Helsinki Center of Economic Research

Discussion Papers

Forecasting Commodity Prices with Nonlinear Models

Valtteri Ahti
University of Helsinki

Discussion Paper No. 268
July 2009

ISSN 1795-0562

HECER – Helsinki Center of Economic Research, P.O. Box 17 (Arkadiankatu 7), FI-00014 University of Helsinki, FINLAND, Tel +358-9-191-28780, Fax +358-9-191-28781,
E-mail info-hecer@helsinki.fi, Internet www.hecer.fi

HECER
Discussion Paper No. 268

Forecasting Commodity Prices with Nonlinear Models

Abstract

The reliance on primary commodity exports posits a weighty problem for a large number of low-income less-developed countries, which is further aggravated by the erratic behaviour of commodity prices. Rational expectations competitive storage models conjecture that storage behaviour engenders nonlinear commodity pricing processes. Motivated by theoretical considerations state of the art nonlinear time series models are put to the task of forecasting an important subset of primary commodities, base metals. Nonlinear models, especially feedforward artificial neural networks, meet their mettle and hence may prove useful tools from a commodity-policy standpoint by providing at least some degree of visibility.

JEL Classification: O1, O13, Q3, G1

Keywords: Developing Countries, Natural Resource, Minerals, Commodity Markets

Valtteri Ahti

Department of Economics
P.O. Box 17
FI-00014 University of Helsinki
FINLAND

Email:
valtteri.ahti@helsinki.fi

1 Introduction

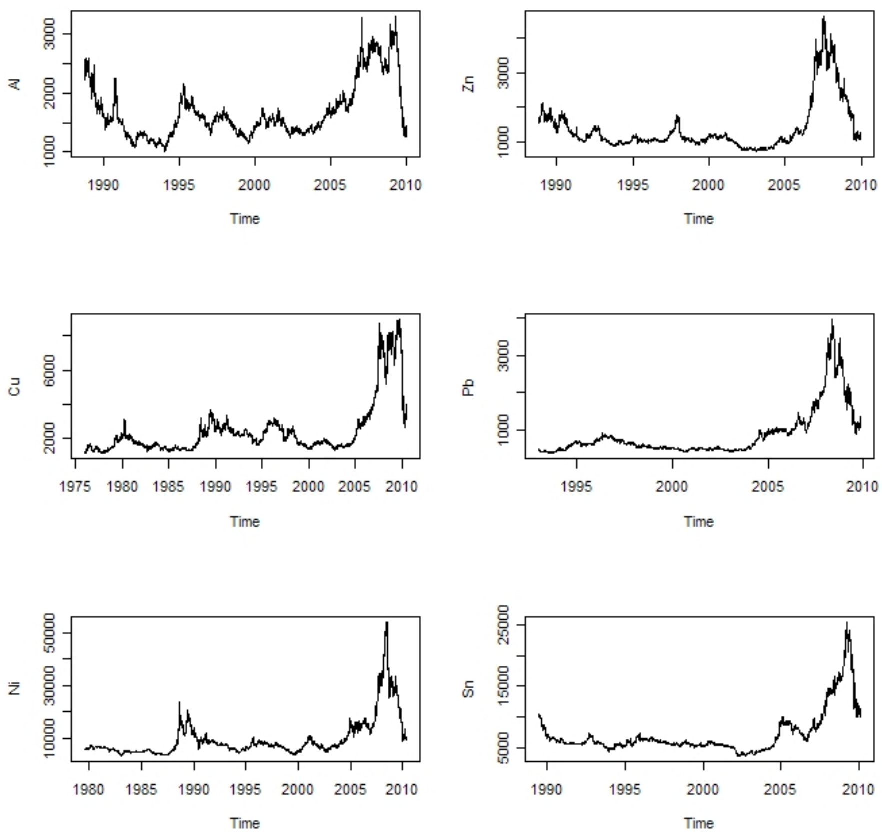
The policy relevance of forecasting primary commodity prices stems from the extreme export dependency of low-income less-developed countries (LDCs) on just a few select commodities. Table 1 reports some remarkable figures on export dependence encountered in Sub-Saharan Africa. Oxfam published a report in 1993 showing that Zambia derived 98 percent of its exports from copper, whilst coffee exports generated 95 percent of Uganda's export revenues. According to the report similar dependency plagues Ethiopia, Ghana, Kenya, Rwanda, Sudan, Tanzania and Zimbabwe with regards to coffee, cotton, cocoa and tobacco. Macroeconomic success in such countries is thus rendered a mere derivative of commodity price developments. Poor commodity price visibility is the implicit culprit for many a macroeconomic problem. To cite one example, the debt problem in LDCs largely traces its roots to unwarranted optimism concerning the future evolution of commodity prices in the late 1970's.

Dependency on primary commodities is an especially hazardous burden due to their extreme unpredictability. Figure 1 displays the evolution of industrial base metals 1976-2009, which are important sub-group of primary commodities. Prices are acutely volatile and exhibit a sequence of sharp peaks and subsequent shallow troughs. During the early stages of the financial crisis there was investor flight to commodities from other asset classes and a subsequent bust, which is apparent towards the end of the sample. To take one example, the price of nickel rose from its historical mean of 9559 to an unprecedented high of 52179 in July 2008. The other base metals followed suite.

Table 1: Export Dependence in Sub-Saharan Africa

	Primary Commodities as a percentage of Total Export Earnings	Individual Commodities as a percentage of Total Export Earnings
Zambia	99.7	copper 98
Rwanda	97.9	coffee 73
Uganda	95	coffee 95
Ethiopia	90	coffee 66
Sudan	88.5	cotton 42
Tanzania	79.3	coffee 40
Ghana	68.5	cocoa 59
Kenya	61.5	coffee 30
Zimbabwe	56.9	tobacco 20
Source: Oxfam (1993)		

The main body of economic theory on commodity prices, the rational expectations competitive storage theory, originates in the work by Gustafson (1958) on the optimal demand for commodity stocks and Muth's (1961) rational expectations hypothesis. The theory was subsequently developed by Samuels-



son (1971), Danthine (1977), Schechtman and Escudero (1977), Kohn (1978), Newbery and Stiglitz (1981, 1982), Scheinkman and Schechtman (1983), Salant (1983), Wright and Williams (1982, 1984), Williams and Wright (1991) and Hart and Kreps (1986). The key feature of the model lies in the inability of competitive speculators to hold negative inventories. The asymmetry in storage behavior feeds through to commodity prices rendering the price process nonlinear. For an excellent exposition on a basic variant of rational expectations competitive storage theory, as well as a structural empirical implementation, see Deaton and Laroque (1992, 1996).

In order to tackle the hypothesized nonlinearity in commodity prices the recent surge in nonlinear time series models is exploited. Specifically, the prevailing nonlinear methodologies are harnessed to the task of forecasting commodity prices. These models consist of two classes of nonlinear models; regime switching models and artificial neural networks. The emergence of regime switching models was heralded by the threshold autoregression model (TAR) and was subsequently followed by the smooth transition autoregression model (STAR). The neural architecture of choice in time series econometrics has been the feed-forward artificial neural network (ANN).

The application of multiple nonlinear models, though cumbersome, is warranted since rational expectations competitive storage models are silent concerning the nature of nonlinearity in commodity prices. At one end of the spectrum lie the rigid TAR models where indicator functions are employed to determine regime switching. Modeling flexibility is significantly augmented by STAR models with their smooth transition function. Ultimate flexibility is attained by employing feed forward neural networks.

The sample is composed of daily observations of London Metal Exchange (LME) metal commodity spot prices 1970-2009: aluminium (Al), copper (Cu), nickel (Ni), zinc (Zn), lead (Pb) and tin (Sn). Weekly and monthly series are constructed from the set of daily observations. On account of non-constant variances and since the series are found to be difference stationary the levels data is transformed into returns form. Forecasts generated by the nonlinear models, autoregressive moving average models (ARMA) and a no change model are evaluated using what Stock and Watson (2008) refer to as pseudo out-of-sample evaluation. Both point and directional accuracy are examined. Point accuracy is appraised by computing root mean squared forecast errors (RMSFE) and p-values from the Diebold-Mariano test of equal predictability (1995). Directional accuracy is assessed by calculating the proportion of times the model was correct in its directional prediction.

Nonlinear models have lower RMSFE and greater directional accuracy than the linear models or the no change model when estimated using weekly and monthly data. The Diebold-Mariano test often corroborates the findings of lower RMSFE vis-à-vis the random walk model, but not usually against the linear model. With respect to daily data all modeling effort is found to be near futile as the difference in point precision and directional accuracy between nonlinear and other models is negligible.

A commodity trading exercise finds that premier nonlinear models generate forecasts that are economically meaningful. Two mechanism are suggested through which nonlinear model could be employed from a policy point of view. One may engage financial markets directly or resort to physical arbitrage by implementing an inventory strategy.

The structure of the paper is the following. Section 2 discusses the metal commodities data set whilst Section 3 describes the nonlinear models. Statistical evaluation is carried out in Section 4 through a pseudo out-of-sample forecasting procedure and economic appraisal is conducted using a commodity trading simulation in Section 5. Finally Section 6 concludes.

2 Metal Prices

Some of the best data on primary commodities is base metals data, since it is listed in the London Metal Exchange (LME). The LME is the predominant metals trading hub and its roots as an international trade forum extend to the late Victorian era. Currently it is the largest pure commodity exchange in Europe and the world's tenth largest futures exchange. Trading features all of the important metals commodities: aluminium, aluminium alloy, copper, nickel, zinc, lead, tin and silver. According to Watkins and McAleer (2004) the LME is used worldwide by producers and consumers as a center for spot, futures and options trading in non-ferrous metals. They fo on to state that the LME offers three primary functions. Firstly, market participants can hedge against the risk of price volatility. Secondly, the LME settlement prices are used as reference prices around the world. Thirdly, the LME offers the services of a global warehouse network for settlements resulting in physical delivery. Price quotes are used as reference prices in base metals trading outside of the exchange and are reported in major financial dailies. For instance, approximately 95 % of the world trade in copper futures takes place in the LME making it the de facto world market price.

The data-set consists of daily observations of metals spot prices for the period 2.1.1976-24.3.2009. The base metals data set contains series of unequal length of all the major metals including aluminium, copper, nickel, zinc, lead and tin. Weekly observations are chosen to be Wednesday prices in order to minimize holiday effects. Monthly observations are constructed by using the final trading day of each month. Both criteria are standard within the empirical finance literature. Table 2 reports summary statistics on the daily spot prices. The latest observation is on 24.3.2009. The series have differing start dates, so that copper is the longest series stretching from 2.1.1976 yielding a total of 8668 daily observations. Lead constitutes the shortest series starting from 1.1.1993 and therefore contributing 4233 daily observations. Metals prices were listed in sterling prior to 1989. Therefore prices reported prior to 1989 are converted into dollar denominated series using the USD-GBP exchange rate.

Natural logarithms are taken of all variables due to variance non-constancy. The Augmented Dickey-Fuller (ADF), Phillips-Perron (PP) and Elliot-Rothenberg-Stock (ERS) tests strongly indicate that all series are nonstationary.[1] First differences of all series are found to be stationary via the same unit root testing protocol. Estimation is therefore conducted on differenced logarithmic data, i.e. on returns data.

The Brock, Dechert and Scheinkman test (Brock, et al. 1996) is arguably one of the most utilized tests for nonlinearity. It was not initially used as a test for nonlinearity, but as a means of detecting nonrandom chaotic dynamics. A large number of studies have found that the BDS test is proficient at detecting nonlinearity against a wide variety of linear and nonlinear alternatives (see Brock et al. 1991, Barnett et al. 1998). Given the number of nonlinear alternatives the BDS test is chosen over tests of linearity against specific forms of nonlinearity such as STAR or ANN type nonlinearity.

[1]In lieu of the extreme regime situated at the end of the sample period and due to anecdotal evidence from metals analysts a structural break does seem an attractive culprit for nonstationarity. The Nyblom Hansen, expF, supF and aveF tests of coefficient stability do indeed reject parameter constancy of the metals price series.

Table 2: Summary Statistics for Daily Metal Commodity Prices

	Al	Cu	Ni	Zn	Pb	Sn
obs	5334	8668	7743	5301	4233	5169
mean	1682	2453	9559	1371	862	7139
sd	470	1589	7574	723	655	3637
min	1018	1130	3157	725	356	3592
max	3291	8982	54150	4619	3978	25498

The BDS test effectively provides a metric for temporal correlation using the so called correlation integral. Given a time series x_t for $t = 1, 2, ..., T$ with m-history $x_t^m = \left(x_t^m, x_{t-1}^m, ..., x_{t-m+1}^m\right)$ the correlation integral at embedding dimension m may be computed as

$$C_{m,\epsilon} = \frac{2}{T_m\left(T_m - 1\right)} \sum_{m \leq s} \sum_{< t \leq T} I\left(x_t^m, x_s^m; \epsilon\right)$$

where $T_m = T - m + 1$ and $I\left(x_t^m, x_s^m; \epsilon\right)$ is the indicator function that takes the value 1 if $|x_{t-i} - x_{s-i}| < \epsilon$ for $i = 0, 1, ..., m-1$. Intuitively, the correlation integral yields the probability that two m-dimensional points are within distance ϵ of each other, i.e. it tests the joint probability

$$\Pr\left(|x_t - x_s| < \epsilon, |x_{t-1} - x_{s-1}| < \epsilon, ..., |x_{t-m+1} - x_{s-m+1}| < \epsilon\right).$$

If x_t are iid the joint probability in the limit should equal

$$C_{1,\epsilon}^m = \Pr\left(|x_t - x_s| < \epsilon\right)^m$$

and the BDS statistic $V_{m,\epsilon}$ may be computed

$$V_{m,\epsilon} = \sqrt{T} \frac{C_{m,\epsilon} - C_{1,\epsilon}^m}{s_{m,\epsilon}}$$

which under moderate regularity conditions converges in distribution to $N\left(0, 1\right)$. The BDS test is applied on the daily returns data and the results are reported in Table 3. The statistics are computed for a range of embedding dimensions $m \in \{2, ..., 5\}$ and ϵ was chosen to be extremely small. The reported test statistics correspond to p-values less than 0.01. The BDS test hence strongly indicates a violation of linearity. The test confirms the guidance of economic theory with respect to the nonlinearity in the equilibrium prices.

3 The Models

3.1 Self exciting threshold autoregression models

Threshold autoregression models (TAR) were proposed by Tong, (1978, 1983) and Tong and Lim (1980). Intuitively the data generating process produces

Table 3: BDS test statistics

	m=2	m=3	m=4	m=5
Al	12.23	12.19	12.15	12.09
Cu	16.75	16.72	16.79	16.75
Ni	20.14	20.18	20.17	20.13
Zn	15.91	15.88	15.96	15.92
Pb	16.30	16.27	16.24	16.28
Sn	18.97	18.95	18.91	18.87

a sequence of distinct linear autoregressions, where a threshold variable determines which linear autoregression is generating the values in question. For a detailed exposition of TAR models see Tong (1995). If the threshold variable is the dependent variable itself the TAR model is designated the self-exciting threshold autoregression (SETAR) model. The textbook (Franses and van Dijk 2002) two regime SETAR model with delay d is defined

$$y_t = (\phi_{0,1} + \phi_{1,1}y_{t-1} + ... + +\phi_{p_1,1}y_{t-p_1})\, I\left[y_{t-d} < c\right]$$

$$+ (\phi_{0,2} + \phi_{1,2}y_{t-1} + ... + +\phi_{p_2,2}y_{t-p_2})\, I\left[y_{t-d} > c\right] + \varepsilon_t.$$

where $I\left[y_{t-d} > c\right]$ is the indicator function with threshold parameter c. Note that the order of the distinct linear autoregressions need not be the same, i.e. $p_1 = p_2$. Indeed different orders of autoregression are considered in the forecasting procedure below. The model adopted in this article is the two regime SETAR with y_{t-1} as the threshold variable on account of its predominance in the forecasting literature. Clements et al. (2003) note a number of empirical applications of SETAR models including Tiao and Tsay (1994) and Potter (1995) with regards to US GNP, Hansen (1997), Montgomery et al. (1998), Rothman (1998) and Koop and Potter (1999) concerning unemployment, Kräger and Kugler (1993), Peel and Speight (1994) and Chappell et al. (1996) to nominal exchange rates, Obstfeld and Taylor (1997) and O'Connell (1998) to real exchange rates, and Pfann, Schotman and Tschernig (1996) to interest rates. The parameters of the SETAR model are estimated using sequential conditional least squares. If the ε_t are normally distributed the estimates are equivalent to maximum likelihood estimates.

3.2 Smooth transition autoregression models

The smooth transition autoregression (STAR) model originates from the work done by Bacon and Watts (1971) and was popularized by Teräsvirta and Anderson (1992) and Teräsvirta (1994). A textbook representation of the STAR model with delay d reads as follows

$$y_t = (\phi_{0,1} + \phi_{1,1}y_{t-1} + ... + +\phi_{p_1,1}y_{t-p_1})\,(1 - G\,(y_{t-d};\gamma, c))$$

$$+ \left(\phi_{0,2} + \phi_{1,2} y_{t-1} + ... + \phi_{p_2,2} y_{t-p_2}\right) G\left(y_{t-d}; \gamma, c\right) + \varepsilon_t.$$

The empirical literature has generally opted for the logistic transition function with delay $d = 1$

$$G\left(y_{t-1}; \gamma, c\right) = \left\{1 + \exp\left(-\gamma \prod_{k=1}^{K} \left(y_{t-1} - c_k\right)\right)\right\}^{-1}$$

i.e. the logistic smooth transition model (LSTAR). As in the case of SETAR models the order of the distinct linear autoregressions need not be the same and different orders of autoregression are considered in the forecasting procedure.

According to Teräsvirta et al. (2005) models with $K = 1$ or $K = 2$ generally outperform more extensive specifications in out-of-sample applications. Setting $K = 2$ implies symmetric behavior for low and high values of the transition variable, where as $K = 1$ generates asymmetric behavior. Only the LSTAR specification with $K = 1$ is retained for analysis given the asymmetry implied by the non-negativity constraint and the already daunting computational burden. Note that this specification nests both the simple linear autoregression model as $\gamma \to 0$ and the two regime SETAR model as $\gamma \to \infty$. The asymmetric property of LSTAR models with $K = 1$ has lead to their use in business cycle applications such as Stock and Watson (1999) and Skalin and Teräsvirta (2002).

The STAR models are estimated using nonlinear least squares (NLS) with the Broyden–Fletcher–Goldfarb–Shanno (BFGS) method. The NLS estimates can be interpreted as maximum likelihood estimates if the ε_t are normally distributed.

3.3 Feed forward artificial neural networks

White (2006) describes neural networks as flexible functional forms emulating the behavior of biological neural systems. An overview of neural networks from an econometric perspective can be found in Kuan and White (1994) or Medeiros et al. (2006). The workhorse of the ANN industry is the single hidden layer feed forward network.

$$y_t = \phi_0 + \phi_1 y_{t-1} + ... + \phi_{p_0} y_{t-p_0} + \sum_{j=1}^{D} \beta_j G\left(\gamma_{0j}, ..., \gamma_{p_j j}, y_{t-1}, ..., y_{t-p_j}\right) + \varepsilon_t$$

where the activation function $G(\cdot)$ is taken to be logistic

$$G\left(\gamma, y_{t-1}, ..., y_{t-p}\right) = \left\{1 + \exp\left(\gamma_0 + \gamma_1 y_{t-1} + ... + \gamma_p y_{t-p}\right)\right\}^{-1}.$$

In neural network nomenclature the terms $\sum G(\cdot)$ are called hidden units within a hidden layer and the parameters γ are referred to as connection strengths. Generally in empirical applications an ANN varies in terms of the total number of specified hidden units D and in the the number of lagged dependent variables in the linear and hidden nonlinear units p_1 , p_2,...,p_D. This paper specifies a

NLS objective function for each ANN specification. The resulting unconstrained nonlinear optimization problem is solved using BFGS.

A chief motivation for ANNs is their interpretation as universal function approximators as discussed in Cybenko (1989), Funahashi (1989), Hornik, Stinchcombe and White (1989) and White (1990). Crudely put the idea is that one is able to approximate any multivariate function by employing a finite linear combination of univariate discriminatory functions. Neural networks employ a class of sigmoidal functions such as the logistic function via the so called hidden units. Sigmoidal functions are a class of discriminatory functions and hence a linear neural network is in effect a linear combination of sigmoidal functions. Therefore one can use neural networks to approximate any unknown multivariate function with arbitrary precision by increasing the number of hidden units in the neural architecture.

4 Pseudo out-of-sample evaluation

The forecasts from the models are subjected to what Stock and Watson (2008) call pseudo out-of sample evaluation. The procedure calls for a rolling window scheme. The length of the estimation window is held fixed by dropping the furthest observation upon the addition of the latest. For example, the window length in the case of daily data is held fixed at 1000 observations throughout the recursive estimation scheme. The parameters of each model are re-estimated after every new observation and based on the optimized parameters a one-step-ahead forecast is computed. By holding the length of the estimation window fixed and by keeping the number of parameters constant statistical testing on forecast errors is enabled. Granger (1993) suggests somewhat heuristically that at least 20 % of the data should be left for out-of-sample forecasting when dealing with nonlinear models. The weekly and monthly window lengths were set such that 30 % of the sample was retained for out-of sample forecast analysis.

The two regime SETAR and LSTAR models use a lag of the dependent variable as the threshold or transition variable. The set of regime switching models can be represented $SETAR(p,q), LSTAR(p,q)$ such that $p, q \in \{1,2,3\}$, where p (q) is the number of lags in the low (high) regime.

The most celebrated neural architecture in the time series literature is the feedforward neural network. The estimated neural networks differ from traditional time series ANNs with one notable exception. The insistence on a linear unit was thought to tie the ANN too closely to the threshold models. The idea here is to offer a more flexible nonlinear alternative to the very inflexible SETAR model and the moderately flexible LSTAR model. Hence the linear unit was discarded and replaced by a mere constant as in Teräsvirta et al. (2005). Akin to the regime switching models the ANNs specified may be represented $ANN(p,q)$ with $p, q \in \{1,2,3\}$, where p is the number of number of hidden units and q the number of lagged variables therein.

Thus for each metal/frequency combination 27 nonlinear models are specified consisting of 9 SETAR, LSTAR and ANN models. The 487 nonlinear models specified for the rolling window estimation scheme with three frequencies posit a nontrivial computational burden.

The benchmark models for an evaluation of nonlinear model are naturally linear models. The linear models include the autoregressive moving average model (ARMA) and the no change model. Additionally a naive futures model was also constructed using three month futures as predictors of future spot prices. The out-of-sample performance of the naive model was far worse than that of the no change model.

One-step-ahead forecast performance is evaluated using both point and directional accuracy. Due to nonstationarity and variance non-constancy the series were transformed from levels into returns and consequently the forecasts are in returns form as well. However, the motivation of this paper is to test a large number of nonlinear models that may then be applied to the policy relevant issue of forecasting commodity prices. Hence prior to computing measures of model accuracy the forecasts are converted back into levels form to ensure that the results are transparent. Point accuracy is determined by computing the root mean squared forecast error (RMSFE) and the Diebold-Mariano test of equal forecast accuracy with a mean squared loss function. Directional accuracy is appraised according to the ratio of times the sign of the forecast was correct (sign). Forecasts from the rolling window scheme were computed for the models for the six base metals and for three frequencies: daily, weekly and monthly.

The pairwise Diebold-Mariano test of equal forecast accuracy is performed to test the significance of observed differences in mean squared forecast errors (MSFE) between linear and nonlinear models. The pairs consisted of the nonlinear and linear model with the lowest MSFE. The nonlinear model is also tested against the random walk model. Since the comparison is between one-step-ahead forecasts there is no need to adopt the modified form of the test proposed by Harvey, Leybourne, and Newbold (1997). Standard asymptotic theory does not apply with regards to the Diebold-Mariano test when the two models are nested as discussed in Clark and McCracken (2001). Teräsvirta et al. (2005) confront a somewhat analogous predicament and defend the use of the Diebold-Mariano test by maintaining that the models are approximations to the same unknown data-generating process.

Tables 4 and 5 relay information on point accuracy for the weekly and monthly frequencies of the base metals data-set. The first column reports the chemical symbol of the given base metal. The second column presents the model that generated the lowest RMSFE along with its RMSFE, whilst the third column contains the linear model that produced the smallest RMSFE and its RMSFE. The fourth column reports the RMSFE from the no change model. The fifth and sixth columns detail the length of the estimation window and the number of forecasts. The seventh and eight columns report p-values from Diebold-Mariano tests of equal forecastability. The premier nonlinear model is pitted against the no change (premier linear model) in the seventh (eight) column. The alternative hypothesis in both tests asserts superior predictability on the part of the nonlinear model.

The set of daily observations proves impervious to the point forecasts derived from both nonlinear and linear models, i.e. the no change model carries the day. Hence no table is provided for the results. Initial computations of RMSFE did provide tentative evidence in favor of a neural architecture, but subsequent Diebold-Mariano tests yielded no indication of forecast superiority against the no change model. Previously envisaged experimentation with the length of the estimation window were aborted given the almost complete lack of statistical support for nonlinear modeling and the considerable computational burden involved in executing the rolling window forecasting procedure with daily data.

The results from the weekly and monthly data-sets stand in stark contrast to the dismal turnout with daily series. Table 4 provides the results from using weekly data and indicates that with respect to all six metal series a nonlinear model provided the lowest RMSFE. Interestingly the four neural networks that yielded the lowest RMSFE have near identical architectures consisting of two or three lags in three hidden units. The two regime switching models have more conservative specifications relative to the neural networks and their linear competitors. All the nonlinear models produce statistically superior forecasts compared to the random walk according to the Diebold-Mariano test. However, only in the case of lead (Pb) does the nonlinear model (ANN) also provide superior statistical predictability relative to the leading linear model.

	Models					Sample		DM	
	Nonlinear	RMSE	Linear	RMSE	RW	Window	Forecasts	RW	LM
Al	ANN(3,2)	69	ARMA(2,1)	69	70	746	319	0.00	0.49
Cu	LSTAR(1,2)	177	ARMA(3,2)	178	187	1212	520	0.00	0.30
Ni	ANN(3,2)	1101	ARMA(3,1)	1107	1131	1083	464	0.01	0.18
Zn	SETAR(1,1)	113	ARMA(3,1)	116	123	741	318	0.00	0.10
Pb	ANN(3,3)	106	ARMA(3,1)	107	111	592	253	0.00	0.00
Sn	ANN(3,2)	543	ARMA(3,0)	558	597	722	310	0.00	0.18

Table 4: Point Accuracy: The Weekly Data-set

The results are similar if statistically diluted with regards to monthly data as shown in Table 5. Once more the nonlinear models generate the lowest RMSFEs across the board. However, the Diebold-Mariano tests are not able to establish nonlinear forecast superiority except with respect to nickel and tin. One in two series favours a neural network with architectures almost identical to those seen with the weekly data set. The SETAR and STAR models opt for more flexible model skeletons than in the weekly case. There seems to be no immediate pattern to the ARMA specifications other than an emphasis on autoregressive parameterization.

The results on directional accuracy are summarized in Table 6. With respect to weekly and monthly observations nonlinear models purport considerable edge over competing model classes. In fact in three-quarters of the cases reported in Table 5 the directional precision of the foremost nonlinear model is at least on par with the premier linear or the no change model. Directional accuracy for nonlinear models in the case of weekly observations hovers around the mid-fifties. Given monthly data the figures surpass the sixty percent threshold in

	Models					Sample		DM	
	Nonlinear	RMSE	Linear	RMSE	RW	Window	Forecasts	RW	LM
Al	ANN(3,3)	130	ARMA(3,2)	136	141	171	73	0.10	0.18
Cu	SETAR(3,1)	382	ARMA(1,0)	402	438	278	119	0.08	0.20
Ni	LSTAR(2,3)	2108	ARMA(5,0)	2153	2433	248	107	0.00	0.10
Zn	SETAR(2,2)	204	ARMA(1,0)	248	265	170	73	0.01	0.02
Pb	ANN(3,2)	242	ARMA(2,0)	252	255	135	58	0.25	0.20
Sn	ANN(3,2)	966	ARMA(3,0)	1023	1115	165	71	0.04	0.10

Table 5: Point Accuracy: The Monthly Data-set

two out of six cases. However, one may ask whether it is plausible that a model may predict the direction of a highly liquid market two-thirds of the time?

	Weekly		Monthly	
Model	sign	Model	sign	
Al	RW	46.7	RW	52.1
	ARMA(5,0)	51.1	ARMA(1,0)	50.7
	LSTAR(3,2)	49.5	LSTAR(3,1)	52.1
Cu	RW	54	RW	50.4
	ARMA(1,0)	54.2	ARMA(3,0)	53.8
	ANN(3,1)	54.6	ANN(1,3)	52.9
Ni	RW	50	RW	56.1
	ARMA(2,1)	51.9	ARMA(4,0)	56.1
	ANN(1,1)	52.6	LSTAR(1,2)	59.8
Zn	RW	55.3	RW	52.1
	ARMA(3,0)	55.3	ARMA(2,0)	56.2
	LSTAR(1,3)	57.2	LSTAR(2,1)	54.8
Pb	RW	56.9	RW	58.6
	ARMA(1,0)	55.3	ARMA(3,0)	43.1
	ANN(2,2)	58.9	SETAR(1,1)	65.5
Sn	RW	51.3	RW	64.8
	ARMA(3,2)	51.6	ARMA(1,0)	62
	ANN(1,1)	54.2	LSTAR(3,1)	64.8

Table 6: Directional Accuracy

5 Commodity Trading

This section constructs a simulated commodity trading framework to gauge the potential economic significance of trading strategies employing nonlinear models to forecast future prices. There are at least two mechanisms through which a representative LDC might use a nonlinear model to conduct or support commodity trading decisions. Firstly, since the metal commodities are listed in a financial exchange the country might simply take a countervailing position using financial instruments such as futures and options as indicated by the nonlinear model. However, such a mechanism relying on financial instruments may be infeasible due to market liquidity constraints or the mere expense of using derivatives.

Rational expectations competitive storage theory implies an alternate mechanism for commodity trading. A country might engage in physical arbitrage via the use of commodity inventories. The optimal level of inventories is implicit in the expected future path of commodity prices, which in turn are implied by the forecasts from a nonlinear model. Such an inventory management strategy takes the familiar form of a (S, s) policy rule, albeit on a macroeconomic scale. Hall and Rust (2000) show that the (S, s) rule is an optimal trading strategy for a commodity speculator. Nesting a nonlinear time series model within a dynamic structural model as in Rust and Hall (2000) is an intriguing idea, but beyond the scope of this paper. Instead this section documents the results of implementing a simple commodity trading strategy using nonlinear models vindicated by their out-of sample directional accuracy as reported in the previous section.

The construct of a simulated trading exercise is by definition somewhat arbitrary. In this instance a developing country decides each week whether to take a long or short position in the LME with respect to a metals commodity. In a real world application such an instrument might be the three month future, which is the most liquid derivative.

The nonlinear model with the highest directional accuracy (see Table 6) is used to determine whether to be long or short in the market each week. The position unravels the following week and a new position is taken based on the one-step-ahead forecast at hand. The entire weekly out-of-sample portion is used to compute the weekly and annualized rates of return from such a strategy. Note that since trading is costly forecasts indicating little or no future change would presumably lead to inaction. Such matters are abstracted from in the trading simulation.

Table 7 reports the results from the trading simulation. The first column reports the metal in question and the second column the nonlinear model that yielded the highest directional accuracy. The third column reports the number of trades and the final column presents the average rate of return from a trade. Standard efficient market literature implies that risk adjusted returns equal the costs of exchange. Hence if the traders are rational risk neutral agents the reported returns should equal trading costs. Given such a view the observed large differences in the weekly rates of return between different metals commodities seem rather puzzling. Since aluminium and copper trades account for almost three quarters of trading volume and given that both markets are very liquid one would expect trading costs from such liquid instruments to be on par. In the same vein, nickel trading, which is the third most traded metal on the exchange yields a rate of return over four times that of aluminium.

As discussed above the chief alternative view on commodity prices is the rational expectations competitive storage hypothesis, which specifically implies a nonlinear equilibrium price process. To some extent then the results of the out-of-sample forecasting exercise and the trading simulations provide evidence for the competitive storage view over the efficient market hypothesis.

The large and varying rates of return from the trading simulation corroborate the out-of-sample forecast results. The statistical evidence from the forecasting exercise is hence augmented by economically significant results from the trading experiment. Naturally the trading simulation only serves to provide a rough metric as to the potential economic value of engaging in commodity markets using nonlinear models. Without statistical tests it is impossible to conclude that the observed rates are indeed positive or different from one another.

Of course from a risk management perspective it may be wise to use export proceeds or other funds to invest in other assets. Oil producing countries such as Norway and some of the Gulf states actively set aside export earnings from oil and manage the funds in international financial markets in a well diversified manner. Especially for countries as dependent on a single commodity as Zambia the role of diversification is difficult to overstate.

	Model	Trades	Weekly Rate
Al	LSTAR(3,2)	519	0.61%
Cu	ANN(3,1)	464	1.64%
Ni	ANN(1,1)	463	2.53%
Zn	LSTAR(1,3)	310	1.14%
Pb	ANN(2,2)	252	2.54%
Sn	ANN(1,1)	309	2.19%

Table 7: Trading Simulation

6 Conclusion

The problem of export dependency on primary commodities in developing countries is exacerbated by the erratic behaviour of commodity prices. Rational expectations competitive storage theory conjectures that commodity prices should follow a nonlinear process. Motivated by theoretical considerations and armed with a veritable explosion in nonlinear model developments an important subgroup of primary commodities, namely industrial base metals, are examined using state of the art nonlinear models.

The simulated out-of-sample results indicate that with respect to weekly and monthly data nonlinear models in all cases produce the lowest forecast errors. However, in some cases the Diebold-Mariano tests cast doubt on whether the observed differences in forecast performance between nonlinear and linear models are statistically significant. A commodity trading framework is constructed to measure whether the forecasts generated using nonlinear models are economically meaningful. The trading exercise reveals significant economic gains from implementing a simple trading rule based on nonlinear models.

Two mechanisms were proposed through which nonlinear models might be incorporated into actual policy making. One might either exploit the emergence of the London Metal Exchange as a highly liquid commodities exchange directly via an appropriate financial position. Or indirectly by relying on physical arbitrage through the optimal deployment of commodity inventories.

A very large number of nonlinear model specifications were estimated in this article. However, one should note that the search over nonlinear specifications was by no means all-inclusive. The fact that nonlinear models uniformly outperformed their linear competitors should be taken as indicative of the forecasting power embedded in nonlinear models vis-à-vis commodity prices. Given that the search over nonlinear models was non-exhaustive further forecast refinement seems more than plausible.

References

[1] Barnett, W.A., Gallant, A.R., Hinich, M.J., Jungeilges, J.A., Kaplan, D.T. and Jensen, M.J., 1998. A single-blind controlled competition among tests for non-linearity and chaos. *Journal of Econometrics* 82, pp. 157–192.

[2] Bacon, D. W. and Watts, D. G.: 1971, Estimating the transition between two intersecting straight lines, *Biometrika* 58, 525-534.

[3] Brock, W., D. A. Hsieh and B. Le Baron 1991, Nonlinear Dynamics, Chaos, and Instability: Statistical Theory and Economic Evidence (MIT Press, Massachusetts and London).

[4] Brock, W., Dechert, D., J. A. Scheinkman and B. Le Baron. 1996, A Test For Independence Based on the Correlation Dimension, *Econometric Reviews*, Vol. 15, pp. 197-235.

[5] Chan, K. S., & Tong, H. 1986. On estimating thresholds in autoregressive models. *Journal of Time Series Analysis*, 7, 178–190.

[6] Clark, T. E., and M. W. McCracken. 2001 Tests of equal forecast accuracy and encompassing for nested models, *Journal of Econometrics*, 105, 85-110.

[7] Clements, M.P., P.H. Franses, J.Smith and van Dijk D. 2003. On SETAR non-linearity and forecasting, *Journal of Forecasting* 22: 359-375.

[8] Chappell D, Padmore J, Mistry P, Ellis C. 1996. A threshold model for the French franc/Deutschmark exchange rate. *Journal of Forecasting* 15: 155–164.

[9] Chowdhury, A.R. 1991. Futures market efficiency: evidence from cointegration tests. *The Journal of Futures Markets* 11, 577–589.

[10] Cybenko, G.V. 1989. Approximation by Superpositions of a Sigmoidal function, *Mathematics of Control, Signals and Systems* vol. 2 no. 4 303-314.

[11] Danthine, J.-P. 1977, Martingale, market efficiency and commodity prices, *European Economic Review*, 10, 1-17.

[12] Diebold, F.X, R.S. Mariano, 1995. Comparing Predictive Accuracy, *Journal of Business and Economic Statistics* 13, 253-263.

[13] Deaton, A. and Laroque, G. 1992. On the behavior of Commodity Prices. *Review of Economic Studies*. 59, pp. 1-24.

[14] Deaton, A. and Laroque, G. 1996. Competitive Storage and Commodity Price Dynamics. *Journal of Political Economy*. 104, pp. 896-923.

[15] Franses P.H, van Dijk D. 2000. Non-Linear Series Models in Empirical Finance. Cambridge University Press: Cambridge.

[16] Funahashi, K.: 1989, On the approximate realization of continuous mappings by neural networks, *Neural Networks* 2, 183192.

[17] Goss, B. A. 1981. The futures pricing function of the London Metal Exchange. *Applied Economics* 13, 133–150.

[18] Goss, B. A. 1983. The semi-strong form efficiency of the London Metal Exchange. *Applied Economics* 15, 681–698.

[19] Gustafson, R. L. 1958. Carry over Levels for Grains. *US Department of Agriculture, Technical Bulletin*, 1178.

[20] Hall, G. and Rust J. 2000. An Empirical Model of Inventory Investment by Durable Commodity Intermediaries. *Carnegie-Rochester Conference Series on Public Policy*, 52, 171-214.

[21] Hansen BE. 1997. Inference in TAR models. *Studies in Nonlinear Dynamics and Econometrics* 2: 1–14.

[22] Hart, O. D. and Kreps, D. M. 1986, Price destabilizing speculation, *Journal of Economy*, 94, 927-952.

[23] Harvey, D., S. Leybourne, and P. Newbold 1997: Testing the Equality of Prediction Mean Squared Errors," *International Journal of Forecasting*, 13, 281{291.

[24] Heaney, R. 2002. Does knowledge of the cost of carry model improve commodity futures price forecasting ability?: A case study using the London Metal Exchange lead contract. *International Journal of Forecasting* 18, 45-65.

[25] Hornik, K., Stinchcombe, M. and White, H.: 1989, Multi-layer feedforward networks are universal approximators, *Neural Networks* 2, 359366.

[26] Hsieh, D. A., & Kulatilaka, N. 1982. Rational expectations and risk premia in futures markets: primary metals at the London Metal Exchange. *The Journal of Finance* 37, 1199–1207.

[27] MacDonald, R., & Taylor, M. P. 1988a. Testing rational expectations and efficiency in the London Metal Exchange. *Oxford Bulletin of Economics and Statistics* 50, 41–52.

[28] MacDonald, R., & Taylor, M. P. 1988b. Metals prices, efficiency and cointegration: some evidence from the London Metal Exchange. *Bulletin of Economic Research* 40, 235–239.

[29] MacDonald, R., & Taylor, M. P. 1989. Rational expectations, risk and efficiency in the London Metal Exchange: an empirical analysis. *Applied Economics* 21, 143–153.

[30] Moore, M. J., & Cullen, U. 1995. Speculative efficiency on the London Metal Exchange. *The Manchester School* 63, 236.

[31] Montgomery AL, Zarnowitz V, Tsay RS, Tiao GC. 1998. Forecasting the U.S. unemployment rate. *Journal of the American Statistical Association* 93: 478–493.

[32] Muth, J. F. 1961. Rational expectations and the theory of price movements, *Econometrica* 29, 315-35.

[33] Kohn, M. 1978, Competitive speculation, *Econometrica*, 46, 1061-1076.

[34] Koop G, Potter SM. 1999. Dynamic asymmetries in US unemployment. *Journal of Business and Economic Statistics* 17: 298–313.

[35] Kolmogorov, AN. On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition, *DokL Akad. Nauk. SSSR*, 114 (1957), 953-956.

[36] Kuan, C.-M. and H. White, 1994. Artificial neural networks: an econometric perspective, *Econometric Reviews* 13, 1-143.

[37] Kräger H, Kugler P. 1993. Non-linearities in foreign exchange markets: a different perspective. *Journal of International Money and Finance* 12: 195–208.

[38] Medeiros, M.C.,Teräsvirta, T. and Rech, G. 2006. Building Neural Network Models for Time Series: A Statistical Approach, *Journal of Forecasting* 25, 49–75.

[39] Newbery, D. M. G. and Stiglitz, J. E. 1981 The Theory of Commodity Price Stabilization: A Study in the Economics of Risk (Oxford: Oxford University Press).

[40] Newbery, D. M. G. and Stiglitz, J. E. 1982, Optimal commodity stockpiling rules, *Oxford Economic Papers*, 34, 403-427.

[41] Teräsvirta, T. and H.M. Anderson, 1992. Characterizing nonlinearities in business cycles using smooth transition autoregressive models, *Journal of Applied Econometrics* 7, S119-S136.

[42] Teräsvirta, Timo & van Dijk, Dick & Medeiros, Marcelo C., 2005. Linear models, smooth transition autoregressions, and neural networks for forecasting macroeconomic time series: A re-examination, *International Journal of Forecasting*, Elsevier, vol. 21(4), pages 755-774.

[43] Teräsvirta, T., 1994. Specification, estimation, and evaluation of smooth transition autoregressive models, *Journal of the American Statistical Association* 89, 208-18.

[44] Tiao GC, Tsay RS. 1994. Some advances in non-linear and adaptive modelling in time-series. *Journal of Forecasting* 13: 109–131.

[45] Tong H. 1978. On a threshold model. In *Pattern Recognition and Signal Processing*, Chen CH. (ed.). Sijhoff and Noordoff: Amsterdam; 101–141.

[46] Tong H. 1983. *Threshold Models in Non-Linear Time Series Analysis.* Springer-Verlag: New York.

[47] Tong, H. 1995. *Non-linear Time Series. A Dynamical System Approach.* Clarendon Press: Oxford. (First published 1990.)

[48] Tong H, Lim KS. 1980. Threshold autoregression, limit cycles and cyclical data. *Journal of The Royal Statistical Society* B 42: 245–292.

[49] Obstfeld M, Taylor AM. 1997. Nonlinear aspects of goods-market arbitrage and adjustment: Heckscher's commodity points revisited. *Journal of the Japanese and International Economies* 11: 441–479

[50] O'Connell PGJ. 1998. Market frictions and real exchange rates. *Journal of International Money and Finance* 17: 71–95.

[51] Peel DA, Speight AEH. 1994. Testing for non-linear dependence in interwar exchange rates. *Weltwirtschaftliches Archiv* 130: 391–417.

[52] Pfann GA, Schotman PC, Tschernig R. 1996. Nonlinear interest rate dynamics and implications for the term structure. *Journal of Econometrics* 74: 149–176.

[53] Potter S. 1995. A nonlinear approach to US GNP. *Journal of Applied Econometrics* 10: 109–125.

[54] Rothman P. 1998. Forecasting asymmetric unemployment rates. *The Review of Economics and Statistics* 80: 164–168.

[55] Salant, S. W. 1983, The vulnerability of price stabilization schemes to speculative attack, *Journal of Political Economy*, 91, 1-38.

[56] Samuelson, P. A. 1971, Stochastic speculative price, *Proceedings of the National Academy of Science*, 68, 335-337.

[57] Schechtman, J. and Escudero, V. 1977, Some results on 'An income fluctuation problem', *Journal of Economic Theory*, 16, 151-166.

[58] Skalin, J., & Teräsvirta, T. 2002. Modeling asymmetries and moving equilibria in unemployment rates. *Macroeconomic Dynamics*, 6, 202–241.

[59] Stock, J. H., & Watson, M. W. 1999. A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series. In R. F. Engle, & H. White (Eds.), *Cointegration, causality and forecasting.* A festschrift in honour of Clive W.J. Granger (pp. 1 – 44). Oxford7 Oxford University Press.

[60] Stock, J. H., & Watson, M. W. 2008. Phillips Curve Inflation Forecasts. *NBER Working Paper* No. W14322. Available at SSRN: http://ssrn.com/abstract=1267553

[61] Sephton, P. S., & D. K. Cochrane, 1990. A note on the efficiency of the London Metal Exchange. *Economic Letters* 33, 341–345.

[62] Sephton, P. S., & D. K. Cochrane, 1991. The efficiency of the London Metal Exchange: another look at the evidence. *Applied Economics* 23, 669–674.

[63] Scheinkman, J. A. and Schechtman, J. 1983, A simple competitive model with production and storage, *Review of Economic Studies*, 50, 427-441.

[64] Watkins C. and McAleer M. 2004. Econometric modelling of non-ferrous metal prices. *Journal of Economic Surveys* 18, 651-701.

[65] White, H.: 1990, Connectionist nonparametric regression: Multilayer feed-forward networks can learn arbitrary mappings, *Neural Networks* 3, 535 550.

[66] Williams, J. C. and Wright, B. D. 1991 Storage and Commodity Markets (Cambridge: Cambridge University Press).

[67] Wright, B. D. and Williams, J. C. 1982, The economic role of commodity storage, *Economic Journal*, 92, 596-614.

[68] Wright, B. D. and Williams, J. C. 1984, The welfare effects on the introduction of storage, *Quarterly Journal of Economics*, 98, 169-192.