HireVAE: An Online and Adaptive Factor Model Based on Hierarchical and Regime-Switch VAE

Zikai Wei, Anyi Rao, Bo Dai and Dahua Lin

Demeure Camille
Partensky Alexandre
Romea Kelvin
Sellami Sami

SUMMARY

Introduction, Problem & Methodology

Goal

Build a Dynamic Factor Model (DFM) using VAE to predict cross-sectional stock returns

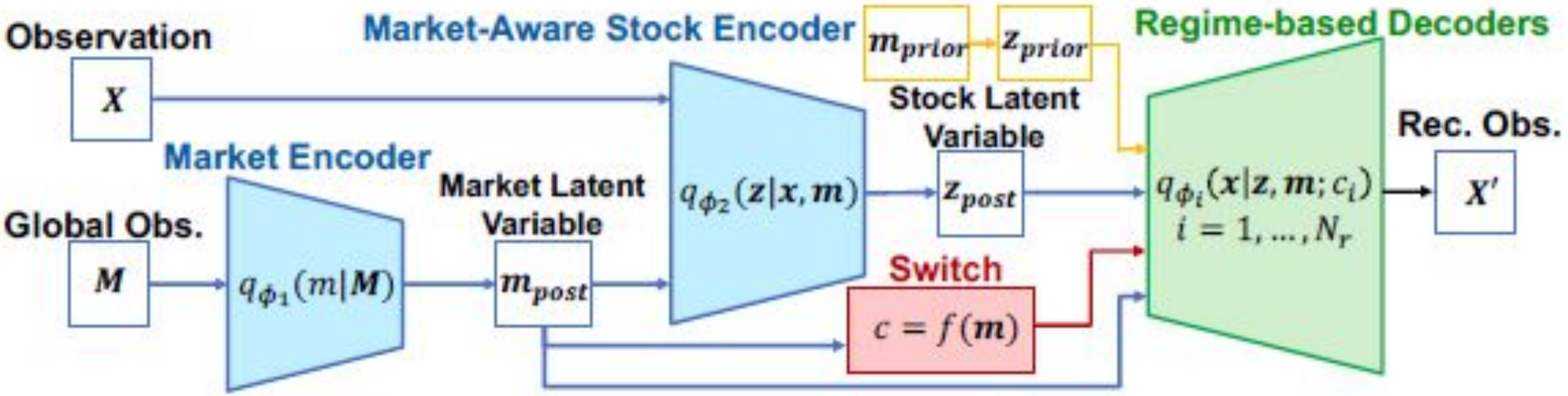
Challenges

1. Traditional factor models fail under abrupt regime shifts.
2. Need regime identification without future information leakage.
3. Factors effective in one regime may become ineffective in another.

Methodology (Core idea)

- Learn hierarchical latent factors (market-level + stock-level) via VAE.
- Market encoder: extracts latent market variable from multimodal features (momentum, volume, volatility).
- Stock encoder: conditions on both stock features and market latent variable to learn stock-wise latent factors.
- Regime learner: online clustering with linear stabilization ensures smooth and consistent regime identification.
- Regime-specific decoders: reconstruct or predict stock returns given latent variables and regime.
- Loss: combines reconstruction, hierarchical KL divergence, and regime differentiation terms.

Model Structure



Loss function

1. **Reconstruction loss** makes sure predictions are accurate
2. **Hierarchical KL loss** keeps priors and posteriors consistent at market & stock level.
3. **Regime separation loss** ensures stable and distinct regime clustering.

$$L_{rec} = -\sum_{r=1}^{N_r} \mathbb{1}_{f(\mathbf{m})}(r) \ln P_{\phi_{dec}}(\hat{\mathbf{y}} = \mathbf{y} | \mathbf{X}, \mathbf{M}, \mathbf{z}, \mathbf{m}; r)$$

$$\begin{aligned} L_{hier} = & \ \mathrm{KL}\left(P_{\phi_{enc}}(\mathbf{m}|\mathbf{M}, \bar{\mathbf{y}}) \,||\, P_{\phi_{prior}}(\mathbf{m}|\mathbf{M}, \bar{\mathbf{y}})\right) + \\ & \ \mathrm{KL}\left(P_{\phi_{enc}}(\mathbf{z}|\mathbf{X}, \mathbf{M}, \mathbf{m}, \mathbf{y}) \,||\, P_{\phi_{prior}}(\mathbf{z}|\mathbf{X}, \mathbf{M}, \mathbf{m})\right) \end{aligned}$$

$$L_{reg} = -\sum_{\substack{i=1,\ldots,N_r \\ j=i+1,\ldots,N_r}} \mathrm{KL}\left(\mathcal{N}(\mu_i, \sigma_i^2) \,||\, \mathcal{N}(\mu_j, \sigma_j^2)\right)$$

$$\boxed{L_{overall} = L_{rec} + L_{hier} + L_{reg}}$$

Experiments & Results

Experiments

- HireVAE achieves higher Rank IC, Rank ICIR, and active returns than baselines (Linear, GRU, MLP, Transformer, VAE, CVAE).
- Adding market index info does not drive the improvement → gains come from the adaptive regime learning, not extra data.
- Online regime switching (HireVAE) outperforms rule-based volatility thresholds and simple clustering in stock prediction & portfolio returns.
- Hierarchical latent space (market + stock factors) clearly improves regime identification and return prediction over flat VAE variants.

Results

Method	IC	Rank IC	Rank ICIR
Linear	0.030	0.031	0.322
GRU	0.046	0.050	0.484
MLP	0.053	0.051	0.537
Trans	0.050	0.040	0.253
GAT	0.029	0.032	0.466
IGAT	0.012	0.018	0.210
DMFM	0.014	0.015	0.480
VAE	0.049	0.059	0.539
CVAE	0.053	0.063	0.628
HiReVAE	**0.058**	**0.066**	**0.734**

Implementation Details

Training Dataset

- **Dataset:** CSI 300 equities (2010–2020) (less than original paper: ~3500 Chinese stocks)
- **Features:** Alpha158 technical indicators
- **Training period:** 2010-01-01 -- 2017-12-31
- **Validation period:** 2018-01-03 -- 2018-12-29
- **Testing period:** 2019-01-02 -- 2020-09-20

Experimental Setup

- **Library:** PyTorch

Evaluation Metrics

Statistical Metrics

- Rank Information Coefficient (Rank IC)
- Rank ICIR (Information Ratio of IC)

Hyperparameters

BATCH_DATES_PER_STEP = 1
EPOCHS = 10
LR = 1e-3
WEIGHT_DECAY = 1e-5
H_MARKET = 16
H_STOCK =32
REGIMES = 3
EMA BETA = 0.97

Implementation Details Market Encoder

Market Encoder

- The Market Encoder is supposed to be fed with market data that can allow it to detect the current regime
- Optimal variables would be variable like market volatility, commodity prices, interest rates, stock bond correlation.
- Here we fed the model with simpler data we simply gave him the mean and standard deviation of the stock cross section.

What the paper does

The paper only mention that they use a conjonction of historical features but do not develop precisely on which features.

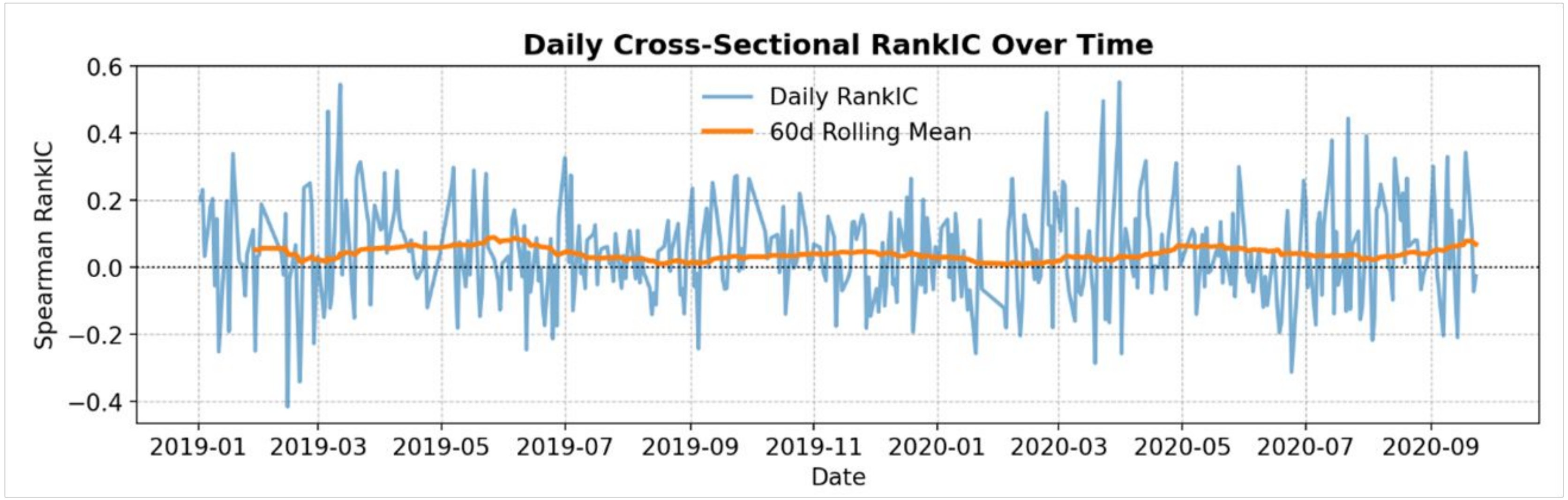
Futher Refinement

Here to have better results it would be interesting to test different set of market features and see how this feature impact the regime switch model, and the quality of our Regime Detection inside our model.

Rank IC

RankIC:
- Rank IC (Information Coefficient): Measures correlation between model's predicted ranks and actual returns (Spearman correlation.
- *RankIC = corr(rank(r_pred), rank(r_true))*

Obtained Rank IC on Test Set:
0.043256
Obtained Rank IC FactorVAE:
0.043089

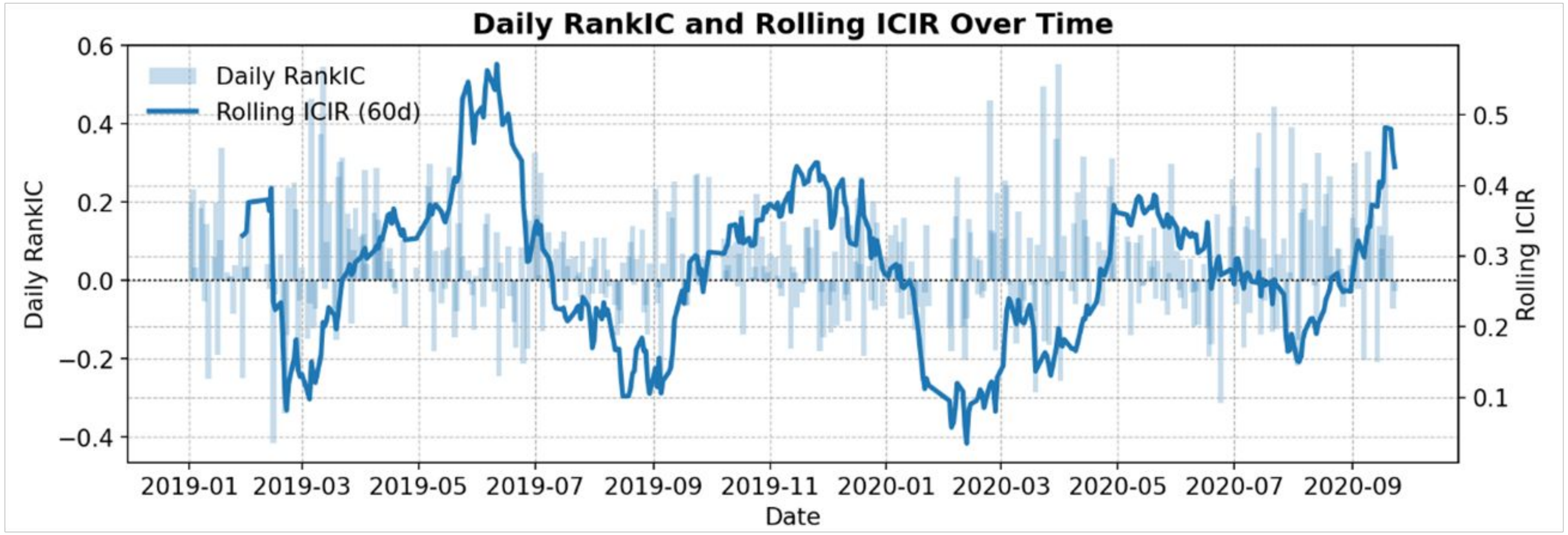


Results analysis:
- Rank IC lower than in the original paper (0.066)
- Coherent with a smaller training dataset

Rank ICIR

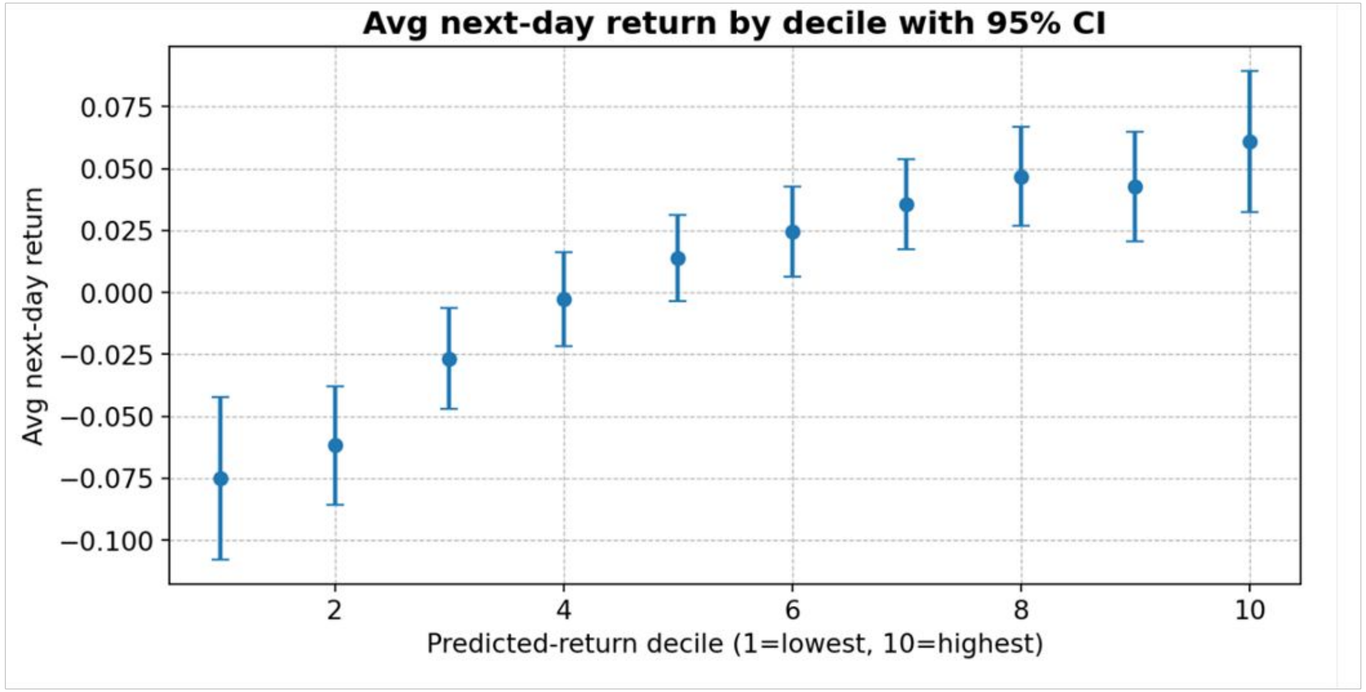
ICIR:
- ICIR (Information Coefficient Information Ratio): Stability of IC over time, computed as mean IC divided by its standard deviation.
- *ICIR = E[IC] / Std(IC)*

Obtained Rank ICIR:
0.294702
Obtained Rank ICIR FactorVAE:
0.317856



Results analysis:
- Rank ICIR lower than in the original paper
- Coherent with a smaller training dataset

Prediction Deciles

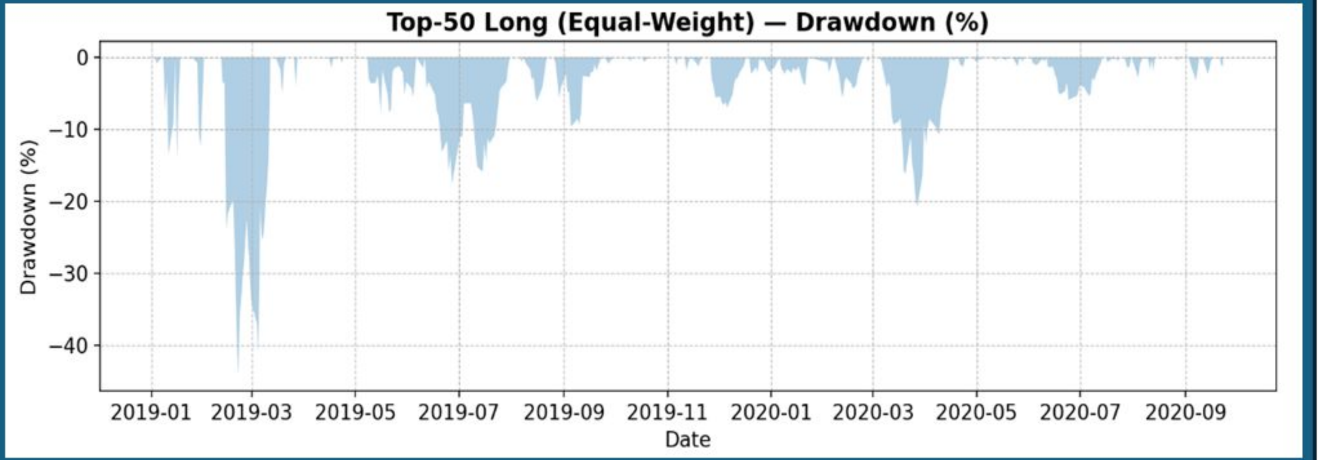


Prediction per decile:

- Nice monotonicity across deciles
- Increased variance at the extreme deciles
- Strong next day return base on that

Opens the door to long short strategy long first names decile short last decile names

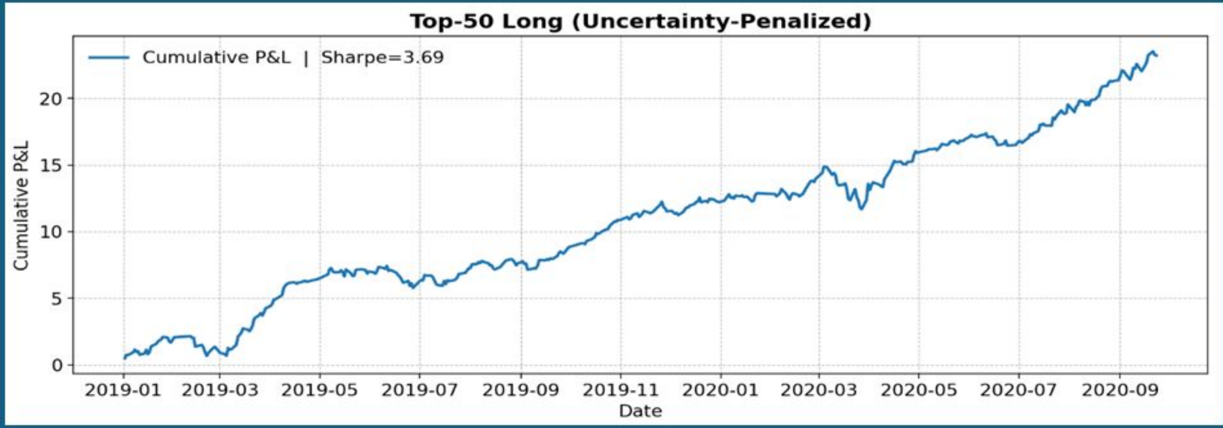
Top50-Long

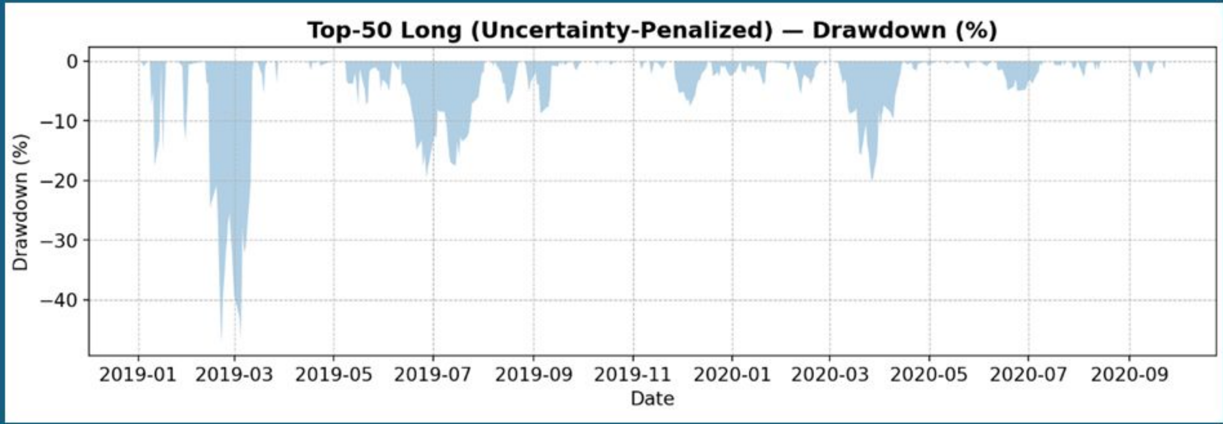


<u>Top50-long</u>:

- Each day long top-50 stocks by predicted return (Mu), 1-day hold
- Equal-weighted, self-financing
- *Results: One important drawdown in our OOS performance but the cumuative return is pretty strong*
- *Very high Sharp*

Top50-drop penalized

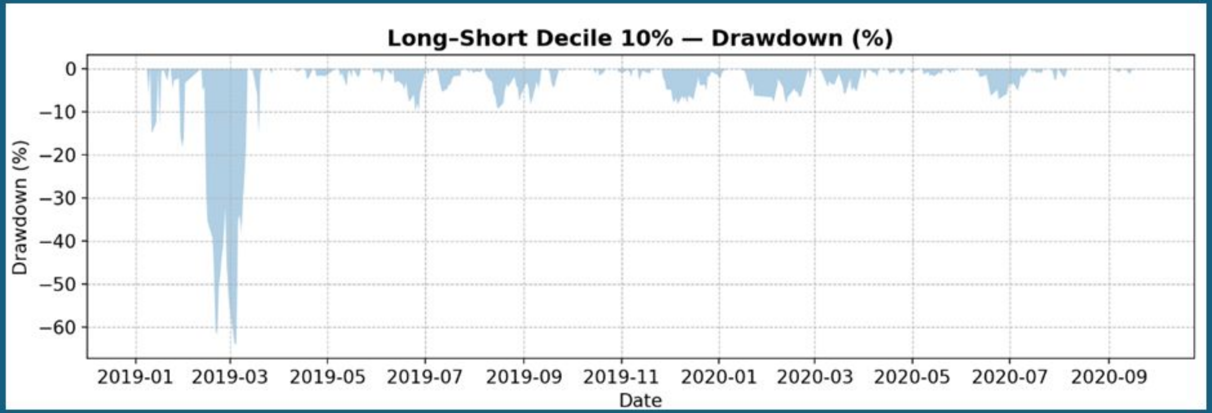




Top50-drop penalized:

- Ranks by risk-adjusted score = Mu − $\lambda \cdot \sigma^2$
- Penalizes stocks with higher predicted risk (σ^2), favors more stable names
- *No evidence on lower Drawdown but we do see a slight increase in the sharp ratio showing that the signal is more stable*

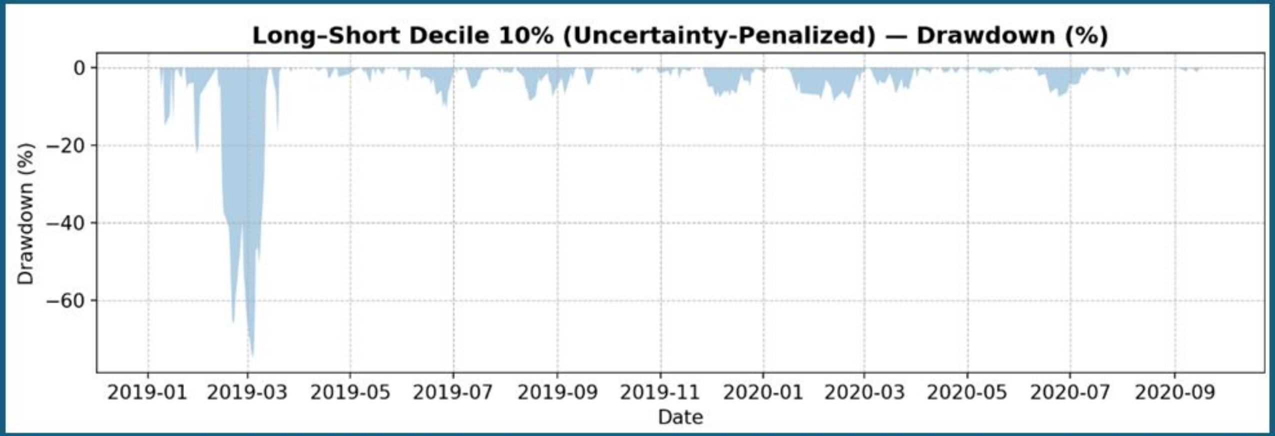
Long-short decile



<u>Long–Short Decile:</u>

- Each day long top **10%** (Mu) and short bottom **10%**
- Market-neutral, equal-weighted
- This Long short significantly boost the returns but the volatility of the strategy increased too, especially we know have a drawdown at 60%.
- But the strategy does perfform relatively wellfor a period such as Covid

Long-short decile penalized



<u>Long–Short Decile penalized:</u>

- Similar to previous long-short but reweights the portfolio according to $\frac{1}{\gamma + \sigma}$
- Here the effect of this penalization is not very clear, the drawdowns did not decrease, and the Sharp is essentially the same

BerkeleyHaas

Limitations and possible improvements

Limitations

- **High computation time** (1 hour per epoch) → couldn't optimize hyperparameters
- Limited Universe only 300 stocks
- No computation of benchmarks
- **Very low interpretability** (black-box model), even thought the nature of this model makes it more interpretable than FactorVAE since it condition on market regimes

Possible Improvements

- **Better computational power**
- Include **more stocks** in the training dataset
- Compare to suggesed benchmarks
- Test for **robustness (test other sample, other dataset…)**

Possible Extension

- Try to Analyse the Market Encoder contribution, test and improve the quality of our Regime Switching model
- Try to interpret more what the model is doing

Thank you!

HireVAE: An Online and Adaptive Factor Model Based on Hierarchical and Regime-Switch VAE

Zikai Wei, Anyi Rao, Bo Dai and Dahua Lin