Assessing Nowcast Accuracy of US GDP Growth in Real Time: The Role of Booms and Busts

Boriss Siliverstovs

Latvijas Banka

boriss.siliverstovs@bank.lv

November 23, 2018

The views expressed here are solely mine and do not necessarily reflect those of the *Latvijas Banka*.

ECONOMETRIC MODEL

FORECAST EVALUATION METRICS FOR MODEL RANKING:

- Aggregated (averaged) measures:

  - (Relative) Differences in RMSFEs

  - Differences in Mean Log-Scores (MLogS)

- Recursive sequences:

  - Cumulated Sum of Squared Forecast Error Difference (CSSFED)

  - Cumulated Sum of Log-Score Difference (CSLogSD)

    Recursive Log Predictive Bayes factors

US DATA AND MODEL SPECIFICATION

RESULTS

- Point forecasts

- Density forecasts

CHOICE OF A BENCHMARK MODEL

ECONOMETRIC MODEL (CCM)

Carriero, A., T. E. Clark, and M. Marcellino (2015). Realtime nowcasting with a Bayesian mixed frequency model with stochastic volatility. Journal of the Royal Statistical Society Series A 178 (4), 837–862.

MIXED-FREQUENCY DATA ARE POOLED IN ONE MODEL

- Quarterly GDP growth
- Monthly economic/financial indicators

STOCHASTIC VOLATILITY

- Time-varying variance of the error term
- Captures changes in forecast uncertainty over time

BAYESIAN ESTIMATION

- Point and density forecasts
- Parameter estimation uncertainty is accounted for
- Solution to the curse of dimensionality: shrinkage to zero

MULTIPLE-INDICATOR U-MIDAS MODEL WITH STOCHASTIC VOLATILITY

ECONOMETRIC MODEL (CCM)

CONDITIONAL MEAN:

$$y_t = X'_{m,t}\beta_m + \varepsilon_{m,t}$$

REGRESSORS $X_{m,t}$:

- Skip-sampled monthly economic/financial indicators
  - One monthly time series $\rightarrow$ three quarterly
    - Jan, Apr, Jul, Oct, Jan,...
    - Feb, May, Aug, Nov, Feb,...
    - Mar, Jun, Sep, Dec, Mar,...
- Specific for every forecast origin
  - Differences in publication lags and release timings

TIME-VARYING VARIANCE OF THE ERROR TERM $\varepsilon_{m,t}$:

$$\varepsilon_{m,t} = \kappa_{m,t}^{0.5}\epsilon_{m,t}, \qquad \epsilon_{m,t} \sim NIID(0,1)$$

$$\ln \kappa_{m,t} = \ln \kappa_{m,t-1} + \nu_{m,t}, \qquad \nu_{m,t} \sim NIID(0,\phi_m)$$

DATA (REAL-TIME VINTAGES) (**1985Q1—2011Q3**)

QUARTERLY GDP GROWTH

12 MONTHLY ECONOMIC/FINANCIAL INDICATORS

Publication lag after the end of month			
	One week	**SP500**	S&P500 index
		TBOND	the 10-year Treasury bond yield
		TBILL	the 3-month Treasury bill rate
		ISM	the ISM index (overall) for manufacturing
		SUPDEL	the ISM index for supplier delivery time
		ORDERS	the ISM index for orders
		HOURS	avg weekly hours of production workers
		EMPLOY	payroll employment
	Two weeks	**CLAIMS**	new claims for unemployment insurance
		RSALES	real retail sales
		IP	industrial production
		STARTS	housing starts

FORECAST SETUP

MODELS:

- (**LRG**) - all 12 indicators
- (**SML**) - five selected indicators

  (ISM, EMPLOY, IP, RSALES, STARTS)

FORECAST SETUP

MODELS:

- (**LRG**) - all 12 indicators
- (**SML**) - five selected indicators

  (ISM, EMPLOY, IP, RSALES, STARTS)

FORECAST TIMING: FOUR FORECAST ORIGINS

- 1st week of the first month in quarter t, FO1
- 1st week of the second month in quarter t, FO2
- 1st week of the third month in quarter t, FO3
- 1st week of the first month in quarter $t + 1$, FO4

FORECAST SETUP

MODELS:

- (**LRG**) - all 12 indicators
- (**SML**) - five selected indicators

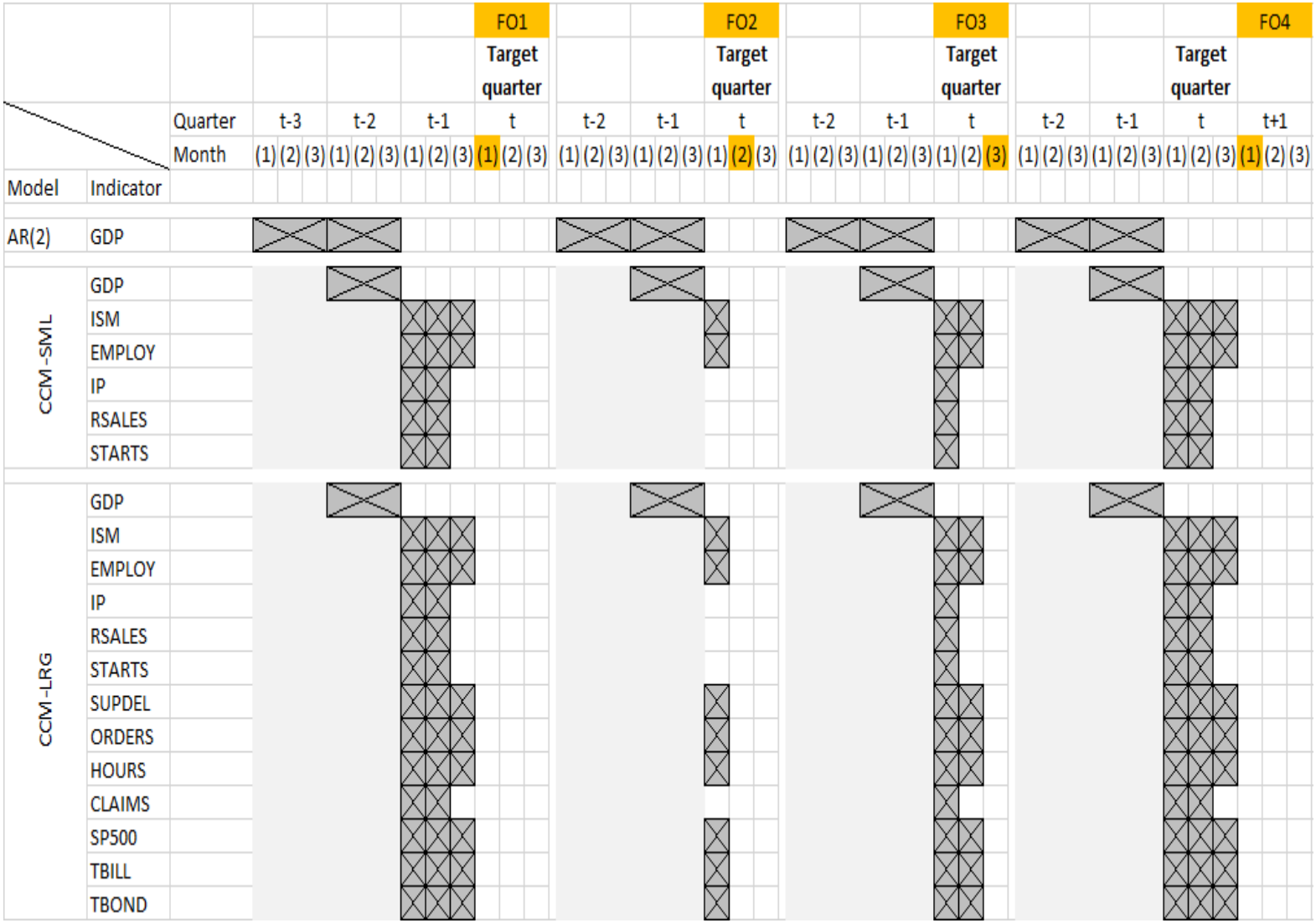
  (ISM, EMPLOY, IP, RSALES, STARTS)

FORECAST TIMING: FOUR FORECAST ORIGINS

- 1st week of the first month in quarter t, FO1
- 1st week of the second month in quarter t, FO2
- 1st week of the third month in quarter t, FO3
- 1st week of the first month in quarter $t + 1$, FO4

ALLOWS ASSESSING INTRAQUARTER INFORMATION FLOW ON FORECAST ACCURACY

Forecasting framework



US APPLICATION: DATA

QUARTERLY GDP GROWTH (SA, ANNUALISED)

- Real-time vintages

12 MONTHLY ECONOMIC/FINANCIAL INDICATORS

- Time-varying variance of the error term
- Captures changes in forecast uncertainty over time

BAYESIAN ESTIMATION

- Point and density forecasts
- Parameter estimation uncertainty is accounted for
- Solution to the curse of dimensionality: shrinkage to zero

FORECAST EVALUATION METRICS

FORECAST EVALUATION (POINT)

RMSFE (negatively oriented)

$$RMSFE = \sqrt{\frac{\sum_{t=\underline{\tau}}^{\overline{\tau}} (y_t - \widehat{y}_t)^2}{T}}$$

RELATIVE RMSFE:

$$RRMSFE_{2/1} = \frac{RMSFE_2 - RMSFE_1}{RMSFE_1}$$

Negative values favour model (2)

CUMULATED SUM OF SQUARED FORECAST ERROR DIFFERENCE

$$CSSFED_{[\underline{\tau},\overline{\tau}],2/1} = \sum_{s=\underline{\tau}}^{t} (e_{2,s}^2 - e_{1,s}^2)$$

for $t \in [\underline{\tau}, \overline{\tau}]$

FORECAST EVALUATION (DENSITY)

MEAN LOG-SCORE (positively oriented)

$$MLogS = \frac{\sum_{t=\underline{\tau}}^{\overline{\tau}} \ln \mathscr{F}_t(y_t)}{T},$$

MEAN LOG-SCORE DIFFERENCE:

$$MLogSD_{2/1} = MLogS_2 - MLogS_1.$$

Positive values favour model (2)

CUMULATED SUM OF LOG-SCORE DIFFERENCE

$$CSLSD_{[\underline{\tau},\overline{\tau}],2/1} =$$

$$= \sum_{s=\underline{\tau}}^{t} (\ln \mathscr{F}_{s,2}(y_s) - \ln \mathscr{F}_{s,1}(y_s))$$

for $t \in [\underline{\tau}, \overline{\tau}]$

Cumulative log predictive Bayes factors (Geweke/Amisano, 2011)

Why use recursive metrics?

Geweke/Amisano (2010, p. 220) state that this way of model comparison "*...shows how individual observations contribute to the evidence in favor of one model over another. For example, it may show that a few observations are pivotal in the evidence strongly favoring one model over another.*"

Why use recursive metrics?

Geweke/Amisano (2010, p. 220) state that this way of model comparison "*...shows how individual observations contribute to the evidence in favor of one model over another. For example, it may show that a few observations are pivotal in the evidence strongly favoring one model over another.*"

THE CSSFED CONTAINS THE SAME INFO ON POINT FORECASTS BUT STILL NOT THAT POPULAR!

Why use recursive metrics?

Geweke/Amisano (2010, p. 220) state that this way of model comparison "*...shows how individual observations contribute to the evidence in favor of one model over another. For example, it may show that a few observations are pivotal in the evidence strongly favoring one model over another.*"

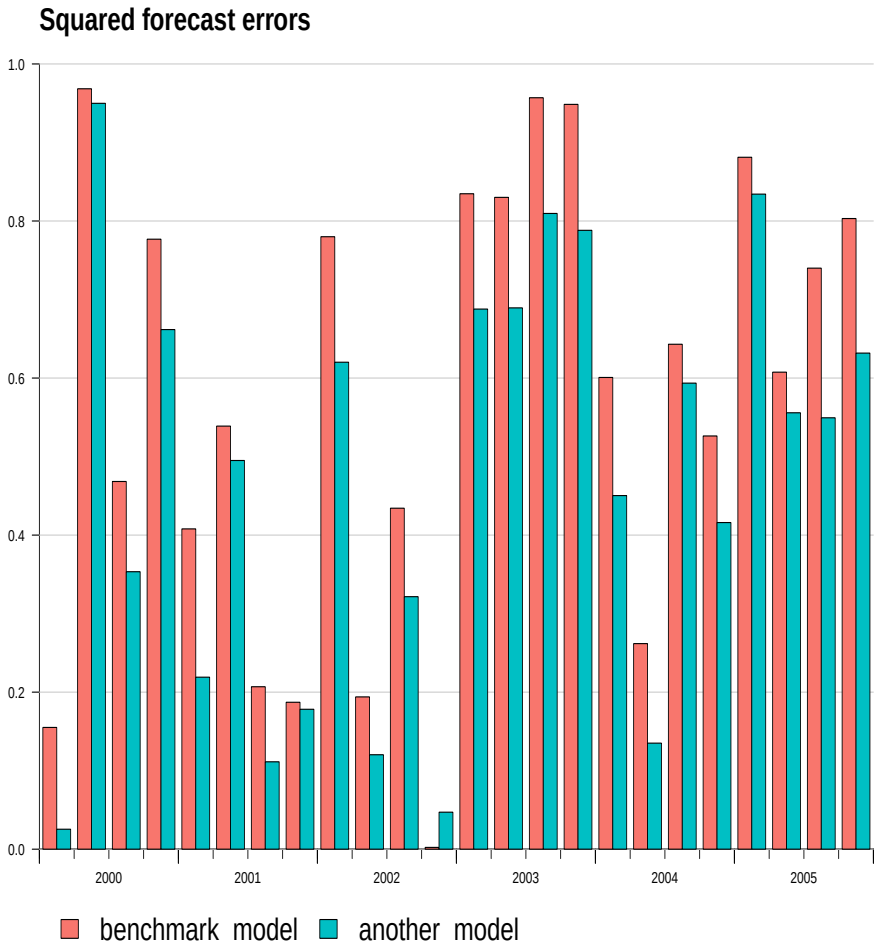
THE CSSFED CONTAINS THE SAME INFO ON POINT FORECASTS BUT STILL NOT THAT POPULAR!

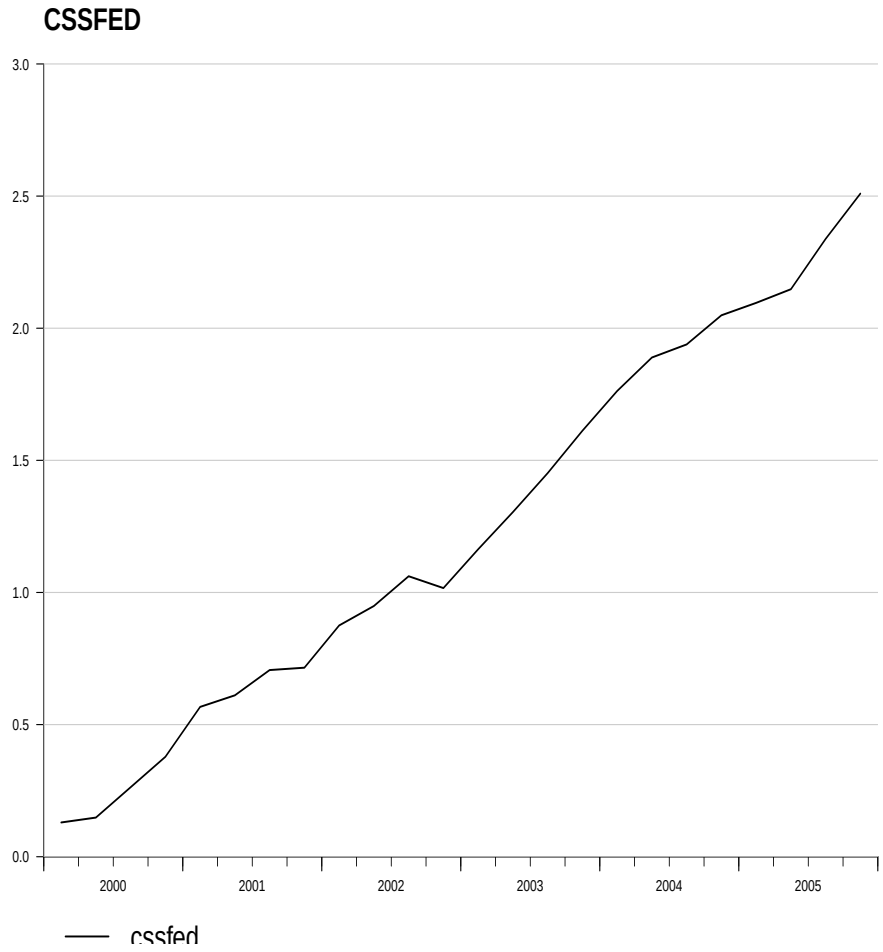
$$CSSFED_{[\underline{\tau},\overline{\tau}],2/1} = \sum_{s=\underline{\tau}}^{t} (e_{2,s}^2 - e_{1,s}^2)$$

INTERESTING PATTERNS:

- upward/downward trends
- horizontal movement
- reversal in trend direction
- jumps

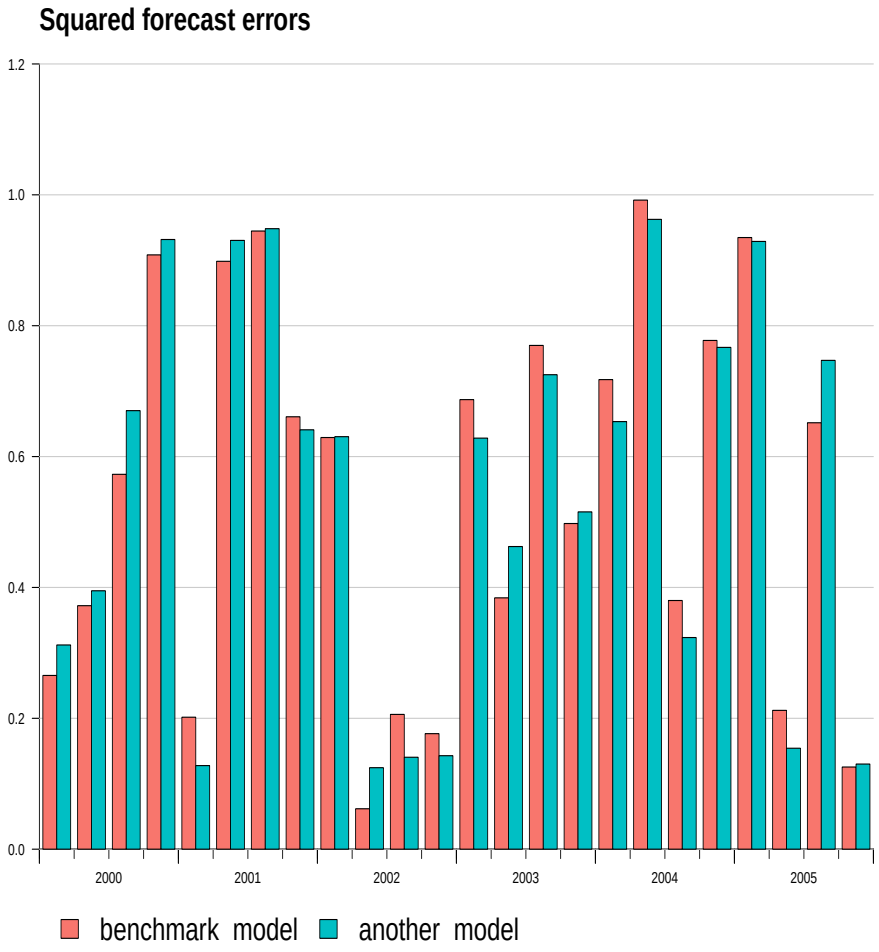
FORECAST EVALUATION: CSSFED

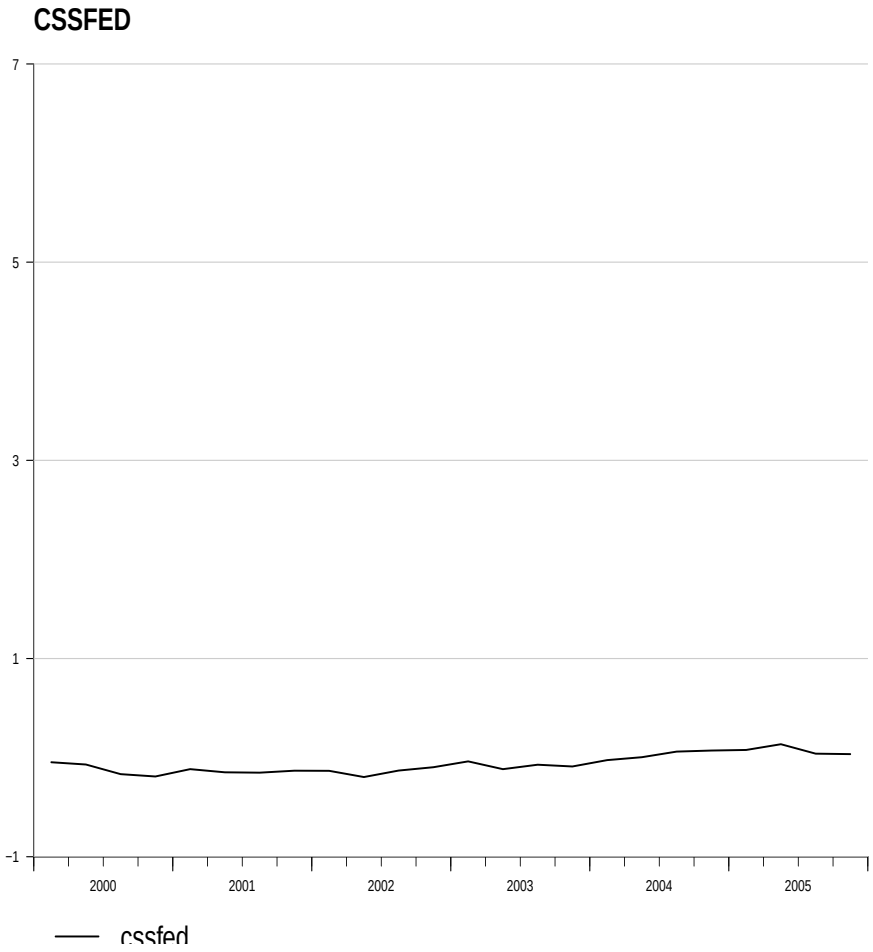




Systematic differences in forecast accuracy
RMSFE(benchmark) = 0.757 and RMSFE(another) = 0.684

FORECAST EVALUATION: CSSFED

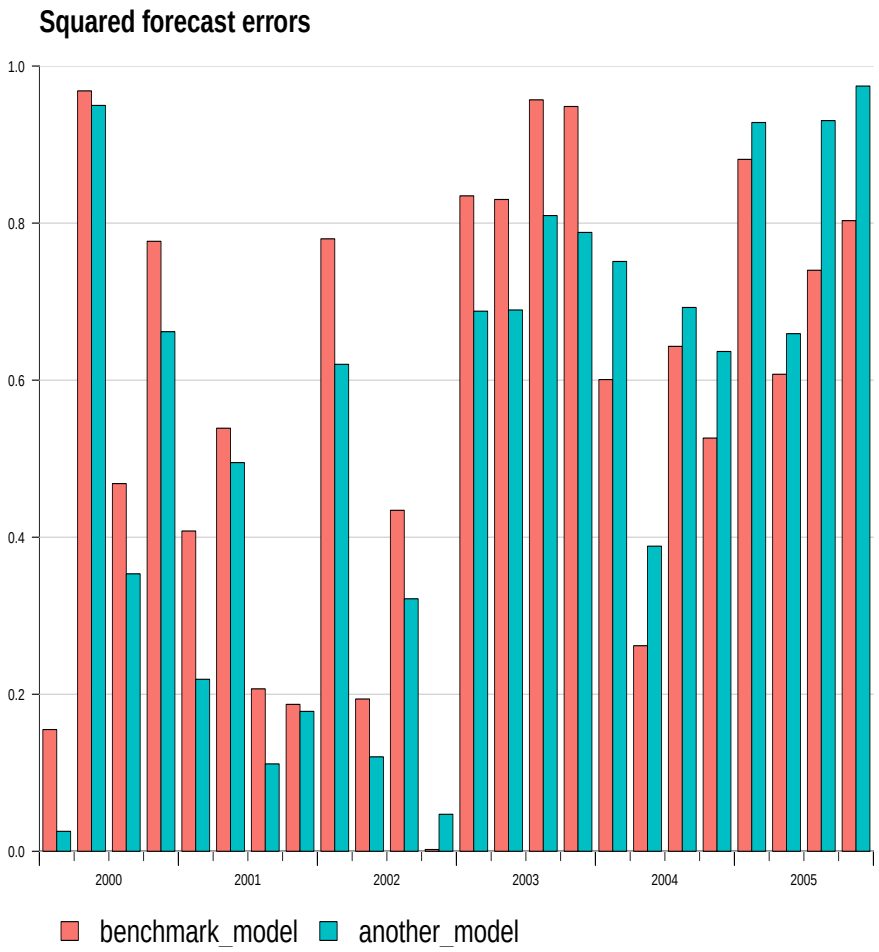


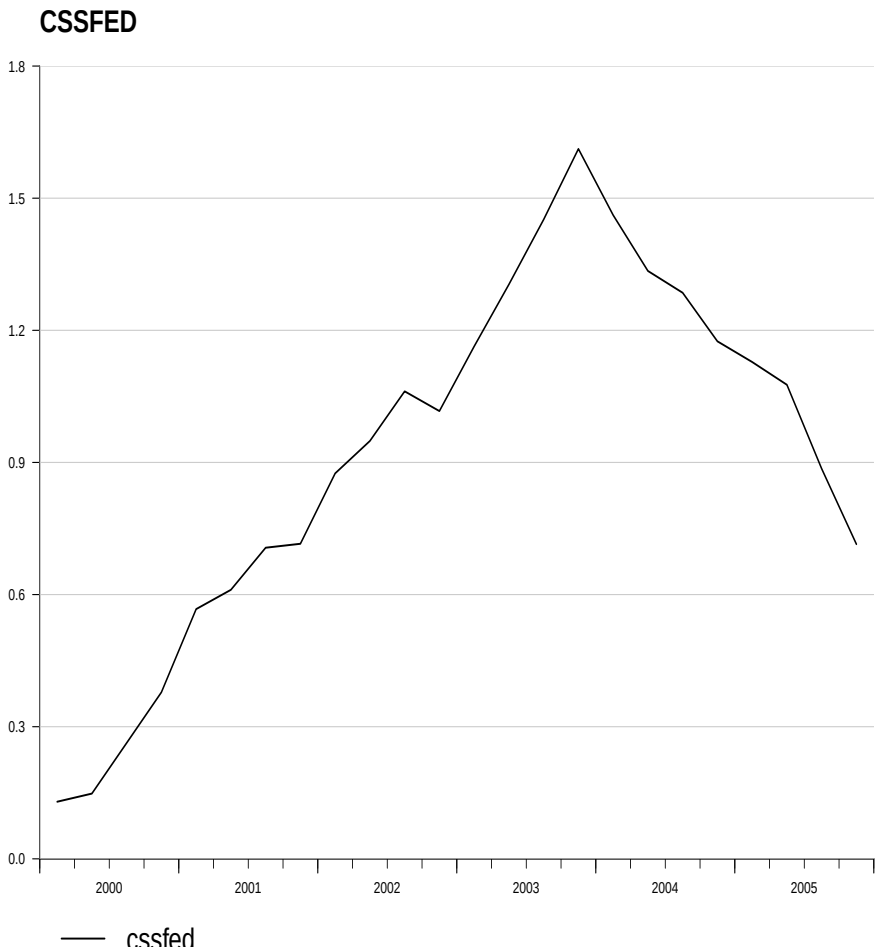


About equal forecast accuracy

RMSFE(benchmark) = 0.737 and RMSFE(another) = 0.736

FORECAST EVALUATION: CSSFED



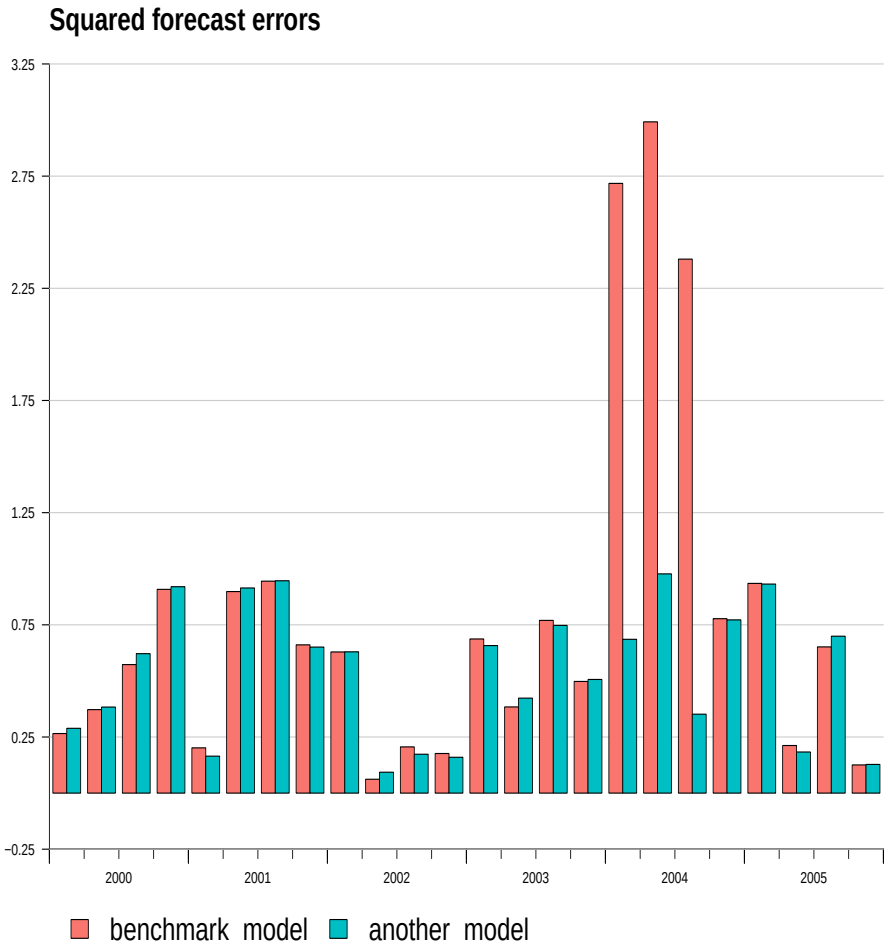


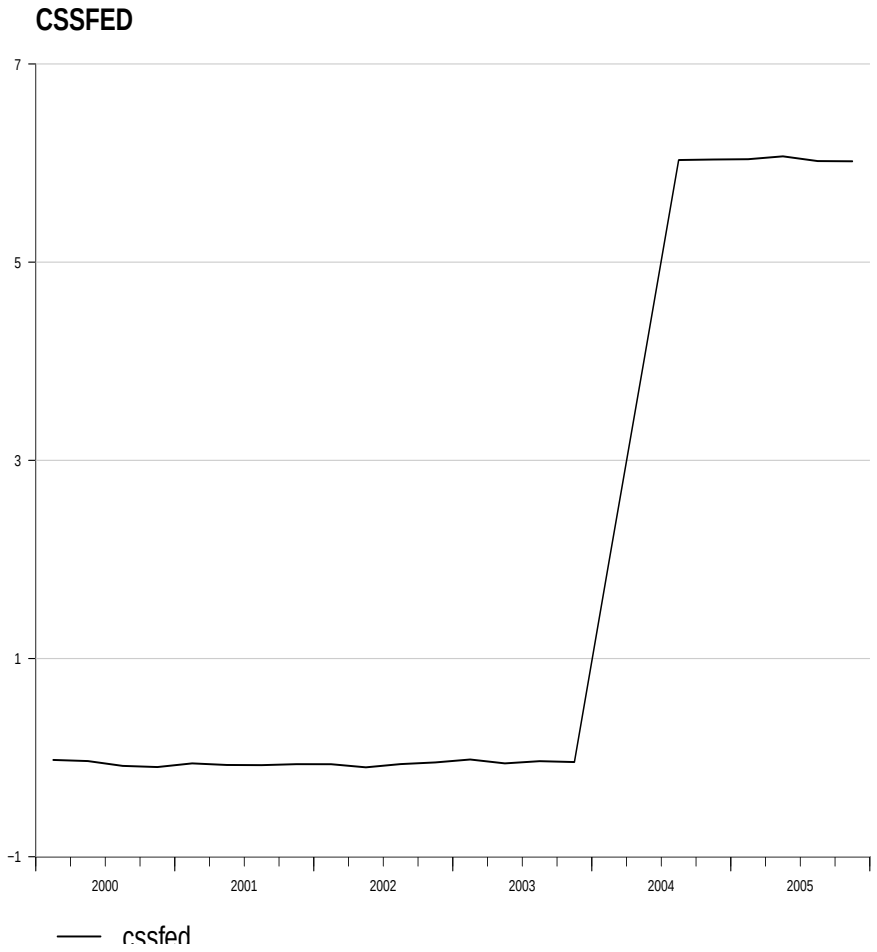
Reversal in relative forecast accuracy

$\Rightarrow$ THE GOOD MODEL TURNS BAD!!!

RMSFE(benchmark) = 0.757 and RMSFE(another) = 0.737

FORECAST EVALUATION: CSSFED





Abrupt (SHORT-LIVED) changes in relative forecast accuracy
RMSFE(benchmark) = 0.89 and RMSFE(another) = 0.736

POINT FORECAST EVALUATION

Table: **Point forecast accuracy**, models of CCM2015: w/o SV

	FO1	FO2	FO3	FO4	FO1	FO2	FO3	FO4
	RMSFE				$RRMSFE_{CCM-XXX/AR2}$			
	Full sample							
AR2	2.237	2.102	2.081	2.074	0.000	0.000	0.000	0.000
CCM-SML	2.082	1.906	1.723	1.605	-0.069	-0.093	-0.172	-0.226
CCM-LRG	2.074	1.831	1.693	1.622	-0.073	-0.129	-0.186	-0.218
	Boom sample							
AR2								
CCM-SML								
CCM-LRG								
	Bust sample							
AR2								
CCM-SML								
CCM-LRG								

Table: **Point forecast accuracy**, models of CCM2015: w/o SV

	FO1	FO2	FO3	FO4	FO1	FO2	FO3	FO4
	RMSFE				$RRMSFE_{CCM-XXX/AR2}$			
	Full sample							
AR2	2.237	2.102	2.081	2.074	0.000	0.000	0.000	0.000
CCM-SML								
CCM-LRG								
	Boom sample							
AR2	1.710	1.705	1.692	1.691	0.000	0.000	0.000	0.000
CCM-SML								
CCM-LRG								
	Bust sample							
AR2	4.339	3.802	3.751	3.726	0.000	0.000	0.000	0.000
CCM-SML								
CCM-LRG								

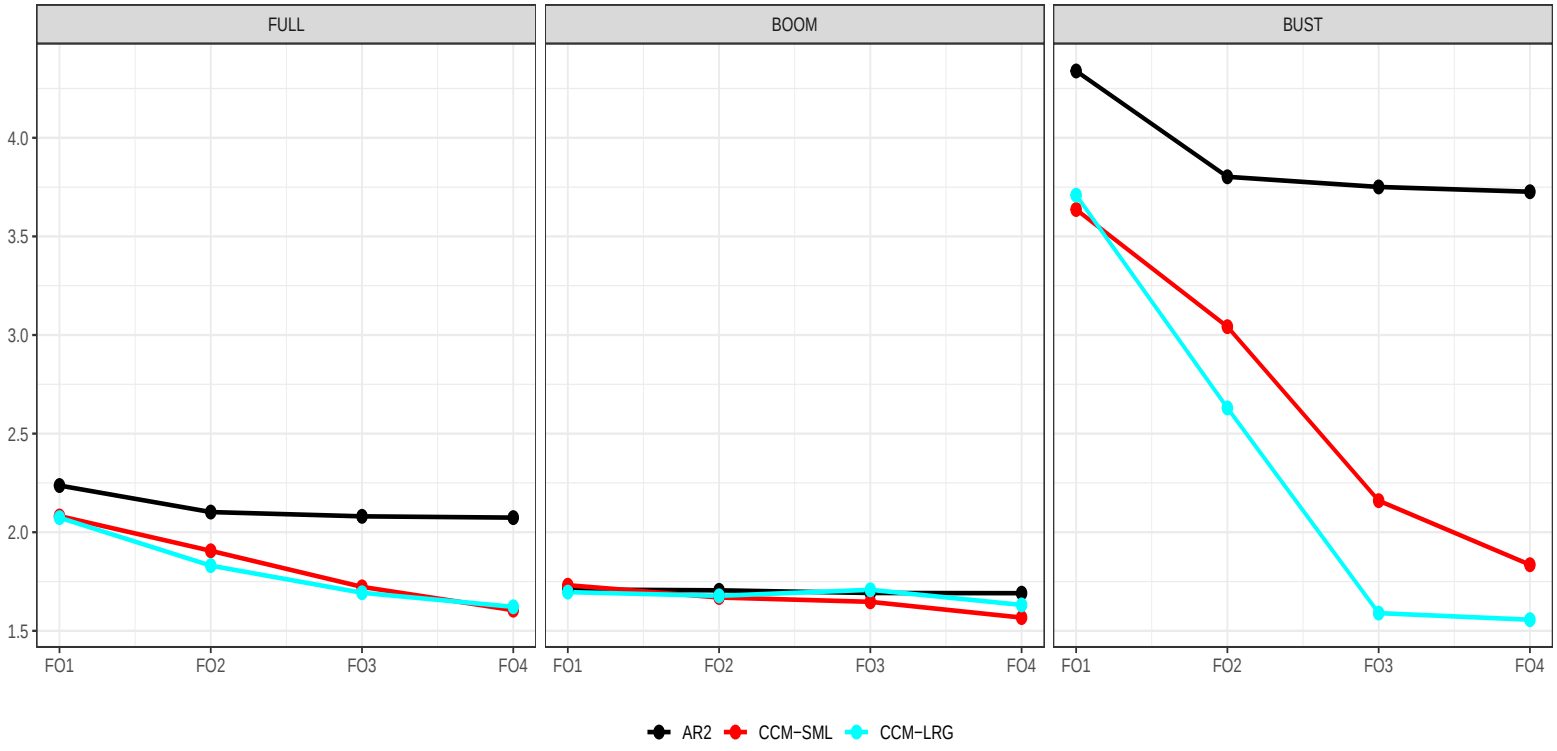
Table: **Point forecast accuracy**, models of CCM2015: w/o SV

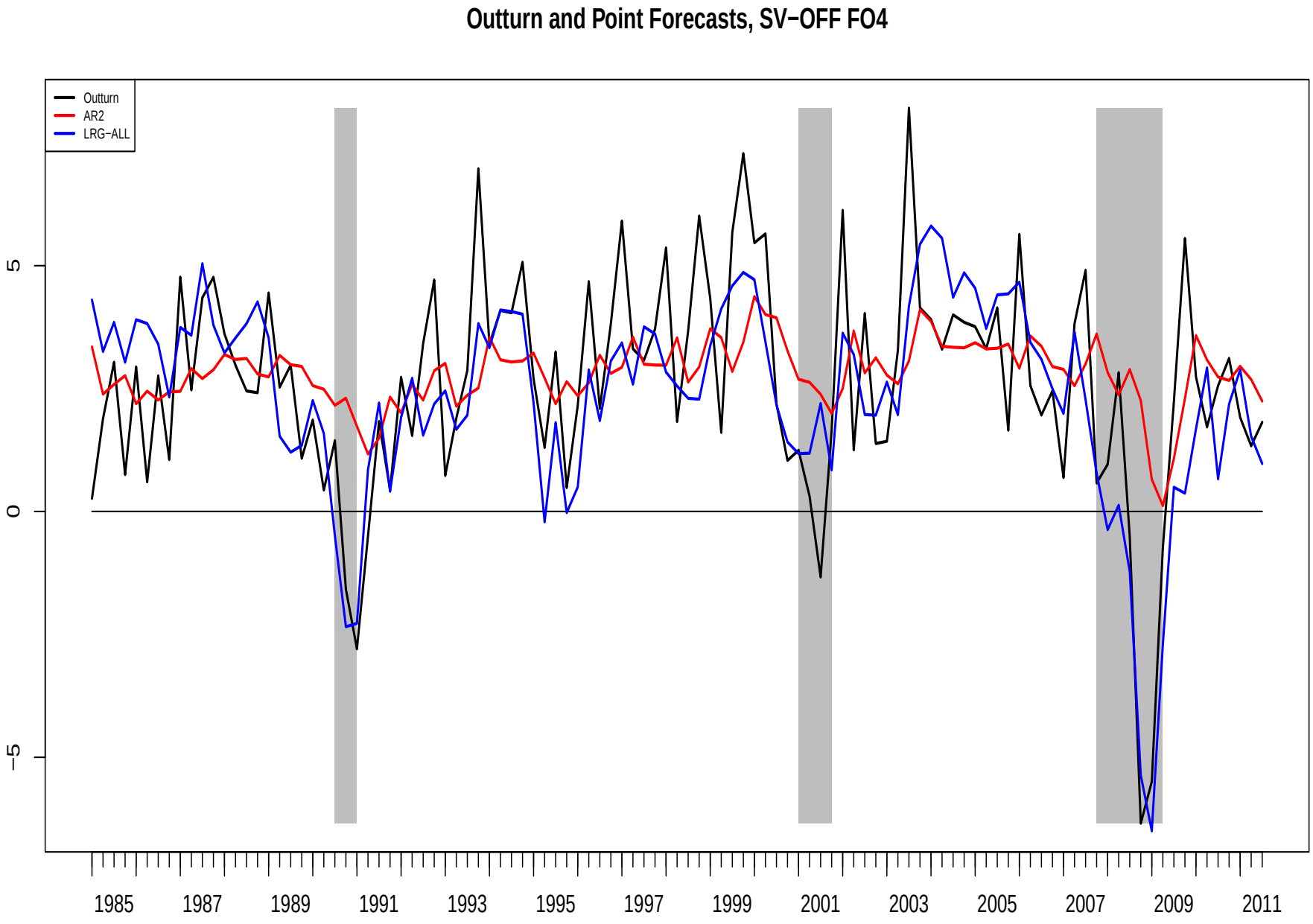
	FO1	FO2	FO3	FO4	FO1	FO2	FO3	FO4
	RMSFE				$RRMSFE_{CCM-XXX/AR2}$			
	Full sample							
AR2								
CCM-SML	2.082	1.906	1.723	1.605	-0.069	-0.093	-0.172	-0.226
CCM-LRG								
	Boom sample							
AR2								
CCM-SML	1.731	1.669	1.647	1.567	0.013	-0.021	-0.027	-0.073
CCM-LRG								
	Bust sample							
AR2								
CCM-SML	3.636	3.042	2.160	1.835	-0.162	-0.200	-0.424	-0.508
CCM-LRG								

Table: **Point forecast accuracy**, models of CCM2015: w/o SV

	FO1	FO2	FO3	FO4	FO1	FO2	FO3	FO4
	RMSFE				$RRMSFE_{CCM-XXX/AR2}$			
	Full sample							
AR2								
CCM-SML								
CCM-LRG	2.074	1.831	1.693	1.622	-0.073	-0.129	-0.186	-0.218
	Boom sample							
AR2								
CCM-SML								
CCM-LRG	1.696	1.678	1.708	1.631	-0.008	-0.016	0.009	-0.035
	Bust sample							
AR2								
CCM-SML								
CCM-LRG	3.709	2.630	1.590	1.556	-0.145	-0.308	-0.576	-0.582

RMSFE


FULL
BOOM
BUST
4.0
3.5
3.0
2.5
2.0
1.5
FO1
FO2
FO3
FO4
AR2
CCM-SML
CCM-LRG


Outturn and Point Forecasts, SV−OFF FO4
Outturn
AR2
LRG−ALL
5
0
−5
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011

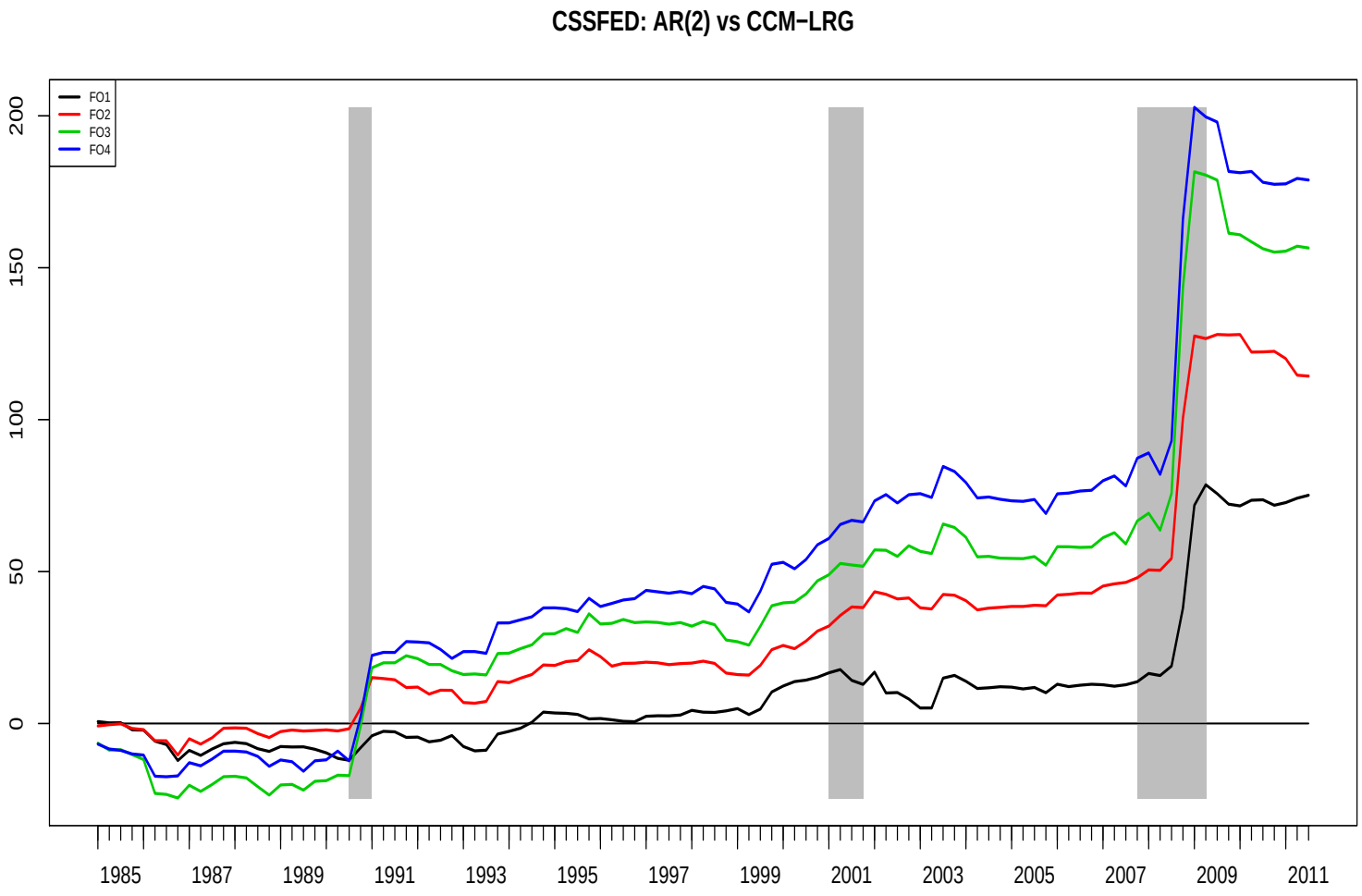


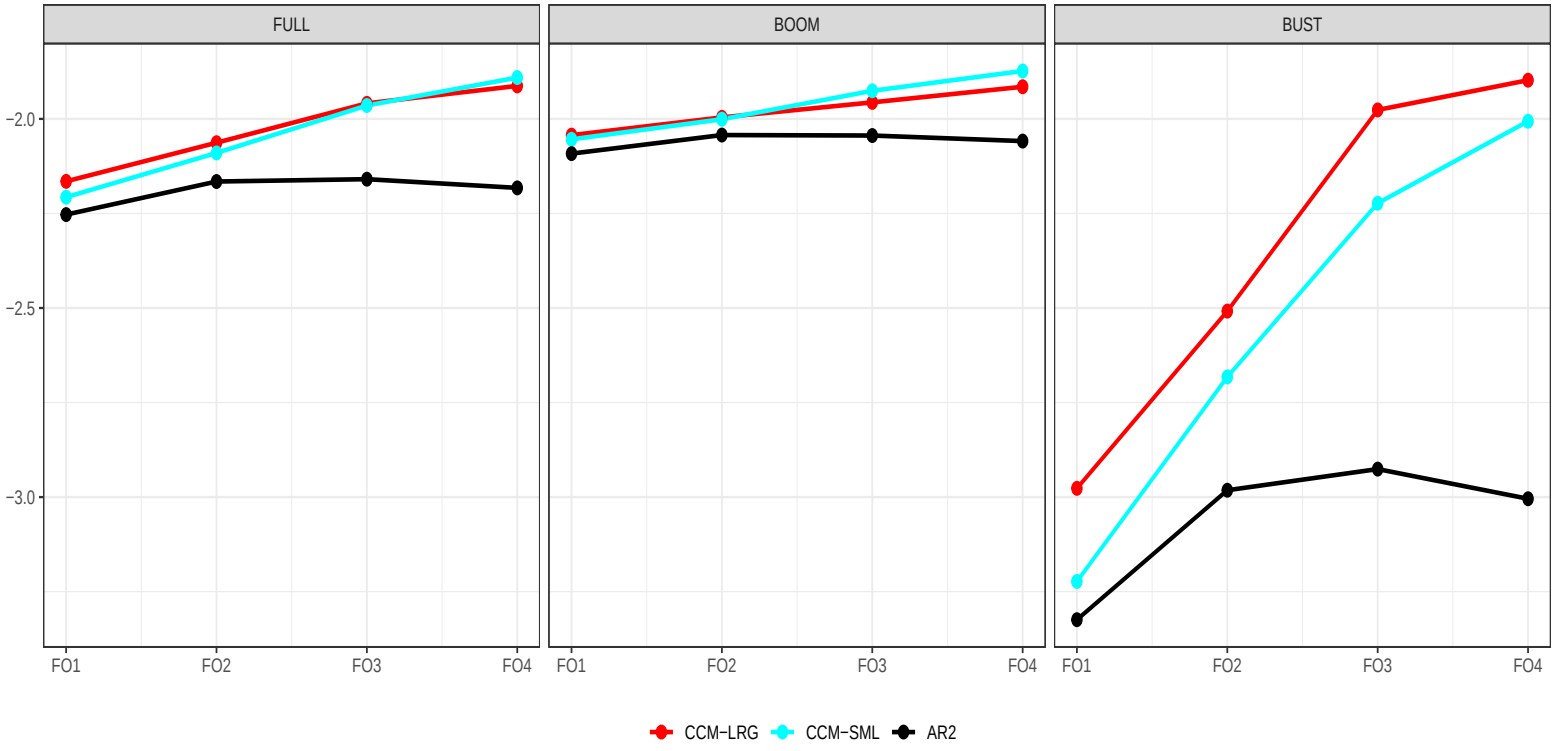
Figure: $\text{CCSFED}_{AR(2)/CCM-LRG}$, w/o SV

DENSITY FORECAST EVALUATION

Table: Density forecast accuracy, w/ SV

	FO1	FO2	FO3	FO4	FO1	FO2	FO3	FO4
	MLogS				MLogSD$_{CCM-XXX-SV/AR2-SV}$			
	Full sample							
AR2-SV	-2.253	-2.166	-2.159	-2.182	0.000	0.000	0.000	0.000
CCM-SML-SV	-2.207	-2.090	-1.965	-1.891	0.046	0.076	0.195	0.292
CCM-LRG-SV	-2.165	-2.063	-1.959	-1.913	0.088	0.103	0.200	0.270
	Boom sample							
AR2-SV	-2.092	-2.043	-2.044	-2.059	0.000	0.000	0.000	0.000
CCM-SML-SV	-2.054	-2.001	-1.926	-1.873	0.038	0.042	0.118	0.185
CCM-LRG-SV	-2.043	-1.996	-1.956	-1.915	0.049	0.047	0.088	0.144
	Bust sample							
AR2-SV	-3.324	-2.982	-2.926	-3.004	0.000	0.000	0.000	0.000
CCM-SML-SV	-3.223	-2.682	-2.223	-2.006	0.101	0.300	0.703	0.998
CCM-LRG-SV	-2.977	-2.508	-1.976	-1.898	0.347	0.474	0.950	1.107

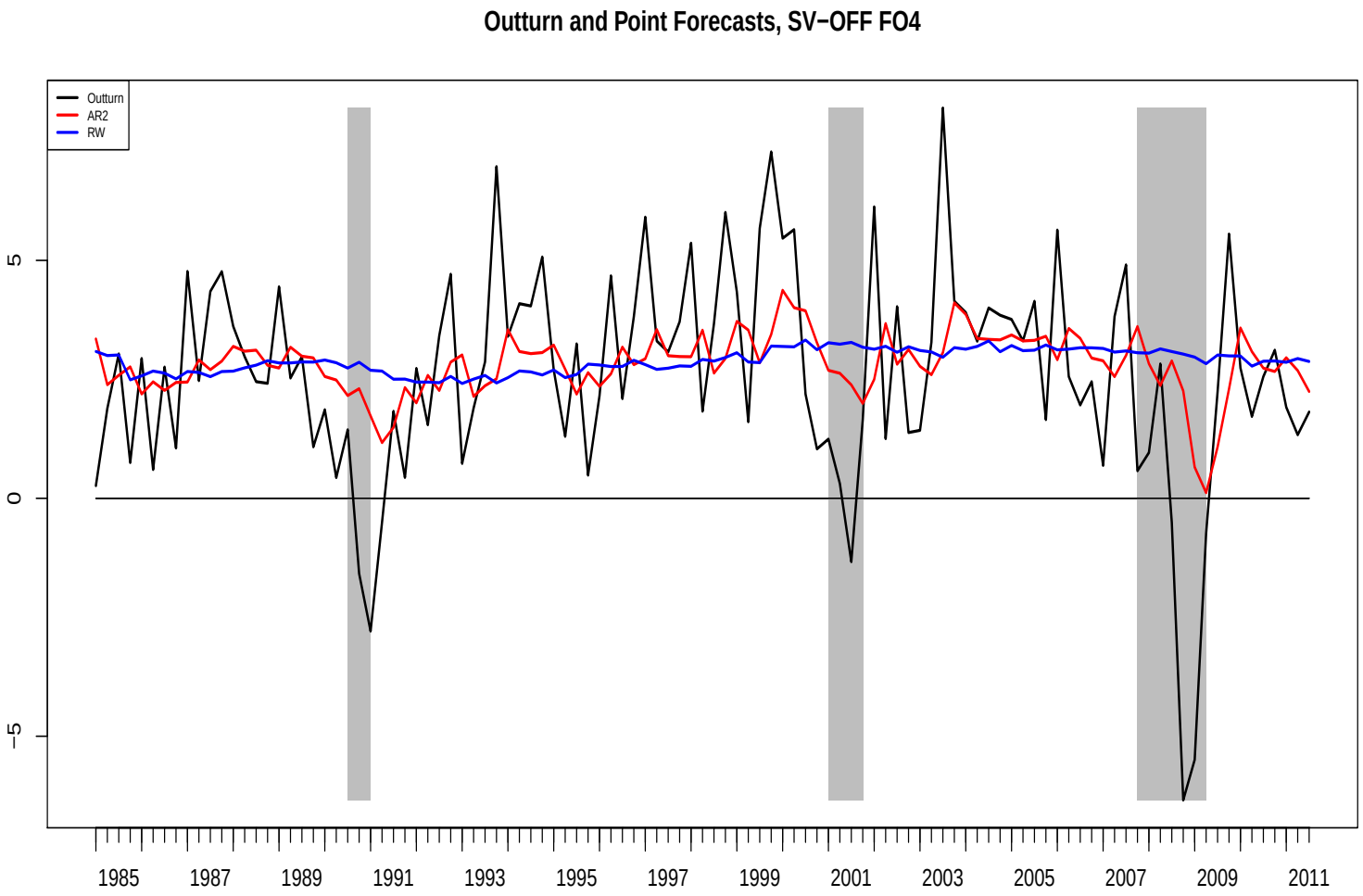
MLogS

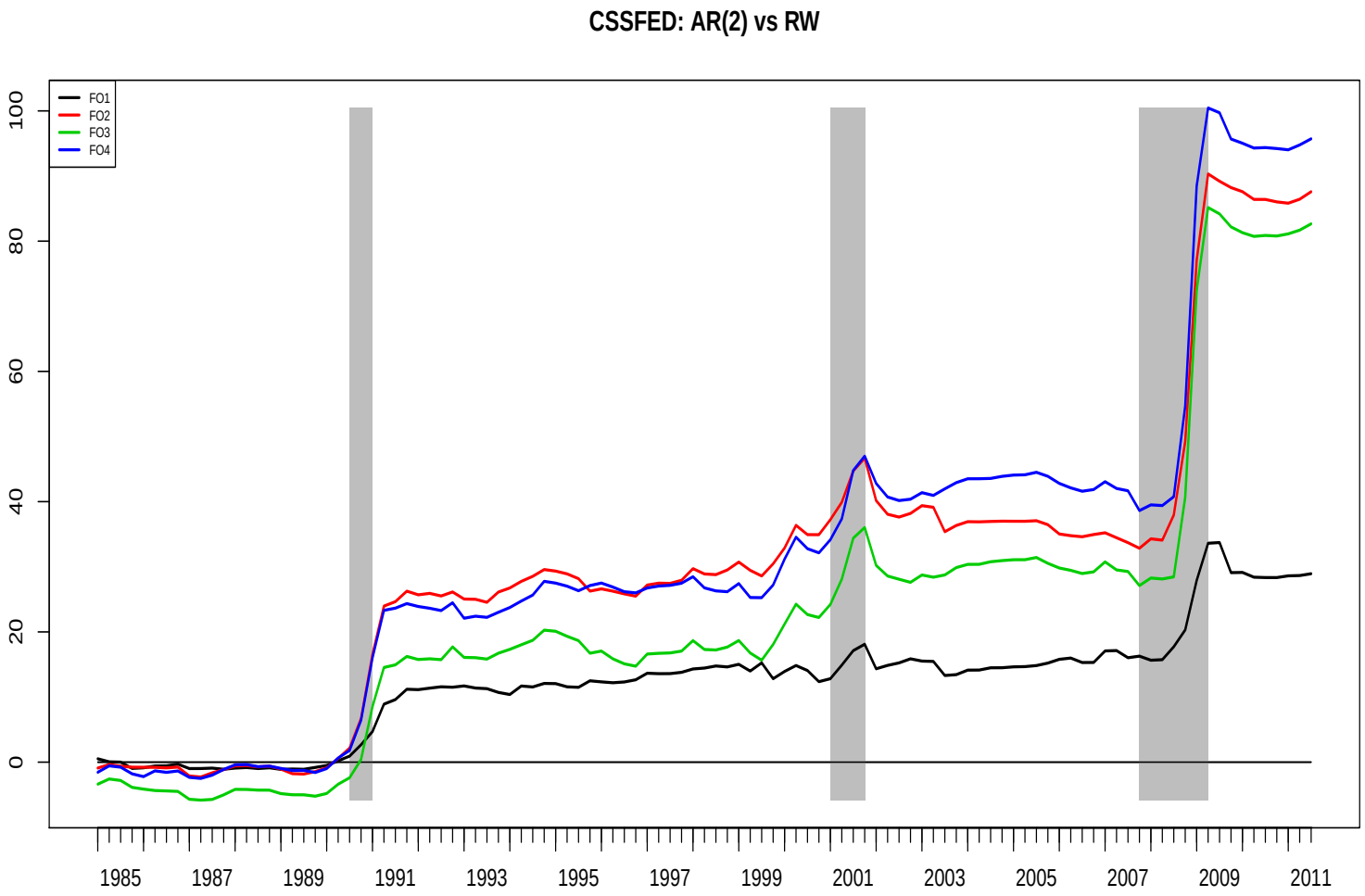


CHOICE OF A BENCHMARK MODEL

Table: Point forecast accuracy, univariate benchmark models: w/o SV

	FO1	FO2	FO3	FO4	FO1	FO2	FO3	FO4
	RMSFE				$\text{RRMSFE}_{RW\,/\,AR(2)}$			
	Full sample							
AR(2)	2.237	2.102	2.081	2.074	0.000	0.000	0.000	0.000
RW	2.296	2.289	2.259	2.280	0.027	0.089	0.086	0.099
	Boom sample							
AR(2)	1.710	1.705	1.692	1.691	0.000	0.000	0.000	0.000
RW	1.713	1.716	1.695	1.712	0.002	0.006	0.002	0.013
	Bust sample							
AR(2)	4.339	3.802	3.751	3.726	0.000	0.000	0.000	0.000
RW	4.562	4.526	4.461	4.499	0.051	0.190	0.189	0.207


Outturn and Point Forecasts, SV−OFF FO4
Outturn
AR2
RW
5
0
−5
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011


CSSFED: AR(2) vs RW
FO1
FO2
FO3
FO4
100
80
60
40
20
0
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011

TAKE-AWAY MESSAGE

We document **asymmetries** in forecasting performance of indicator-augmented models across US business cycle phases

- in relative magnitude wrt the univariate benchmark models
- (and often) in absolute magnitude

Ignoring the boom/bust asymmetries leads to

- **exaggeration** of forecasting performance of sophisiticated models in expansions
- **understating** their relative forecasting performance in recessions

TAKE-AWAY MESSAGE

In expansions a historical mean (or AR(2)) model is as **good** as sophisticated models

TAKE-AWAY MESSAGE

In expansions a historical mean (or AR(2)) model is as **good** as sophisticated models

Chauvet/Potter (2013): "... by using a simple univariate ... linear autoregressive model of GDP growth, one would have gotten in real time as good as forecasts during expansions than any other model and the professional forecasters. These simple models exhibit a very good ability to track the mean of GDP growth during normal times."

TAKE-AWAY MESSAGE

In expansions a historical mean (or AR(2)) model is as **good** as sophisticated models

Chauvet/Potter (2013): "... by using a simple univariate ... linear autoregressive model of GDP growth, one would have gotten in real time as good as forecasts during expansions than any other model and the professional forecasters. These simple models exhibit a very good ability to track the mean of GDP growth during normal times."

Improvement in forecast accuracy over the benchmark models attributed to sophisticated models for the **full sample** is due to:

- Significant worsening of forecasting ability of benchmark models during recessions compared to expansions
- Drop in forecasting ability of benchmark models during recessions is *more pronounced* than that of sophisticated models