Testing for State-Dependent Predictive Ability

Sebastian Fossati

University of Alberta

IAAE 2016

Macroeconomic Forecasting

Empirical fact: **predictability is unstable over time**.

- many individual indicators exhibit predictive content for output growth and inflation but only sporadically
- Stock and Watson (2003), Rossi and Sekhposyan (2010), Rossi (2013), and Granziera and Sekhposyan (2014), among others

New (but related) empirical fact: **predictability varies across states**.

- Phillips curve forecasts of inflation rate: Dotsey et al. (2015), Gibbs (2015)
- output growth: Chauvet and Potter (2013)
- stock returns: Rapach et al. (2010), Henkel et al.(2011), Dangl et al.(2012)
- bond excess returns: Gargano et al. (2014)
- commodity prices: Gargano and Timmermann (2014)

Macroeconomic Forecasting

Empirical fact: **predictability is unstable over time**.

- many individual indicators exhibit predictive content for output growth and inflation but only sporadically
- Stock and Watson (2003), Rossi and Sekhposyan (2010), Rossi (2013), and Granziera and Sekhposyan (2014), among others

New (but related) empirical fact: **predictability varies across states**.

- Phillips curve forecasts of inflation rate: Dotsey et al. (2015), Gibbs (2015)
- output growth: Chauvet and Potter (2013)
- stock returns: Rapach et al. (2010), Henkel et al.(2011), Dangl et al.(2012)
- bond excess returns: Gargano et al. (2014)
- commodity prices: Gargano and Timmermann (2014)

Macroeconomic Forecasting

Empirical fact: **predictability is unstable over time**.

- many individual indicators exhibit predictive content for output growth and inflation but only sporadically
- Stock and Watson (2003), Rossi and Sekhposyan (2010), Rossi (2013), and Granziera and Sekhposyan (2014), among others

New (but related) empirical fact: **predictability varies across states**.

- Phillips curve forecasts of inflation rate: Dotsey et al. (2015), Gibbs (2015)
- output growth: Chauvet and Potter (2013)
- stock returns: Rapach et al. (2010), Henkel et al.(2011), Dangl et al.(2012)
- bond excess returns: Gargano et al. (2014)
- commodity prices: Gargano and Timmermann (2014)

This Paper

Current approach: use **exogenously provided shift dates**.

- NBER recession dates for the US economy

This paper: A test for comparing the out-of-sample forecasting performance of two competing models with state-dependent predictive content.

Main features:

- forecast loss differences modeled using Markov-switching process
- econometrician not required to observe when underlying states shift
- test equal and constant predictive ability vs. state-dependent ability
- heteroskedasticity and autocorrelation consistent (HAC) test

This Paper

Current approach: use **exogenously provided shift dates**.

- NBER recession dates for the US economy

This paper: A test for comparing the out-of-sample forecasting performance of two competing models with state-dependent predictive content.

Main features:

- forecast loss differences modeled using Markov-switching process
- econometrician not required to observe when underlying states shift
- test equal and constant predictive ability vs. state-dependent ability
- heteroskedasticity and autocorrelation consistent (HAC) test

This Paper

Current approach: use **exogenously provided shift dates**.

- NBER recession dates for the US economy

This paper: A test for comparing the out-of-sample forecasting performance of two competing models with state-dependent predictive content.

Main features:

- forecast loss differences modeled using Markov-switching process
- econometrician not required to observe when underlying states shift
- test equal and constant predictive ability vs. state-dependent ability
- heteroskedasticity and autocorrelation consistent (HAC) test

This Paper

Current approach: use **exogenously provided shift dates**.

- NBER recession dates for the US economy

This paper: A test for comparing the out-of-sample forecasting performance of two competing models with state-dependent predictive content.

Main features:

- forecast loss differences modeled using Markov-switching process
- econometrician not required to observe when underlying states shift
- test equal and constant predictive ability vs. state-dependent ability
- heteroskedasticity and autocorrelation consistent (HAC) test

This Paper

Current approach: use **exogenously provided shift dates**.

- NBER recession dates for the US economy

This paper: A test for comparing the out-of-sample forecasting performance of two competing models with state-dependent predictive content.

Main features:

- forecast loss differences modeled using Markov-switching process
- econometrician not required to observe when underlying states shift
- test equal and constant predictive ability vs. state-dependent ability
- heteroskedasticity and autocorrelation consistent (HAC) test

This Paper

Current approach: use **exogenously provided shift dates**.

- NBER recession dates for the US economy

This paper: A test for comparing the out-of-sample forecasting performance of two competing models with state-dependent predictive content.

Main features:

- forecast loss differences modeled using Markov-switching process
- econometrician not required to observe when underlying states shift
- test equal and constant predictive ability vs. state-dependent ability
- heteroskedasticity and autocorrelation consistent (HAC) test

Main Results

Available tests (Diebold-Mariano, Giacomini-White, Giacomini-Rossi, etc.):

- can have low power when predictability varies across economic states
- even when they reject, can lead to incorrect inference

Proposed test:

- performs well with unequal but constant performance
- exhibits more power when predictability varies across states
- estimates when underlying states shift
- HAC estimators work well

Main Results

Available tests (Diebold-Mariano, Giacomini-White, Giacomini-Rossi, etc.):

- can have low power when predictability varies across economic states
- even when they reject, can lead to incorrect inference

Proposed test:

- performs well with unequal but constant performance
- exhibits more power when predictability varies across states
- estimates when underlying states shift
- HAC estimators work well

Main Results

Available tests (Diebold-Mariano, Giacomini-White, Giacomini-Rossi, etc.):

- can have low power when predictability varies across economic states
- even when they reject, can lead to incorrect inference

Proposed test:

- performs well with unequal but constant performance
- exhibits more power when predictability varies across states
- estimates when underlying states shift
- HAC estimators work well

Main Results

Available tests (Diebold-Mariano, Giacomini-White, Giacomini-Rossi, etc.):

- can have low power when predictability varies across economic states
- even when they reject, can lead to incorrect inference

Proposed test:

- performs well with unequal but constant performance
- exhibits more power when predictability varies across states
- estimates when underlying states shift
- HAC estimators work well

Main Results

Available tests (Diebold-Mariano, Giacomini-White, Giacomini-Rossi, etc.):

- can have low power when predictability varies across economic states
- even when they reject, can lead to incorrect inference

Proposed test:

- performs well with unequal but constant performance
- exhibits more power when predictability varies across states
- estimates when underlying states shift
- HAC estimators work well

Environment

Objective: compare sequences of h-step ahead out-of-sample forecasts for y_t obtained from two competing models.

Set-up (Giacomini and White, 2006):

- sample of size T, in-sample portion R, out-of-sample portion P
- out-of-sample forecast loss differences $\Delta L_t(\hat{\delta}_{t-h,R}, \hat{\gamma}_{t-h,R})$
- parameters estimated using rolling scheme
- loss differences as observed data

Null hypothesis of **equal predictive ability** (Diebold and Mariano, 1995):

$$H_0^{(1)} : E\left[\Delta L_t(\hat{\delta}_{t-h,R}, \hat{\gamma}_{t-h,R})\right] = 0 \text{ for all } t = R+h, \ldots, T$$

Environment

Objective: compare sequences of h-step ahead out-of-sample forecasts for y_t obtained from two competing models.

Set-up (Giacomini and White, 2006):

- sample of size T, in-sample portion R, out-of-sample portion P
- out-of-sample forecast loss differences $\Delta L_t(\hat{\delta}_{t-h,R}, \hat{\gamma}_{t-h,R})$
- parameters estimated using rolling scheme
- loss differences as observed data

Null hypothesis of **equal predictive ability** (Diebold and Mariano, 1995):

$$H_0^{(1)} : E\left[\Delta L_t(\hat{\delta}_{t-h,R}, \hat{\gamma}_{t-h,R})\right] = 0 \text{ for all } t = R + h, \ldots, T$$

Environment

Objective: compare sequences of h-step ahead out-of-sample forecasts for y_t obtained from two competing models.

Set-up (Giacomini and White, 2006):

- sample of size T, in-sample portion R, out-of-sample portion P
- out-of-sample forecast loss differences $\Delta L_t(\hat{\delta}_{t-h,R}, \hat{\gamma}_{t-h,R})$
- parameters estimated using rolling scheme
- loss differences as observed data

Null hypothesis of **equal predictive ability** (Diebold and Mariano, 1995):

$$H_0^{(1)} : E\left[\Delta L_t(\hat{\delta}_{t-h,R}, \hat{\gamma}_{t-h,R})\right] = 0 \text{ for all } t = R+h, \dots, T$$

State-Dependent Loss Differences

Model out-of-sample forecast loss differences using **Markov-switching** mean plus noise **regression**:

$$\Delta L_t(\hat{\delta}_{t-h,R}, \hat{\gamma}_{t-h,R}) = \mu_{s_t} + \sigma_{s_t} u_t$$

where s_t $(= 0, 1)$ is an unobserved two-state first-order Markov process with transition probabilities

$$\text{Prob}(s_t = j \mid s_{t-1} = i) = p_{ij} \quad i, j = 0, 1$$

and u_t is an unobservable moving-average (MA) process with zero mean and non-zero autocorrelations up to lag $h - 1$.

Next, estimate $\theta = (\mu_0, \mu_1, \sigma_0, \sigma_1, p_{00}, p_{11})'$ by quasi maximum likelihood (Q-ML) following Hamilton (1989), Hamilton (1990), or Kim and Nelson (1999).

Predictive Ability Tests

Null hypothesis of **equal and constant predictive ability**: $\mu_0 = \mu_1 = 0$.

Use **Wald test** statistic:

$$\text{SD-Wald} = P\big(R_0\hat{\theta}\big)'\big(R_0\hat{\Omega}R_0'\big)^{-1}\big(R_0\hat{\theta}\big)$$

where $R_0 = (I_2, 0_2, 0_2)$ and $\hat{\Omega}$ a consistent estimator of Ω.

Under the null hypothesis, SD-Wald $\xrightarrow{d} \chi^2(2)$ as $P \to \infty$.

Predictive Ability Tests

Null hypothesis of **equal and constant predictive ability**: $\mu_0 = \mu_1 = 0$.

Use **Wald test** statistic:

$$\text{SD-Wald} = P\left(R_0\hat{\theta}\right)'\left(R_0\hat{\Omega}R_0'\right)^{-1}\left(R_0\hat{\theta}\right)$$

where $R_0 = (I_2, 0_2, 0_2)$ and $\hat{\Omega}$ a consistent estimator of Ω.

Under the null hypothesis, SD-Wald $\xrightarrow{d} \chi^2(2)$ as $P \to \infty$.

Estimators of Ω

Most common estimators of Ω:

- **Hessian** estimator as in Hamilton (1989)
- **Outer-product** estimator as in Hamilton (1996)

But HAC estimator of the covariance matrix is required:

1. multi-step forecasts (Diebold and Mariano, 1995)
2. instabilities (Morley and Rabah, 2014; Martins and Perron, 2014)

Alternative estimator: **'sandwich'** estimator (Hayashi, 2000).

- suggested in Hamilton (1996) to calculate robust standard errors with serially uncorrelated scores
- HAC estimators work well with autocorrelated scores

Estimators of Ω

Most common estimators of Ω:

- **Hessian** estimator as in Hamilton (1989)
- **Outer-product** estimator as in Hamilton (1996)

But HAC estimator of the covariance matrix is required:

1. multi-step forecasts (Diebold and Mariano, 1995)
2. instabilities (Morley and Rabah, 2014; Martins and Perron, 2014)

Alternative estimator: **'sandwich'** estimator (Hayashi, 2000).

- suggested in Hamilton (1996) to calculate robust standard errors with serially uncorrelated scores
- HAC estimators work well with autocorrelated scores

Estimators of Ω

Most common estimators of Ω:

- **Hessian** estimator as in Hamilton (1989)
- **Outer-product** estimator as in Hamilton (1996)

But HAC estimator of the covariance matrix is required:

1. multi-step forecasts (Diebold and Mariano, 1995)
2. instabilities (Morley and Rabah, 2014; Martins and Perron, 2014)

Alternative estimator: **'sandwich'** estimator (Hayashi, 2000).

- suggested in Hamilton (1996) to calculate robust standard errors with serially uncorrelated scores
- HAC estimators work well with autocorrelated scores

Unequal but Constant Performance

Data generating process:

$$y_t = \beta x_t + \sigma_\varepsilon \varepsilon_t$$
$$x_t = \phi x_{t-1} + \sigma_\nu \nu_t$$

- ε_t and ν_t are i.i.d.$N(0,1)$, $\phi = .5$, $\sigma_\varepsilon = \sigma_\nu = 1$

The time-t one-step ahead forecasts of y_{t+1}:

$$\hat{f}^1_{t,R} = 0$$
$$\hat{f}^2_{t,R} = \hat{\beta}_{t,R} x_{t+1}$$

- $\hat{\beta}_{t,R}$ the in-sample estimate of β, x_{t+1} known at time t
- same DGP as Giacomini and White (2006) and Martins and Perron (2014)

Different Performance in Different States

Data generating process:

$$\begin{aligned} y_t &= -\beta s_t + \sigma_\varepsilon \varepsilon_t \\ x_t &= \delta s_t + \sigma_\nu \nu_t \end{aligned}$$

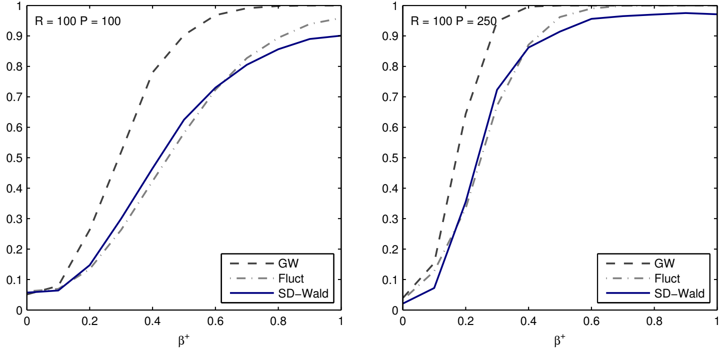
- ε_t and ν_t are i.i.d.$N(0,1)$, $\delta = 1$, $\sigma_\varepsilon = \sigma_\nu = 0.5$

The time-t one-step ahead forecasts of y_{t+1}:

$$\begin{aligned} \hat{f}^1_{t,R} &= 0 \\ \hat{f}^2_{t,R} &= \hat{\gamma}_{t,R} x_{t+1} \end{aligned}$$

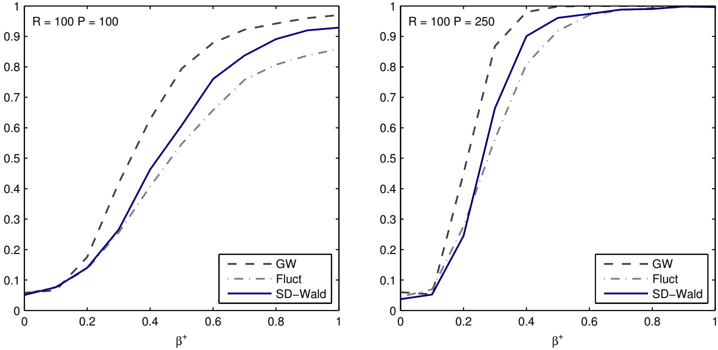
- $\hat{\gamma}_{t,R}$ the in-sample rolling estimate of γ, x_{t+1} known at time t

Figure 3: Unequal but Constant Performance



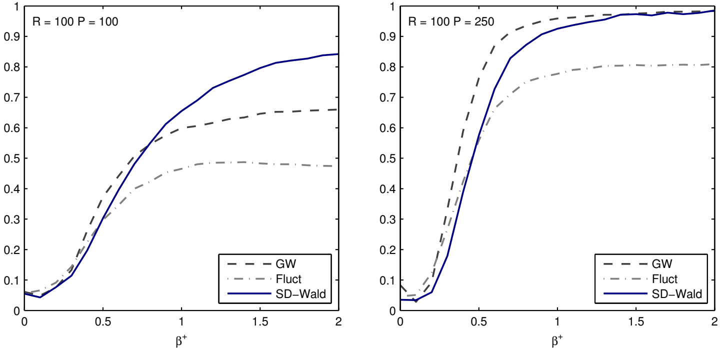
- empirical power (rejection frequencies) of the GW, Fluctuation ($m = .3P$), and SD-Wald tests

Figure 4: Different Performance in Different States (1)



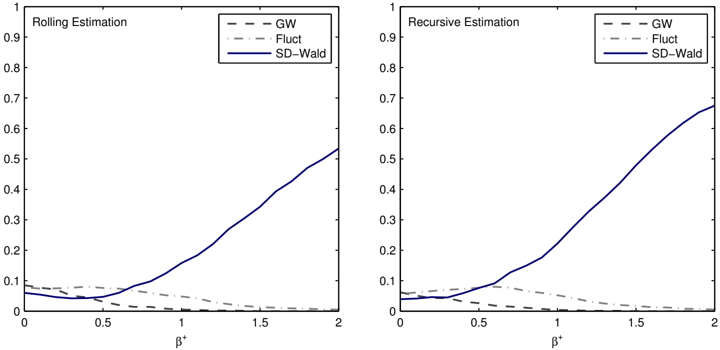
- empirical power (rejection frequencies) of the GW, Fluctuation ($m = .3P$), and SD-Wald tests with $p_{00} = p_{11} = 0.8$

Figure 5: Different Performance in Different States (2)



- empirical power (rejection frequencies) of the GW, Fluctuation ($m = .3P$), and SD-Wald tests with $p_{00} = 0.90$ and $p_{11} = 0.75$

Figure 6: Different Performance in Different States (3)



- empirical power (rejection frequencies) of the GW, Fluctuation ($m = .3P$), and SD-Wald tests with NBER recession dates, $R = 100$, and $P = 124$

Forecasting Output

Chauvet and Potter (2013): most output growth forecasting models exhibit a similar performance during economic expansions but **one model performs significantly better during recessions**.

Two competing forecasting models for real GDP growth:

1. Benchmark model: $\hat{f}^1_{t,R} = \text{AR}(2)$
2. 1 + DF-MS model: $\hat{f}^2_{t,R} = \text{AR}(2) + \text{DF} + \text{prob of recession}$

Other details:

- one-step ahead out-of-sample forecasts
- parameters estimated using rolling window
- forecast comparisons using real-time data

Forecasting Output

Chauvet and Potter (2013): most output growth forecasting models exhibit a similar performance during economic expansions but **one model performs significantly better during recessions**.

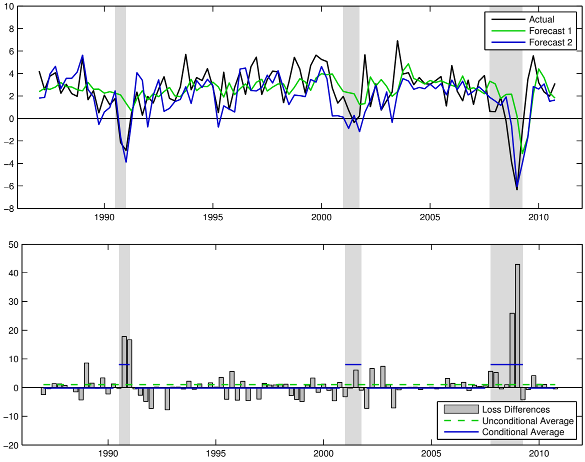
Two competing forecasting models for real GDP growth:

1. Benchmark model: $\hat{f}^1_{t,R} = \text{AR}(2)$
2. 1 + DF-MS model: $\hat{f}^2_{t,R} = \text{AR}(2) + \text{DF} + \text{prob of recession}$

Other details:

- one-step ahead out-of-sample forecasts
- parameters estimated using rolling window
- forecast comparisons using real-time data

Figure 7: Real GDP Growth and Forecasts



- real GDP growth and forecasts (top), loss differences (bottom)

Table 4: Loss Difference Statistics for GDP Growth

	OOS	OOS_0	OOS_1
Observations	96	82	14
Average	1.032	-0.159	8.005
Standard Dev.	6.388	3.019	13.422
AR(1)	0.279**	-0.136	0.121
GW	1.582	-0.477	2.232**
GW(HAC)	1.360		
Fluctuation(HAC, $m = .1P$)	3.517**		
Fluctuation(HAC, $m = .3P$)	2.232		
SD-Wald(HAC)	16.484**		

Notes: Significance of the AR(1) coefficients is tested based on the asymptotic result $\sqrt{T}\hat{\rho}_1 \xrightarrow{d} N(0,1)$. The 5% (10%) critical value is 1.96 (1.645) for a two-sided GW test, 3.012 (2.766) for a two-sided Fluctuation test with a rolling window size of $m = .3P$, and 3.393 (3.170) for $m = .1P$. The 5% (10%) critical value for a SD-Wald test is 5.99 (4.61). ** (*) denotes rejection of the null hypothesis at the 5% (10%) level.

Table 4: Loss Difference Statistics for GDP Growth

	OOS	OOS_0	OOS_1
Observations	96	82	14
Average	1.032	-0.159	8.005
Standard Dev.	6.388	3.019	13.422
AR(1)	0.279**	-0.136	0.121
GW	1.582	-0.477	2.232**
GW(HAC)	1.360		
Fluctuation(HAC, $m = .1P$)	3.517**		
Fluctuation(HAC, $m = .3P$)	2.232		
SD-Wald(HAC)	16.484**		

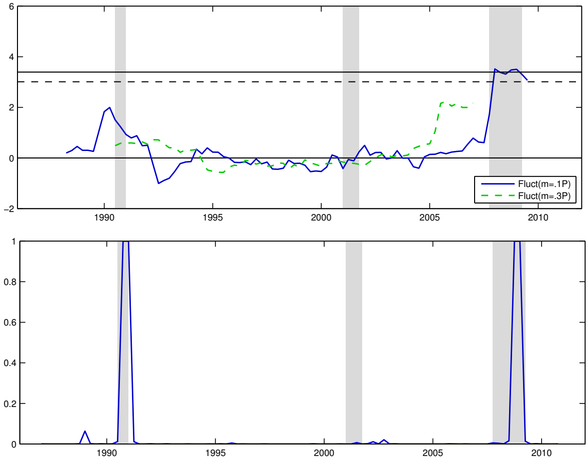
Notes: Significance of the AR(1) coefficients is tested based on the asymptotic result $\sqrt{T}\hat{\rho}_1 \overset{d}{\to} N(0,1)$. The 5% (10%) critical value is 1.96 (1.645) for a two-sided GW test, 3.012 (2.766) for a two-sided Fluctuation test with a rolling window size of $m = .3P$, and 3.393 (3.170) for $m = .1P$. The 5% (10%) critical value for a SD-Wald test is 5.99 (4.61). ** (*) denotes rejection of the null hypothesis at the 5% (10%) level.

Table 4: Loss Difference Statistics for GDP Growth

	OOS	OOS_0	OOS_1
Observations	96	82	14
Average	1.032	-0.159	8.005
Standard Dev.	6.388	3.019	13.422
AR(1)	0.279**	-0.136	0.121
GW	1.582	-0.477	2.232**
GW(HAC)	1.360		
Fluctuation(HAC, $m = .1P$)	3.517**		
Fluctuation(HAC, $m = .3P$)	2.232		
SD-Wald(HAC)	16.484**		

Notes: Significance of the AR(1) coefficients is tested based on the asymptotic result $\sqrt{T}\hat{\rho}_1 \xrightarrow{d} N(0,1)$. The 5% (10%) critical value is 1.96 (1.645) for a two-sided GW test, 3.012 (2.766) for a two-sided Fluctuation test with a rolling window size of $m = .3P$, and 3.393 (3.170) for $m = .1P$. The 5% (10%) critical value for a SD-Wald test is 5.99 (4.61). ** (*) denotes rejection of the null hypothesis at the 5% (10%) level.

Figure 8: Test Statistics and Probabilities of $s_t = 1$



- Fluctuation test statistics (top), probabilities of $s_t = 1$ (bottom)

Appendix

An Illustrative Example

Data generating process:

$$\begin{aligned} y_t &= -\beta s_t + \sigma_\varepsilon \varepsilon_t \\ x_t &= \delta s_t + \sigma_\nu \nu_t \end{aligned}$$

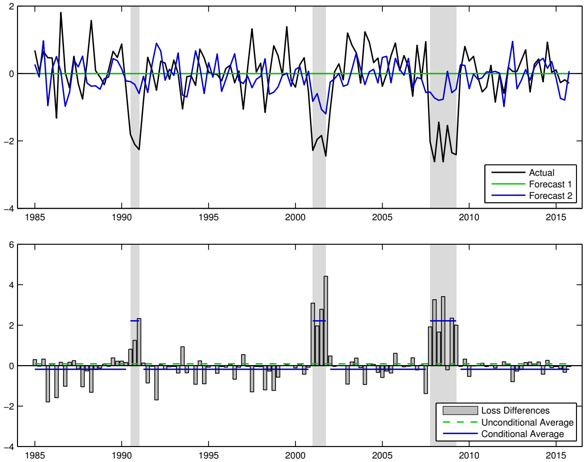
- ε_t and ν_t are i.i.d. $N(0,1)$, $\beta = 2$, $\delta = 1$, $\sigma_\varepsilon = \sigma_\nu = 0.5$
- s_t = actual NBER recession dates 1960Q1–2015Q4 ($T = 224$)

The time-t **one-step ahead forecasts** of y_{t+1}:

$$\begin{aligned} \hat{f}^1_{t,R} &= 0 \\ \hat{f}^2_{t,R} &= \hat{\gamma}_{t,R} x_{t+1} \end{aligned}$$

- $\hat{\gamma}_{t,R}$ in-sample rolling estimate of γ, $R = 100$, $P = 124$

Figure 1: Simulated Time Series and Forecasts



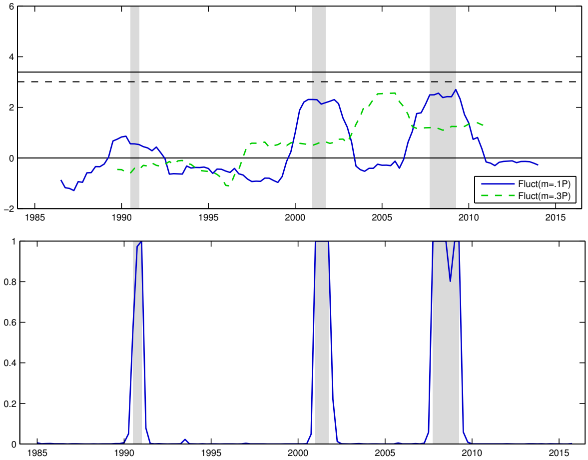
- simulated time series and forecasts (top), loss differences (bottom)

Table 1: Loss Difference Statistics

	OOS	OOS_0	OOS_1
Observations	124	110	14
Average	0.092	-0.179	2.217
Standard Dev.	0.969	0.492	1.169
AR(1)	0.407**	-0.137	-0.005
GW	1.056	-3.809**	7.096**
GW(HAC)	0.683		
Fluctuation(HAC, $m = .1P$)	2.704		
Fluctuation(HAC, $m = .3P$)	2.568		

Notes: Significance of the AR(1) coefficients is tested based on the asymptotic result $\sqrt{T}\hat{\rho}_1 \xrightarrow{d} N(0,1)$. The 5% (10%) critical value is 1.96 (1.645) for a two-sided GW test, 3.012 (2.766) for a two-sided Fluctuation test with a rolling window size of $m = .3P$, and 3.393 (3.170) for $m = .1P$. ** $(^{*})$ denotes rejection of the null hypothesis at the 5% (10%) level.

Figure 2: Test Statistics and Probabilities of $s_t = 1$



- Fluctuation test statistics (top), probabilities of $s_t = 1$ (bottom)

HAC Estimators

Sequences of loss differences ΔL_t generated assuming:

- $\mu_0 = -2$, $\mu_1 = 2$, $\sigma_0^2 = \sigma_1^2 = 1$, and $p_{00} = p_{11} = 0.8$

The error term is the serially correlated MA(1) process:

$$u_t = (1 + \theta^2)^{-1/2}(1 + \theta L)\varepsilon_t$$

- $\varepsilon_t \sim \text{i.i.d.} N(0, 1)$

Set up:

- $T = 50, 100, 250$ and $\theta = 0, 0.5, 0.9$
- Q-ML estimates obtained using EM algorithm (Hamilton, 1990)
- covariance matrix of $\hat{\theta}$ estimated using:
  - Hessian (H) and outer-product (OP) estimators
  - 'sandwich' estimator using Bartlett (NW), Parzen (PK), and quadratic spectral (QS) kernels

Table 2: Small sample properties of ML estimates of $\hat{\mu}_0$

	θ	Bias	SD	H	OP	NW	PK	QS
$T = 100$	0	0.006	0.159	0.154	0.156	0.163	0.162	0.163
				0.933	0.938	0.946	0.942	0.945
	0.5	0.015	0.220	0.161	0.156	0.202	0.210	0.210
				0.846	0.838	0.896	0.909	0.910
	0.9	0.010	0.233	0.163	0.157	0.211	0.219	0.220
				0.824	0.817	0.904	0.908	0.913
$T = 250$	0.9	-0.001	0.135	0.102	0.098	0.129	0.134	0.134
				0.874	0.856	0.932	0.942	0.943

Notes: The covariance matrix of $\hat{\theta}$ is estimated using the Hessian (H) estimator, the outer product (OP) estimator, and the 'sandwich' estimator implemented using three kernels: the Bartlett kernel (NW), the Parzen kernel (PK), and the quadratic spectral (QS) kernel.

Size and Power

$\beta = 0$: the competing forecasting models are equally accurate in the population, then the smaller model would be preferable in the finite sample.

To have forecasting models **equally accurate in the finite sample**:

$$E\left[\left(y_{t+1} - \hat{f}^1_{t,R}\right)^2\right] = E\left[\left(y_{t+1} - \hat{f}^2_{t,R}\right)^2\right]$$

Size:

- Set-up 1: $\beta^0 \approx 1/\sqrt{R4/3}$
- Set-ups 2 & 3: $\beta^0 \approx \sqrt{\delta^2 \left[1 - \frac{\sigma^2_v}{R(\delta^2 p_1 + \sigma^2_v)}\right]^{-1} \frac{\sigma^2_\varepsilon}{R(\delta^2 p_1 + \sigma^2_v)}}$

Power:

- $\beta = \beta^0 + \beta^+$ with $\beta^+ = 0, .1, \ldots, 2$

Size and Power

$\beta = 0$: the competing forecasting models are equally accurate in the population, then the smaller model would be preferable in the finite sample.

To have forecasting models **equally accurate in the finite sample**:

$$E\left[\left(y_{t+1} - \hat{f}^1_{t,R}\right)^2\right] = E\left[\left(y_{t+1} - \hat{f}^2_{t,R}\right)^2\right]$$

Size:

- Set-up 1: $\beta^0 \approx 1/\sqrt{R4/3}$
- Set-ups 2 & 3: $\beta^0 \approx \sqrt{\delta^2 \left[1 - \frac{\sigma^2_\nu}{R(\delta^2 p_1 + \sigma^2_\nu)}\right]^{-1} \frac{\sigma^2_\varepsilon}{R(\delta^2 p_1 + \sigma^2_\nu)}}$

Power:

- $\beta = \beta^0 + \beta^+$ with $\beta^+ = 0, .1, \ldots, 2$

Table 3: Empirical Size under Quadratic Loss

R	P	GW	Fluct	SD-W	GW	Fluct	SD-W
		Set-up 1			Set-up 2		
50	50	0.044	0.043	0.084	0.053	0.056	0.093
	100	0.055	0.048	0.044	0.057	0.040	0.045
	250	0.031	0.037	0.034	0.076	0.044	0.055
100	50	0.059	0.043	0.103	0.047	0.053	0.097
	100	0.038	0.045	0.053	0.067	0.062	0.049
	250	0.038	0.051	0.030	0.064	0.030	0.046
250	50	0.054	0.052	0.118	0.075	0.068	0.114
	100	0.061	0.065	0.056	0.052	0.063	0.052
	250	0.047	0.071	0.041	0.073	0.062	0.045
					Set-up 3		
100	124				0.077	0.075	0.052

Notes: Set-up 1 is unequal but constant performance. Set-up 2 is $p_{00} = p_{11} = 0.8$. Set-up 3 is s_t = actual NBER recession dates 1960Q1–2015Q4 ($T = 224$). Nominal size 0.05.