**A Comparison of Multistep Commodity Price Forecasts Using Direct and Iterated Smooth Transition Autoregressive Methods**

**David Ubilava[[1]](#footnote-1)**

*School of Economics, University of Sydney*

**Abstract**

The smooth transition autoregressive (STAR) modeling framework has gained popularity in commodity price analysis due to its ability to capture essential features of complex dynamics. This study addresses the questions of whether the improved in-sample fit of STAR models results in more accurate forecasts compared to linear autoregressive models, and whether direct or iterated multistep STAR methods yield more accurate multistep forecasts. In the STAR framework, either a bootstrap simulation is necessary to numerically approximate iterated multistep forecasts, or a range of horizon-specific STAR models need to be estimated to generate direct multistep forecasts. The associated computational trade-off underscores the need for a better understanding of advantages one method may have over another. Based on the analysis of 25 agricultural and non-agricultural commodity prices, this study finds that even when the STAR models appear to well approximate complex commodity price dynamics, they offer little advantage, and indeed, in most instances present as inferior alternatives to the basic autoregressive framework for multistep commodity price forecasting.

**Keywords**: Commodity Prices; Direct Forecasts; Iterated Forecasts; Multistep Forecasts; Smooth Transition Autoregression.

**JEL Classification**: C53; E37; Q02

**Introduction**

Smooth transition autoregressive (STAR) models have been successfully deployed to capture essential features of complex commodity price dynamics (e.g., Holt and Craig, 2006; Balagtas and Holt, 2009; Enders and Holt, 2012; Ubilava, 2012; Hood and Dorfman, 2015; Ubilava, 2018). This line of empirical research is important and relevant, as it allows to examine nonlinearities in commodity prices that can arise due to the presence of transactions costs in spatial and temporal arbitrage, the market structure, or simply because of peculiarities of production processes.

These previous studies have demonstrated the improved in-sample fit of STAR models compared to their linear counterparts. A question remains whether this specific nonlinear modeling framework can help achieve more accurate intermediate--term forecasts. To that end, Enders and Pascalau (2015) proposed a pre-testing procedure to investigate the potential benefit of multistep forecasting from STAR-type models. They concluded that, in a direct multistep setting, when a pre-test strongly rejects the null of linearity in favor of the STAR-type alternative, the STAR model will likely yield more accurate forecasts than the AR model. This pre-testing framework, moreover, lends itself to a broader research question on relative accuracy of direct and iterated multistep forecasts (e.g., Chevillon and Hendry, 2005; Marcellino et al., 2006; Chevillon, 2007).

This study focuses on two research questions: (i) whether the improved in-sample fit of STAR models manifests into a more accurate multistep forecasts of commodity prices; and (ii) whether the direct multistep price forecasts are more accurate than the iterated multistep price forecasts, specifically those from STAR-type models. In addressing these questions, the finding of this study can benefit market participants who make their decisions based on prices that are set to be realized several months into the future. This group certainly involves agricultural producers and policymakers not only in high income countries, but also, especially, in developing and emerging economies that either heavily rely on primary commodity exports or are susceptible to global commodity price shocks.

Direct and iterated multistep forecasting methods differ from each other in that the iterated method fits a one-step model, thus minimizing the squares of one-step-ahead residuals, to generate a multistep forecast by recursively substituting the intermediate forecasts into the projection up to horizon *h*, whereas the direct method fits an *h*-step model, thus minimizing the squares of *h*-step-ahead residuals, which directly generates a multistep ahead forecast at the desired horizon. Theory suggests that while iterated multistep forecasts are more efficient when the (one-step) model is correctly specified, direct multistep forecasts can be more robust to model misspecification (Weiss, 1991; Stock and Watson, 1998).

The foregoing trade-off, usually examined in a linear context, extends to STAR-type nonlinear models. This nonlinear modeling framework, moreover, adds an additional layer of trade-offs. On the one hand, an iterated method necessitates a numerical approximation to generate multistep forecasts from nonlinear models, because a simple extrapolation, commonly applied to generate forecasts from linear models, biases the results (e.g., Teräsvirta et al., 2010). On the other hand, a direct method requires selection and estimation of the preferred nonlinear specification at each considered horizon, which in and of itself brings on a computational burden (e.g., Teräsvirta, 2006).

The present study contributes to the literature by creating empirical evidence of forecast accuracy for commodity prices from STAR-type direct and iterated multistep methods, as well as of these for nonlinear methods and their linear counterparts. Monthly time series of 25 agricultural and non-agricultural commodities, ranging the January 1980 – December 2020 period, are analysed. The data are obtained from online portals of the World Band and the International Monetary Fund. Statistical computing and graphics is done in a freely available language and environment, R version 4.1.0. In general, the results of this empirical exercise suggest that (i) the autoregressive methods offer more accurate forecasts than the STAR methods even when the linearity tests of Teräsvirta (1994), including its multistep variant of Enders and Pascalau (2015), point to STAR-type dynamics in the time series; and (ii) there is a mixed bag of results when comparing iterated and direct forecast methods.

The overall takeaway from the present study is that while STAR modeling framework is well suited to unveil nonlinear intricacies of commodity price series, as suggested by previous studies and confirmed by the current one, when it comes to multistep forecasting, the framework offers little advantage, and indeed, in most instances appears to be an inferior alternative to the basic autoregressive framework. This takeaway should not be generalized to all nonlinear models, however. For any linear specification there are virtually unlimited nonlinear specifications at a forecaster's disposal. This study considers just one, albeit an increasingly popular, nonlinear specification. That the considered nonlinear specification does not yield more accurate price forecasts than its linear counterpart is not an indication that others do not. That is an avenue for future research.

**The Smooth Transition Autoregression**

The inception of STAR type econometric models is attributed to Bacon and Watts (1971), who pioneered the concept of a smooth transition regression. Subsequently, Chan and Tong (1986) advocated the use of this model in a time series context. Luukkonen et al. (1988) and Teräsvirta (1994) introduced and developed smooth STAR modeling and testing frameworks. A brief outline of the key characteristics of a STAR model follows.

To begin, consider a linear autoregressive model of order *p*, AR(p):

(1)

where is a realization of a covariance-stationary time series in period *t*; , *i=0,…,p*, are parameters defining the dynamic properties of the model; and is a white noise process.

The linearity restriction in the foregoing model can be relaxed in many ways. One such way is the smooth transition autoregressive modeling framework of Teräsvirta (1994). This framework introduces a specific type of regime--dependency in the model:

(2)

where is the transition function bounded by zero and one. The value of this function is attained from the transition variable (), and the smoothness () and centrality () parameters. While any covariance-stationary variable can be used as the transition variable, often the lagged dependent variable is applied in practice.[[2]](#footnote-2) That is, , , where *d* is a positive integer, sometimes referred to as the delay factor. Such model falls within the family of self-exciting autoregressive models. Indeed, the model converges to the self-exciting threshold autoregression when . The non-negative smoothness parameter determines the speed-of-transition between the regimes, wherein the regimes are centred around the centrality parameter. For the low values of , the transition between the regimes is gradual. As , depending on the functional form of , the switch between the regimes becomes instantaneous and the model converges to a threshold autoregression, if the transition function is *logistic*, or the model reduces to a linear autoregression, in the case of the *exponential* transition function.

Indeed, logistic and exponential transition functions are the two popular choices for STAR-type regime-dependent modeling. The functional forms of these transition functions are given by:

(logistic) (3)

(exponential) (4)

The popularity of these functions has its economic rationale. The sigmoid-shaped logistic function is suitable in situations where asymmetries in autoregressive dynamics in relation to the transition variable are suspected. For example, when the price stickiness in one direction is the underlying feature of price dynamics. In this case, the centrality parameter is the inflection point of the transition function. The inverted bell-shaped exponential function is useful for situations where nonlinearity in dynamics is linked to the deviation of from the centrality parameter. Such dynamics may be linked with the role transaction costs play in the way prices adjust to market shocks (Goodwin et al., 2011). Both these functions are normalized by the sample standard deviation of the transition variable (), which deems the smoothness parameter unit-free.

Whether a logistic or an exponential transition function is better suited to approximate the regime-dependent dynamics of the time series, or whether the time series is, indeed, characterized by a STAR-type nonlinearity, are testable hypotheses. The tests need to be carried out in an auxiliary regression setting, as proposed by Luukkonen et al. (1988), to circumvent the unidentified nuisance parameter issue, better known as the Davies’ problem (Davies, 1977, 1987). The auxiliary regression is the result of the Taylor series expansion of the transition function around that, in effect, interacts the polynomials of the transition variable with the lagged dependent variables of the linear autoregressive model. That is:

(5)

where combines the original error term, and the approximation error due to the Taylor series expansion. The null of linearity is equivalent to the joint hypothesis test of . The framework is also suited to decide on the type of the transition function. The test against the logistic STAR is equivalent to tests of and , while the test against the exponential STAR is equivalent to a test of .

The testing framework is similar for multistep models, with an exception that the vector of the lagged dependent variables, on the right-hand-side, becomes , where *h* is the horizon for which a multistep forecast is generated. See Enders and Pascalau (2015) for further details of this testing procedure.

**Direct and Iterated Multistep Forecasts**

For a time series , a one-step-ahead point forecast—regardless of the econometric model (i.e., linear vs. nonlinear), or the forecasting method (i.e., direct vs. iterated)—is given by:

, (6)

where is the functional form of our best guess of the true model, is the information set, and is the vector of parameter estimates.

If is a linear model, multistep iterated point forecasts can be obtained recursively using naive extrapolation, i.e., by substituting forecasts from the preceding horizons into the model. For example, in the case of a linear autoregressive model, an $h$-step-ahead iterated point forecast is:

, (7)

where for , , and where , are the parameter estimates from regressing on .

When , is nonlinear, the foregoing naive extrapolation could lead to biased multistep forecasts (e.g., Teräsvirta, 2006). We can avoid this bias via a numerical approximation using a bootstrap procedure. The procedures is that in each iteration, the forecast path is ‘disturbed’ with a sequence of bootstrapped forecast errors, , which are sampled (with replacement) from the residuals of the estimated model. A multistep iterated point forecast from a nonlinear model then is the mean of the array of forecast paths at a given horizon. For example, in the case of a smooth transition autoregressive model, an h-step-ahead point forecast is:

, where

, (8)

*B* is the total number of bootstrap iterations. is the transition function where, depending on the horizon length and the delay factor, the transition variable is either the realized lagged value of the dependent variable or its forecast. That is,

The direct forecasting method is applicable both for linear and nonlinear models. For a linear autoregressive model, an *h*-step-ahead direct forecast is:

, (9)

where , are the parameter estimates from regressing on . For a smooth transition autoregressive type of nonlinear model, an *h*-step-ahead direct forecast is:

, (10)

where , as before, is the transition function, bounded by zero and one, and which depends on the transition variable, which is observed at the time when the forecast is made. With direct forecasts, to obtain point forecasts, there is no need to generate empirical distributions of the forecast (although the option is available if desired).

**Model Selection and Forecasting**

Commodity price data were sourced from the online portals of the World Bank and the International Monetary Fund. After discarding the series with missing (or repeated) observations during the January 1980 – December 2020 period, 53 commodity price series were retained. Not all these prices exhibited STAR-type nonlinearities, however. Thus, retained were 25 commodity prices that presented evidence of STAR-type nonlinear dynamics, based on Teräsvirta (1994)’s linearity test, applied in each estimation window (see below). As a result, the list consists of several important commodity groups such as grains and oilseeds, timber, fertilizers, and industrial and rare metals. Figure 1 illustrates the log-transformed nominal price series (in what follows, *y* denotes the log-transformed nominal price, simply referred to as ‘price’).

**Text, letter

Description automatically generated**

**Figure 1: Time series of the commodity prices used in the analysis**

Presented are log-transformed monthly nominal price series, obtained from the online portals of the World Bank (https://www.worldbank.org/en/research/commodity-markets) and the International Monetary Fund (https://www.imf.org/en/Research/commodity-prices), spanning the January 1980 – December 2020 period. Appendix Table A1 offers the brief description of these commodity prices.

The study applies the first-differenced price series (i.e., the prices are assumed to be integrated of order one). For each commodity and each forecast horizon, in each estimation window the autoregressive lag, *p*, is based on the Schwartz Information Criterion. The estimation window excludes the data from the first 24 months, which are reserved for lag selection and for testing linearity against multistep STAR models. In each estimation window the null hypothesis of linearity is tested against STAR alternatives using lagged dependent variables as candidate transition variables, i.e., , where is the first-difference operator, and where the maximum allowed delay factor is set to the selected autoregressive order. This is to decide on the horizon-specific suitable transition variable, as well as the form of a transition function—logistic or exponential—as discussed in the previous section. Appendix Table A2 presents the summary of the model specifications across estimation windows.

**Table 1: Direct multistep STAR-type nonlinearity**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Commodity** | **Horizon** | | | | | | | | | | | |
|  | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10* | *11* | *12* |
| Bananas | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.67 | 1.00 | 0.00 | 0.00 | 0.30 |
| Oranges | 1.00 | 0.31 | 0.08 | 0.63 | 0.71 | 0.91 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 |
| Cotton | 1.00 | 1.00 | 0.03 | 1.00 | 1.00 | 0.00 | 0.00 | 0.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Rice | 1.00 | 1.00 | 1.00 | 0.95 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Lamb | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.78 | 0.18 | 0.45 | 0.00 | 0.00 |
| Hides | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Fishmeal | 1.00 | 0.94 | 1.00 | 1.00 | 1.00 | 0.91 | 0.78 | 0.62 | 0.59 | 0.14 | 0.08 | 0.21 |
| Groundnut Oil | 1.00 | 1.00 | 1.00 | 1.00 | 0.90 | 0.82 | 0.77 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Groundnuts | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Olive Oil | 1.00 | 0.01 | 0.59 | 0.33 | 0.06 | 0.05 | 0.00 | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 |
| Palm Oil | 1.00 | 0.91 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Rapeseed Oil | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.16 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Soybean Oil | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.86 | 0.00 | 0.00 | 0.00 | 0.23 | 1.00 |
| Rubber | 1.00 | 1.00 | 1.00 | 1.00 | 0.68 | 0.00 | 0.00 | 0.00 | 0.01 | 0.45 | 0.97 | 1.00 |
| Hard Sawnwood | 1.00 | 0.59 | 0.00 | 0.02 | 1.00 | 1.00 | 0.51 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Plywood | 1.00 | 0.74 | 1.00 | 0.00 | 0.46 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Soft Logs | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.55 | 1.00 | 0.90 |
| Soft Sawnwood | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.00 | 1.00 | 0.00 | 0.56 | 0.68 | 0.00 |
| Diammonium Phosphate | 1.00 | 1.00 | 0.44 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Urea | 1.00 | 1.00 | 0.80 | 0.97 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Aluminum | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.99 | 1.00 | 1.00 | 1.00 |
| Gold | 1.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.89 | 0.00 | 0.00 | 0.00 | 0.94 | 0.34 | 0.37 |
| Lead | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Platinum | 1.00 | 1.00 | 1.00 | 0.00 | 0.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Uran | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |

The table entries indicate the proportions of estimation windows where the null hypotheses of linearity were rejected at 5 percent significance level. A total of 87 estimation windows are considered. The first estimation window spans the January 1982 – October 2012 period. The next estimation window spans the January 1982 – November 2012 period. All the subsequent estimation windows are expanded in a similar fashion, so that the very last estimation window spans the January 1982 – December 2019 period. Each estimation window yields forecasts for the subsequent twelve months. For horizon 1, the hypotheses are associated with Teräsvirta (1994)’s linearity test; for horizons 2 to 12, the hypotheses are associated with its multistep variant, i.e., Enders and Pascalau (2015)’s linearity pre-test.

Regarding multistep STAR models, a clarifying note is in order. While in all one-step specifications, the null of linearity is rejected in favour of the STAR alternative (recall, only such price series are retained in the dataset), there are instances when the null of linearity is not rejected in favour of the multistep STAR alternative. This is not surprising, and indeed is, to some extent, expected. Table 1 presents the proportions of probability values that are less than 0.05 over all considered expanding windows. The probability values are of the null hypothesis of linearity as per Teräsvirta (1994), for horizon 1, and its Enders and Pascalau (2015) multistep variant, for horizons 2 through 12. For the sake of generating forecasts, even when the null hypotheses of linearity are not rejected for horizons 2 through 12, the more likely STAR model (i.e., logistic or exponential) is estimated, nonetheless.

The first estimation window uses 80 percent of the series for model estimation, thus retaining the remaining 20 percent of the series for out-of-sample evaluation of the one-to-twelve-months-ahead forecasts. To generate the sequence of forecasts, the expanding window approach is applied. Thus, the first estimation window spans the January 1982 – October 2012 period, yielding forecasts for months of November 2012 through October 2013. Similarly, the second estimation window spans the January 1982 – November 2012 period, yielding forecasts for December 2012 through November 2013. All the subsequent estimation windows are expanded in a similar fashion, the very last estimation window yielding forecasts for months of January 2020 through December 2020.

The expanding window scheme, coupled with the model selection routine, presents the set-up that closely mimics a forecaster’s behaviour. Specifically, a forecaster considers the available information set to decide on the most suitable model to generate short and intermediate term forecasts. Over time, as new data enters the information set, the forecaster re-specifies their model as needed. The forecasting routine in this analysis takes into the account such adjustments.

Note that the goal is to assess forecast accuracy of the series in levels, but the modeled dependent variable is in first-differences. To account for this, following Marcellino et al. (2006), we should work with slightly modified versions of equations (7) and (8) for iterated forecasts, and of equations (9) and (10) for direct forecasts. In the case of the iterated method, first the forecast path of the first-differences up to horizon *h* is generated, then an h-step-ahead forecast   is calculated where , are obtained iteratively, as described above. For the direct method, an *h*-step-ahead forecast becomes: , where the last term is the point forecast of , which, in the case of a linear model, for example, is obtained by first regressing on , and then applying the parameter estimates of this regression on .

**Forecast Evaluation and Discussion**

For each commodity price series, and for each considered horizon, two sets of nonlinear forecasts, and , are obtained from iterated and direct STAR specifications. Iterated multistep forecasts from a one-step STAR model are generated using 5,000 bootstrap projections as outlined in the previous section. The horizon-specific averages of these projections form the iterated multistep forecasts. Direct multistep forecasts are generated from horizon-specific STAR models (for which no bootstrapping is necessary). In addition, two sets of linear forecasts, and , using AR specifications (i.e., the linear counterparts of the previous two forecasts) are generated.

Forecast accuracy is evaluated under the assumption of a quadratic loss—a widely accepted loss function, particularly as it relates to the implicit criterion for the in-sample fitting of the data (i.e., minimizing the sum of squared residuals). Specifically, for a given forecast, , the out-of-sample forecast error is (here, and in what follows, the method and model superscripts are omitted for ease of notation). The quadratic loss leads to the root mean squared forecast error (RMSFE), given by:

where denotes the period at which the first forecast is made, and denotes the period at which the last forecast is made. The relative accuracy of forecasts from two competing methods (or models) is assessed by comparing the respective RMSFE measures. Specifically, the ratios of the RMSFE measures of the competing forecast pairs are obtained. Thus, the values below unity imply the method in the numerator is more accurate, and vice versa.

To gain further insights about the extent (or, rather, the statistical significance) of the relative accuracy of competing methods, horizon-specific tests for the equal predictive accuracy as per Diebold and Mariano (1995), using Harvey et al. (1997)–modified test statistics are performed. In addition, as a single measure of multi-horizon accuracy, the average superior predictive ability method of Quaedvlieg (2021) for each multistep horizon (i.e., for horizons 2 to 12) is also obtained.

*Does STAR method lead to more accurate multistep forecasts?*

Table 2 presents the RMSFE ratios of the iterated STAR method relative to the iterated AR method. Overall, there is no evidence that the nonlinear modeling benefits price forecasting. Indeed, if we only focus on commodities with statistically significant relative accuracy measures (i.e., cotton, hides, rubber, and lead, among just a few others), there is strong evidence that AR models outperform STAR models in generating accurate multistep forecasts (the only exception is platinum prices at long horizons where the STAR forecasts appear to be more accurate than the AR forecasts). That STAR models do not yield more accurate forecasts than their AR counterparts is not an unusual finding, which accords with the existing literature. The possible explanations include parameter non-constancy in the time series, as well as misspecification and over-fitting of the nonlinear model (Teräsvirta, 2006).

Table 3 presents the RMSFE ratios of the direct STAR method relative to the direct AR method. The general point, made in the context of iterated methods, is valid in this setting as well: when we focus on statistically significant relative accuracy measures, there is strong evidence that direct AR forecasts outperform direct STAR forecasts. The commodities in which this advantage is observed is different from the previous group, however. Specifically, the direct AR method yields more accurate multistep forecasts than the direct STAR method for bananas, rapeseed and soybean oils, and rubber, among a select few other commodities. Whereas the iterated AR method outperforms the iterated STAR method in generating more accurate price forecasts of cotton, hides, rubber, and lead.

*Does direct method lead to more accurate multistep forecasts?*

First, consider forecast accuracy of the iterated STAR method relative to the direct STAR method. Table 4 presents the results together with the statistical significance identifiers. Overall, the iterated STAR method appears to be more accurate than the direct STAR method in multistep commodity price forecasting (e.g., rice, groundnuts, olive oil, soybean oil, platinum), but there are also instances where the opposite is the case (e.g., cotton, rubber).

**Table 2: Relative forecast accuracy of iterated STAR vs iterated AR**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Commodity** | **Horizon** | | | | | | | | | | | |
|  | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10* | *11* | *12* |
| Bananas | 1.05 | 1.10\* | 1.13 | 1.14 | 1.10 | 1.05 | 1.01 | 1.00 | 0.99 | 1.00 | 1.01 | 1.01 |
| Oranges | 1.02 | 0.99 | 0.98 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.01 | 1.01 | 1.01 | 1.01 |
| Cotton | 1.03 | 1.04\*† | 1.04\*† | 1.04\*† | 1.04\*† | 1.04\*† | 1.03† | 1.02† | 1.02 | 1.01 | 1.01 | 1.01 |
| Rice | 1.06\* | 1.02 | 1.01 | 1.00 | 1.00 | 0.99 | 0.99 | 1.00 | 0.99 | 0.99 | 0.99 | 0.99 |
| Lamb | 0.99 | 1.00 | 1.00 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 |
| Hides | 0.97 | 0.99 | 1.00 | 1.00 | 1.01 | 1.02 | 1.02\* | 1.03\*† | 1.04\*† | 1.05\*† | 1.05\*† | 1.05\*† |
| Fishmeal | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.98 | 0.98 | 0.99 | 0.99 | 0.99 | 0.99 |
| Groundnut Oil | 1.02 | 1.02 | 1.01 | 1.01 | 1.01 | 1.01 | 1.02 | 1.02\* | 1.02 | 1.02 | 1.02 | 1.01 |
| Groundnuts | 0.99 | 0.99 | 0.99 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Olive Oil | 1.03 | 1.01 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.99 | 1.00 | 1.00 | 1.00 |
| Palm Oil | 1.04\* | 1.04\*† | 1.01† | 1.00 | 1.00 | 1.01 | 1.01 | 1.02 | 1.02 | 1.01 | 1.01 | 1.01 |
| Rapeseed Oil | 0.98 | 0.99 | 1.00 | 1.00 | 1.00 | 1.01 | 1.01 | 1.01 | 1.01 | 1.00 | 1.00 | 0.99 |
| Soybean Oil | 0.99 | 1.01 | 1.01 | 1.00 | 1.00 | 1.00 | 1.01 | 1.01 | 1.00 | 1.00 | 1.00 | 1.00 |
| Rubber | 1.04\* | 1.03† | 1.02 | 1.03 | 1.03\*† | 1.04\*† | 1.04\*† | 1.05\*† | 1.06\*† | 1.05\*† | 1.05† | 1.05† |
| Hard Sawnwood | 1.00 | 1.00 | 0.99 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Plywood | 1.00 | 0.98 | 0.99 | 0.99 | 0.99 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Soft Logs | 1.01 | 1.01 | 1.00 | 1.00 | 0.99 | 1.00 | 0.99 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 |
| Soft Sawnwood | 0.97 | 0.99 | 0.99 | 1.00 | 1.00 | 1.01 | 1.00 | 1.00 | 0.99 | 0.99 | 0.99 | 0.99 |
| Diammonium Phosphate | 1.01 | 1.00 | 1.00 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.01 | 1.01 |
| Urea | 1.02 | 1.02 | 1.03 | 1.02\* | 1.02 | 1.02 | 1.02 | 1.02 | 1.04 | 1.05 | 1.05† | 1.05 |
| Aluminum | 1.04\* | 1.03\*† | 1.01 | 1.00 | 1.00 | 1.00 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 |
| Gold | 0.96 | 0.98 | 1.00 | 1.00 | 0.99 | 0.99 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Lead | 1.04 | 1.02\*† | 1.03\*† | 1.03\*† | 1.02† | 1.02\*† | 1.03\*† | 1.02\*† | 1.02† | 1.02† | 1.02 | 1.02 |
| Platinum | 1.02 | 1.01 | 1.01\*† | 1.01† | 1.00 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99\* | 0.99\* | 0.98\* |
| Uran | 1.02 | 0.99 | 0.97 | 0.95 | 0.94 | 0.98 | 1.02 | 1.05 | 1.06 | 1.05 | 1.04 | 1.03 |

The entries are RMSFE ratios of the iterated STAR method relative to the iterated AR method. Thus, the values below unity imply that the forecasts from the iterated STAR method are more accurate than those from the iterated AR method. \* denotes 5% statistical significance of the loss differential of the given horizon, and † denotes 5% statistical significance of the sum of loss differentials up to a given horizon, as per Quaedvlieg (2021). Both tests are based on the heteroskedasticity and autocorrelation consistent t distribution of the Harvey et al. (1997)–modified Diebold and Mariano (1995) statistic.

Next, consider the forecast accuracy of the iterated STAR method relative to the direct AR method. This is of specific interest to the present study as one of the main arguments for the use of a direct method to generate forecasts is when the one--step model is possibly mis-specified. Table 5 presents this comparison. One key take-away from these results is that when the null of equal forecast accuracy is rejected, the direct AR method typically generates more accurate multistep forecasts than the iterated STAR method, even though the STAR-type nonlinearity is the characterizing feature of these price dynamics. There are some exceptions to this general observation. At relatively short horizons (up to four steps ahead) the iterated STAR method offers more accurate price forecasts of orange; at longer horizons, the iterated STAR method offers more accurate (albeit not statistically significantly) price forecasts of lumber and wood commodities.

**Table 3: Relative forecast accuracy of direct STAR vs direct AR**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Commodity** | **Horizon** | | | | | | | | | | | |
|  | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10* | *11* | *12* |
| Bananas | 1.05 | 1.20\*† | 1.17\*† | 1.08† | 1.06 | 1.07 | 1.04 | 1.20\*† | 1.09† | 1.02† | 1.04† | 1.01† |
| Oranges | 1.02 | 1.00 | 0.99 | 0.96\* | 0.97 | 0.97 | 0.99 | 1.00 | 0.99 | 0.99 | 1.02\* | 1.01 |
| Cotton | 1.03 | 1.02† | 1.00 | 0.99 | 0.98 | 0.98 | 0.99 | 1.00 | 1.03 | 1.04 | 1.04 | 1.04 |
| Rice | 1.06\* | 0.99 | 0.99 | 1.01 | 0.99 | 0.99 | 1.00 | 1.02 | 1.03 | 1.05 | 1.06 | 1.04 |
| Lamb | 0.99 | 0.99 | 0.99 | 0.98 | 0.98 | 0.98 | 0.97\* | 0.98 | 0.99 | 0.99 | 1.00 | 0.99 |
| Hides | 0.97 | 1.00 | 0.98 | 1.01 | 1.01 | 0.99 | 0.99 | 1.00 | 1.00 | 1.01 | 1.01 | 1.00 |
| Fishmeal | 0.99 | 0.98 | 0.97 | 0.97 | 0.97 | 0.97 | 0.96 | 0.96 | 0.96 | 0.97 | 0.98 | 0.97 |
| Groundnut Oil | 1.02 | 1.01 | 1.01 | 1.00 | 0.98 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 1.00 | 1.00 |
| Groundnuts | 0.99 | 1.00 | 1.00 | 1.02 | 1.02 | 1.03 | 1.04 | 1.03 | 1.04 | 1.05 | 1.03 | 1.02 |
| Olive Oil | 1.03 | 1.00 | 1.01 | 1.00 | 1.03 | 1.01 | 1.02 | 1.02 | 1.01 | 1.01 | 1.03 | 1.02 |
| Palm Oil | 1.04\* | 1.02† | 1.03† | 1.00 | 1.02 | 1.01 | 1.02\* | 1.01 | 1.02 | 0.99 | 1.01 | 1.02 |
| Rapeseed Oil | 0.98 | 0.99 | 1.01 | 1.01 | 1.03\*† | 1.02\*† | 1.03\*† | 1.02† | 1.03\*† | 1.04† | 1.05† | 1.05 |
| Soybean Oil | 0.99 | 1.02 | 1.03 | 1.04\*† | 1.04 | 1.05 | 1.05\* | 1.05 | 1.06 | 1.06\* | 1.07\* | 1.07\*† |
| Rubber | 1.04\* | 1.02 | 1.02 | 1.02 | 1.02 | 1.02\*† | 1.02\*† | 1.03\*† | 1.02† | 0.99† | 1.02† | 1.02† |
| Hard Sawnwood | 1.00 | 0.98 | 0.99 | 0.99 | 1.01 | 1.02 | 0.97 | 0.95 | 0.95 | 0.95 | 0.95 | 0.95 |
| Plywood | 1.00 | 0.98 | 1.00 | 1.02 | 0.99 | 1.02 | 1.01 | 1.01 | 1.01 | 1.00 | 1.01 | 1.03\* |
| Soft Logs | 1.01 | 1.02 | 1.00 | 1.01 | 1.00 | 1.01 | 1.01 | 0.99 | 1.01 | 0.98 | 1.03 | 0.97 |
| Soft Sawnwood | 0.97 | 1.00 | 1.01 | 1.05 | 1.02 | 1.03\* | 1.00 | 1.00 | 1.05 | 1.00 | 1.04 | 1.02\* |
| Diammonium Phosphate | 1.01 | 1.01 | 1.01 | 1.03 | 1.02 | 1.02 | 1.03 | 1.00 | 1.01 | 1.01 | 1.01 | 1.01 |
| Urea | 1.02 | 1.03† | 1.01 | 1.01 | 1.06\*† | 1.03 | 1.02 | 1.00 | 1.02 | 1.02 | 1.00 | 1.00 |
| Aluminum | 1.04\* | 1.03† | 1.01 | 1.03 | 1.03† | 1.01 | 1.05\*† | 1.04† | 0.99 | 0.99 | 1.04 | 0.99 |
| Gold | 0.96 | 0.99 | 1.00 | 1.00 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 |
| Lead | 1.04 | 1.02\*† | 0.99 | 0.99 | 1.00 | 1.03 | 1.03 | 0.99 | 0.97 | 0.97 | 0.95 | 0.96 |
| Platinum | 1.02 | 1.01 | 1.03\*† | 1.03† | 0.99 | 1.00 | 1.01 | 1.02 | 1.01 | 1.01 | 1.00 | 1.01 |
| Uran | 1.02 | 0.94 | 0.94 | 1.00 | 0.96 | 0.98 | 1.01 | 1.03 | 1.02 | 1.02 | 1.04 | 1.03 |

The entries are RMSFE ratios of the direct STAR method relative to the direct AR method. Thus, the values below unity imply that the forecasts from the direct STAR method are more accurate than those from the direct AR method. \* denotes 5% statistical significance of the loss differential of the given horizon, and † denotes 5% statistical significance of the sum of loss differentials up to a given horizon, as per Quaedvlieg (2021). Both tests are based on the heteroskedasticity and autocorrelation consistent t distribution of the Harvey et al. (1997)–modified Diebold and Mariano (1995) statistic.

**Table 4: Relative accuracy of the iterated STAR method vs direct STAR method**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Commodity** | **Horizon** | | | | | | | | | | | |
|  | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10* | *11* | *12* |
| Bananas | - | 0.92\*† | 0.96 | 1.05\* | 1.02 | 0.99 | 1.00 | 0.85 | 0.92 | 0.98 | 0.98 | 1.04 |
| Oranges | - | 0.96 | 0.95 | 1.00 | 1.01 | 1.02 | 1.00 | 1.00 | 1.00 | 1.00 | 0.96 | 0.96 |
| Cotton | - | 1.02 | 1.05 | 1.05\* | 1.08\*† | 1.09\*† | 1.07† | 1.04† | 1.02† | 1.00 | 0.98 | 0.99 |
| Rice | - | 1.02 | 1.02 | 1.00 | 1.02 | 1.01 | 1.00 | 0.98 | 0.98 | 0.96\* | 0.95\* | 0.97 |
| Lamb | - | 1.00 | 1.02 | 1.01 | 0.99 | 0.98 | 0.99 | 1.00 | 0.99 | 0.99 | 0.99 | 0.99 |
| Hides | - | 0.98 | 1.01 | 0.99 | 1.00 | 1.03 | 1.03 | 1.03 | 1.04 | 1.04 | 1.04 | 1.05 |
| Fishmeal | - | 1.01 | 1.02 | 1.01 | 1.01 | 1.02 | 1.02 | 1.03 | 1.02 | 1.02 | 1.02 | 1.03 |
| Groundnut Oil | - | 1.01 | 1.01 | 1.02 | 1.05 | 1.05 | 1.05 | 1.05 | 1.05 | 1.04 | 1.04 | 1.04 |
| Groundnuts | - | 0.99 | 0.99 | 0.98 | 0.98 | 0.96\* | 0.96\*† | 0.96† | 0.96 | 0.95\*† | 0.96† | 0.97 |
| Olive Oil | - | 1.00 | 0.98 | 0.98 | 0.97\* | 0.98 | 0.97 | 0.96 | 0.97\*† | 0.96† | 0.96† | 0.96 |
| Palm Oil | - | 1.01 | 0.98 | 1.01 | 0.99 | 1.00 | 0.99 | 1.00 | 0.99 | 1.01 | 0.99 | 0.99 |
| Rapeseed Oil | - | 0.99 | 0.99 | 0.98 | 0.97 | 0.98 | 0.98 | 0.99 | 0.98 | 0.97 | 0.95 | 0.94 |
| Soybean Oil | - | 0.98 | 0.98 | 0.98 | 0.99 | 0.97 | 0.98 | 0.98 | 0.97 | 0.96 | 0.95\* | 0.95\* |
| Rubber | - | 1.01 | 1.00 | 1.00 | 1.02\*† | 1.03\*† | 1.03\*† | 1.03† | 1.04\*† | 1.05\*† | 1.04† | 1.04 |
| Hard Sawnwood | - | 1.01 | 1.00 | 0.98 | 0.97 | 0.96 | 1.01 | 1.04\* | 1.03 | 1.04\* | 1.03 | 1.04 |
| Plywood | - | 0.99 | 0.98 | 0.96 | 0.98 | 0.96 | 0.96 | 0.97 | 0.96 | 0.97 | 0.97 | 0.94 |
| Soft Logs | - | 1.00 | 1.00 | 0.99 | 1.00 | 0.99 | 0.99 | 0.98 | 0.98 | 1.00 | 0.97 | 1.01 |
| Soft Sawnwood | - | 0.99 | 0.98 | 1.01 | 1.07 | 1.04 | 1.04 | 1.04 | 0.99 | 1.03 | 0.98 | 1.01 |
| Diammonium Phosphate | - | 1.00 | 1.01 | 1.00 | 0.99 | 0.99 | 0.97 | 1.02 | 1.00 | 1.01 | 1.01 | 1.02 |
| Urea | - | 0.98 | 1.02 | 1.03 | 1.00 | 1.01 | 1.02 | 1.02 | 1.03 | 1.04 | 1.04 | 1.03 |
| Aluminum | - | 0.99 | 0.99 | 0.98 | 0.98 | 0.99 | 0.95 | 0.96 | 1.00 | 0.99 | 0.95\* | 1.00 |
| Gold | - | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.01 | 1.01 | 1.01 | 1.01 |
| Lead | - | 1.00 | 1.04 | 1.04 | 1.02 | 0.99 | 0.99 | 1.03 | 1.06 | 1.05 | 1.06 | 1.06 |
| Platinum | - | 0.99 | 0.98\*† | 0.98† | 1.00 | 0.98\* | 0.98\* | 0.98\*† | 0.98† | 0.98\*† | 0.98† | 0.98\*† |
| Uran | - | 1.07 | 1.05 | 0.99 | 1.00 | 0.99 | 0.98 | 0.99 | 1.02 | 1.02 | 1.01 | 1.01 |

The entries are RMSFE ratios of the iterated STAR method relative to the direct STAR method. Thus, the values below unity imply that the forecasts from the iterated STAR method are more accurate than those from the direct STAR method. \* denotes 5% statistical significance of the loss differential of the given horizon, and † denotes 5% statistical significance of the sum of loss differentials up to a given horizon, as per Quaedvlieg (2021). Both tests are based on the heteroskedasticity and autocorrelation consistent t distribution of the Harvey et al. (1997)–modified Diebold and Mariano (1995) statistic.

Finally, consider the forecast accuracy of the iterated AR method relative to the direct AR method. Table 6 presents this comparison. In most cases, there is very little difference between the two methods. Notable exceptions include prices of oranges for which the iterated AR model offers more accurate forecasts, and prices of cotton and groundnut oil, for which the direct AR method appears to be more suitable as a forecasting tool.

**Table 5: Relative accuracy of the iterated STAR method vs direct AR method**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Commodity** | **Horizon** | | | | | | | | | | | |
|  | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10* | *11* | *12* |
| Bananas | 1.05 | 1.10\*† | 1.13\*† | 1.13† | 1.09 | 1.06 | 1.04 | 1.02 | 1.00 | 1.00 | 1.02 | 1.05 |
| Oranges | 1.02 | 0.97 | 0.94\*† | 0.96\*† | 0.98 | 0.99 | 0.99 | 1.00 | 1.00 | 0.99 | 0.98 | 0.97 |
| Cotton | 1.03 | 1.05\*† | 1.05\*† | 1.04\*† | 1.06\*† | 1.07\*† | 1.06\*† | 1.05† | 1.05† | 1.04† | 1.02† | 1.03† |
| Rice | 1.06\* | 1.02 | 1.01 | 1.01 | 1.01 | 1.00 | 1.01 | 1.00 | 1.01 | 1.01 | 1.01 | 1.00 |
| Lamb | 0.99 | 1.00 | 1.00 | 0.99 | 0.96 | 0.96 | 0.97 | 0.97 | 0.98 | 0.98 | 0.99 | 0.99 |
| Hides | 0.97 | 0.98 | 0.99 | 1.00 | 1.01 | 1.02 | 1.02\* | 1.03\* | 1.04\*† | 1.05\*† | 1.05\*† | 1.06\*† |
| Fishmeal | 0.99 | 0.99 | 1.00 | 0.99 | 0.99 | 0.99 | 0.98 | 0.98 | 0.99 | 0.99 | 1.00 | 0.99 |
| Groundnut Oil | 1.02 | 1.02 | 1.02 | 1.02 | 1.03 | 1.04\*† | 1.04\*† | 1.04† | 1.04† | 1.04† | 1.04 | 1.04 |
| Groundnuts | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.99 | 0.99 |
| Olive Oil | 1.03 | 1.00 | 0.99 | 0.99 | 0.99 | 0.99 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 |
| Palm Oil | 1.04\* | 1.03\*† | 1.01† | 1.01 | 1.01 | 1.02 | 1.01 | 1.01 | 1.01 | 1.00 | 1.00 | 1.01 |
| Rapeseed Oil | 0.98 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.01 | 1.01 | 1.00 | 1.00 | 0.99 | 0.99 |
| Soybean Oil | 0.99 | 1.00 | 1.02 | 1.02 | 1.03 | 1.02 | 1.03 | 1.03 | 1.03 | 1.03 | 1.02 | 1.02 |
| Rubber | 1.04\* | 1.03\*† | 1.03† | 1.03\*† | 1.04\*† | 1.05\*† | 1.06\*† | 1.06\*† | 1.06\*† | 1.05† | 1.06\*† | 1.06† |
| Hard Sawnwood | 1.00 | 1.00 | 0.98 | 0.97 | 0.97 | 0.98 | 0.98 | 0.99 | 0.98 | 0.98 | 0.98 | 0.98 |
| Plywood | 1.00 | 0.98 | 0.98 | 0.98 | 0.97 | 0.98 | 0.97 | 0.97 | 0.98 | 0.97 | 0.98 | 0.97 |
| Soft Logs | 1.01 | 1.01 | 1.00 | 0.99 | 1.00 | 1.00 | 1.00 | 0.97 | 0.98 | 0.99 | 1.00 | 0.98 |
| Soft Sawnwood | 0.97 | 0.99 | 0.99 | 1.06 | 1.09 | 1.08 | 1.04 | 1.05 | 1.04 | 1.02 | 1.01 | 1.04 |
| Diammonium Phosphate | 1.01 | 1.01 | 1.02 | 1.02 | 1.01 | 1.01 | 1.00 | 1.01 | 1.02 | 1.02 | 1.02 | 1.03 |
| Urea | 1.02 | 1.01 | 1.03\* | 1.04\*† | 1.05† | 1.03 | 1.05 | 1.02 | 1.05† | 1.06† | 1.04† | 1.03† |
| Aluminum | 1.04\* | 1.02† | 1.00 | 1.02 | 1.01 | 1.00 | 1.00 | 1.00 | 0.99 | 0.99 | 0.99 | 0.99 |
| Gold | 0.96 | 0.98 | 1.00 | 1.00 | 0.99 | 0.99 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Lead | 1.04 | 1.02\*† | 1.04 | 1.03 | 1.01 | 1.02\* | 1.02 | 1.03 | 1.03 | 1.01 | 1.01 | 1.01 |
| Platinum | 1.02 | 1.00 | 1.01 | 1.00 | 1.00 | 0.99 | 0.99 | 1.00 | 0.99 | 0.99 | 0.98 | 0.98 |
| Uran | 1.02 | 1.00 | 0.99 | 0.99 | 0.96 | 0.97 | 0.99 | 1.03 | 1.04 | 1.04 | 1.05 | 1.04 |

The entries are RMSFE ratios of the iterated STAR method relative to the direct AR method. Thus, the values below unity imply that the forecasts from the iterated STAR method are more accurate than those from the direct AR method. \* denotes 5% statistical significance of the loss differential of the given horizon, and † denotes 5% statistical significance of the sum of loss differentials up to a given horizon, as per Quaedvlieg (2021). Both tests are based on the heteroskedasticity and autocorrelation consistent t distribution of the Harvey et al. (1997)–modified Diebold and Mariano (1995) statistic.

*Discussion*

The foregoing analysis indicates that in most instances the autoregressive methods offer more accurate forecasts than the STAR-type methods, and that the results are ambiguous or inconclusive when comparing iterated and direct forecast methods. For ease of illustration, Figure 2 presents the graphical summary of the previously presented relative accuracy measures.

**Table 6: Relative accuracy of the iterated AR method vs direct AR method**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Commodity** | **Horizon** | | | | | | | | | | | |
|  | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10* | *11* | *12* |
| Bananas | - | 1.00 | 1.00 | 0.99 | 0.99 | 1.01 | 1.02 | 1.02 | 1.01 | 1.00 | 1.01 | 1.04 |
| Oranges | - | 0.98 | 0.95\*† | 0.97 | 0.98 | 0.99 | 0.99 | 0.99 | 0.99 | 0.98† | 0.97 | 0.96 |
| Cotton | - | 1.00 | 1.00 | 1.00 | 1.01\* | 1.03\*† | 1.02\*† | 1.02† | 1.03\*† | 1.02† | 1.02† | 1.02 |
| Rice | - | 0.99 | 1.00 | 1.00 | 1.01 | 1.01 | 1.01 | 1.01 | 1.02 | 1.03 | 1.03 | 1.01 |
| Lamb | - | 1.00 | 1.00 | 1.00 | 0.97 | 0.97 | 0.98 | 0.98 | 0.98 | 0.99 | 0.99 | 0.99 |
| Hides | - | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Fishmeal | - | 1.00 | 1.00 | 1.00 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.01 | 1.00 |
| Groundnut Oil | - | 1.00 | 1.01 | 1.01\*† | 1.02\*† | 1.02† | 1.02 | 1.02 | 1.02 | 1.02 | 1.03 | 1.03 |
| Groundnuts | - | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.99 |
| Olive Oil | - | 0.99 | 0.99 | 0.98\*† | 0.99 | 0.99 | 0.98 | 0.99 | 0.98 | 0.98 | 0.99 | 0.99 |
| Palm Oil | - | 0.99 | 1.00 | 1.01 | 1.01 | 1.01 | 1.00 | 1.00 | 0.99 | 0.99 | 0.99 | 1.00 |
| Rapeseed Oil | - | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Soybean Oil | - | 1.00 | 1.01 | 1.01 | 1.02 | 1.02 | 1.02 | 1.02 | 1.02 | 1.02 | 1.02 | 1.02 |
| Rubber | - | 1.00 | 1.00 | 1.00 | 1.00 | 1.02 | 1.01 | 1.01 | 1.00 | 0.99 | 1.00 | 1.01 |
| Hard Sawnwood | - | 1.00 | 0.99 | 0.97 | 0.97 | 0.98 | 0.98 | 0.99 | 0.98 | 0.98 | 0.98 | 0.98 |
| Plywood | - | 1.00 | 1.00 | 1.00 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.97 | 0.97 | 0.96 |
| Soft Logs | - | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.01 | 0.98 | 0.98 | 0.99 | 1.00 | 0.98 |
| Soft Sawnwood | - | 1.00 | 1.00 | 1.06 | 1.09 | 1.06 | 1.04 | 1.05 | 1.04 | 1.03 | 1.02 | 1.05 |
| Diammonium Phosphate | - | 1.01 | 1.02 | 1.03 | 1.01 | 1.01 | 1.00 | 1.01 | 1.01 | 1.02 | 1.02 | 1.02 |
| Urea | - | 0.99 | 1.01 | 1.02 | 1.04 | 1.02 | 1.02 | 1.00 | 1.01 | 1.01 | 0.99 | 0.99 |
| Aluminum | - | 0.99 | 0.99 | 1.02 | 1.01 | 1.00 | 1.01 | 1.01\* | 1.00 | 1.00 | 1.00 | 1.00 |
| Gold | - | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Lead | - | 1.00 | 1.00 | 1.00 | 0.99 | 1.00 | 0.99\* | 1.00 | 1.00 | 0.99 | 0.99 | 1.00 |
| Platinum | - | 1.00 | 1.00 | 1.00 | 1.00† | 0.99 | 1.00 | 1.01 | 1.00 | 1.00 | 1.00 | 1.00 |
| Uran | - | 1.00 | 1.02 | 1.04 | 1.02 | 1.00 | 0.97 | 0.97 | 0.98 | 0.99 | 1.01 | 1.01 |

The entries are RMSFE ratios of the iterated AR method relative to the direct AR method. Thus, the values below unity imply that the forecasts from the iterated AR method are more accurate than those from the direct AR method. \* denotes 5% statistical significance of the loss differential of the given horizon, and † denotes 5% statistical significance of the sum of loss differentials up to a given horizon, as per Quaedvlieg (2021). Both tests are based on the heteroskedasticity and autocorrelation consistent t distribution of the Harvey et al. (1997)–modified Diebold and Mariano (1995) statistic.

Several features become apparent from these graphs. First, the direct STAR to direct AR and the iterated STAR to direct STAR relative accuracy measures appear as the mirror images of each other, suggesting that the notable discrepancies in these relative accuracy measures are likely driven by the direct STAR forecasts. That said, there are instances when the direct STAR forecasts are more accurate than direct AR and iterated STAR methods (e.g., in the case of fishmeal, groundnut oil, and to some extent cotton and lead), and there are also instances when the opposite is the case (e.g., groundnuts, olive oil, rapeseed oil, soybean oil, among others).

Diagram, engineering drawing, schematic

Description automatically generated

**Figure 2: Relative forecast accuracy among considered method pairs**

iSTAR and dSTAR denote iterated and direct STAR methods; iAR and dAR denote iterated and direct AR methods. RMSFE Ratio is the ratio of the root mean square forecast errors of the competing method pairs. For example, iSTAR/iAR denotes the RMSFE of the iterated STAR method relative to the RMSFE of the iterated AR method. The values less than unity indicate better forecast accuracy of the method featured in the numerator of the ratio. For example, in the case of Bananas, we observe that at relatively short horizons both iterated and direct AR methods yield more accurate forecasts than the iterated or direct STAR methods, at longer horizons, except for the direct STAR method, all methods offer similar forecast accuracy. For each horizon, the RMSFE are obtained from 87 forecast errors. At a given horizon, a forecast error is the difference between the realized value of the log-transformed price and its point forecast.

Second, the iterated STAR to iterated AR, the iterated STAR to direct AR, and iterated AR to direct AR relative accuracy measures, often, tend to co-move across horizons within commodities. These ratios do not deviate from unity compared to the previously discussed ratios, and typically, they remain on the same side of unity, suggesting that a one-step model generates comparably accurate short term and intermediate term forecasts. Some exceptions include platinum with somewhat more accurate short term direct forecasts but relatively more accurate iterated forecasts at longer horizons, and uranium with more accurate short term iterated forecasts but relatively more accurate direct multistep forecasts at longer horizons.

Diagram, engineering drawing

Description automatically generated

**Figure 3: Commodity price series and their multistep forecasts in 2018**

The series and forecasts are for log-transformed prices. The graphs illustrate the snapshots of the commodity price series and their forecasts during the twelve calendar months of 2018. In all instances, the forecasts are made using information set up to and including December 2017. That is, the January forecast are one-step-ahead forecasts, the February forecasts are two-step-ahead forecasts, and so on until the December forecasts, which are twelve-step-ahead forecasts. These forecasts are obtained using iterated and direct STAR methods, denoted by iSTAR and dSTAR, as well as their linear autoregressive counterparts, denoted by iAR and dAR, respectively.

The foregoing discussion indicates that depending on a commodity, any given method can outperform others, and that there is no clear *ex ante* evidence that would facilitate a recommendation of a given forecasting method for a given commodity group. To elaborate on the point, Figure 3 illustrates a snapshot of one-to-twelve step ahead forecasts for months January through December of 2018 from the four considered methods. Granted this is just a single snapshot, in conjunction with the previously provided evidence, several additional observations can be made. First, all four models tend to yield comparable forecasts, and these forecasts can be visibly off the mark, in the wake of considerably fluctuating commodity prices. Second, there is still considerable difference in multistep forecasts, particularly between the direct STAR method and the other three methods, that can result in notable differences in forecast accuracy measures, as documented above.

**Conclusion**

Short and intermediate term changes in prices can considerably impact revenues of producers and distributors of primary commodities. Perhaps more importantly, commodity price shocks can have sizeable macroeconomic consequences on emerging and developing economies (e.g., Gelos and Ustyugova, 2017; Drechsel and Tenreyro, 2018), and can impact directly as well as indirectly households within these economies (e.g., Ivanic and Martin, 2008). To that end, a decision maker’s ability to predict commodity price trends with some degree of accuracy can be of paramount importance.

A range of tools are usually available to a decision maker to analyse the data and make predictions. These tools, or models can range from overly simplistic to very complex. The present study considers two sets of specific models, linear autoregressions and a form of regime-dependent autoregressions, smooth transition autoregressions, and empirically investigates possible advantages the latter may offer for intermediate term commodity price forecasting. Specifically, the study addresses questions of whether the STAR framework can offer superior forecast accuracy relative to the basic autoregressive framework, and whether an iterated STAR method may be preferred over a direct STAR method.

From an array of primary commodity prices, sourced from the World Bank and the International Monetary Fund data portals, the study focuses on 25 series that feature STAR-type nonlinear dynamics. Based on out-of-sample forecast evaluation, the study finds that while STAR models may well approximate nonlinear dynamics of commodity price series, when it comes to multistep forecasting, the framework offers little advantage, and in most instances appears to be an inferior alternative to the basic autoregressive framework. While for select commodities a case can be made that a nonlinear method yields more accurate forecasts, general recommendation—based on the results of this study—is that linear methods are to be seen as a safer approach for their multistep forecasting, particularly when forecasts are made for longer horizons.

This finding is useful, particularly considering the popularity of STAR-type models in commodity price analysis. A caveat is that the conclusion here refers to this one specific family of nonlinear models. That the considered nonlinear specification does not yield more accurate forecasts does not imply that others will not. Even within the family of regime-dependent models, addressing shifts and switches in the price series (as in Enders and Holt, 2012, for example) can be one of many potential avenues for future research.

**References**

Bacon, D., and D. Watts (1971). Estimating the Transition between Two Intersecting Straight Lines. Biometrika 58 (3), 525–534.

Balagtas, J.V., and M.T. Holt (2009). The Commodity Terms of Trade, Unit Roots, and Nonlinear Alternatives: A Smooth Transition Approach. American Journal of Agricultural Economics 91 (1), 87–105.

Chan, K., and H. Tong (1986). On Estimating Thresholds in Autoregressive Models. Journal of Time Series Analysis 7 (3), 179–190.

Chevillon, G. (2007). Direct Multi-Step Estimation and Forecasting. Journal of Economic Surveys 21 (4), 746–785.

Chevillon, G., and D.F. Hendry (2005). Non-Parametric Direct Multi-Step Estimation for Forecasting Economic Processes. International Journal of Forecasting 21 (2), 201–218.

Davies, R. (1977). Hypothesis Testing when a Nuisance Parameter is Present only under the Alternative. Biometrika 64 (2), 247–254.

Davies, R. (1987). Hypothesis Testing when a Nuisance Parameter is Present only under the Alternative. Biometrika 74 (1), 33–43.

Diebold, F., and R. Mariano (1995). Comparing Predictive Accuracy. Journal of Business & Economic Statistics 13 (3), 253–263.

Drechsel, T., and S. Tenreyro (2018). Commodity Booms and Busts in Emerging Economies. Journal of International Economics 112, 200–218.

Enders, W., and M.T. Holt (2012). Sharp Breaks or Smooth Shifts? An Investigation of the Evolution of Primary Commodity Prices. American Journal of Agricultural Economics 94 (3), 659–673.

Enders, W., and R. Pascalau (2015). Pretesting for Multi-step-ahead Exchange Rate Forecasts with STAR Models. International Journal of Forecasting 31 (2), 473–487.

Gelos, G., and Y. Ustyugova (2017). Inflation Responses to Commodity Price Shocks–How and Why Do Countries Differ? Journal of International Money and Finance 72, 28–47.

Goodwin, B.K., M.T. Holt, and J.P. Prestemon (2011). North American Oriented Strand Board Markets, Arbitrage Activity, and Market Price Dynamics: A Smooth Transition Approach. American Journal of Agricultural Economics 93 (4), 993–1014.

Harvey, D., S. Leybourne, and P. Newbold (1997). Testing the Equality of Prediction Mean Squared Errors. International Journal of Forecasting 13 (2), 281–291.

Holt, M.T., and L.A. Craig (2006). Nonlinear Dynamics and Structural Change in the U.S. Hog-Corn Cycle: A Time-Varying STAR Approach. American Journal of Agricultural Economics 88 (1), 215–233.

Hood, H.B., and J.H. Dorfman (2015). Examining Dynamically Changing Timber Market Linkages. American Journal of Agricultural Economics 97 (5), 1451–1463.

Ivanic, M. and W. Martin (2008). Implications of Higher Global Food Prices for Poverty in Low-Income Countries. Agricultural Economics 39, 405–416.

Luukkonen, R., P. Saikkonen, and T. Teräsvirta (1988). Testing Linearity Against Smooth Transition Autoregressive Models. Biometrika 75 (3), 491–499.

Marcellino, M., J.H. Stock, and M.W. Watson (2006). A Comparison of Direct and Iterated Multistep AR Methods for Forecasting Macroeconomic Time Series. Journal of Econometrics 135 (1), 499–526.

Quaedvlieg, R. (2021). Multi–Horizon Forecast Comparison. Journal of Business & Economic Statistics 39 (1), 40–53.

Stock, J. H., and M.W. Watson (1998). A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series. NBER Working Paper No. 6607.

Teräsvirta, T. (1994). Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models. Journal of the American Statistical Association 89 (425), 208–218.

Teräsvirta, T. (2006). Forecasting Economic Variables with Nonlinear Models. In Handbook of Economic Forecasting, Volume 1, Chapter 8, pp. 413–457. Elsevier.

Teräsvirta, T., D. Tjøstheim, and C.W.J. Granger (2010). Modelling Nonlinear Economic Time Series. Advanced Texts in Econometrics. Oxford University Press.

Ubilava, D. (2012). El Niño, La Niña, and World Coffee Price Dynamics. Agricultural Economics 43 (1), 17–26.

Ubilava, D. (2018). The Role of El Niño Southern Oscillation in Commodity Price Movement and Predictability. American Journal of Agricultural Economics 100 (1), 239–263.

Weiss, A.A. (1991). Multi-Step Estimation and Forecasting in Dynamic Models. Journal of Econometrics 48 (1-2), 135–149.

**Appendix Tables**

**Table A1: Details of the commodity prices used in the analysis**

|  |  |  |
| --- | --- | --- |
| *Commodity* | *Source* | *Details* |
| Bananas | IMF | Bananas, Central American and Ecuador, FOB U.S. Ports, US$/mt |
| Oranges | WB | Oranges (Mediterranean exporters) navel, European Union indicative import price, CIF Paris, US$/kg |
| Cotton | IMF | Cotton, Cotton Outlook `A Index', Middling 1-3/32 inch staple, CIF Liverpool, US cents/lb |
| Rice | IMF | Rice, 5% broken milled white rice, Thailand nominal price quote, US$/mt |
| Lamb | IMF | Lamb, frozen carcass Smithfield London, US cents/lb |
| Hides | IMF | Hides, Heavy native steers, over 53 pounds, wholesale dealer's price, US, Chicago, fob Shipping Point, US cents/lb |
| Fishmeal | IMF | Fishmeal, Peru Fish meal/pellets 65% protein, CIF, US$/mt |
| Groundnut Oil | WB | Groundnut oil, U.S. crude, FOB South-East beginning January 1999; previously any origin, CIF Rotterdam, US$/mt |
| Groundnuts | WB | Groundnuts, from January 1999, Runners 40/50, CFR N.W. Europe; previously, Runners 40/50 shelled basis, CIF Rotterdam, US$/mt |
| Olive Oil | IMF | Olive Oil, extra virgin less than 1% free fatty acid, ex-tanker price U.K., US$/mt |
| Palm Oil | WB | Palm oil (Malaysia), from December 2001, RBD, CIF Rotterdam; previously Malaysia 5%, CIF N.W. Europe, bulk, nearest forward, US$/mt |
| Rapeseed Oil | IMF | Rapeseed oil, crude, fob Rotterdam, US$/mt |
| Soybean Oil | WB | Soybean oil, from January 1999, Dutch crude degummed, FOB NW Europe; previously crude, FOB ex-mill Netherlands, nearest forward, US$/mt |
| Rubber | WB | Rubber (Asia) RSS3 grade, from January 2004, Singapore Commodity Exchange Ltd (SICOM) nearby contract; from January 2000 to December 2003, Singapore RSS1; previously Malaysia RSS1, US$/kg |
| Hard Sawnwood | IMF | Hard Sawnwood, Dark Red Meranti, select and better quality, C&F U.K port, US$/m$^3$ |
| Plywood | WB | Plywood (Africa and Southeast Asia), Lauan, 3-ply, extra, 91cm$times$182cm$times$4mm, wholesale price, spot Tokyo, US cents/sheet |
| Soft Logs | IMF | Soft Logs, Average Export price from the U.S. for Douglas Fir, US$/m$^3$ |
| Soft Sawnwood | IMF | Soft Sawnwood, average export price of Douglas Fir, U.S. Price, US$/m$^3$ |
| Diammonium Phosphate | WB | Diammonium Phosphate (DAP), spot, FOB US Gulf, US$/mt |
| Urea | IMF | US Gulf NOLA Urea Granular Spot Price, US$/mt |
| Aluminum | IMF | Aluminum, 99.5% minimum purity, LME spot price, CIF UK ports, US$/mt |
| Gold | WB | Gold (UK), 99.5% fine, London afternoon fixing, average of daily rates, US$/oz |
| Lead | IMF | Lead, 99.97% pure, LME spot price, CIF European Ports, US$/mt |
| Platinum | WB | Platinum (UK), 99.9% refined, London afternoon fixing, US$/oz |
| Uran | IMF | Uranium, NUEXCO, Restricted Price, Nuexco exchange spot, US$/lb |

**Table A2: Summary of model specifications across estimation windows**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Commodity* | *p* | *logistic* | *d* |  |
| Bananas | 10.00 | 1.00 | 8.00 | 1.00 |
| Oranges | 9.47 | 0.53 | 1.47 | 0.00 |
| Cotton | 2.00 | 1.00 | 1.00 | 0.00 |
| Rice | 3.00 | 1.00 | 1.00 | 0.00 |
| Lamb | 3.78 | 0.99 | 1.99 | 0.00 |
| Hides | 3.00 | 1.00 | 1.00 | 0.00 |
| Fishmeal | 2.00 | 0.00 | 1.00 | 0.00 |
| Groundnut Oil | 2.00 | 1.00 | 1.00 | 0.00 |
| Groundnuts | 2.00 | 1.00 | 1.00 | 0.00 |
| Olive Oil | 2.00 | 0.00 | 1.00 | 0.00 |
| Palm Oil | 5.00 | 1.00 | 1.00 | 0.00 |
| Rapeseed Oil | 1.00 | 1.00 | 1.00 | 0.00 |
| Soybean Oil | 3.00 | 1.00 | 2.00 | 0.00 |
| Rubber | 2.00 | 1.00 | 1.00 | 0.00 |
| Hard Sawnwood | 2.00 | 0.00 | 1.00 | 0.00 |
| Plywood | 2.00 | 1.00 | 1.00 | 1.00 |
| Soft Logs | 2.00 | 1.00 | 1.00 | 0.00 |
| Soft Sawnwood | 2.17 | 1.00 | 1.00 | 0.17 |
| Diammonium Phosphate | 2.00 | 0.00 | 1.00 | 1.00 |
| Urea | 3.75 | 1.00 | 1.00 | 0.00 |
| Aluminum | 2.00 | 0.00 | 1.00 | 0.00 |
| Gold | 1.06 | 0.94 | 1.00 | 0.40 |
| Lead | 2.00 | 0.00 | 1.00 | 0.00 |
| Platinum | 2.00 | 1.00 | 1.00 | 0.00 |
| Uran | 6.00 | 0.70 | 1.00 | 1.00 |

The table entries are the averages over 87 expanding windows. *p* denotes the autoregressive lag length. *logistic*, which denotes the logistic STAR function, indicates the fraction of logistic STAR function applied over the considered 87 expanding windows; thus, the value of 1 indicates the logistic STAR function whereas the value of 0 indicates the exponential STAR function across all considered windows (recall, the null of linearity is rejected in all windows). *d* denotes the delay factor of the lagged dependent variable used as the transition variable, which is in all instances, and where is a positive integer. indicates the fraction of estimation windows where the null of parameter constancy was rejected as per Teräsvirta (1994)’s test.

1. david.ubilava@sydney.edu.au. Data and replication material available at https://github.com/dubilava/multistep. [↑](#footnote-ref-1)
2. In principle, any covariance-stationary variable, whether already a regressor in the linear specification or not, can serve as the transition variable. To that end, the possibilities of a STAR-type nonlinear specification are virtually unlimited. This is an important caveat of the present analysis, which only considers a concrete type of nonlinear specification, and a specific subset of candidate transition variables. That said, the type of nonlinearity considered by this study is a generalization of the self-exciting threshold autoregressive (SETAR) framework—a widely applied technique to address complex dynamics in economic time series, including commodity price series. Nonetheless, a way to think about this limitation of the proposed framework is that the considered nonlinear specification is merely an approximation of the true potentially complex dynamics of commodity price series. A corollary is that the results of this study offer a ‘lower bound’ of forecast accuracy that can be achieved by nonlinear modeling of the considered commodity price series. [↑](#footnote-ref-2)