Empirical Asset Pricing with Probability Forecasts

Songrun He[1], Linying Lv[1], and Guofu Zhou[1]

[1]Washington University in St. Louis

Abstract

We study probability forecasts in the context of cross-sectional asset pricing with a large number of firm characteristics. Empirically, we find that a simple probability forecast model can surprisingly perform as well as a sophisticated probability forecast model, all of which deliver long-short portfolios whose Sharpe ratios are comparable to those of the widely used return forecasts. Moreover, we show that combining probability forecasts with return forecasts yields superior portfolio performance versus using each type of forecast individually, suggesting that probability forecasts provide valuable information beyond return forecasts for our understanding of the cross-section of stock returns.

Why Probability Forecast?

- Focusing on probability unifies both risk and return.
- The probability of outperformance is linked to the Information Ratio (IR).
- Consider probability of a stock outperforming the market. Under normal assumption: $R_{i,t+1} \sim N\left(\mu_{it}, (\sigma^{i}_{t+1|t})^2\right)$, and $R^{mkt}_{t+1} \sim N\left(\mu_{it}, (\sigma^{mkt}_{t+1|t})^2\right)$.

- Probability of outperformance can be expressed as:

$$Prob_t\left(R_{i,t+1} - R^{mkt}_{t+1} > 0\right) = \Phi\left(\frac{\mu_{it} - \mu^{mkt}_{t}}{\sigma_{t+1|t}}\right).$$

  - Applying CDF function to IR.
  - If probability can be estimated with low error, sorting on probability is equivalent to sorting on IR.
  - Time-varying $\sigma_{t+1|t}$ leads to additional predictability in probability.

- The argument extends to factor models.

$$Prob_t\left(R_{i,t+1} - R_{f,t} > \beta'_{it}F_{t+1}\right) = \Phi\left(\frac{\alpha_{it}}{IdioVol_{i,t}}\right).$$

  - Probability is an increasing function of IR relative to the factor model.
  - The IR can also be mapped to t-stat of α in time-series regression.

Methodology and Data

- Forecasting target: the probability of outperforming a benchmark

$$y_{i,t} = I\{R_{i,t} > R^{bench}_{t}\}$$

- Objective functions:
  - Linear probability model: mean-squared error loss:

$$L(\theta) = 1/NT \, \Sigma_{i=1}^{N} \Sigma_{t=1}^{T} \left(y_{i,t+1} - g(z_{i,t};\theta)\right)^2.$$

  - Logit probability model: cross-entropy loss:

$$L(\theta) = -1/NT \, \Sigma_{i=1}^{N} \Sigma_{t=1}^{T} \left(y_{i,t+1} \log g\left(z_{i,t};\theta\right) + (1 - y_{i,t+1}) \log(1 - g(z_{i,t};\theta))\right).$$

- Prediction models with different complexity:
  - Linear probability models: OLS, PLS
  - Logit probability model: Logistic Regression, Neural Networks with 1 to 5 layers.

- Data:
  - Monthly stock returns from CRSP; 94 firm-level characteristics from Green, Hand, and Zhang (2017) and Gu, Kelly, and Xiu (2020).
  - Training, validation, and test specifications following Gu et al. (2020)
  - Out-of-sample testing from January 1987 to December 2020.

Empirical Performance

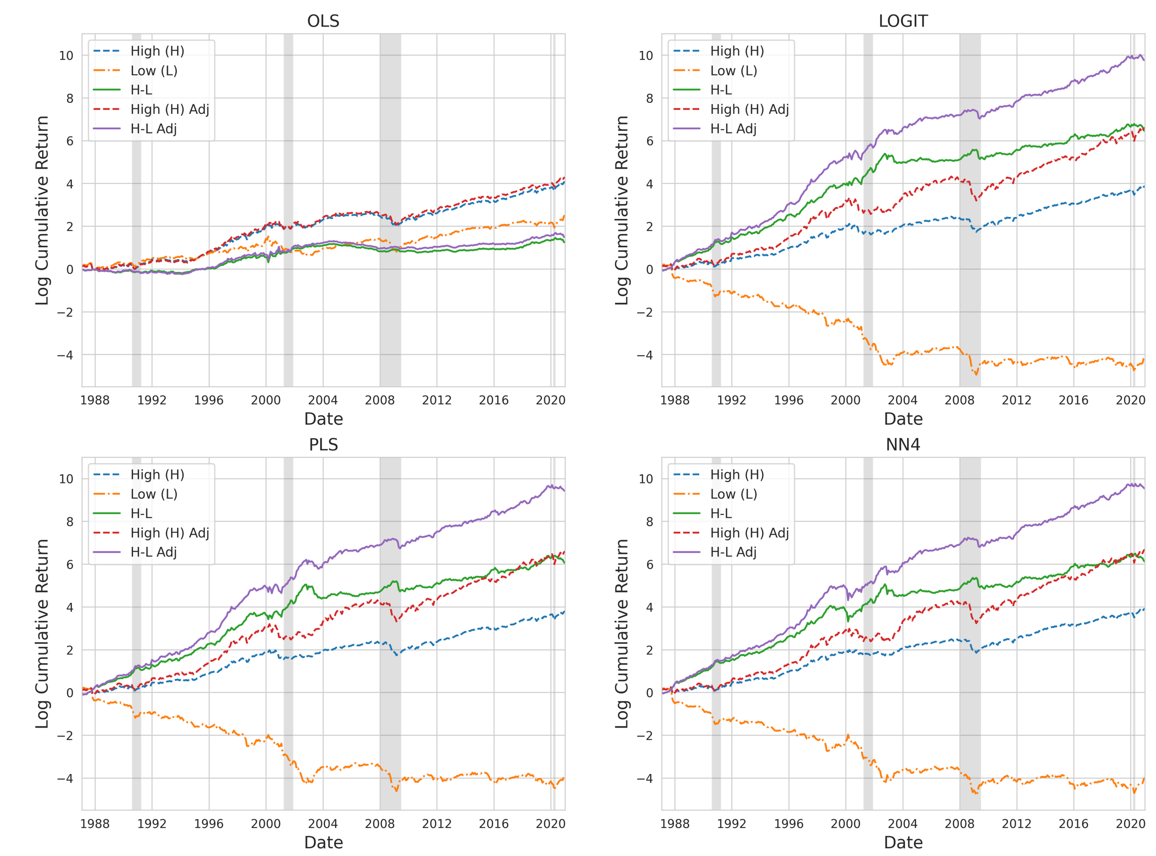


Figure 1. Cumulative Portfolio Returns Sorted on Probability Forecast

- Conditional volatility plays an important role in determining probability.
  - Sorting on probability generates portfolios with different volatility.
  - We adjust the σ of decile portfolios in realtime to have the same σ as decile 1:

$$r^{adj}_{j,t+1} = r_{j,t+1}\, \sigma_{1,t}/\sigma_{j,t} \quad \text{for } j = 1, \ldots, 10.$$

- Performance of variance adjusted decile portfolio sorted on probability forecasts:

	OLS					Logit				
	$\widehat{Prob}$	$Prob$	Mean	SD	SR	$\widehat{Prob}$	$Prob$	Mean	SD	SR
Low (L)	-0.02	0.45	0.74	5.01	0.51	0.38	0.39	-0.69	7.77	-0.31
2	0.14	0.44	0.31	5.48	0.19	0.41	0.43	0.14	7.73	0.06
3	0.29	0.44	0.10	5.20	0.07	0.43	0.44	0.48	7.72	0.21
4	0.38	0.44	0.34	5.25	0.23	0.45	0.45	0.62	7.64	0.28
5	0.42	0.45	0.53	5.10	0.36	0.46	0.46	0.79	7.63	0.36
6	0.44	0.46	0.51	5.15	0.35	0.47	0.47	0.93	7.71	0.42
7	0.47	0.47	0.63	5.04	0.43	0.48	0.48	1.23	7.49	0.57
8	0.49	0.48	0.73	4.98	0.51	0.49	0.48	1.46	7.40	0.69
9	0.52	0.48	0.89	4.87	0.63	0.50	0.49	1.57	7.36	0.74
High (H)	0.58	0.49	1.17	4.86	0.84	0.53	0.51	1.92	7.45	0.89
H-L	-	-	0.43	3.66	0.41	-	-	2.61	6.19	1.46
	PLS					NN4				
	$\widehat{Prob}$	$Prob$	Mean	SD	SR	$\widehat{Prob}$	$Prob$	Mean	SD	SR
Low (L)	0.38	0.39	-0.62	7.78	-0.28	0.37	0.39	-0.64	8.03	-0.27
⋮			⋮					⋮		
High (H)	0.53	0.51	1.91	7.48	0.89	0.50	0.50	1.95	7.70	0.88
H-L	-	-	2.54	6.44	1.36	-	-	2.59	6.74	1.33

Table 1. Probability Forecasts Decile Portfolio Performance

Probability vs Expected Return Forecasts

- Do probability forecasts contain incremental information relative to expected return forecasts?
- Construct the following portfolios:
  - Probability forecast long-short portfolios
  - Expected return forecast long-short portfolios from Gu, Kelly, and Xiu (2020).
  - Combination of the two portfolios
- Probability LS portfolio has low correlation with expected return LS portfolio. Combining both leads to significant higher SR.

	OLS	Logit	PLS	NN4
Corr	0.25	0.34	0.34	0.33
Panel A: Probability Forecast				
SR	0.41	1.46	1.36	1.33
t_α	2.15	5.72	5.60	5.52
Panel B: Expected Return Forecast (NN4)				
SR	1.43	1.43	1.43	1.43
t_α	4.45	4.45	4.45	4.45
Panel C: 1/N Combination of Probability and Expected Return Forecasts				
SR	1.33	1.76	1.71	1.69
t_α	4.36	5.65	5.58	5.73
Panel D: Mean-variance Combination of Probability and Expected Return Forecasts				
SR	1.38	1.73	1.68	1.65
t_α	4.56	5.37	5.19	5.15

Table 2. Combining Probability Forecasts with Expected Return Forecasts

- Which variables matter for probability vs expected return forecasts?
  - Shapley value decomposition of the prediction model.

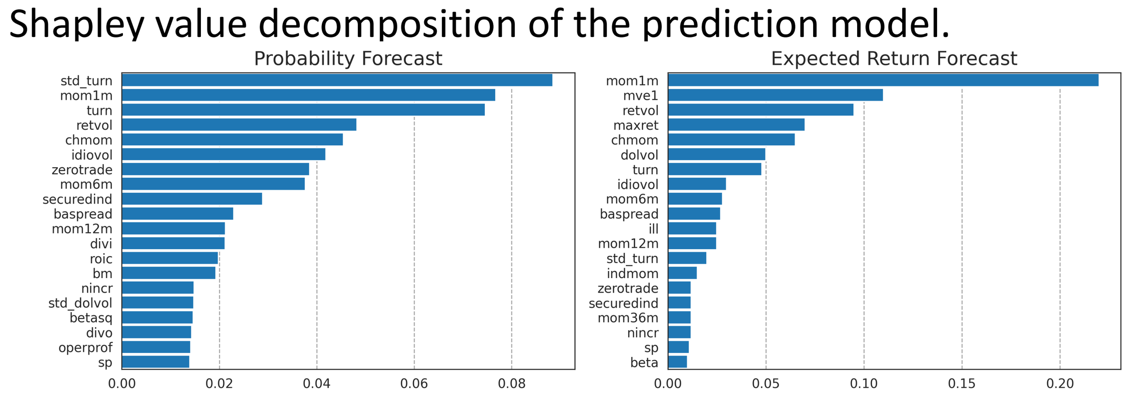


Figure 2. Variable Importance for Probability vs Expected Return Forecasts

- Consistent results if we consider probability of outperforming factor models.

Forecasting and Managing Tail Risks

- Consider the probability that the stock will have return smaller than -20% in the next month.

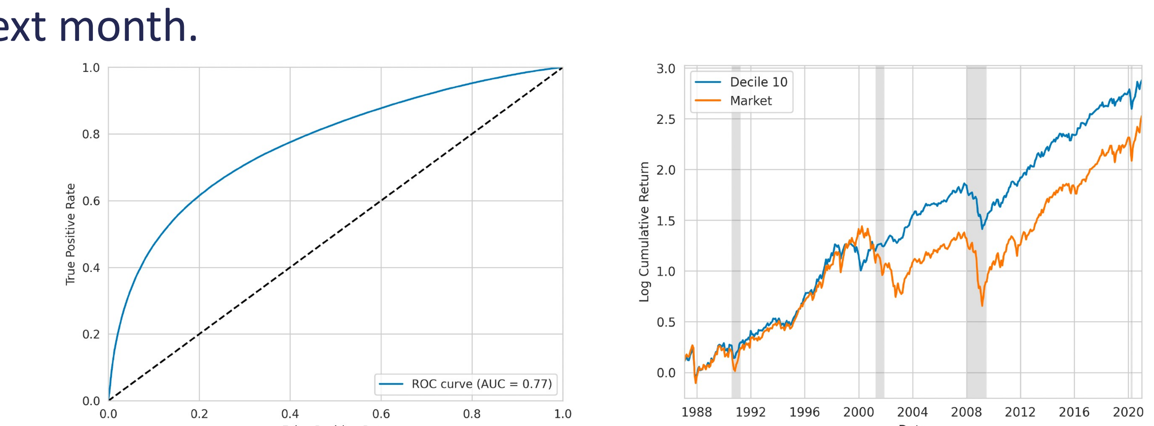


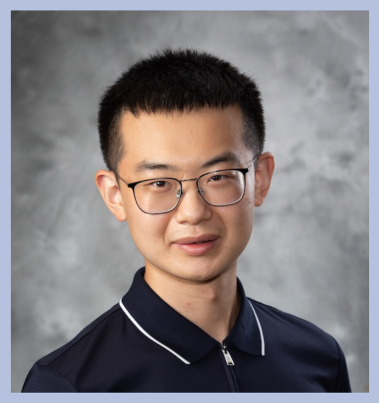
Figure 3. Prediction Accuracy and Economic Gains from Tail Risk Forecasts

- The tail risk measure remains significant under a wide array of controls from tail risk literature.
- Economic gains: decile portfolios with lowest tail risk has MaxDD 33% lower than market and SR 45% higher than market.

Conclusion

- Probability forecasts offer a valuable alternative to E[R] forecasts.
- Simple prediction model works well for probability forecasts.
- Combining both leads to superior portfolio performance.
- Probability forecasts generate large economic gains for managing tail risk.

Listen to our paper podcast here!



Contact

Songrun He
h.songrun@wustl.edu
https://www.linkedin.com/in/songrun-he-735a67140/