THE UNIVERSITY OF TEXAS AT AUSTIN
McCOMBS SCHOOL OF BUSINESS

Regret-based Selection

David Puelz (UT Austin)
Carlos M. Carvalho (UT Austin)
P. Richard Hahn (Chicago Booth)

May 27, 2017

Two problems

1. Asset pricing: What are the fundamental dimensions (risk factors) of the financial market?

Two problems

1. Asset pricing: What are the fundamental dimensions (risk factors) of the financial market?

2. Investing: Among thousands of choices, which passive funds should I invest in?

Two problems

1. Asset pricing: What are the fundamental dimensions (risk factors) of the financial market?

2. Investing: Among thousands of choices, which passive funds should I invest in?

How are they connected?

The context for this talk

Both problems can be studied using variable selection techniques from statistics.

Separating priors from utilities

Our view: Subset selection is a decision problem. We need a suitable loss function, **not** a more clever prior.

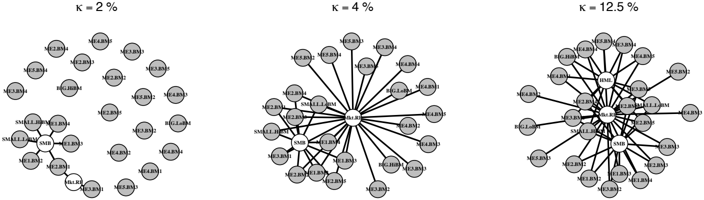
Separating priors from utilities

Our view: Subset selection is a decision problem. We need a suitable loss function, **not** a more clever prior.

This leads us to think of selection in a "post-inference world" by comparing models based on regret.

Where we are headed ...

- Risk factor selection in SUR models



- Sparse dynamic portfolios

Date	DIA	IWD	IWB	IWN	IWM	IYR
	Dow Jones	Value	Large	Small	Small value	Real estate
2002	-	21.6	23.7	2.55	2.83	49.3
2003	-	18.4	18.2	-	-	63.3
2004	-	14.1	22.3	-	-	63.6
2005	-	31.2	35.2	-	-	33.6
2006	-	32.7	40.6	-	-	26.7

Regret-based selection: Primitives

Let d be a decision, λ be a complexity parameter, Θ be a vector of model parameters, and $\tilde{Y}$ be future data.

1. Loss function $\mathcal{L}(d, \tilde{Y})$ – measures utility.

2. Complexity function $\Phi(\lambda, d)$ – measures sparsity.

3. Statistical model $\Pi(\Theta)$ – characterizes uncertainty.

4. Regret tolerance κ – characterizes degree of comfort from deviating from a "target decision" (in terms of posterior probability).

Regret-based selection: Procedure

- Optimize expected loss (1) + complexity (2). The expectation is over $p(\tilde{Y}, \Theta \mid \mathbf{Y})$ (3).

- Calculate regrets versus a target d^* for decisions indexed by λ.

  $\rightarrow \;\; \rho(d_\lambda, d^*, \tilde{Y}) = \mathcal{L}(d_\lambda, \tilde{Y}) - \mathcal{L}(d^*, \tilde{Y})$

- Select d^*_λ as the decision satisfying the regret tolerance.

  $\rightarrow \;\; \pi_\lambda = \mathbb{P}[\rho(d_\lambda, d^*, \tilde{Y}) < 0]$

  $\rightarrow$ Select d^*_λ s.t. $\pi_{d^*_\lambda} > \kappa$ (3,4)

Which risk factors matter?

The Factor Zoo (Cochrane, 2011)

- Market
- Size
- Value
- Momentum
- Short and long term reversal
- Betting against β
- Direct profitability
- Dividend initiation
- Carry trade
- Liquidity
- Quality minus junk
- Investment
- Leverage
- ...

The Factor Zoo (Cochrane, 2011)

- Market
- Size
- Value
- Momentum
- Short and long term reversal
- Betting against β
- Direct profitability
- Dividend initiation
- Carry trade
- Liquidity
- Quality minus junk
- Investment
- Leverage
- ...

The setup for determining important factors

Let the return on test assets be R, and the return on factors be F.

$$R = \gamma F + \epsilon,\ \epsilon \sim N(0, \Psi)$$

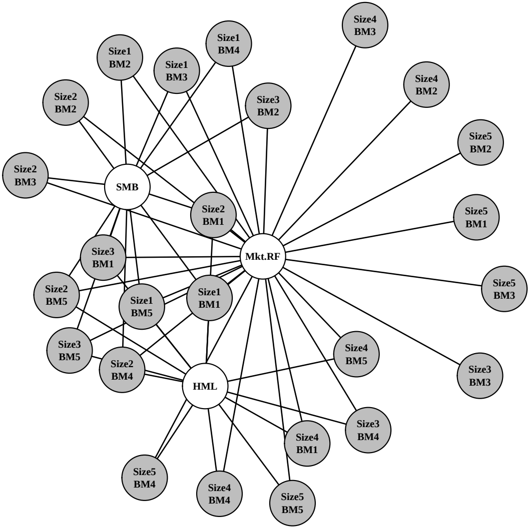
Primitives:

1. Loss: $\mathcal{L}(\gamma, \tilde{R}, \tilde{F}) = -\log p(\tilde{R}|\tilde{F})$
2. Complexity: $\Phi(\lambda, \gamma) = \lambda \|\gamma\|_1$.
3. Model: $R|F$ with normal errors and conjugate g-priors and F via gaussian linear latent factor model.
4. Regret tolerance: Let's consider several κ's.

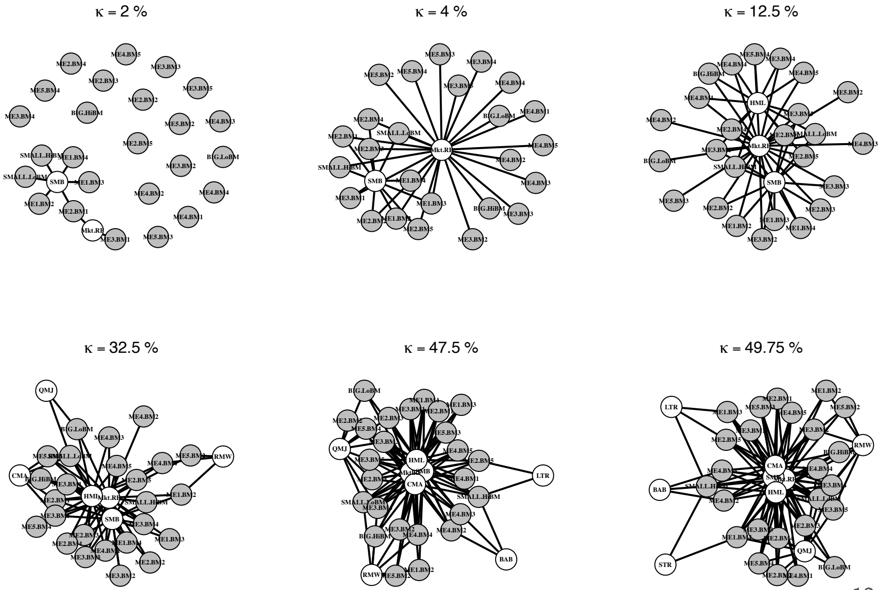
Assume the target is the $\lambda = 0$ model.

Factor selection graph ($\kappa = 12.5\%$)

R: 25 Fama-French portfolios, F: 10 factors from finance literature



Selected graphs under different regret tolerances κ



Passive Investing

∃ thousands of investment opportunities

SPDR

STATE STREET GLOBAL ADVISORS

iShares®

The world's No.1 ETF provider

BLACKROCK

Vanguard®

The setup for sparse passive investing

- Let $\tilde{R}_t$ be a vector of N future asset returns.
- Let w_t be the portfolio weight vector (decision) at time t.
- We use the log cumulative growth rate for our utility!

Primitives:

1. Loss: $-\log\left(1 + \sum_{k=1}^{N} w_t^k \tilde{R}_t^k\right)$
2. Complexity: $\lambda_t \|w_t\|_1$
3. Model: DLM for $\tilde{R}_t$ parameterized by (μ_t, Σ_t)
4. Regret tolerance: $\kappa = 55\%$.

Assume the target is fully invested (dense) portfolio.

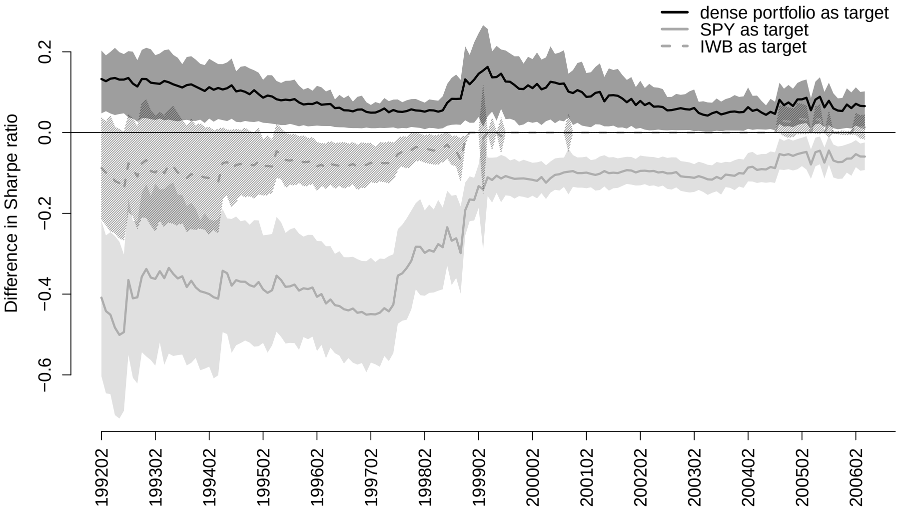
Static regret tolerance → dynamic portfolio decisions

Data: Returns on 25 ETFs from 1992-2016. $\kappa = 55\%$ decision.

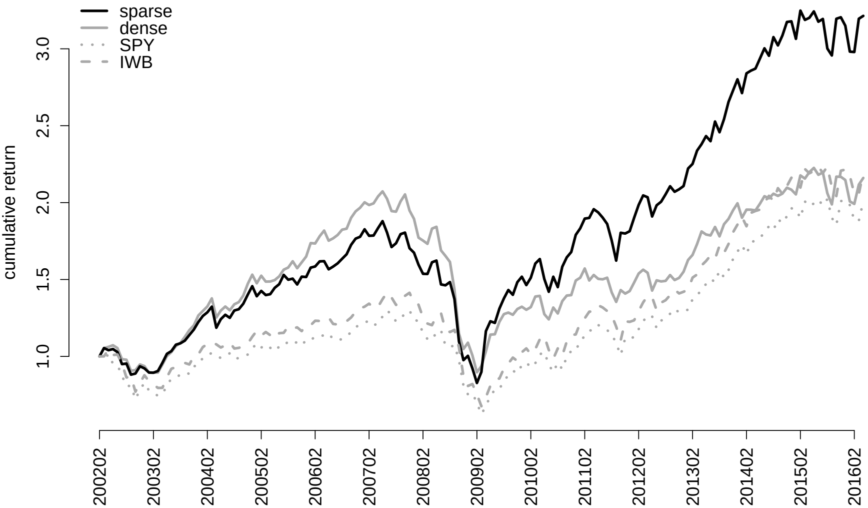
Date	DIA	IWD	IWB	IWN	IWM	IYR
	Dow Jones	Value	Large	Small	Small value	Real estate
2002	-	21.6	23.7	2.55	2.83	49.3
2003	-	18.4	18.2	-	-	63.3
2004	-	14.1	22.3	-	-	63.6
2005	-	31.2	35.2	-	-	33.6
2006	-	32.7	40.6	-	-	26.7
2007	-	41.8	38.4	-	-	19.8
2008	-	43.8	39.3	-	-	16.9
2009	-	-	100	-	-	-
2010	-	-	100	-	-	-
2011	-	-	100	-	-	-
2012	-	-	100	-	-	-
2013	-	-	100	-	-	-
2014	-	-	100	-	-	-
2015	100	-	-	-	-	-
2016	86.7	-	9.59	-	-	3.72

Ex ante "$SR_{\text{target}} - SR_{\text{decision}}$" evolution

Data: Returns on 25 ETFs from 1992-2016. $\kappa = 55\%$ decision.



Ex post performance of the $\kappa = 55\%$ decision



Last slide

- Passive investing and factor selection for asset pricing models approached using new variable selection technique.

- **Utility functions can enforce inferential preferences that are not prior beliefs.**

- *Variable selection in SUR models with random predictors*. Bayesian Analysis (2017).
  Sparse dynamic portfolios with regret-based selection. Submitted (2017).

- Thanks!

Extra slides

A complicated posterior!

$$\tilde{R}_t^i = (\beta_t^i)^T \tilde{R}_t^F + \epsilon_t^i, \qquad \epsilon_t^i \sim N(0, 1/\phi_t^i),$$
$$\beta_t^i = \beta_{t-1}^i + w_t^i, \qquad w_t^i \sim \mathrm{T}_{n_{t-1}^i}(0, W_t^i),$$
$$\beta_0^i \mid D_0 \sim \mathrm{T}_{n_0^i}(m_0^i, C_0^i),$$
$$\phi_0^i \mid D_0 \sim \mathrm{Ga}(n_0^i/2, d_0^i/2),$$
$$\beta_t^i \mid D_{t-1} \sim \mathrm{T}_{n_{t-1}^i}(m_{t-1}^i, R_t^i), \qquad R_t^i = C_{t-1}^i/\delta_\beta,$$
$$\phi_t^i \mid D_{t-1} \sim \mathrm{Ga}(\delta_\epsilon n_{t-1}^i/2, \delta_\epsilon d_{t-1}^i/2),$$

$$\tilde{R}_t^F = \mu_t^F + \nu_t \qquad \nu_t \sim \mathrm{N}(0, \Sigma_t^F),$$
$$\mu_t^F = \mu_{t-1}^F + \Omega_t \qquad \Omega_t \sim \mathrm{N}(0, W_t, \Sigma_t^F),$$
$$(\mu_0^F, \Sigma_0^F \mid D_0) \sim \mathrm{NW}_{n_0}^{-1}(m_0, C_0, S_0),$$
$$(\mu_t^F, \Sigma_t^F \mid D_{t-1}) \sim \mathrm{NW}_{\delta_F n_{t-1}}^{-1}(m_{t-1}, R_t, S_{t-1}), \qquad R_t = C_{t-1}/\delta_c,$$

Dynamic regret-based selection

Assume N asset returns follow the model: $\tilde{R}_t \sim \Pi(\mu_t, \Sigma_t)$

- Specifically, let the covariates be the five Fama-French factors, $R_t^F \sim N(\mu_t^F, \Sigma_t^F)$, so that:

$$\mu_t = \beta_t^T \mu_t^F$$
$$\Sigma_t = \beta_t \Sigma_t^F \beta_t^T + \Psi_t$$

- **Given μ_t and Σ_t, make portfolio decision for time $t+1$.**

Seemingly unrelated regressions

Replace R with generic response vector Y and F with generic covariate vector X:

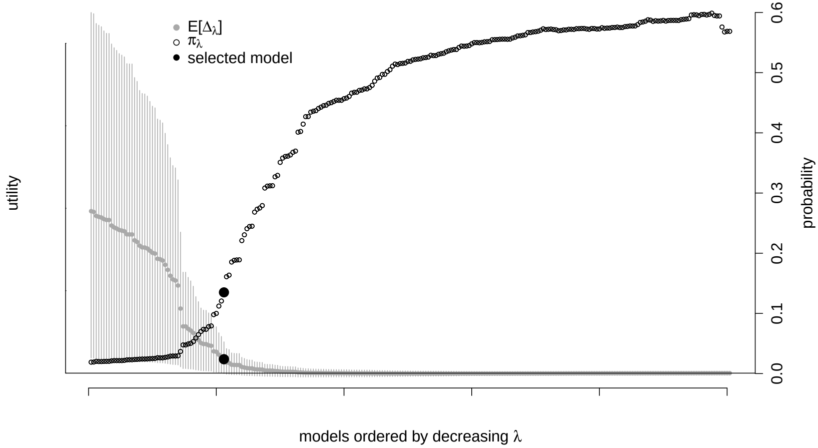
$$R \rightsquigarrow Y \text{ and } F \rightsquigarrow X$$

$$Y_j = \beta_{j1}X_1 + \cdots + \beta_{jp}X_p + \epsilon_j, \quad \boldsymbol{\epsilon} \sim \mathrm{N}(0, \Psi), \quad j = 1, \cdots, q$$

The proposed framework permits variable selection in SUR models with random predictors!

Posterior summary plot

$$\Delta_\lambda \equiv \mathcal{L}(\gamma^*_\lambda, \Theta, \tilde{R}, \tilde{F}) - \mathcal{L}(\gamma^*_0, \Theta, \tilde{R}, \tilde{F}), \qquad \pi_\lambda \equiv \mathrm{P}(\Delta_\lambda < 0)$$

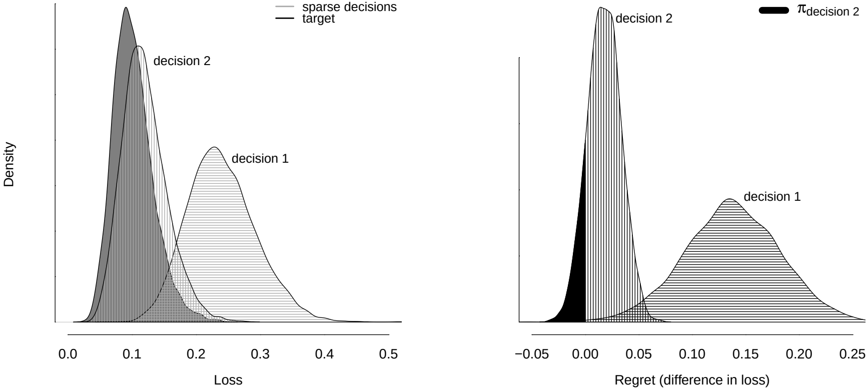


π_λ = probability that λ-model is no worse than the dense model.

Regret-based selection: Illustration

d_λ : sparse decisions, d^* : target decision.

$\pi_\lambda = \mathbb{P}[\rho(d_\lambda, d^*, \tilde{Y}) < 0]$: probability of not regretting λ-decision.



Ex ante regret evolution

Data: Returns on 25 ETFs from 1992-2016. $\kappa = 55\%$ decision.