametric portfolio;
  - Equally-weighted portfolios;
  - Value-weighted portfolios.

Empirical application

Out-of-sample evaluation: methodology

- Portfolios are re-balanced on a monthly basis.
- Parametric and semiparametric portfolio coefficients are estimated using a expanding window with initial length of $T = 120$ months.
- We compute for the $L - T = 458$ out-of-sample months the following quantities:
  - Average monthly returns;
  - Standard deviation of average returns;
  - Sharpe ratio;
  - Portfolio turnover ;
  - p-values for the differences in Shape ratio and portfolio variance using the bootstrap test of Ledoit and Wolf (2008).

Empirical application

Out-of-sample evaluation: results

	Mean Return (%)	Std. Dev. (%)	Sharpe Ratio
	Transaction cost = 0 b.p.		
Semiparametric	0.256	0.161**	1.339***
Parametric	0.198	0.163**	0.962**
Equally-weighted	0.141	0.185	0.544
Value-weighted	0.092	0.149***	0.343**

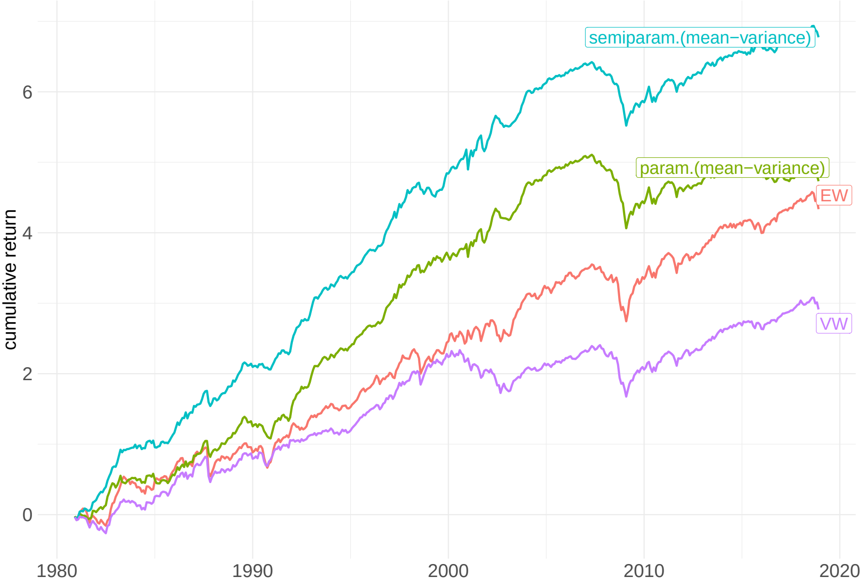
Empirical application

Out-of-sample evaluation: results

	Mean Return (%)	Std. Dev. (%)	Sharpe Ratio
	Transaction cost = 50 b.p.		
Semiparametric	0.192	0.160**	0.944**
Parametric	0.138	0.163**	0.598
Equally-weighted	0.132	0.184	0.494
Value-weighted	0.088	0.149***	0.317***

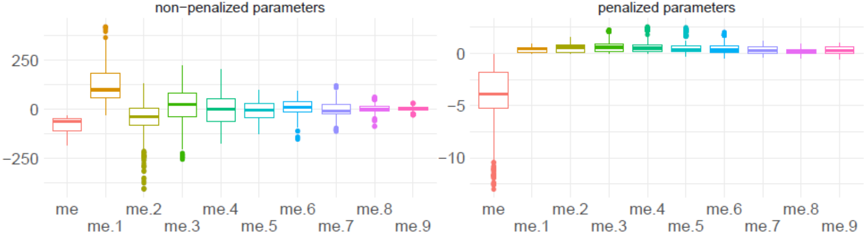
Semiparametric portfolio policies

Cumulative portfolio returns

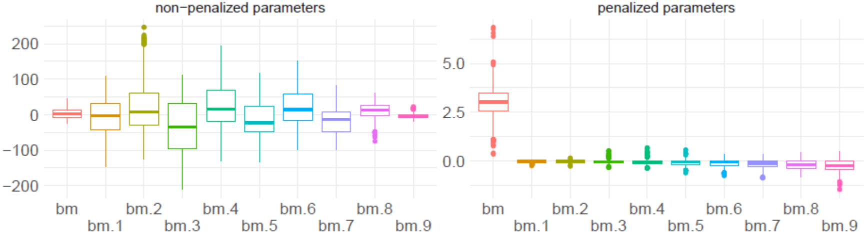


Semiparametric portfolio policies

Estimated parameters



Panel (a): Market equity



Panel (b): Book−to−market

Empirical application

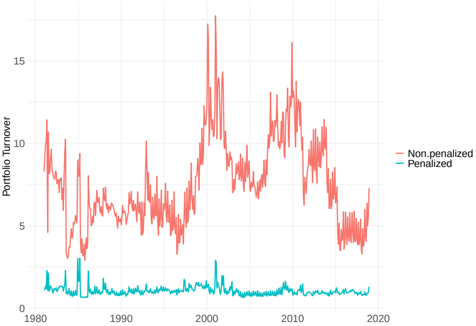
Out-of-sample evaluation: penalization scheme

	Mean Return (%)	Std. Dev. (%)	Sharpe Ratio	Turnover	Mean Return (%)	Std. Dev. (%)	Sharpe Ratio
	Panel A: Transaction cost = 0 b.p.				Panel B: Transaction cost = 50 b.p.		
Non-penalized semiparametric portfolios	0.590	0.221***	2.474***	7.492	0.118	0.214**	0.359

Empirical application

Out-of-sample evaluation: penalization scheme

	Mean Return (%)	Std. Dev. (%)	Sharpe Ratio	Turnover	Mean Return (%)	Std. Dev. (%)	Sharpe Ratio
	Panel A: Transaction cost = 0 b.p.				Panel B: Transaction cost = 50 b.p.		
Non-penalized semiparametric portfolios	0.590	0.221***	2.474***	7.492	0.118	0.214**	0.359



Concluding remarks

- We have proposed an alternative formulation of the linear parametric portfolio policies that allows for a nonlinear relationship between portfolio weights and characteristics.

- The novel approach is based on the notion of penalized splines and allows us to specify a set of nonlinear (piece-wise linear) additive function that maps characteristics and portfolio weights.
  - The linear specification is a special case of our approach.

- Our empirical application reveals that allowing for a more flexible relation between asset characteristics and portfolio weights translates into portfolios with better out-of-sample performance.

References I

Michael W Brandt, Pedro Santa-Clara, and Rossen Valkanov. Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. *The Review of Financial Studies*, 22(9):3411–3447, 2009.

Andrew Y Chen and Tom Zimmermann. Open source cross-sectional asset pricing. *SSRN Working Paper*, 2020.

Victor DeMiguel, Lorenzo Garlappi, and Raman Uppal. Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *The review of Financial studies*, 22(5):1915–1953, 2009.

Victor DeMiguel, Alberto Martin-Utrera, Francisco J Nogales, and Raman Uppal. A transaction-cost perspective on the multitude of firm characteristics. *The Review of Financial Studies*, 33(5):2180–2222, 2020.

Eugene F Fama and Kenneth R French. Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 33(1):3–56, 1993.

Eugene F Fama and Kenneth R French. A five-factor asset pricing model. *Journal of financial economics*, 116(1): 1–22, 2015.

Guanhao Feng, Stefano Giglio, and Dacheng Xiu. Taming the factor zoo: A test of new factors. *The Journal of Finance*, 75(3):1327–1370, 2020.

Joachim Freyberger, Andreas Neuhierl, and Michael Weber. Dissecting characteristics nonparametrically. *The Review of Financial Studies*, 33(5):2326–2377, 2020.

Shihao Gu, Bryan Kelly, and Dacheng Xiu. Autoencoder asset pricing models. *Journal of Econometrics*, 222(1): 429–450, 2021.

Kewei Hou, Chen Xue, and Lu Zhang. Replicating anomalies. *The Review of Financial Studies*, 33(5):2019–2133, 2020.

Philippe Jorion. Bayes-stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis*, 21 (03):279–292, 1986.

Raymond Kan and Guofu Zhou. Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis*, 42(03):621–656, 2007.

O. Ledoit and M. Wolf. Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance*, 15(5):850–859, 2008.

References II

David Ruppert. Selecting the number of knots for penalized splines. *Journal of computational and graphical statistics*, 11(4):735–757, 2002.

David Ruppert, Matt P Wand, and Raymond J Carroll. *Semiparametric regression*. Cambridge University Press, 2003.

Charles Stein. Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In *Proceedings of the Third Berkeley symposium on mathematical statistics and probability*, volume 1, pages 197–206, 1956.

Jun Tu and Guofu Zhou. Markowitz meets talmud: A combination of sophisticated and naive diversification strategies. *Journal of Financial Economics*, 99(1):204–215, 2011.