Commodity prices forecasting using autoregressive nonlinear models

Zuzanna Karolak

Warsaw School of Economics

15.09.2017

Plan

1. Introduction and motivation

Plan

1. Introduction and motivation
2. Round-up of forecasting methods

Plan

1. Introduction and motivation
2. Round-up of forecasting methods
3. Data and competition design

Plan

1. Introduction and motivation
2. Round-up of forecasting methods
3. Data and competition design
4. Results

Plan

1. Introduction and motivation
2. Round-up of forecasting methods
3. Data and competition design
4. Results
5. Conclusions

Introduction and motivation

More accurate forecasts could:

- help to pursue political and economic goals at the country level
- support investment and business decisions at the firm level

Introduction and motivation

More accurate forecasts could:

- help to pursue political and economic goals at the country level
- support investment and business decisions at the firm level

However

- are commodity prices forecastable?
- can we consistently beat no-change forecast?

Introduction and motivation

More accurate forecasts could:

- help to pursue political and economic goals at the country level
- support investment and business decisions at the firm level

However

- are commodity prices forecastable?
- can we consistently beat no-change forecast?

What if ...

- we allow for nonlinearities in the forecasting model?

Introduction and motivation

Aim of the study:

- understand the price dynamics of several primary commodities
- verify whether nonlinear methods help to improve forecast accuracy for commodity prices

Introduction and motivation

Aim of the study:

- understand the price dynamics of several primary commodities
- verify whether nonlinear methods help to improve forecast accuracy for commodity prices

Contribution:

- more comprehensive study as more models and commodities are covered

Round-up of forecasting methods

Competitiors:

- threshold model (self-exciting threshold autoregression model - *SETAR*)
- transition model (logistic smooth transition model - *LSTAR*)
- autoregressive Markov regime switching models ($MS - AR$)
- autoregressive feedforward neural network ($AR - NN$)

Round-up of forecasting methods

Competitors:

- threshold model (self-exciting threshold autoregression model - *SETAR*)
- transition model (logistic smooth transition model - *LSTAR*)
- autoregressive Markov regime switching models ($MS - AR$)
- autoregressive feedforward neural network ($AR - NN$)

Benchmarks:

- random walk
- ARIMA class model

Two regime $SETAR(p_1, p_2, d)$ model with delay d and autoregressive parameters p_1, p_2 is defined as follows (Franses and Dijk, 2000):

$$\begin{aligned} y_t = & (\phi_{0,1} + \phi_{1,1} y_{t-1} + ... + \phi_{p_1,1} y_{t-p_1}) I[y_{t-d} \leq c] + \\ & (\phi_{0,2} + \phi_{1,2} y_{t-1} + ... + \phi_{p_1,2} y_{t-p_2}) I[y_{t-d} > c] + \epsilon_t \end{aligned} \quad (1)$$

where c is the value of threshold parameter.

Two regime $SETAR(p_1, p_2, d)$ model with delay d and autoregressive parameters p_1, p_2 is defined as follows (Franses and Dijk, 2000):

$$\begin{aligned} y_t = & (\phi_{0,1} + \phi_{1,1} y_{t-1} + \dots + \phi_{p_1,1} y_{t-p_1}) I[y_{t-d} \leq c] + \\ & (\phi_{0,2} + \phi_{1,2} y_{t-1} + \dots + \phi_{p_1,2} y_{t-p_2}) I[y_{t-d} > c] + \epsilon_t \end{aligned} \tag{1}$$

where c is the value of threshold parameter.

Two regime $LSTAR(p_1, p_2, d)$ model with delay d and autoregressive parameters p_1, p_2 is specified as follows (Franses and Dijk, 2000):

$$\begin{aligned} y_t = & (\phi_{0,1} + \phi_{1,1} y_{t-1} + \dots + \phi_{p_1,1} y_{t-p_1})(1 - G(y_{t-d}, \gamma, c)) + \\ & (\phi_{0,2} + \phi_{1,2} y_{t-1} + \dots + \phi_{p_1,2} y_{t-p_2}) G(y_{t-d}, \gamma, c) + \epsilon_t \end{aligned} \tag{2}$$

where $G()$ is a logistic transition function.

The $MS-AR(s)$ model, where s is a number of lags, is described as follows Franses and Dijk, 2000):

$$y_t = \sum_{k=1}^{K} S_{tk} (\sum_{s=1}^{p} \phi_{k,s} t_{t-s} + \epsilon_t) \tag{3}$$

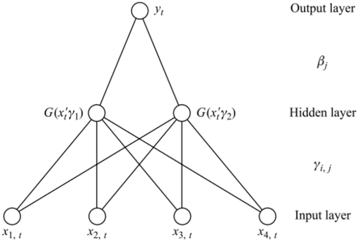
where $S_{tk} = 1$ if if the state variable $Q_t = k$ and $S_{tk} = 0$ otherwise.

The distribution of the hidden state process is given by the transition probability matrix: $p_{jk} = P(Q_t = k | Q_{t-1} = j)$

Single hidden layer feedforward specification with D hidden units and p autoregressive lags is described as follows (Franses and Dijk, 2000):

$$y_t = \phi_0 + \sum h = 1^D \beta_j G(\gamma_{0,h}, ..., \gamma_{p,h}, y_{t-1}, ..., y_{t-p}) + \epsilon_t, \qquad (4)$$

where $G()$ is an activation function and parameters γ are connection weights between the $h - th$ hidden unit with all input units.



Data and competition design

1. Three groups of commodities are considered:
   - energy commodities: Brent crude oil, WTI crude oil, natural gas, coal,
   - metals: gold, silver, aluminum, platinum, zinc, copper,
   - agriculture goods: coffee, maize, wheat, soybeans, sugar.

Data and competition design

1. Three groups of commodities are considered:
   - energy commodities: Brent crude oil, WTI crude oil, natural gas, coal,
   - metals: gold, silver, aluminum, platinum, zinc, copper,
   - agriculture goods: coffee, maize, wheat, soybeans, sugar.
2. Recursive forecasting scheme first vintage 15 years of data from the period 1987:5 2002:4

Data and competition design

1. Three groups of commodities are considered:
   - energy commodities: Brent crude oil, WTI crude oil, natural gas, coal,
   - metals: gold, silver, aluminum, platinum, zinc, copper,
   - agriculture goods: coffee, maize, wheat, soybeans, sugar.
2. Recursive forecasting scheme first vintage 15 years of data from the period 1987:5 2002:4
3. Forecast horizon for each sample from 1 to 12 months

Data and competition design

1. Three groups of commodities are considered:
    - energy commodities: Brent crude oil, WTI crude oil, natural gas, coal,
    - metals: gold, silver, aluminum, platinum, zinc, copper,
    - agriculture goods: coffee, maize, wheat, soybeans, sugar.
2. Recursive forecasting scheme first vintage 15 years of data from the period 1987:5 2002:4
3. Forecast horizon for each sample from 1 to 12 months
4. Accuracy measure RMSE

Data and competition design

1. Three groups of commodities are considered:
   - energy commodities: Brent crude oil, WTI crude oil, natural gas, coal,
   - metals: gold, silver, aluminum, platinum, zinc, copper,
   - agriculture goods: coffee, maize, wheat, soybeans, sugar.
2. Recursive forecasting scheme first vintage 15 years of data from the period 1987:5 2002:4
3. Forecast horizon for each sample from 1 to 12 months
4. Accuracy measure RMSE
5. Two different specifications for each method

Summary statistics and nonlinearity test

	mean	median	minimum	maximum	variance	kurtosis	skewness	BDS test p-value
Brent oil	0,001	0,004	-0,296	0,450	0,008	2,218	-0,114	0,000
Wti oil	0,001	0,006	-0,314	0,366	0,007	1,801	-0,239	0,000
Gas	0,000	-0,003	-0,407	0,478	0,017	1,198	0,033	0,004
Coal	0,001	-0,002	-0,318	0,361	0,003	7,997	0,368	0,000
Gold	0,001	-0,002	-0,127	0,158	0,001	1,519	0,391	0,002
Silver	0,000	-0,005	-0,216	0,186	0,004	1,255	-0,032	0,000
Aluminum	-0,001	-0,003	-0,330	0,176	0,003	3,754	-0,597	0,000
Platinum	-0,001	0,002	-0,283	0,231	0,003	5,152	-0,641	0,000
Zinc	0,001	0,001	-0,277	0,236	0,004	1,655	-0,435	0,001
Copper	0,002	0,004	-0,340	0,248	0,004	3,364	-0,265	0,000
Sugar	0,000	0,002	-0,312	0,279	0,006	0,787	0,112	0,000
Wheat	-0,001	-0,003	-0,209	0,228	0,004	1,899	0,457	0,000
Soybeans	-0,001	0,001	-0,246	0,210	0,003	2,753	-0,210	0,004
Coffee	-0,002	-0,008	-0,252	0,375	0,004	3,858	0,651	0,012
Maize	0,000	0,000	-0,247	0,293	0,003	3,533	-0,011	0,013

The BDS test was developed by Brock, Dechert and Scheinkman (1987). The null hypothesis assumes independent and identical distribution (iid) againsed a nonlinear structure.

Results
Energy commodities

h	RW RMSE	ARIMA(1,0,1)	SETAR(2,2,2)	SETAR(1,1,1)	LSTAR(2,2,2)	LSTAR(1,1,1)	MS-AR(1)	MS-AR(2)	NN (2,1,6)	NN (1,1,4)
Oil Brent										
1	0,090	0,984	0,997	0,993	1,004	0,991	0,978	0,99	1,058	0,996
2	0,143	**0,994***	1,013	1	1,025	1,006	1,039	1,11	1,049	1,007
3	0,186	**0,999****	1,03	1,01***	1,038	1,013	1,105	1,231	1,098	1,002
6	0,273	1,011***	1,045	1,036***	1,051	1,039	1,217	1,608*	1,271	1,034
9	0,322	1,021***	1,039	1,068***	1,049	1,056	1,417	1,971**	1,355	1,067
12	0,355	1,021***	1,046	1,095***	1,046	1,058	1,544	2,304**	1,336	1,074
Oil WTI										
1	0,088	0,964	0,99	0,969	0,983	1,492	0,963	8,958***	1,02	1,151
2	0,142	0,976	1,005	0,985	1,032	1,225	1,032	5,528***	1,043	1,267
3	0,188	0,998	1,068	1,024**	1,13	1,1	1,13	4,232**	1,142	1,45
6	0,275	1,016*	1,177	1,086***	1,478	1,044**	1,308	3,056	1,368	1,878
9	0,321	1,028***	1,24	1,152***	2,019	1,073*	1,515	2,843	1,575	2,247
12	0,349	1,026***	1,305	1,22***	3,157	1,085*	1,66	2,852	1,732	2,64
Gas										
1	0,127	1,016	1,049	1,016	1,039	1,019	1,006	1,048	1,135	1,092
2	0,183	1,01**	1,022	1,019	1,031	1,013**	1,007	1,028	1,123	1,086
3	0,225	1,011***	1,022	1,024**	1,019**	1,018***	1,033	1,078	1,14	1,064
6	0,321	1,014***	1,024*	1,039***	1,032***	1,024***	1,087	1,15	1,154***	1,092*
9	0,384	1,026***	1,067**	1,097***	1,062***	1,054***	1,206	1,301	1,17***	1,097***
12	0,403	1,033**	1,115*	1,148***	1,098***	1,081***	1,328	1,457	1,221***	1,084***
Coal										
1	0,071	0,972	**0,982****	**0,958****	0,993	**0,964***	**0,964***	0,975	1,127	1,016
2	0,116	1	**0,999****	0,949	1,009	0,973	0,96	0,988	1,22	1,068
3	0,156	1,026	0,996	0,961	1,058	0,98	1	1,017	1,312	1,112**
6	0,252	1,078	1,031	1,032	1,306	1,012	1,065	1,098	1,552**	1,241**
9	0,330	1,112	1,077	1,087	1,95	1,032	1,16	1,199	1,741***	1,446**
12	0,380	1,113	1,09	1,111	3,519	1,04*	1,248	1,294*	1,867***	1,662*

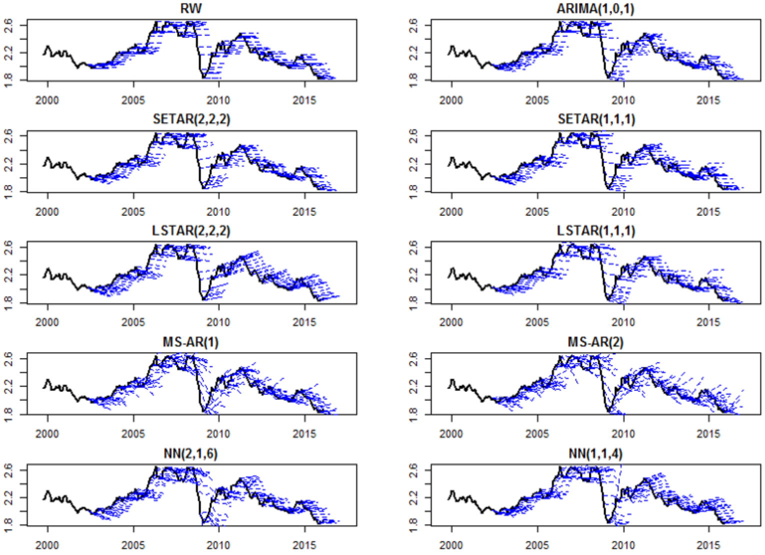
The table shows the ratios of the RMSE from a given model in comparison to the RW benchmark. Asterisks ***, ** and * denote,the significance levels of the two-tailed Diebold-Mariano test.

Results
Metal commodities

	h	RW RMSE	ARIMA(1,0,1)	SETAR(2,2,2)	SETAR(1,1,1)	LSTAR(2,2,2)	LSTAR(1,1,1)	MS-AR(1)	MS-AR(2)	NN (2,1,6)	NN (1,1,4)
Gold											
	1	0,04	1,002	1,018	1,005	1,033	1,013	1,005***	1,198**	1,059	1,107
	3	0,075	1,021	1,044	1,031	1,05	1,012***	1,102***	1,244***	1,083**	1,081**
	6	0,115	1,039	1,059*	1,039	1,058	1,007***	1,202***	1,395**	1,046***	1,045**
	9	0,145	1,057	1,083*	1,087	1,076	1,023***	1,265***	1,615**	1,047**	1,049**
	12	0,17	1,073	1,117	1,114	1,099	1,032**	1,325***	1,783**	1,063**	1,059**
Silver											
	1	0,074	0,981	0,996	1,002**	1,008	1	**0,988*** **	1,007***	1,082	1,302
	3	0,143	1,012**	1,039**	1,026***	1,053	1,008***	1,026	1,038	1,067	1,973
	6	0,212	1,024*	1,032***	1,064***	1,036	1,026***	1,087	1,07	1,048	2,769
	9	0,264	1,038	1,037**	1,104***	1,045	1,049***	1,132	1,112	1,034	3,326
	12	0,301	1,052	1,041**	1,147***	1,058	1,071***	1,187	1,148	1,039	3,852
Aluminum											
	1	0,052	**0,982***	0,967	0,985	0,99	0,979	0,992	1,013	1,102	1,017
	3	0,108	0,994	0,979	**0,999*** **	0,993	1,002	0,984	1,014	1,088	1,048
	6	0,173	1,018**	1,02	1,005***	1,017*	1,023	1,021	1,047	1,133*	1,09
	9	0,208	1,028***	1,043*	1,007***	1,033***	1,047	1,079	1,123*	1,178**	1,11**
	12	0,224	1,038***	1,066	1,014***	1,054***	1,074	1,144*	1,187***	1,201**	1,117**
Platinum											
	1	0,058	0,978	0,996	0,963	NA	NA	1,006	0,943	1,149	0,983
	3	0,126	1,023	1,108	0,994	NA	NA	1,268	1,143	1,71	1,459
	6	0,193	1,038	1,191***	1,007***	NA	NA	1,688	1,679	2,293*	1,847
	9	0,224	1,044	1,234***	1,017***	NA	NA	2,228	2,183***	2,446***	2,206***
	12	0,241	1,053	1,277***	1,039***	NA	NA	2,649	2,667***	2,346***	2,082***
Zinc											
	1	0,071	0,959	0,965	0,962	NA	0,978	1,096***	10,331***	0,984	1,02
	3	0,151	0,986	0,988	0,985***	NA	**0,988*** **	1,226***	5,034***	0,987	1,009***
	6	0,239	0,995	0,993	1,006***	NA	1,01***	1,409***	3,529***	1,015*	1,029***
	9	0,316	1,012	0,991	1,03***	NA	1,033***	1,588***	2,976***	1,022***	1,047***
	12	0,379	1,022	0,999	1,05***	NA	1,052***	1,72***	2,746***	1,019**	1,065***
Copper											
	1	0,072	0,915	0,976	0,936	0,964	21,314***	0,921	13,191***	1,021	0,98
	3	0,161	**0,974*** **	1,017***	0,991***	1,025***	9,476***	1,013	5,899	1,143	1,058*
	6	0,245	1,002***	1,002***	1,019***	**0,986*** **	6,225**	1,139	3,937	1,238	1,152
	9	0,297	1,02***	1,002***	1,055***	**0,973*** **	5,127**	1,302	3,382	1,336	1,38
	12	0,331	1,022***	1,003***	1,062***	**0,969*** **	4,589*	1,376	3,16	1,356	1,558

Results
Agriculture commodities

h	RW RMSE	ARIMA(1,0,1)	SETAR(2,2,2)	SETAR(1,1,1)	LSTAR(2,2,2)	LSTAR(1,1,1)	MS-AR(1)	MS-AR(2)	NN (2,1,6)	NN (1,1,4)
Sugar										
1	0,075	0,949	0,974	0,974	0,963	0,981	1,004***	0,972	0,994	0,99
3	0,155	0,997	1,031***	1,027*	1,013	1,028***	1,115**	1,129	1,056***	1,04**
6	0,231	1,01	1,034***	1,06**	1,023***	1,059***	1,336**	1,257	1,079***	1,055***
9	0,283	1,022	1,049***	1,153***	1,043***	1,107***	1,552***	1,433	1,131***	1,09***
12	0,305	1,029	1,073***	1,244***	1,06***	1,154***	1,783***	1,646	1,2***	1,125***
Wheat										
1	0,070	0,968	0,995	0,982	1,043	0,986	0,979	1,689***	1,042	1,024
3	0,141	1,001	1,045	1,008	1,104	1,007	1,042	1,246	1,098	1,083
6	0,204	1,006	1,048	1,031*	1,108	1,015	1,134	1,293	1,061	1,026
9	0,249	1,022**	1,054	1,063*	1,093	1,028	1,297*	1,472	1,043	1,056
12	0,281	1,03***	1,062	1,094*	1,084	1,034	1,413**	1,664	1,032	1,102
Soybeans										
1	0,058	0,975	0,994	0,975	1,049	0,98	0,983	1,406	1,065	1,01
3	0,10	1,004	1,025	1,001	1,03	1,009	1,024	1,07	1,175	1,058
6	0,181	1,012	1,012***	1,005	1,032**	1,005	1,078	1,066	1,181**	1,052
9	0,217	1,029	1,028***	1,026	1,105***	1,023*	1,199	1,181	1,263***	1,037
12	0,229	1,037	1,041***	1,033	1,395***	1,022*	1,282	1,265	1,286***	1,027
Coffee										
1	0,057	0,986	**0,999****	**0,995***	1,001**	0,995	1,03**	0,988	1,056*	1,062
3	0,114	1,028**	1,053***	1,044***	1,033***	1,027***	1,227**	1,115	1,091***	1,081***
6	0,166	1,072***	1,126**	1,116***	1,07***	1,092***	1,607**	1,382	1,154***	1,142***
9	0,199	1,093***	1,152*	1,168***	1,089**	1,137***	1,945**	1,549	1,204***	1,177***
12	0,231	1,119***	1,195	1,224***	1,106	1,18***	2,218*	1,708	1,261***	1,215**
Maize										
1	0,064	0,992	0,993	0,989	0,984	0,994	1,032	1,583	1,2	1,159
3	0,128	0,992	0,992	1,003	0,982	0,983	1,097	1,243	1,494	1,283
6	0,196	1,012	0,998	1,075	1,007**	1,003	1,327**	1,43	1,297	1,284
9	0,235	1,02**	0,995	1,135	1,016**	1,003	1,54***	1,649**	1,201	1,319
12	0,265	1,025**	0,995	1,173	1,023**	1,009	1,691***	1,836**	1,166	1,403


RW
ARIMA(1,0,1)
SETAR(2,2,2)
SETAR(1,1,1)
LSTAR(2,2,2)
LSTAR(1,1,1)
MS-AR(1)
MS-AR(2)
NN(2,1,6)
NN(1,1,4)
1.8
2.2
2.6
2000
2005
2010
2015

Conclustions

- The study shows that outperforming *RW* forecasts is a difficult task

Conclustions

- The study shows that outperforming *RW* forecasts is a difficult task
- For some commodities and forecasting horizons nonlinear methods might increase accuracy

Conclustions

- The study shows that outperforming *RW* forecasts is a difficult task
- For some commodities and forecasting horizons nonlinear methods might increase accuracy
- No method was proven to consistently beat *RW* model