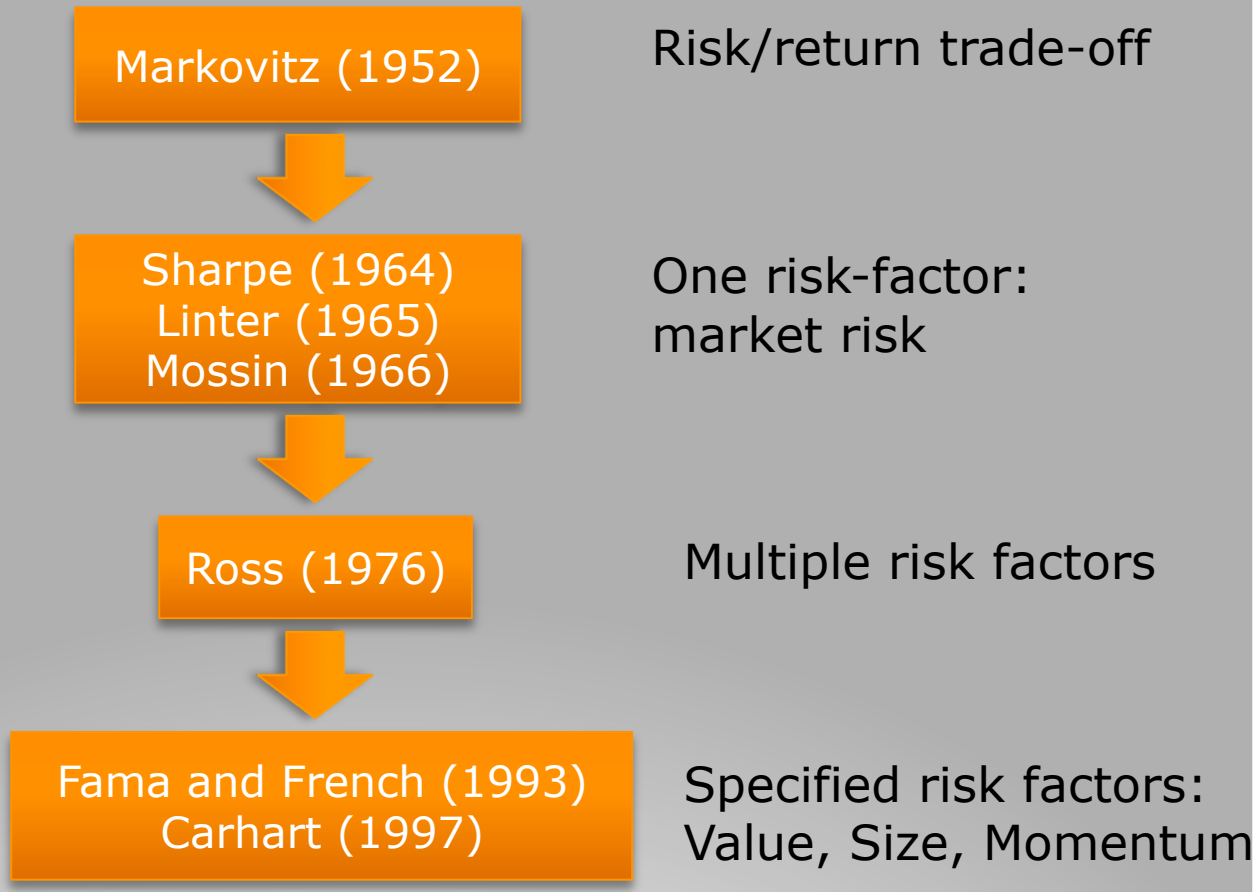
Alternative models to portfolio selection in the Brazilian market

FinanceHub, 13th/Dec/2021

Presented by:
Helder Palaro

Key papers(1)


Markovitz (1952)
Risk/return trade-off
Sharpe (1964)
Linter (1965)
Mossin (1966)
One risk-factor:
market risk
Ross (1976)
Multiple risk factors
Fama and French (1993)
Carhart (1997)
Specified risk factors:
Value, Size, Momentum

Key papers (2)

Cochrane (2011)

Many papers since 1993 about different risk factors. Factor "zoo".

Kelly et al (2019)

IPCA Model: Latent risk factors estimated by PCA, characteristics as instrumental variables.

Gu et al (2020)
Gu et al (2021)

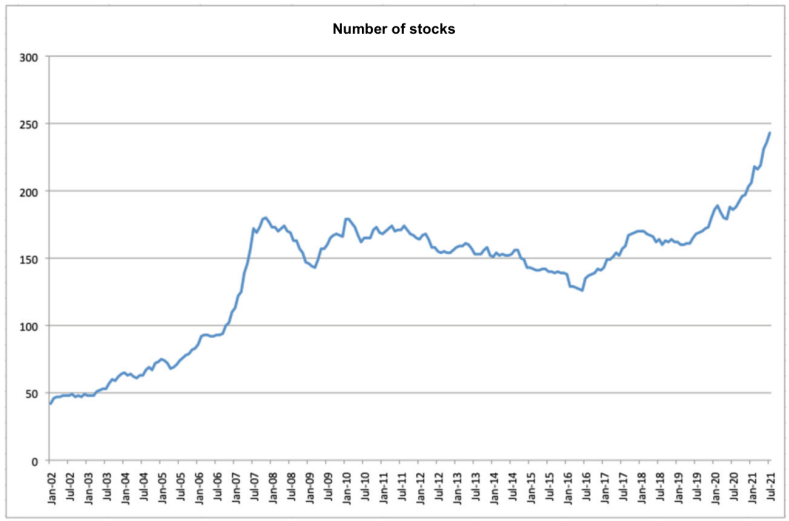
Machine Learning methods.

Rubesam (2019)

Machine Learning for Brazilian stocks.

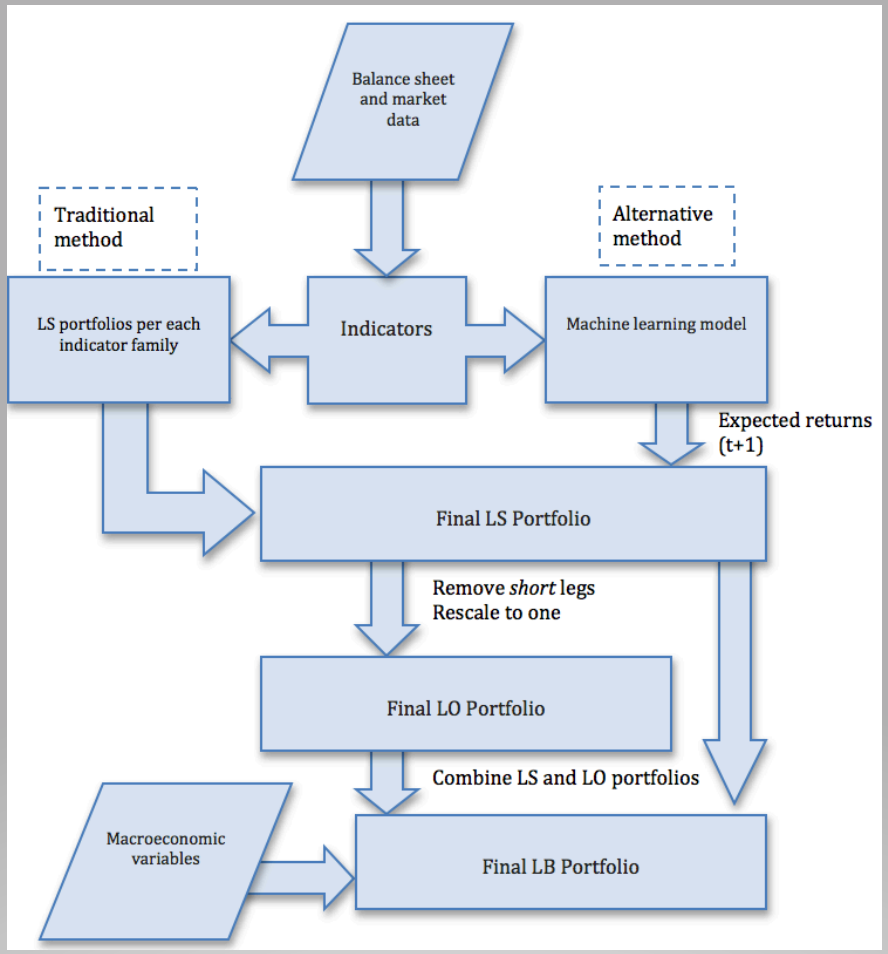
Data

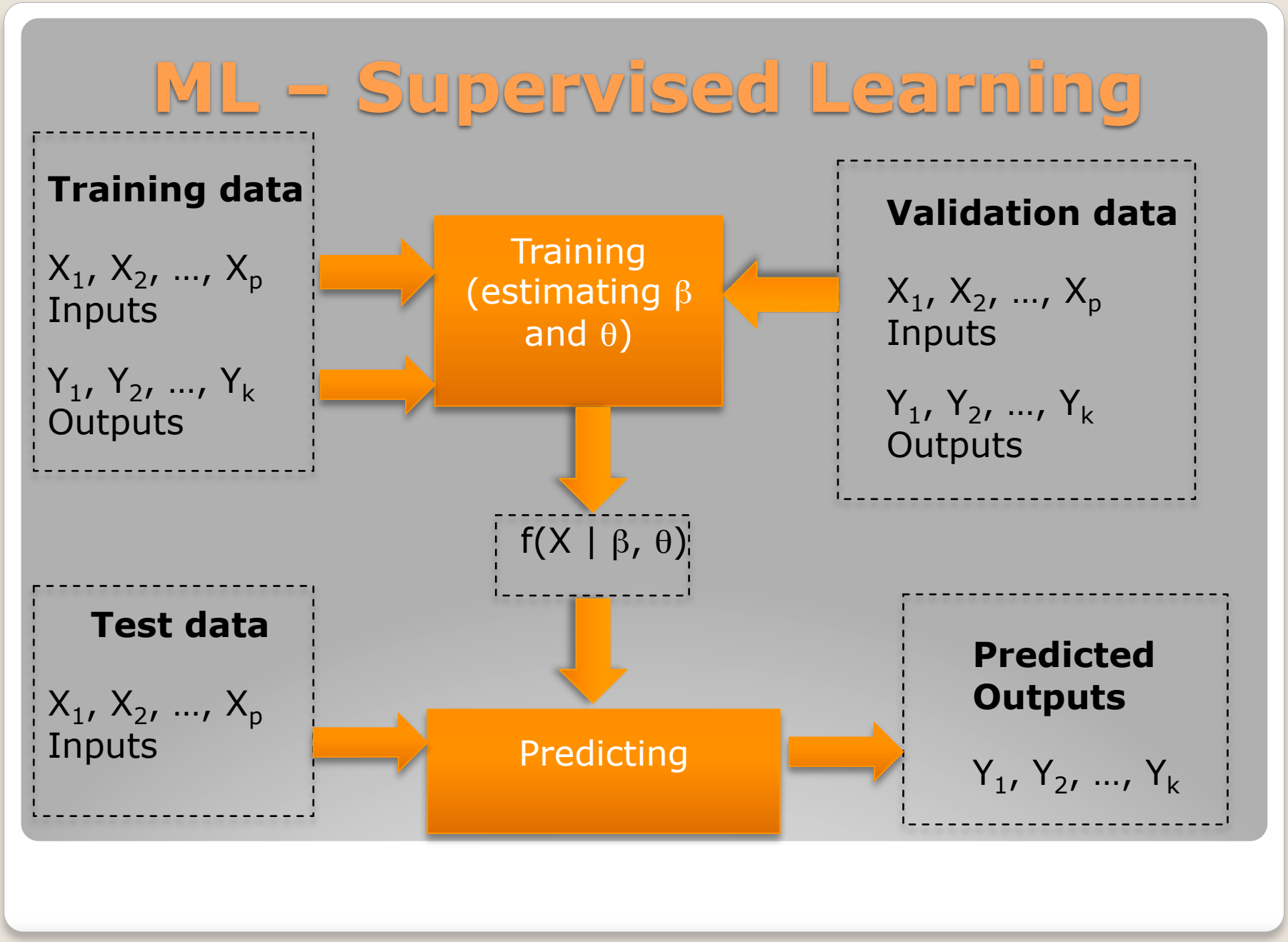
- 413 Brazilian stocks in the January 2003-July 2021 period.
- Liquidity filter
- 107 fundamental and technical indicators



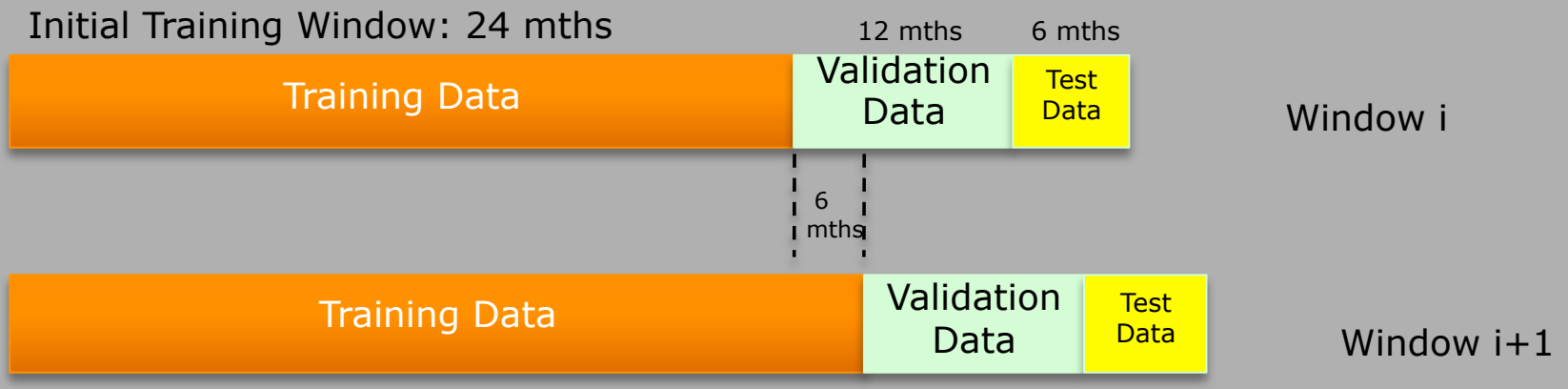
Family	Number of Risk Factors
Value	14
Momentum	23
Growth	31
Quality	32
Low risk	7
Total	107

Portfolio Construction




ML – Supervised Learning
Training data
$X_1, X_2, \ldots, X_p$ Inputs
$Y_1, Y_2, \ldots, Y_k$ Outputs
Training (estimating β and θ)
Validation data
$X_1, X_2, \ldots, X_p$ Inputs
$Y_1, Y_2, \ldots, Y_k$ Outputs
f(X | β, θ)
Test data
$X_1, X_2, \ldots, X_p$ Inputs
Predicting
Predicted Outputs
$Y_1, Y_2, \ldots, Y_k$

ML – Sampling



$X =$

Stock 1: $X_1, X_2, \ldots, X_p$

Stock 2: $X_1, X_2, \ldots, X_p$

...

Stock n: $X_1, X_2, \ldots, X_p$

$y =$

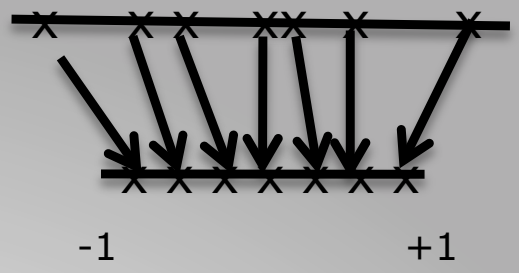
Stock 1: r

Stock 2: r

...

Stock n: r

Transform Xs and y:



ML Models

1) OLS (Ordinary Least Squares)

$$Y = X\beta + \varepsilon,$$
$$\varepsilon \sim N(0, \sigma^2).$$

$$L_{OLS}(\hat{\beta}) = \sum_{i-1}^{n}(y_i - x_i'\hat{\beta})^2 = ||y - X\hat{\beta}||^2$$

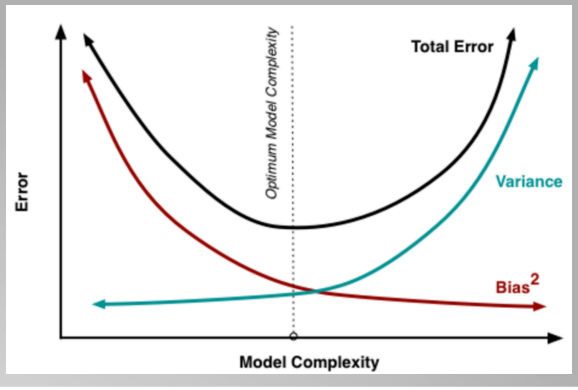
$$\hat{\beta}_{OLS} = (X'X)^{-1}(X'Y)$$

2) Ridge regression

$$L_{ridge}(\hat{\beta}) = \sum_{i-1}^{n}(y_i - x_i'\hat{\beta})^2 + \lambda\sum_{j-1}^{m}\hat{\beta}_j^2 = ||y - X\hat{\beta}||^2 + \lambda||\hat{\beta}||^2.$$

$$\hat{\beta}_{ridge} = (X'X + \lambda I)^{-1}(X'Y)$$

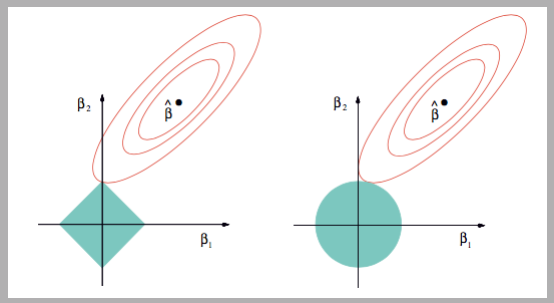
Hyperparameter λ is the level of penalization.
$\lambda \in [0.001, 1000]$



ML Models

3) Lasso

$$L_{lasso}(\hat{\beta}) = \sum_{i-1}^{n}(y_i - x_i'\hat{\beta})^2 + \lambda \sum_{j-1}^{m}|\hat{\beta}_j|.$$



Hyperparameter λ is the level of penalization.
$\lambda \in [0.000001, 0.1]$

- Lasso can set some coefficient to zero, Ridge cannot.

- In Ridge, the coefficients of correlated predictors are similar, in Lasso one of the correlated predictors has a larger coefficient, while the rest are zeroed.

- Lasso tends to do well if there is a small number of significant parameters. Ridge when there are many.

ML Models

4) PCR

M Principal Components: m = 1, 2, …, M

$$\mathbf{z}_m = \mathbf{X}v_m$$

Regress y on Z. Since z_i are orthogonal, just add the univariate regressions.

$$\hat{\mathbf{y}}^{\text{pcr}}_{(M)} = \bar{y}\mathbf{1} + \sum_{m=1}^{M} \hat{\theta}_m \mathbf{z}_m$$

Rewrite as coefficients on X:

$$\hat{\beta}^{\text{pcr}}(M) = \sum_{m=1}^{M} \hat{\theta}_m v_m$$

5) PLS

Similar to PCR, but it uses y as well as X. In the construction of z_m, the inputs are weighted by the strength of univariate effect on y.

PCR	Hyperparameter is the number of principal components *pc*.
PLS	*pc* ∈ {2, 3, ..., 20}

ML Models

6) IPCA

$$r_{i,t+1} = \beta_{i,t} f_{t+1} + \epsilon_{i,t+1}$$

$$\beta_{i,t} = z'_{i,t} \Gamma_\beta + \nu_{\beta,i,t}$$

N assets *r*
K factors *f*
L characteristics *z*
T data points *t*

As in Fama-French, risk premia are still determined by exposures to risk factors *f,* but as in PCA, these factors are considered latent.

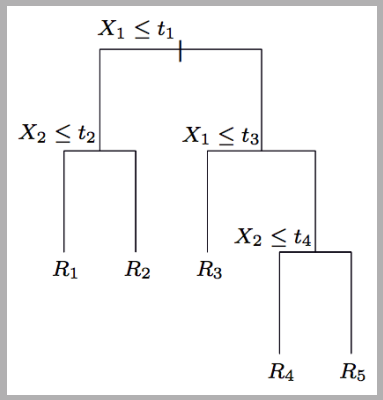
Asset characteristics *z* (P/E, P/B, etc) serve as instrumental variables to the time-varying conditional loadings β.

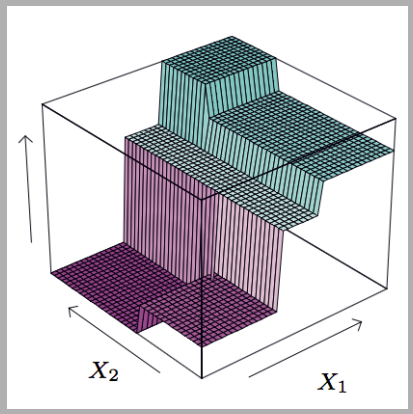
The mapping Γ from characteristics *z* to loadings β is fixed over time and across individuals.

ML Models

7) Random Forests

CART: Find split points and variables which minimize sum of squares.



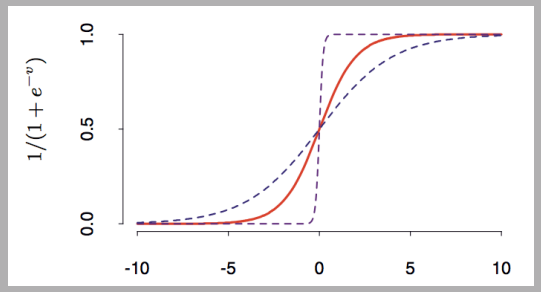


- Random forest is based on idea of bagging (bootstrap aggregation).
-We draw B bootstrap samples from the original sample. For each sample, we subsample m <= p variables, and grow a tree to the data.
- Finally, we average the predictions.

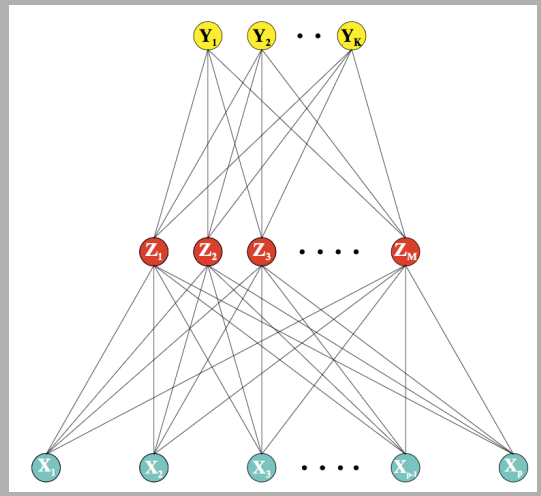
Hyperparameter is the number of variables to be sub-sampled for the decision tree construction. The initial guess is the number of indicators divided b 3. We build 100 trees. Nodesize controls the number of nodes in each tree = 5.

ML Models

8 and 9) Neural Networks



$$Z_m = \sigma(\alpha_{0m} + \alpha_m^T X),\ m = 1, \ldots, M$$
$$T_k = \beta_{0k} + \beta_k^T Z,\ k = 1, \ldots, K,$$
$$Y_k = g_k(T),\ k = 1, \ldots, K,$$



Neural Network 1	One intermediate layer with 4 neurons. Hyperparameter *thr* is the stopping point for the algorithm. *thr* $\in [0.01, 0.2]$. We train 50 neural networks, and then average them to reduce the variability.
Neural Network 2	Similar to above, but we have 3 intermediate layers, with 4, 2 and 1 neurons, respectively.

ML Models

10) Comb4

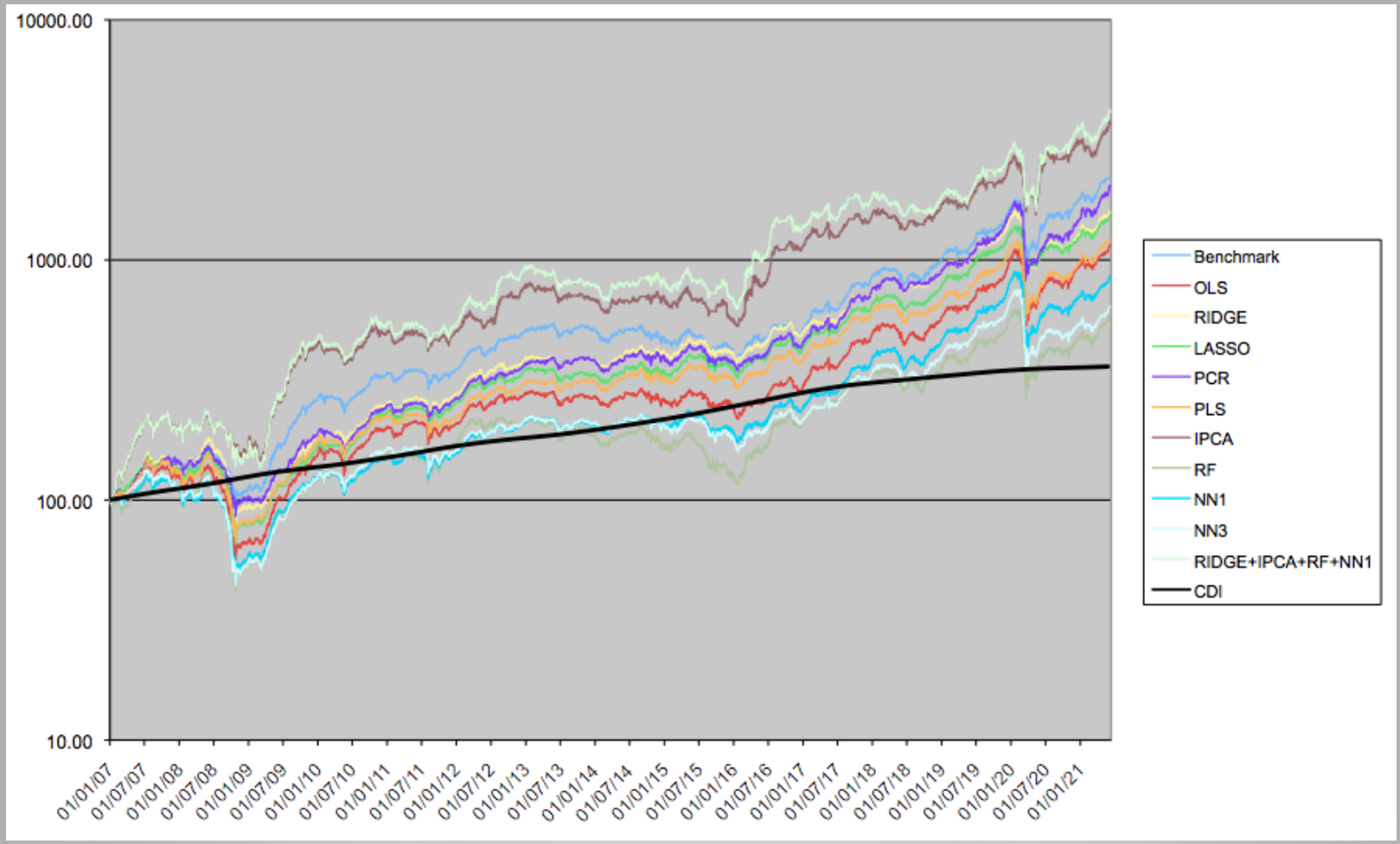
Simple average between Ridge, IPCA, Random Forests and Neural Network 1.

Portfolio LO

	Bench	OLS	RIDGE	LASSO	PCR	PLS	IPCA
Average return p.a.	23.84%	19.31%	21.27%	20.67%	22.85%	19.64%	28.20%
Volatility p.a.	20.51%	22.24%	20.42%	19.07%	21.05%	21.20%	25.38%
Information Ratio[4]	1.16	0.87	1.04	1.08	1.09	0.93	1.11
Hit Ratio	56.75%	55.85%	56.33%	55.90%	55.90%	56.54%	47.15%
Max Daily Loss	-14.75%	-16.40%	-16.43%	-12.24%	-17.04%	-16.96%	-12.14%
Max Drawdown	46.78%	61.35%	57.35%	59.93%	53.22%	56.82%	43.36%
Average Recover Time	11	28	18	18	15	17	5
Maximum Recover Time	40	28	18	18	16	17	38

	Random Forests	Neural Network 1	Neural Network 2	Comb4
Average return p.a.	15.50%	17.24%	15.18%	29.64%
Volatility p.a.	26.08%	22.45%	22.21%	27.92%
Information Ratio	0.59	0.77	0.68	1.06
Hit Ratio	55.66%	55.71%	55.38%	56.66%
Max Daily Loss	-17.62%	-17.37%	-16.91%	-13.81%
Max Drawdown	70.95%	64.47%	67.03%	66.30%
Average Recover Time	53	29	35	35
Maximum Recover Time	53	29	35	35

Portfolio LO

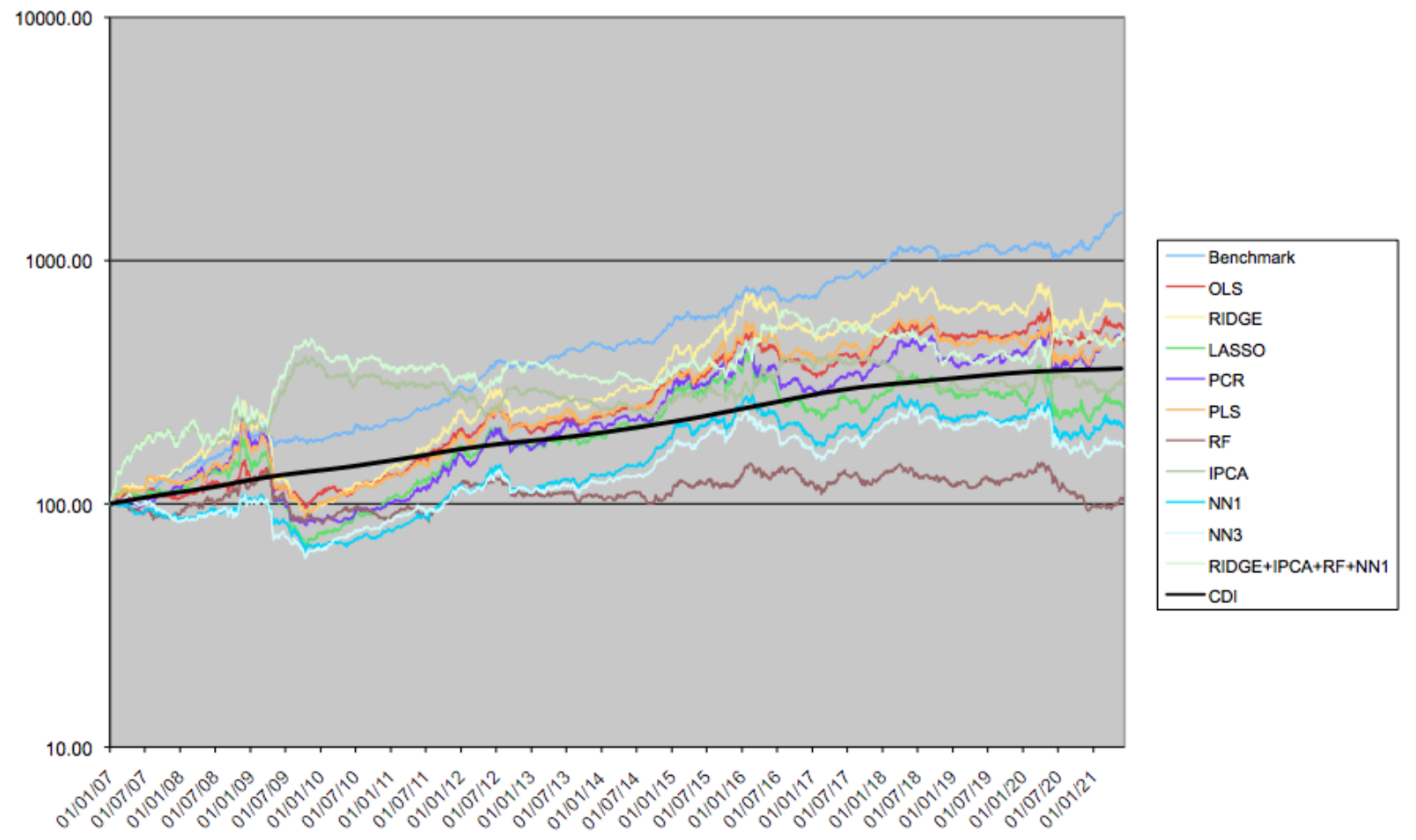


Portfolio LS

	Bench	OLS	RIDGE	LASSO	PCR	PLS	IPCA
Average return p.a.	20.02%	13.28%	15.03%	8.55%	12.91%	12.79%	9.83%
Volatility p.a.	11.20%	16.19%	18.96%	19.35%	19.50%	18.77%	18.97%
Information Ratio	1.79	0.82	0.79	0.44	0.66	0.68	0.10
Hit Ratio	55.32%	53.85%	53.42%	52.63%	52.71%	53.23%	44.65%
Max Daily Loss	-4.98%	-7.80%	-8.84%	-9.15%	-10.09%	-9.54%	-6.72%
Max Drawdown	18.29%	36.66%	66.38%	64.23%	65.82%	59.55%	43.85%
Average Recover Time	19	29	41	41	69	40	87
Maximum Recover Time	31	29	41	64	69	40	87

	Random Forests	Neural Network 1	Neural Network 2	Comb4
Average return p.a.	1.38%	6.58%	5.30%	13.13%
Volatility p.a.	14.21%	15.48%	15.63%	20.65%
Information Ratio	0.52	0.43	0.34	0.64
Hit Ratio	52.82%	54.04%	53.23%	55.42%
Max Daily Loss	-6.03%	-7.38%	-7.18%	-7.32%
Max Drawdown	51.93%	42.68%	45.56%	42.97%
Average Recover Time	133	34	54	32
Maximum Recover Time	133	38	54	38

Portfolio LS

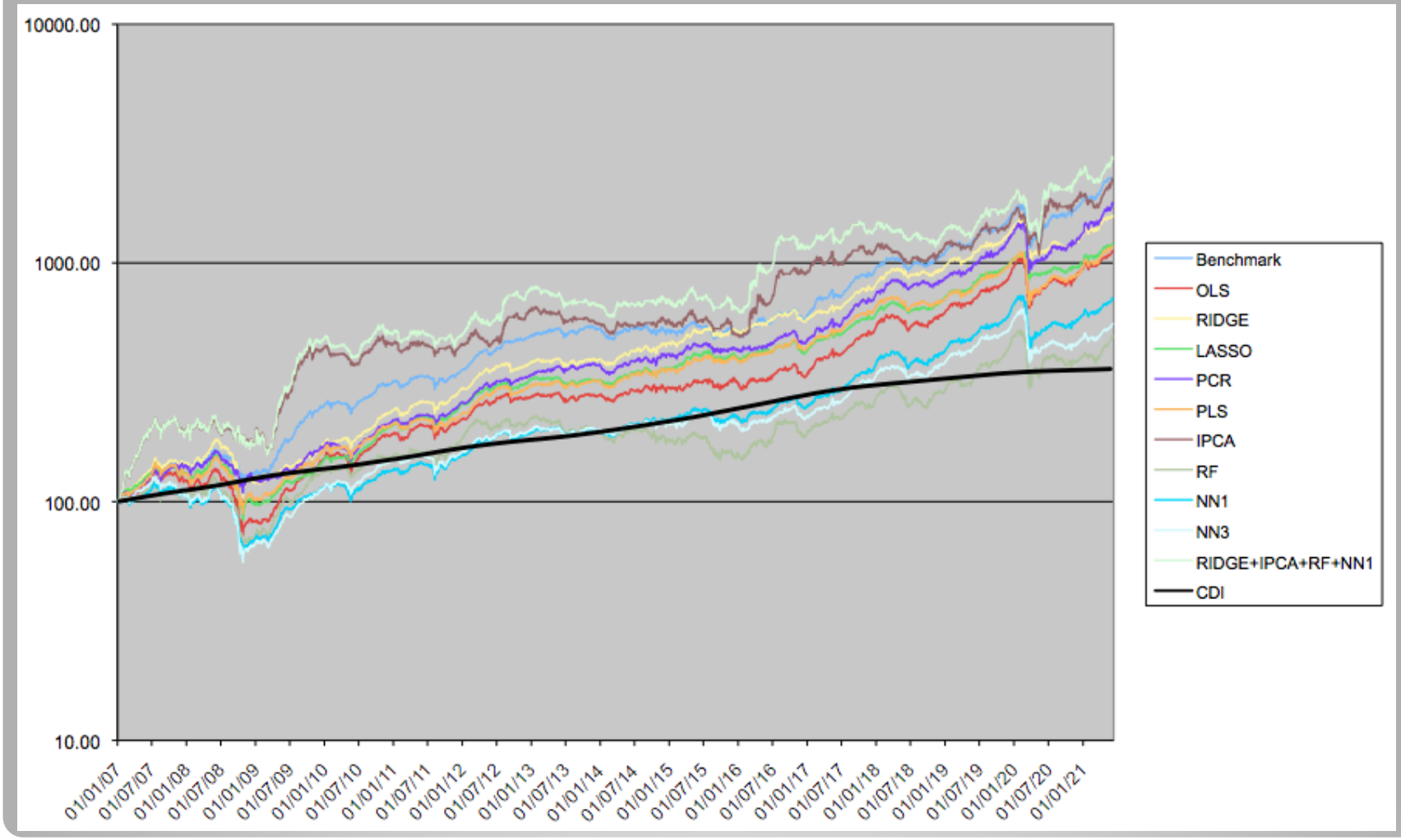


Portfolio LB

	Bench	OLS	RIDGE	LASSO	PCR	PLS	IPCA
Average return p.a.	23.20%	18.10%	20.37%	18.27%	20.98%	18.36%	23.54%
Volatility p.a.	15.68%	16.48%	14.89%	13.79%	15.51%	15.68%	21.17%
Information Ratio	1.48	1.10	1.37	1.33	1.35	1.17	1.11
Hit Ratio	56.90%	55.33%	56.78%	56.33%	55.49%	56.19%	46.53%
Max Daily Loss	-10.70%	-11.11%	-10.88%	-7.78%	-11.32%	-11.54%	-9.17%
Max Drawdown	35.54%	48.03%	42.71%	47.06%	38.32%	42.00%	36.63%
Average Recover Time	10	27	19	24	11	18	5
Maximum Recover Time	11	27	19	24	17	18	38

	Random Forests	Neural Network 1	Neural Network 2	Comb4
Average return p.a.	12.82%	14.88%	13.16%	25.40%
Volatility p.a.	19.34%	16.56%	16.45%	23.06%
Information Ratio	0.66	0.90	0.80	1.10
Hit Ratio	55.30%	56.00%	55.52%	57.07%
Max Daily Loss	-11.26%	-11.77%	-11.35%	-10.40%
Max Drawdown	62.38%	52.66%	56.09%	54.66%
Average Recover Time	52	36	36	35
Maximum Recover Time	52	36	36	35

Portfolio LB



Correlation - LB

Correlation Matrix	Bench	OLS	RIDGE	LASSO	PCR	PLS	IPCA	RF	NN1	NN2	Comb4
Bench	1.00	0.87	0.86	0.83	0.86	0.86	0.68	0.81	0.86	0.85	0.69
OLS		1.00	0.90	0.86	0.85	0.92	0.57	0.80	0.94	0.92	0.58
RIDGE			1.00	0.92	0.89	0.94	0.53	0.75	0.89	0.87	0.54
LASSO				1.00	0.84	0.89	0.51	0.73	0.85	0.84	0.52
PCR					1.00	0.89	0.52	0.72	0.83	0.81	0.53
PLS						1.00	0.53	0.77	0.90	0.89	0.54
IPCA							1.00	0.58	0.55	0.55	0.96
Random Forests								1.00	0.84	0.85	0.60
Neural Network 1									1.00	0.96	0.57
Neural Network 2										1.00	0.57
Comb4											1.00

Top 5 indicators

OLS	RIDGE	LASSO	PCR	PLS
Slope_50	RSI_14	DowVol_1y	AccRet_5	AccRet_5
UpVol_1y	AccRet_5	Vol_1y	IndustryRel5	IndustryRel5
P/Book Value	IndustryRel5	R1y	Vol_1y	Vol_1y
NetEquity_G2Q	UpVol_1Y	Mom_10_40wk	R1y	R1y
LogMarketCap	R6m	AccRet_250_20	UpVol_1y	UpVol_1y
IPCA	**Random Forests**	**Neural Network 1**	**Neural Network 2**	
EVSales	AccRet_10	FCFOA_3Y	DownVol_3m	
DividendYield	MoneyFlow5	DownVol_6m	EBITOA_3Y	
EBITGrowth2Q	Slope_50	ROIC_Mean3Y	AccRet_120	
IndustryRel20	MACD	Vol_1y	RSI_14	
P/BookValue	IndustryRel5	AccRet_250_20	UpVol_1y	

Conclusion

-For long-short portfolios, the traditional method of portfolio construction is still the best.

- For the long-only and long-biased cases, we believe there is a justification for an experimental allocation to the alternative models.

- The main candidates for this allocation are the IPCA model, the linear models with penalization, or the combination of four models.

- Given that the IPCA model has the best interpretability and theoretical foundation, we believe it is the best candidate.