Jensen's Inequality and the Success of Linear Prediction Pools

Fabian Krüger

University of Konstanz

Ifo-Bundesbank workshop,
June 2, 2012

Motivation

Economists are increasingly interested in *probabilistic* forecasting

- "Fan charts" issued by central banks
- (European) Survey of Professional Forecasters

Much recent work on constructing & evaluating probabilistic forecasts.

Particular focus: **Linear prediction pools**.

- Wallis, 2005; Hall & Mitchell, 2007; Gneiting & Ranjan, 2010, 2011; Jore, Mitchell & Vahey, 2010; Kascha & Ravazzolo, 2010; Clements & Harvey, 2011; Geweke & Amisano, 2011

$\Rightarrow$ Good performance of linear pools in terms of the log score (LS) criterion.

Contribution of this paper: Analyze linear pools under the quadratic and continuous ranked probability scores

- Important robustness check since these scoring rules are sensible alternatives to LS
- To date, little evidence on this issue, at least for continuous settings

Main findings:

Contribution of this paper: Analyze linear pools under the quadratic and continuous ranked probability scores

- Important robustness check since these scoring rules are sensible alternatives to LS
- To date, little evidence on this issue, at least for continuous settings

Main findings:

- Generally: Good performance of pools carries over to two other scoring rules

Contribution of this paper: Analyze linear pools under the quadratic and continuous ranked probability scores

- Important robustness check since these scoring rules are sensible alternatives to LS
- To date, little evidence on this issue, at least for continuous settings

Main findings:

- Generally: Good performance of pools carries over to two other scoring rules
- Theory: Success of linear pools partly by construction of the scoring rules (concavity!)

Contribution of this paper: Analyze linear pools under the quadratic and continuous ranked probability scores

- Important robustness check since these scoring rules are sensible alternatives to LS
- To date, little evidence on this issue, at least for continuous settings

Main findings:

- Generally: Good performance of pools carries over to two other scoring rules
- Theory: Success of linear pools partly by construction of the scoring rules (concavity!)
- Simulations: Simple, misspecified pools may be hard to distinguish from the true model

Contribution of this paper: Analyze linear pools under the quadratic and continuous ranked probability scores

- Important robustness check since these scoring rules are sensible alternatives to LS
- To date, little evidence on this issue, at least for continuous settings

Main findings:

- Generally: Good performance of pools carries over to two other scoring rules
- Theory: Success of linear pools partly by construction of the scoring rules (concavity!)
- Simulations: Simple, misspecified pools may be hard to distinguish from the true model
- Empirics: Pools very attractive when there is no "single best model"

Scoring Rules

Tools for evaluating density forecasts:

1. Probability Integral Transforms (PIT; Rosenblatt 1952, Diebold et al 1998, 1999)
2. Scoring Rules (Winkler 1969, Gneiting & Raftery 2007)

- Focus on (proper) Scoring Rules here.
- Assign score $S(y, f(\cdot)) \in \mathbb{R}$ when $f(\cdot)$ is the density forecast and $y \in \mathbb{R}$ materializes.

Scoring Rule # 1: **Log Score** (Good 1952)

$$LS(y, f(\cdot)) = \ln f(y). \quad (1)$$

- Simple
- Related to ML, Kullback-Leibler divergence
- Local (Bernardo 1979)
- Infinite penalty for tail events (Selten 1998)

Scoring Rule # 2: **Quadratic Score** (Brier 1951)

$$QS(y, f(\cdot)) = 2f(y) - \int f^2(z)dz. \quad (2)$$

- Continuous form of famous Brier score for discrete events
- Neutral (Selten 1998)
- Numerically more stable than LS

Scoring Rule # 3: **Continuous Ranked Probability Score** (Winkler & Matheson 1976)

$$CRPS(y, f(\cdot)) \;=\; -\int (F(z) - \mathbb{I}_{(z \geq y)})^2 dz \qquad (3)$$

- $F(\cdot)$ is the c.d.f. implied by $f(\cdot)$
- Sensitive to distance
- Generalizes to absolute error if $f(\cdot)$ is a point forecast
- Difficult to evaluate when $f(\cdot)$ is non-normal

Which rule to pick?

- All rules are proper, i.e. maximized in expectation by true model.
- May (easily) give different rankings of misspecified models.
- Different properties (see above), but no consensus on which properties are desirable
- I look at all three rules in the following.

JI and Linear Pools

Linear Pool (Wallis, 2005)

$$f_c(Y) = \sum_{i=1}^{n} \omega_i f_i(Y), \tag{4}$$

weights ω_i positive and sum to one.
Mean and variance given by

$$\begin{aligned} \mu_c &= \sum_{i=1}^{n} \omega_i \mu_i, \\ \sigma_c^2 &= \sum_{i=1}^{n} \omega_i \sigma_i^2 + \sum_{i=1}^{n} \omega_i (\mu_i - \mu_c)^2, \end{aligned}$$

where μ_i and σ_i^2 are the mean and variance of model i.

Following proposition implies that success of pools is partly by construction of the scoring rules

- Extends results by Kascha & Ravazzolo (2010) in the context of the log score
- Simillar results by McNees (1992) and Manski (2011) for point forecasts and squared error loss.

Proposition

Consider a linear pool as defined in Equation (4) and the three scoring rules defined in (1) to (3). Then, if an outcome $Y = y$ materializes,

$$LS(y, \sum_{i=1}^{n} \omega_i f_i(\cdot)) \geq \sum_{i=1}^{n} \omega_i LS(y, f_i(\cdot)), \quad (5)$$

$$QS(y, \sum_{i=1}^{n} \omega_i f_i(\cdot)) \geq \sum_{i=1}^{n} \omega_i QS(y, f_i(\cdot)), \quad (6)$$

$$CRPS(y, \sum_{i=1}^{n} \omega_i f_i(\cdot)) \geq \sum_{i=1}^{n} \omega_i CRPS(y, f_i(\cdot)). \quad (7)$$

What does this mean?

What does this mean?

- Let $\omega_i = \frac{1}{n} \; \forall \; i$. Then the score of the pool is **necessarily** higher than the average score of the n components.

What does this mean?

- Let $\omega_i = \frac{1}{n} \; \forall \; i$. Then the score of the pool is **necessarily** higher than the average score of the n components.
- E.g. with $n = 2$, combination closer to the better of the two models.

What does this mean?

- Let $\omega_i = \frac{1}{n} \; \forall \; i$. Then the score of the pool is **necessarily** higher than the average score of the n components.
- E.g. with $n = 2$, combination closer to the better of the two models.
- This holds for
  - Each ex-post outcome $y \in \mathbb{R}$.
  - Each set of densities $f_i(\cdot); i = 1, \ldots, n$.

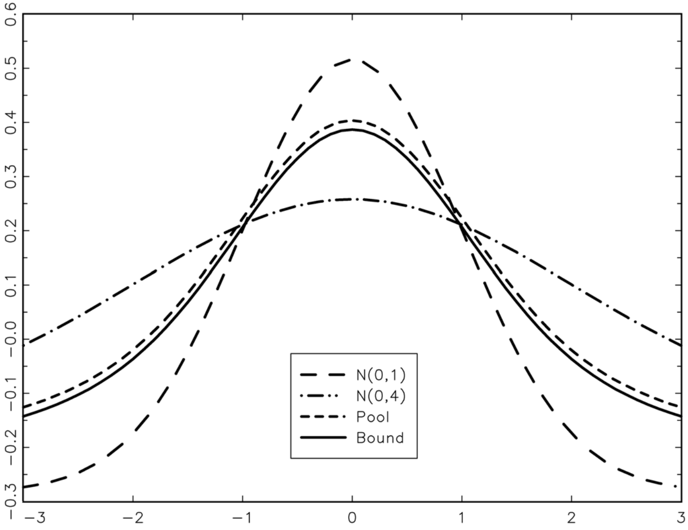
What does this mean?

- Let $\omega_i = \frac{1}{n} \; \forall \; i$. Then the score of the pool is **necessarily** higher than the average score of the n components.
- E.g. with $n = 2$, combination closer to the better of the two models.
- This holds for
  - Each ex-post outcome $y \in \mathbb{R}$.
  - Each set of densities $f_i(\cdot); i = 1, \ldots, n$.
- Attractive property since predicting the relative performance of n models is usually very hard

What does this mean?

- Let $\omega_i = \frac{1}{n} \; \forall \; i$. Then the score of the pool is **necessarily** higher than the average score of the n components.
- E.g. with $n = 2$, combination closer to the better of the two models.
- This holds for
  - Each ex-post outcome $y \in \mathbb{R}$.
  - Each set of densities $f_i(\cdot)$; $i = 1, \ldots, n$.
- Attractive property since predicting the relative performance of n models is usually very hard
  - Pool is "always on the side that's winning"

Example: Quadratic score of two Gaussian densities



Corollary

Let $f_0(Y)$ be the true density of Y, and denote by $ELS_0(f(\cdot))$ the expected log score of a predictive density $f(\cdot)$, with respect to the true density $f_0(\cdot)$. Then,

$$ELS_0(\sum_{i=1}^{n} \omega_i f_i(\cdot)) \geq \sum_{i=1}^{n} \omega_i ELS_0(f_i(\cdot)), \tag{8}$$

and analogous relations hold for the quadratic and continuous ranked probability scores.

- **No matter what the true model is**, the expected score of the combination is at least as high as the average expected score of the components.
- Hence combination pays off from an ex-ante perspective.

Further results in paper:

- Sharpness of bounds somewhat different across scoring rules
  - Stochastic for LS; deterministic for QS and CRPS
- Lower bounds do not necessarily hold for logarithmic and beta-transformed pools
  - Argument in favor of linear pools vs these methods

Restrictions of preceding evidence

1. Simple one-shot scenario $\neq$ multi-period setup used in time series contexts
2. Focus on **lower bounds**
    - Shows that pools perform well in a “worst case” sense
    - However, up to now no evidence on efficiency
    - Tackle this in a simulation study & empirically

Aim: Analyze the efficiency loss of linear pools relative to the true model

- From strict propriety, it is clear that this efficiency loss exists
- Not straightforward to quantify since numerical score differences are hard to interpret
- My approach: Adopt the perspective of a researcher who tests for equal predictive ability (Diebold & Mariano 1995, Giacomini & White 2006)
- Many (few) rejections $\leftrightarrow$ large (small) efficiency loss of linear pool

Simulation setup

Simulation setup

- Use a set of $n = 5$ densities from Gaussian autoregressions

Simulation setup

- Use a set of $n = 5$ densities from Gaussian autoregressions
  - Models calibrated to different subperiods of a US macro series

Simulation setup

- Use a set of $n = 5$ densities from Gaussian autoregressions
  - Models calibrated to different subperiods of a US macro series
    - Model 1: 1960–2011,..., Model 5: 2000 – 2011

Simulation setup

- Use a set of $n = 5$ densities from Gaussian autoregressions
  - Models calibrated to different subperiods of a US macro series
    - Model 1: 1960–2011,. . . , Model 5: 2000 – 2011
    - Series: CPI inflation, IP growth, changes in TBILL rate, change in unemployment rate

Simulation setup

- Use a set of $n = 5$ densities from Gaussian autoregressions
  - Models calibrated to different subperiods of a US macro series
    - Model 1: 1960–2011,. . . , Model 5: 2000 – 2011
    - Series: CPI inflation, IP growth, changes in TBILL rate, change in unemployment rate
- The true process i^* corresponds to one of the five individual models

Simulation setup

- Use a set of $n = 5$ densities from Gaussian autoregressions
  - Models calibrated to different subperiods of a US macro series
    - Model 1: 1960–2011,. . . , Model 5: 2000 – 2011
    - Series: CPI inflation, IP growth, changes in TBILL rate, change in unemployment rate
- The true process i^* corresponds to one of the five individual models
- Test for EPA between the true model $f_{i^*,t-1}(\cdot)$ and the equally weighted (EW) pool $f_{EW,t-1}(\cdot)$

Simulation setup

- Use a set of $n = 5$ densities from Gaussian autoregressions
  - Models calibrated to different subperiods of a US macro series
    - Model 1: 1960–2011,..., Model 5: 2000 – 2011
    - Series: CPI inflation, IP growth, changes in TBILL rate, change in unemployment rate
- The true process i^* corresponds to one of the five individual models
- Test for EPA between the true model $f_{i^*,t-1}(\cdot)$ and the equally weighted (EW) pool $f_{EW,t-1}(\cdot)$
- Of course, the null of EPA is false since

$$\frac{1}{n}\sum_{i=1}^{n} ES_0(f_{i,t-1}(\cdot)) \leq ES_0(f_{EW,t-1}(\cdot)) < ES_0(f_{i^*,t-1}(\cdot)),$$

where ES_0 denotes unconditional expectation w.r.t. the true model.

Average EPA rejection rates (%) by scoring rule & series

		CPI	INDPRO	TBILL	UNEMP
LS	$T = 120$	23.29	9.93	73.17	9.86
	$T = 360$	47.44	15.99	95.95	17.47
QS	$T = 120$	14.42	7.11	68.15	7.50
	$T = 360$	30.21	11.03	94.75	11.97
CRPS	$T = 120$	17.21	9.11	66.09	8.46
	$T = 360$	35.93	13.85	91.23	14.46

Detailed Tables

The simulation study is intentionally biased *against* linear pools

- Assumption: True process coincides with one of the individual models
- Of course, this need not hold in practice
- Next turn to empirical application with unknown true process

Empirics

- Use four monthly US macro series also used in simulation study
- Data between January 1985 and November 2011 (= 323 obs.) used as evaluation period; earlier obs. used for model estimation
- Component models are 8 (V)AR specifications which differ with respect to
  1. Estimation sample (short vs long rolling window)
     - ⋆ C.f. simulation study
  2. System variables
- Simple equally weighted pool in addition to individual models

Meta view: Rank of EW among the 9 models

		CPI	INDPRO	TBILL	UNEMP
LS	h=1	1	1	1	1
	h=3	1	1	1	1
	h=6	2	1	4	1
QS	h=1	2	1	5	1
	h=3	5	1	5	1
	h=6	5	1	5	1
CRPS	h=1	2	1	5	1
	h=2	2	2	5	2
	h=6	5	3	5	3

Detailed Tables

Giacomini-White (2006) EPA tests of each component model vs EW:
$8 \times 36 = 288$ comparisons

Giacomini-White (2006) EPA tests of each component model vs EW:
$8 \times 36 = 288$ comparisons

- Nr of cases in which an individual model beats EW: 11

Giacomini-White (2006) EPA tests of each component model vs EW:
$8 \times 36 = 288$ comparisons

- Nr of cases in which an individual model beats EW: 11
- Nr of cases in which EW beats an individual model: 118

Giacomini-White (2006) EPA tests of each component model vs EW:
$8 \times 36 = 288$ comparisons

- Nr of cases in which an individual model beats EW: 11
- Nr of cases in which EW beats an individual model: 118
- Again, results for EW especially good under LS

Instability of real-world data plays into the hands of EW

Instability of real-world data plays into the hands of EW

- At each evaluation date t, there may be a number of better models than EW

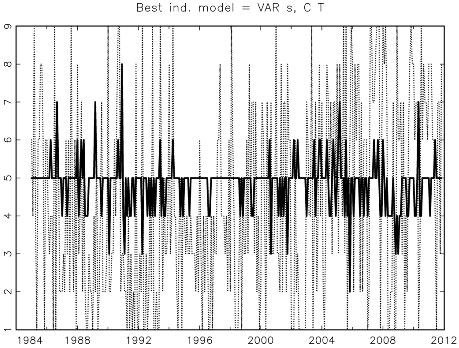
Instability of real-world data plays into the hands of EW

- At each evaluation date t, there may be a number of better models than EW
- However, it rarely occurs that *a particular model* is constantly better than EW

Instability of real-world data plays into the hands of EW

- At each evaluation date *t*, there may be a number of better models than EW
- However, it rarely occurs that *a particular model* is constantly better than EW

Example: Period-by-period ranks of EW (bold) and best individual model



Discussion

Discussion

- A number of recent papers have pointed to negative features of linear pools
  - Hora (2004), Gneiting and Ranjan (2010, 2011)
  - "When the component models are correctly dispersed, the linear pool is too dispersed"
  - Diagnosis of dispersion via PITs

Discussion

- A number of recent papers have pointed to negative features of linear pools
  - Hora (2004), Gneiting and Ranjan (2010, 2011)
  - "When the component models are correctly dispersed, the linear pool is too dispersed"
  - Diagnosis of dispersion via PITs
- These results do not contradict my positive findings on linear pools

Discussion

- A number of recent papers have pointed to negative features of linear pools
  - Hora (2004), Gneiting and Ranjan (2010, 2011)
  - "When the component models are correctly dispersed, the linear pool is too dispersed"
  - Diagnosis of dispersion via PITs
- These results do not contradict my positive findings on linear pools
- In fact, overdispersion results and lower bounds (this paper) both point to the conservative character of linear pools
  - Sacrifice sharpness for good "worst-case properties"
  - May be a good deal in turbulent times

Thank You!

		Consumer Price Index					*Industrial Production*				
	True model (i^*)	1	2	3	4	5	1	2	3	4	5
	$\frac{1}{n}\sum_{i=1}^{n} ES_0(f_{i,t-1}(\cdot))$	-0.066	-0.118	-0.047	-0.052	-0.314	-1.110	-1.080	-1.032	-1.002	-1.105
	$ES_0(f_{EW,t-1}(\cdot))$	-0.048	-0.096	-0.032	-0.038	-0.283	-1.104	-1.075	-1.027	-0.998	-1.100
LS	$ES_0(f_{i^*,t-1}(\cdot))$	-0.031	-0.076	-0.027	-0.022	-0.247	-1.100	-1.072	-1.026	-0.992	-1.092
	Rej. ($T = 120$)	0.208	0.174	0.191	0.328	0.263	0.043	0.033	0.117	0.186	0.118
	Rej. ($T = 360$)	0.440	0.434	0.267	0.551	0.679	0.106	0.074	0.123	0.264	0.232
	$\frac{1}{n}\sum_{i=1}^{n} ES_0(f_{i,t-1}(\cdot))$	1.094	1.040	1.114	1.109	0.862	0.384	0.396	0.416	0.429	0.386
	$ES_0(f_{EW,t-1}(\cdot))$	1.114	1.062	1.131	1.123	0.879	0.386	0.398	0.417	0.430	0.388
QS	$ES_0(f_{i^*,t-1}(\cdot))$	1.130	1.082	1.135	1.140	0.910	0.388	0.399	0.418	0.433	0.391
	Rej. ($T = 120$)	0.104	0.109	0.081	0.213	0.215	0.049	0.038	0.066	0.102	0.100
	Rej. ($T = 360$)	0.252	0.283	0.116	0.359	0.500	0.092	0.070	0.066	0.150	0.173
	$\frac{1}{n}\sum_{i=1}^{n} ES_0(f_{i,t-1}(\cdot))$	-0.145	-0.153	-0.142	-0.143	-0.180	-0.413	-0.401	-0.382	-0.370	-0.411
	$ES_0(f_{EW,t-1}(\cdot))$	-0.143	-0.150	-0.140	-0.142	-0.178	-0.411	-0.400	-0.381	-0.369	-0.410
$CRPS$	$ES_0(f_{i^*,t-1}(\cdot))$	-0.141	-0.147	-0.140	-0.140	-0.175	-0.410	-0.399	-0.381	-0.368	-0.407
	Rej. ($T = 120$)	0.118	0.141	0.088	0.262	0.252	0.043	0.038	0.098	0.145	0.132
	Rej. ($T = 360$)	0.304	0.374	0.135	0.465	0.519	0.091	0.080	0.097	0.195	0.230

Table 4: Simulation results. Horizontal blocks represent the log score (LS), quadratic score (QS) and cumulative ranked probability score ($CRPS$). In each block, the first three rows are simulation estimates of the quantities in Equation (16). All estimates are averages over 10000 Monte Carlo samples, each of which is 480 periods long. The fourth and fifth rows are rejection frequencies of the null hypothesis in (18), for two different sample sizes. The frequencies are computed over 10000 Monte Carlo samples. A truncation lag of four is used for the Newey and West (1987) estimator. Columns represent different true processes, calibrated to different subsamples of CPI inflation and industrial production. See Table 3 for details on calibration.

	True model (i^*)	*Treasury Bill Rate* 1	2	3	4	5	*Unemployment Rate* 1	2	3	4	5
LS	$\frac{1}{n}\sum_{i=1}^{n} ES_0(f_{i,t-1}(\cdot))$	-1.175	-1.445	-1.514	-0.010	-0.024	0.273	0.290	0.326	0.413	0.362
	$ES_0(f_{EW,t-1}(\cdot))$	-0.621	-0.736	-0.767	0.061	0.048	0.280	0.296	0.332	0.417	0.367
	$ES_0(f_{i^*,t-1}(\cdot))$	-0.570	-0.657	-0.681	0.299	0.275	0.286	0.298	0.332	0.424	0.378
	Rej. ($T = 120$)	0.436	0.594	0.634	0.998	0.997	0.064	0.039	0.058	0.184	0.149
	Rej. ($T = 360$)	0.860	0.962	0.975	1.000	1.000	0.157	0.070	0.064	0.279	0.303
QS	$\frac{1}{n}\sum_{i=1}^{n} ES_0(f_{i,t-1}(\cdot))$	0.450	0.368	0.349	1.232	1.211	1.535	1.561	1.617	1.766	1.676
	$ES_0(f_{EW,t-1}(\cdot))$	0.587	0.506	0.488	1.365	1.344	1.543	1.569	1.625	1.773	1.684
	$ES_0(f_{i^*,t-1}(\cdot))$	0.659	0.603	0.590	1.573	1.535	1.552	1.572	1.626	1.783	1.702
	Rej. ($T = 120$)	0.424	0.570	0.594	0.919	0.901	0.073	0.054	0.055	0.096	0.097
	Rej. ($T = 360$)	0.835	0.947	0.957	1.000	1.000	0.130	0.063	0.059	0.155	0.192
CRPS	$\frac{1}{n}\sum_{i=1}^{n} ES_0(f_{i,t-1}(\cdot))$	-0.254	-0.279	-0.286	-0.121	-0.124	-0.103	-0.102	-0.098	-0.090	-0.095
	$ES_0(f_{EW,t-1}(\cdot))$	-0.245	-0.270	-0.277	-0.113	-0.116	-0.103	-0.101	-0.098	-0.089	-0.095
	$ES_0(f_{i^*,t-1}(\cdot))$	-0.241	-0.264	-0.270	-0.101	-0.104	-0.103	-0.101	-0.098	-0.089	-0.094
	Rej. ($T = 120$)	0.306	0.460	0.541	0.999	0.998	0.080	0.034	0.069	0.142	0.099
	Rej. ($T = 360$)	0.711	0.905	0.945	1.000	1.000	0.151	0.055	0.072	0.202	0.243

Table 5: Simulation results (continued). True processes calibrated to the treasury bill rate and the unemployment rate. See Table 4 for details.

Back

Consumer Price Index

Model	LS	QS	CRPS	SE
		$h=1$		
AR^s	$0.02^{22.9}$	$1.49^{75.8}$	$-0.12^{63.3}$	$0.06^{76.2}$
AR^l	$-0.09^{2.5}$	$1.36^{0.0}$	$-0.13^{0.0}$	$0.06^{2.3}$
$VAR^s_{C,I}$	$0.00^{10.3}$	$1.46^{53.5}$	$-0.13^{7.8}$	$0.06^{4.7}$
$VAR^l_{C,I}$	$-0.11^{0.6}$	$1.33^{0.0}$	$-0.13^{0.0}$	$0.06^{0.7}$
$VAR^s_{C,T}$	$0.04^{39.1}$	$1.47^{78.7}$	$-0.13^{40.8}$	$0.06^{76.1}$
$VAR^l_{C,T}$	$-0.09^{1.0}$	$1.36^{0.0}$	$-0.13^{0.0}$	$0.06^{4.4}$
$VAR^s_{C,U}$	$-0.01^{6.7}$	$1.46^{41.5}$	$-0.13^{2.8}$	$0.06^{2.9}$
$VAR^l_{C,U}$	$-0.12^{0.8}$	$1.33^{0.0}$	$-0.13^{0.0}$	$0.06^{0.0}$
EW	0.06	1.48	−0.12	0.06
(Rank)	(1)	(2)	(2)	(2)
		$h=3$		
AR^s	$-0.10^{58.3}$	$1.39^{41.7}$	$-0.13^{46.6}$	$0.08^{82.2}$
AR^l	$-0.27^{3.5}$	$1.20^{0.0}$	$-0.14^{0.0}$	$0.08^{0.5}$
$VAR^s_{C,I}$	$-0.09^{77.6}$	$1.37^{70.4}$	$-0.14^{75.3}$	$0.08^{40.3}$
$VAR^l_{C,I}$	$-0.28^{2.1}$	$1.18^{0.0}$	$-0.15^{0.0}$	$0.08^{0.1}$
$VAR^s_{C,T}$	$-0.09^{68.9}$	$1.38^{64.3}$	$-0.14^{99.2}$	$0.08^{70.2}$
$VAR^l_{C,T}$	$-0.28^{2.1}$	$1.17^{0.0}$	$-0.15^{0.0}$	$0.08^{0.0}$
$VAR^s_{C,U}$	$-0.09^{68.0}$	$1.37^{74.8}$	$-0.14^{58.9}$	$0.08^{31.0}$
$VAR^l_{C,U}$	$-0.29^{1.6}$	$1.17^{0.0}$	$-0.15^{0.0}$	$0.08^{0.0}$
EW	−0.08	1.36	−0.14	0.08
(Rank)	(1)	(5)	(2)	(1)
		$h=6$		
AR^s	$-0.12^{85.8}$	$1.34^{51.0}$	$-0.14^{45.6}$	$0.08^{82.2}$
AR^l	$-0.31^{2.4}$	$1.13^{0.0}$	$-0.15^{0.0}$	$0.08^{0.3}$
$VAR^s_{C,I}$	$-0.11^{98.4}$	$1.34^{53.6}$	$-0.14^{43.4}$	$0.08^{65.8}$
$VAR^l_{C,I}$	$-0.33^{1.4}$	$1.10^{0.0}$	$-0.15^{0.0}$	$0.08^{0.0}$
$VAR^s_{C,T}$	$-0.11^{96.6}$	$1.34^{58.2}$	$-0.14^{57.8}$	$0.08^{93.0}$
$VAR^l_{C,T}$	$-0.33^{1.2}$	$1.09^{0.0}$	$-0.15^{0.0}$	$0.08^{0.0}$
$VAR^s_{C,U}$	$-0.11^{99.7}$	$1.34^{47.9}$	$-0.14^{39.3}$	$0.08^{61.8}$
$VAR^l_{C,U}$	$-0.33^{1.4}$	$1.11^{0.0}$	$-0.15^{0.0}$	$0.08^{0.0}$
EW	−0.11	1.32	−0.14	0.08
(Rank)	(2)	(5)	(5)	(5)

Industrial Production

Model	LS	QS	CRPS	SE
		$h=1$		
AR^s	$-1.04^{3.2}$	$0.51^{20.4}$	$-0.35^{0.9}$	$0.44^{1.1}$
AR^l	$-1.02^{2.1}$	$0.51^{31.8}$	$-0.33^{79.9}$	$0.39^{62.0}$
$VAR^s_{I,C}$	$-1.07^{0.6}$	$0.49^{1.0}$	$-0.36^{0.1}$	$0.46^{1.1}$
$VAR^l_{I,C}$	$-1.05^{6.7}$	$0.50^{3.1}$	$-0.34^{11.5}$	$0.41^{43.3}$
$VAR^s_{I,T}$	$-1.01^{22.2}$	$0.51^{51.8}$	$-0.34^{34.4}$	$0.41^{28.2}$
$VAR^l_{I,T}$	$-1.05^{5.8}$	$0.50^{12.0}$	$-0.34^{28.6}$	$0.41^{43.4}$
$VAR^s_{I,U}$	$-1.02^{9.7}$	$0.49^{1.3}$	$-0.35^{1.4}$	$0.43^{5.7}$
$VAR^l_{I,U}$	$-1.04^{0.9}$	$0.50^{1.4}$	$-0.34^{5.2}$	$0.41^{46.0}$
EW	-0.98	0.52	-0.33	0.40
(Rank)	(1)	(1)	(1)	(2)
		$h=3$		
AR^s	$-1.05^{11.9}$	$0.49^{27.7}$	$-0.35^{2.9}$	$0.43^{2.4}$
AR^l	$-1.04^{9.6}$	$0.48^{15.8}$	$-0.33^{71.7}$	$0.37^{23.9}$
$VAR^s_{I,C}$	$-1.04^{21.1}$	$0.49^{18.4}$	$-0.35^{0.4}$	$0.42^{0.2}$
$VAR^l_{I,C}$	$-1.07^{7.9}$	$0.48^{3.1}$	$-0.34^{23.6}$	$0.41^{57.7}$
$VAR^s_{I,T}$	$-1.03^{31.9}$	$0.49^{37.7}$	$-0.34^{4.4}$	$0.42^{2.5}$
$VAR^l_{I,T}$	$-1.08^{6.1}$	$0.48^{10.2}$	$-0.34^{40.6}$	$0.40^{77.3}$
$VAR^s_{I,U}$	$-1.04^{15.3}$	$0.48^{13.7}$	$-0.35^{0.2}$	$0.43^{0.0}$
$VAR^l_{I,U}$	$-1.08^{3.2}$	$0.48^{7.7}$	$-0.34^{18.1}$	$0.40^{47.3}$
EW	-1.01	0.50	-0.34	0.40
(Rank)	(1)	(1)	(2)	(2)
		$h=6$		
AR^s	$-1.14^{9.4}$	$0.47^{11.0}$	$-0.37^{4.7}$	$0.49^{8.8}$
AR^l	$-1.12^{5.3}$	$0.47^{9.4}$	$-0.35^{52.8}$	$0.42^{66.0}$
$VAR^s_{I,C}$	$-1.10^{56.1}$	$0.48^{46.9}$	$-0.35^{12.6}$	$0.43^{15.9}$
$VAR^l_{I,C}$	$-1.11^{9.4}$	$0.47^{34.7}$	$-0.35^{69.4}$	$0.41^{26.3}$
$VAR^s_{I,T}$	$-1.09^{80.2}$	$0.48^{56.8}$	$-0.35^{33.8}$	$0.43^{58.7}$
$VAR^l_{I,T}$	$-1.12^{7.0}$	$0.47^{19.7}$	$-0.35^{94.2}$	$0.41^{28.5}$
$VAR^s_{I,U}$	$-1.10^{48.0}$	$0.48^{39.6}$	$-0.35^{8.5}$	$0.44^{10.8}$
$VAR^l_{I,U}$	$-1.14^{1.1}$	$0.46^{5.3}$	$-0.35^{12.5}$	$0.43^{41.8}$
EW	-1.08	0.48	-0.35	0.43
(Rank)	(1)	(1)	(3)	(4)

Treasury Bill Rate

Model	LS	QS	CRPS	SE
		$h=1$		
AR^s	$-0.02^{58.9}$	$1.61^{0.4}$	$-0.12^{34.8}$	$0.04^{24.2}$
AR^l	$-0.23^{0.0}$	$1.11^{0.0}$	$-0.15^{0.0}$	$0.05^{0.0}$
$VAR^s_{T,C}$	$-0.02^{56.0}$	$1.61^{0.5}$	$-0.12^{58.4}$	$0.04^{5.6}$
$VAR^l_{T,C}$	$-0.26^{0.0}$	$1.07^{0.0}$	$-0.15^{0.0}$	$0.04^{1.6}$
$VAR^s_{T,I}$	$0.04^{99.4}$	$1.64^{0.1}$	$-0.12^{22.6}$	$0.04^{15.4}$
$VAR^l_{T,I}$	$-0.24^{0.0}$	$1.08^{0.0}$	$-0.14^{0.0}$	$0.04^{20.0}$
$VAR^s_{T,U}$	$-0.07^{45.0}$	$1.60^{0.7}$	$-0.12^{88.1}$	$0.04^{2.1}$
$VAR^l_{T,U}$	$-0.24^{0.0}$	$1.09^{0.0}$	$-0.14^{0.0}$	$0.04^{0.1}$
EW	0.04	1.49	−0.12	0.04
(Rank)	(1)	(5)	(5)	(1)
		$h=3$		
AR^s	$-0.21^{45.7}$	$1.40^{5.3}$	$-0.13^{50.3}$	$0.05^{39.8}$
AR^l	$-0.32^{0.0}$	$0.99^{0.0}$	$-0.16^{0.0}$	$0.06^{0.1}$
$VAR^s_{T,C}$	$-0.21^{47.6}$	$1.40^{2.8}$	$-0.13^{52.1}$	$0.05^{39.1}$
$VAR^l_{T,C}$	$-0.35^{0.0}$	$0.96^{0.0}$	$-0.16^{0.0}$	$0.05^{21.8}$
$VAR^s_{T,I}$	$-0.11^{79.5}$	$1.44^{0.1}$	$-0.13^{5.9}$	$0.05^{77.1}$
$VAR^l_{T,I}$	$-0.34^{0.0}$	$0.98^{0.0}$	$-0.16^{0.0}$	$0.05^{47.4}$
$VAR^s_{T,U}$	$-0.21^{48.2}$	$1.42^{0.5}$	$-0.13^{51.8}$	$0.05^{32.1}$
$VAR^l_{T,U}$	$-0.34^{0.0}$	$0.98^{0.0}$	$-0.16^{0.0}$	$0.05^{77.7}$
EW	−0.08	1.30	−0.13	0.05
(Rank)	(1)	(5)	(5)	(3)
		$h=6$		
AR^s	$-0.10^{99.2}$	$1.39^{0.7}$	$-0.13^{35.6}$	$0.05^{21.4}$
AR^l	$-0.34^{0.0}$	$0.97^{0.0}$	$-0.16^{0.0}$	$0.06^{0.5}$
$VAR^s_{T,C}$	$-0.09^{81.5}$	$1.39^{0.5}$	$-0.13^{23.7}$	$0.05^{30.7}$
$VAR^l_{T,C}$	$-0.36^{0.0}$	$0.96^{0.0}$	$-0.16^{0.0}$	$0.05^{72.0}$
$VAR^s_{T,I}$	$-0.09^{82.0}$	$1.39^{0.7}$	$-0.13^{27.7}$	$0.05^{44.4}$
$VAR^l_{T,I}$	$-0.36^{0.0}$	$0.96^{0.0}$	$-0.16^{0.0}$	$0.05^{80.6}$
$VAR^s_{T,U}$	$-0.09^{81.2}$	$1.38^{0.9}$	$-0.13^{24.4}$	$0.05^{37.2}$
$VAR^l_{T,U}$	$-0.36^{0.0}$	$0.96^{0.0}$	$-0.16^{0.0}$	$0.05^{95.8}$
EW	−0.10	1.28	−0.14	0.05
(Rank)	(4)	(5)	(5)	(2)

Unemployment Rate

Model	LS	QS	CRPS	SE
		$h=1$		
AR^s	$0.38^{2.1}$	$1.78^{0.4}$	$-0.09^{1.2}$	$0.03^{4.0}$
AR^l	$0.41^{14.9}$	$1.87^{61.5}$	$-0.09^{99.0}$	$0.02^{78.2}$
$VAR^s_{U,C}$	$0.37^{1.2}$	$1.79^{0.9}$	$-0.09^{0.2}$	$0.03^{0.2}$
$VAR^l_{U,C}$	$0.36^{3.3}$	$1.80^{0.7}$	$-0.09^{2.7}$	$0.03^{6.7}$
$VAR^s_{U,I}$	$0.40^{7.8}$	$1.80^{1.8}$	$-0.09^{2.8}$	$0.03^{7.9}$
$VAR^l_{U,I}$	$0.41^{11.4}$	$1.85^{30.2}$	$-0.09^{39.9}$	$0.02^{51.7}$
$VAR^s_{U,T}$	$0.39^{21.0}$	$1.86^{39.0}$	$-0.09^{11.0}$	$0.03^{7.1}$
$VAR^l_{U,T}$	$0.34^{3.8}$	$1.80^{3.7}$	$-0.09^{6.6}$	$0.03^{9.6}$
EW	0.44	1.89	-0.09	0.02
(Rank)	(1)	(1)	(1)	(2)
		$h=3$		
AR^s	$0.38^{21.9}$	$1.80^{1.9}$	$-0.09^{6.9}$	$0.03^{14.7}$
AR^l	$0.40^{48.2}$	$1.86^{81.3}$	$-0.09^{60.7}$	$0.02^{42.1}$
$VAR^s_{U,C}$	$0.38^{8.4}$	$1.82^{3.7}$	$-0.09^{0.0}$	$0.03^{0.0}$
$VAR^l_{U,C}$	$0.37^{9.2}$	$1.84^{24.0}$	$-0.09^{37.9}$	$0.02^{53.6}$
$VAR^s_{U,I}$	$0.38^{8.0}$	$1.81^{2.0}$	$-0.09^{0.1}$	$0.03^{0.1}$
$VAR^l_{U,I}$	$0.38^{2.4}$	$1.84^{14.3}$	$-0.09^{79.3}$	$0.02^{42.4}$
$VAR^s_{U,T}$	$0.40^{30.5}$	$1.84^{24.5}$	$-0.09^{3.0}$	$0.03^{1.9}$
$VAR^l_{U,T}$	$0.36^{6.3}$	$1.83^{20.4}$	$-0.09^{36.3}$	$0.02^{60.3}$
EW	0.41	1.87	-0.09	0.02
(Rank)	(1)	(1)	(2)	(3)
		$h=6$		
AR^s	$0.33^{9.5}$	$1.77^{2.5}$	$-0.09^{1.3}$	$0.03^{2.2}$
AR^l	$0.35^{29.3}$	$1.82^{68.5}$	$-0.09^{74.6}$	$0.03^{38.9}$
$VAR^s_{U,C}$	$0.35^{20.4}$	$1.80^{7.8}$	$-0.09^{0.1}$	$0.03^{0.0}$
$VAR^l_{U,C}$	$0.35^{15.2}$	$1.82^{60.8}$	$-0.09^{65.0}$	$0.03^{23.3}$
$VAR^s_{U,I}$	$0.34^{14.0}$	$1.79^{4.3}$	$-0.09^{0.0}$	$0.03^{0.0}$
$VAR^l_{U,I}$	$0.35^{7.4}$	$1.81^{22.5}$	$-0.09^{69.7}$	$0.03^{39.9}$
$VAR^s_{U,T}$	$0.36^{48.5}$	$1.80^{9.5}$	$-0.09^{1.8}$	$0.03^{2.6}$
$VAR^l_{U,T}$	$0.35^{6.5}$	$1.80^{16.6}$	$-0.09^{59.9}$	$0.03^{51.5}$
EW	0.37	1.83	-0.09	0.03
(Rank)	(1)	(1)	(3)	(5)