Journal of Agricultural and Applied Economics, 41,1(April 2009):253–270

The Asymmetric Cycling of U.S. Soybeans and Brazilian Coffee Prices: An Opportunity for Improved Forecasting and Understanding of Price Behavior

Octavio A. Ramirez

The behavior of agricultural commodity markets can arguably result in markedly asymmetric price cycles, that is, downward cycles of substantially different length and breadth than upward cycles. This study assesses whether asymmetric-cycle models can enhance the understanding of the dynamics and provide for a better forecasting of U.S. soybeans and Brazilian coffee prices. The forecasts from asymmetric cycle models are found to be substantially mode precise than those obtained from standard autoregressive models. The asymmetric cycle models also provide useful insights on the markedly different dynamics of the upward versus the downward cycles exhibited by the prices of these two commodities.

Key Words: asymmetric cycles, coffee prices, soybean prices, threshold autoregressive models, time series forecasting

JEL Classifications: C22, C53

Basic economic theory suggests that the behavior of agricultural commodity prices over time is governed by shifts supply and demand. The impact of those shifts on prices depends on the elasticities characterizing the supply and demand relations. In a long-run dynamic framework, multi-period impacts will also be affected by other factors such as storability, irreversibility of capital investment, and the ease with which farmers can switch into producing alternative crops.

In the case of most agricultural commodities, it can be argued that the demand is relatively more stable than supply. Systematic exogenous demand shifters such as population and demographic changes, disposable income, and so on, usually exhibit smooth gradual trends; and random demand shocks tend not to be very pronounced. Economic slowdowns, however, can cause prolonged albeit not very pronounced downward shifts in demand, particularly when international, developing country, markets are involved. Given that the supply for raw agricultural commodities exhibits a relatively low elasticity, even in a multi-period time frame, such demand shifts can result in several consecutive years of generally depressed prices.

Supply can be substantially more volatile. In the case of soybeans and other field crops, for example, unusually adverse climatic or pest conditions affecting key producing regions such as the Midwestern U.S. can effect large, although short-lived, negative supply shocks. Under an inelastic demand, such shocks can

Octavio Ramirez is professor and head, Department of Agricultural and Applied Economics, University of Georgia, Athens, GA.

result in substantial upward price spikes. Since it may take a few years for stocks to replenish to acceptable levels, those spikes can spill over and cause several consecutive years of high prices. Positive supply shocks of comparable magnitude, however, are highly unlikely, as evidenced by the significant left-skewness of Midwest yield distributions identified by numerous authors (Ker and Coble; Moss and Shonkwiler; Nelson and Preckel; Ramirez; Taylor). Thus, supply-induced downward price spikes of similar scale are not expected. The duration of low price cycles, however, can be exacerbated by the irreversibility of the capital investments made during good price periods in order to increase production capacity, particularly if there is a high correlation with the prices of alternative crops. A final argument for asymmetric cycling behavior in the case of soybeans is the fact that, for several decades, agricultural policies have affected the functioning of the market during low price periods.

A somewhat analogous situation is observed in the case of coffee prices. The two main differences are that coffee is a perennial crop and that coffee prices are not substantially affected by agricultural price protection policies. As in the case of soybeans, coffee yields are fairly stable on the up side. However, worldwide supply has been subject to considerable shortages several times during the last few decades, mainly due to freezes damaging the crop coming out of Brazil, which is a major producer. It is hypothesized that this supply effect is responsible for the extreme spikes that are periodically observed in coffee prices. The fact that Brazilian production usually takes more than one year to fully recover from a severe freeze would then cause several consecutive years of high prices.

Although there are no substantial agricultural policies supporting coffee prices on the downside, the irreversibility of capital investment is more pronounced than in the case of soybeans because coffee is a perennial crop. Plantations are costly to establish and have to be considered a long-term (15- to 30-year) investment. In addition, in many countries, growing coffee is a family tradition and producers will not consider alternative crops, even in the long run. Thus, when a few years of very high prices allure the planting of additional areas and the replanting of very old nearly unproductive lots, the inevitable result is a long period of below-average prices. This behavior is evidenced by the finding that the probability distribution of world coffee prices in any particular year is substantially right skewed (Ramirez and Somarriba).

Similar arguments to those previously made in the case of soybean and coffee can be advanced about many other important agricultural commodities. Thus, there is no reason to expect that agricultural time series variables such as commodity cash or futures prices, crop acreage, and so on, exhibit symmetric cycles as assumed in the standard autoregressive models. Therefore, models that can account for asymmetric cycling behavior could be very useful to provide a better understanding of the intertemporal dynamics and for an improved forecasting of agricultural time series variables.

The basic tools for modeling asymmetric cycling behavior in time series variables were introduced to the econometrics literature by Tong, who labeled them threshold autoregressive (TAR) models. In essence, in TAR models the autoregressive parameters are allowed to switch values over time as lagged dependent variable observations cross one or more thresholds. The most common TAR model specification exhibits two sets of parameters that apply depending on whether the previous value of the dependent variable is below or above a certain threshold. This implies that the characteristics, that is, the length and breadth, of the upward and downward cycles can be substantially different, or in other words, that the cycles can be asymmetric. In the extreme, there could be upward but no downward cycling behavior or vice versa.

Researchers have explored the use of TAR models in a variety of nonforecasting applications. Petrucelli and Woolford illustrate the estimation and use of a first order threshold autoregressive [TAR(1)] model. Tsay focuses on the testing for TAR processes, while Brockwell, Liu, and Tweedie investigate the existence of stationary TAR moving average processes. Chang evaluates the consistency and limiting distribution of the least squares estimator of a TAR model. Balke and Fomby propose an approach for testing for cointegration in the presence of TAR rather than AR processes.

Recent applications of TAR models include Granger and Lee's investigation of production, sales and inventory relationships using multicointegration and nonsymmetric error-correction models; Potter's analysis of the changes in real U.S. GNP; Bradley and Jansen's crosscountry evaluation of business cycle dynamics; Obstfeld and Taylor's analysis of purchasing power parity and the law of one price under imperfect arbitrage in the presence of transaction costs and uncertainty; and Goodwin and Piggott's evaluation of dairy price linkages among corn and soybean markets in North Carolina. Such research corroborates the potential importance of TAR models in the analysis of agricultural time series.

In light of this background, the main objectives of this article are (a) to develop better time-series forecasting models and add to the understanding of the cyclical behavior and price movements in the U.S. soybean and Brazilian coffee markets, and (b) to evaluate the ability of TAR models to explain the behavior of these two price series in comparison to standard AR models. A collateral result of these objectives is an assessment of the potential of TAR models to provide a better understanding of the inter-temporal dynamics and to improve the forecasting of agricultural prices in general. In the pursuit of its primary objectives, this paper also deals with a methodological issue: the estimation and forecasting from TAR models with systematic components, that is, with explanatory variables. Suitable estimation procedures and the formulas to obtain one-, two-, and three-period-ahead predictions from such TAR models are derived.

The TAR models of quarterly U.S. soybeans future prices and Brazilian coffee spot prices estimated using the proposed procedures are found to render substantially more precise within- and out-of-sample forecasts than the standard AR models. The estimated TAR models also provide useful insights on the markedly different dynamics of the upward versus the downward cycles exhibited by U.S. soybeans and Brazilian coffee prices.

Specification and Estimation of TAR Models

The basic TAR models used in the econometrics literature are specified as follows:

$$
\begin{aligned}
&y_t = \rho_{0p} + \rho_{1p}y_{t-1} + \ldots + \rho_{kp}y_{t-k} \\
(1)\quad &+ \nu_t \quad \text{if } y_{t-d} \geq \text{TR}, \quad \text{and} \quad y_t = \rho_{0n} + \rho_{1n}y_{t-1} \\
&+ \ldots + \rho_{kn}y_{t-k} + \nu_t \quad \text{if } y_{t-d} < \text{TR},
\end{aligned}
$$

where ρ_{ip} and ρ_{in} ($i = 0,\ldots,k$) are autoregressive parameters, v_t is an independently and identically distributed (iid) random variable, TR is a threshold parameter, and d is called the delay lag. This specification is also known as the self-exciting threshold autoregressive (SETAR) model.

More complicated TAR structures with multiple thresholds have also been explored in the econometrics literature (Balke and Fomby; Enders and Granger; Enders and Siklos).

As discussed by Hansen (1997), the model defined in Equation (1) can be estimated by the least squares (LS) method which, under the auxiliary assumption that v_t is $N(0,\sigma^2)$, is equivalent to maximum likelihood (ML). Since the residuals in that model are a nonlinear and discontinuous function of TR, its parameters have to be estimated by sequential conditional LS (SCLS). Under the simplifying assumption that $d = 1$, that is, that the threshold in Equation (1) is defined in relation to y_{t-1}, this method involves computation of the autoregressive parameters and the resulting residual sum of squares (RSS) for select alternative values of TR.

Specifically, for each TR value, the observations are divided into two sets depending on whether $y_{t-1} \geq$ TR or $y_{t-1} <$ TR and ordinary least squares (OLS) is used to compute the corresponding sets of autoregressive coefficients (ρ_{0p}, $\rho_{1p},\ldots,\rho_{kp}$ and ρ_{0n}, $\rho_{1n},\ldots,\rho_{kn}$). The overall model's RSS is then obtained by adding the RSS from those two OLS regressions. Note that, since only $T - k$ observations can be included in them, there will be at most $T - k$ different TR values that will change the sets of observations going into each of the two regressions, that is, the sets will only change if

TR is sufficiently increased to equal the value of the next highest y_{t-1} ($t = k + 1,\ldots,T$) observation. Therefore, the previously described process only has to be conducted for TR = y_{t-1}, $t = k + 1,\ldots,T$. The estimated model with the highest overall RSS across those TR values is assured to have satisfied the conditions for LS (and ML) estimation.

This paper considers the following alternative threshold autoregressive specification for the modeling of U.S. soybeans and Brazilian coffee prices:

$$(2) \qquad y_t = \boldsymbol{x}_\mathbf{t}\boldsymbol{\beta} + e_t,$$

where:

$$e_t = \phi_{1p}e_{t-1} + \phi_{2p}e_{t-2} + \ldots + \phi_{kp}e_{t-k} + \nu_t \quad \text{if } e_{t-1} \geq \text{TR}, \quad \text{and}$$
$$e_t = \phi_{1n}e_{t-1} + \phi_{2n}e_{t-2} + \ldots + \phi_{kn}e_{t-k} + \nu_t \quad \text{if } e_{t-1} < \text{TR},$$

where y_t is the dependent variable of interest, $\mathbf{x_t}$ is a $1 \times m$ vector of exogenous variables, $\boldsymbol{\beta}$ is an $m \times 1$ vector of intercept and slope parameters, v_t is an iid random variable, and TR is the threshold. Equation (2) follows standard practice in the specification of time series models with systematic components, that is, to define the autoregressive process in relation to the error term.

This alternative TAR model specification allows for two different autocorrelation regimes to apply depending on the value of the error term (e_t) during the previous time period. The occurrence of an error (e_{t-1}) greater than or equal to TR prompts the regime implied by the "p" set of autocorrelation parameters $\boldsymbol{\phi}_\mathbf{p} = [\phi_{1p}\ \phi_{2p} \ldots \phi_{kp}]$, while an error that is less than TR sets off the alternative "n" regime implied by $\boldsymbol{\phi}_\mathbf{n} = [\phi_{1n}\ \phi_{2n} \ldots \phi_{kn}]$. This allows for an asymmetric cycling behavior of the error term.

Note that this specification is different from the traditional TAR model [Equation (1)] in two aspects of empirical importance: the inclusion of a systematic component ($\boldsymbol{x}_\mathbf{t}\boldsymbol{\beta}$) and the defining of the threshold in relation to the error term, not the lagged values of the dependent variable. This alternative threshold definition is relevant because, in the presence of a systematic component, it is reasonable to argue that, at least in some applications, a potential switch from one autoregressive regime to another could be related to the magnitude of the error term rather than the dependent variable. In other words, the regime switch might be defined by where the dependent variable value is in relation to its systematic component rather than in reference to a constant. An intuitively appealing threshold is TR = 0, which implies that the cycling behavior of the error term (and thus of the dependent variable) is different when the dependent variable is above its unconditional expected value ($\boldsymbol{x}_\mathbf{t}\boldsymbol{\beta}$) than when it is below it.

Estimation of the Proposed TAR Model

As in the traditional TAR model [Equation (1)], estimation of the proposed threshold autoregressive with systematic component (TARSC) model is based on the least squares method which, under the auxiliary assumption that v_t is $N(0,\sigma^2)$, is also equivalent to maximum likelihood (ML). The residual sum of squares to be minimized in this case is:

$$(3) \qquad \text{RSS} = \sum_{t=k+1}^{T} (r_t)^2$$

where, from Equation (2):

$$r_t = y_t - \boldsymbol{x}_\mathbf{t}\boldsymbol{\beta} - \phi_{1p}e_{t-1} - \phi_{2p}e_{t-2} - \ldots - \phi_{kp}e_{t-k} \quad \text{if } e_{t-1} \geq \text{TR}, \quad \text{and}$$
$$r_t = y_t - \boldsymbol{x}_\mathbf{t}\boldsymbol{\beta} - \phi_{2n}e_{t-2} - \phi_{2n}e_{t-2} - \ldots - \phi_{kn}e_{t-k} \quad \text{if } e_{t-1} < \text{TR}.$$

The estimation problem in this case is characterized by the fact that the RSS is discontinuous with respect to both TR and $\boldsymbol{\beta}$. Specifically, in certain regions of the $\boldsymbol{\beta}$ space, an infinitesimal change in one of the parameters in that vector will cause one of the residuals ($e_{t-1} = y_{t-1} - \boldsymbol{x}_\mathbf{t-1}\boldsymbol{\beta}$) to transition from being below to being above TR, or vice versa, which will switch the set of autoregressive parameter values used to compute r_t and result on a discrete shift in the RSS. Because of these discontinuities, iterative search methods of numerical optimization (Judge et al., pp. 951–979) are generally ineffective in finding the function's global minimum.

Note, however, that when TR and $\boldsymbol{\beta}$ are fixed, the values of the autoregressive parameters that conditionally minimize the RSS can be found by ordinary least squares (OLS). This suggests that, as in the case of the traditional TAR model (Hansen 1997), estimation can be accomplished by Sequential Conditional Least Squares (SCLS). In addition, note that e_{t-1} ($t = 2, \ldots, T$) can be computed for any fixed value of $\boldsymbol{\beta}$. Therefore, as in the traditional TAR case, there are at most $T - k$ TR values that are relevant for estimation (TR $= e_{t-1}$, $t = k + 1, \ldots, T$).

Given the previously discussed characteristics of the minimization problem, the SCLS procedure to estimate the parameters of the proposed TARSC model involves a grid search of the $\boldsymbol{\beta}$ space and limited subsets of the TR space combined with least-squares estimation of $\boldsymbol{\phi_p} = \{\phi_{1p}\ \phi_{2p}\ \ldots\ \phi_{kp}\}$ and $\boldsymbol{\phi_n} = \{\phi_{1n}\ \phi_{2n}\ \ldots\ \phi_{kn}\}$. The proposed steps for estimating a TARSC model are:

(a) Select the grid on the $\boldsymbol{\beta}$ space over which the search is to be conducted. Guidance for narrowing down the hyperspace for the initial grid is provided in the next section.
(b) For each $\boldsymbol{\beta}$ on the grid, compute $e_{t-1} = y_{t-1} - \mathbf{x_{t-1}}\boldsymbol{\beta}$ ($t = k + 1, \ldots, T$) and the select set of $T - k$ TR values that are relevant for estimation (TR $= e_{t-1}$, $t = k + 1, \ldots, T$).
(c) For each relevant TR value, divide the errors ($e_t = y_t - \mathbf{x_t}\boldsymbol{\beta}$) into two sets according to whether or not e_{t-1} is less that TR and estimate $\boldsymbol{\phi_p}$ and $\boldsymbol{\phi_n}$ by OLS regressions of those two sets of errors on their lagged values (e_{t-j}, $j = 1, \ldots, k$).
(d) Using Equation (3), compute the RSS for each $\boldsymbol{\beta}$ on the grid and its corresponding set of TR values from (b) and $\boldsymbol{\phi_p}$ and $\boldsymbol{\phi_n}$ estimates from (c) above, for a total of $T-k$ RSS values. Choose the TR and associated $\boldsymbol{\phi_p}$ and $\boldsymbol{\phi_n}$ estimates that result in the lowest RSS for that particular $\boldsymbol{\beta}$.
(e) Repeat the process for all $\boldsymbol{\beta}$ values on the grid. The $\boldsymbol{\beta}$, TR, $\boldsymbol{\phi_p}$, and $\boldsymbol{\phi_n}$ values that yield the lowest overall RSS are the LS estimates for those parameters.

A simpler OLS-based procedure to estimate the proposed TARSC model is also evaluated in this study. This method involves computing $\boldsymbol{\beta}$ by OLS, dividing this OLS model's errors ($e_t = y_t - \mathbf{x_t}\boldsymbol{\beta}$) into two sets according to whether or not e_{t-1} exceeds a particular TR value, and estimating $\boldsymbol{\phi_p}$ and $\boldsymbol{\phi_n}$ by OLS regressions of those two sets of errors on their lagged values (e_{t-j}, $j = 1, \ldots, k$). This process is repeated over the set of $T - k$ TR values that are relevant for estimation. As in the previous procedure, the "optimal" TR and autoregressive parameter vector estimates are the ones corresponding to the lowest overall RSS. Note that this method is in fact an application of Hansen's standard TAR model estimation procedure to the OLS residuals.

Because of what is known about LS estimation of standard autoregressive models with systematic components (Judge et al., pp 275–349), it is not expected that this method will minimize the RSS [Equation (3)] in the case of TARSC models. However, given the significant computational requirements of the procedure that guarantees RSS minimization, it is important to explore the statistical performance of this simpler alternative.

Estimation and Forecasting Performance

Results from Monte Carlo experiments evaluating the finite sample estimation and forecasting performance of the proposed LS method, the simpler OLS-based procedure, and standard AR models are presented in Appendix A. These results lead to the following general conclusions:

(1) Although the proposed LS-based method (LS-TARSC) produces biased estimates for the intercept (β_0) and autocorrelation coefficients (ϕ_{1p}, ϕ_{2p}, ϕ_{1n}, and ϕ_{2n}), it appears to be a consistent estimator for all of those parameters. The degree of bias steadily decreases with sample size (T) and, in all cases evaluated, the percentage bias is negligible at $T = 10{,}000$.
(2) The standard method for estimating AR models apparently produces biased and inconsistent estimates for the intercept of a TARSC model.
(3) Both the proposed LS and the standard AR method seem to produce unbiased

and consistent estimates for the slope parameter (β_1). However, LS appears to be a more efficient estimator for this parameter. Estimation efficiency differences range from 7% to over 100%.

(4) The forecasts obtained using the formulas provided in Appendix B seem to be unbiased both within and out of sample, although biased LS estimates for the intercept and autoregressive parameters are used to compute them. Interestingly, the predictions from the AR models, obtained using standard formulas, appear to be unbiased as well.

(5) Although AR models can be used to approximate TARSC processes, these approximations are generally far from perfect. The average R^2s values of the AR models are 3% to 24% lower than those obtained when using LS-estimated TARSC models. Moderate to relatively high differences in forecasting precision, as measured by the root mean square of the within- and out-of-sample forecast errors, between the AR and LS-TARSC models are also found. These differences range from 3% to 65% and average approximately 20% for both the one- and the two-period-ahead out-of-sample forecasts, and are somewhat smaller in the case of the three-period-ahead predictions (Appendix A).

(6) The simpler OLS-based alternative for the estimation of TARSC models (OLS-TARSC) also produces biased and inconsistent estimates for the intercept and the autocorrelation parameters. As in the case of the AR model, however, the OLS estimate for the slope parameter seems to be unbiased and consistent. Depending on the magnitude of the bias on intercept estimation, the forecasting precision of the model appears to be somewhere in between that of the standard AR and the LS-TARSC.

(7) As expected, the differences in the root mean square of the one-period-ahead within sample forecasting errors corresponding to the AR, OLS-TARSC, and LS-TARSC models (WFE in Appendix A) are a good relative indicator of the differences in one-, two-, and three-period-ahead out-of-sample forecasting precision (FE1, FE2, and FE3 in Appendix A) across the models.

The previously discussed conclusions hint to the potential of a third estimation procedure for TARSC models. Since both OLS and standard AR methods seem to yield unbiased and consistent estimates for the slope parameters in the systematic component, a modification of the proposed LS procedure in which those parameters are set to their OLS or AR model estimates in the grid search should be considered. The grid search then only has to involve the intercept and its corresponding $T - k$ TR values that are relevant for estimation, which entails simpler programming and low computational demands regardless of the number of slope parameters.

The predictions from the models estimated using this restricted-search LS procedure (RLS-TARSC) appear to be as efficient as those from the full-search LS-TARSC, that is, there is no perceivable difference in these two models' forecasting precision as measured by the R^2, WFE, FE1, FE2, and FE3 (Appendix A). While, in a few cases, the average minimum RSS value is slightly higher for the RLS-TARSC, the averages of the parameter estimates for β_0, ϕ_{1p}, ϕ_{2p}, ϕ_{1n}, and ϕ_{2n} are practically identical. The only difference between these two procedures seems to be in the efficiency of estimation of the slope parameter(s). Therefore, although applied researchers now have access to substantial computational capacity, the RLS-TARSC might be an attractive alternative when estimating TARSC models with a large number of explanatory variables, particularly if efficiency in the estimation of the slope parameters is not a concern. In addition to being simpler to program, depending on the number of explanatory variables involved, the RLS-TARSC can reduce estimation time in relation to the LS-TARSC from several minutes or hours to just a few seconds. Alternatively, the RLS-TARSC intercept and TR estimates can be used as a mid starting point to reduce the scope of the initial grid search in the LS-TARSC procedure.

The Asymmetric Cycling U.S. Soybean and Brazilian Coffee Prices

In this section, TARSC and standard AR models of U.S. soybean futures and Brazilian coffee spot (New York) prices are estimated and compared, with data spanning over the past four decades. The TARSC models are estimated using the previously discussed LS-, RLS-, and OLS-based methods. Both price series are found to be stationary according to the augmented Dickey-Fuller unit root test ($\alpha = 0.10$). All models are initially specified with five autoregressive error term lags and a systematic component ($\boldsymbol{x_t\beta}$) consisting of an intercept and a linear trend. The AR models are estimated utilizing the preprogrammed Gauss 6.0 ARIMA procedure. The LS-, RLS-, and OLS-TARSC models are estimated using Gauss 6.0 code developed as part of this research. All Gauss programs used in these applications will be made available upon request.

In the case of coffee prices, the AR(5) and LS-TARSC(5) models reach minimum RSS values of 72,165.19 and 44,896.40, respectively. Box-Pierce statistics of 10.91 and 16.13 do not reject the null hypothesis that the autocorrelation coefficients between these models' disturbances and their first 20 lags are jointly equal to zero ($\alpha = 0.25$), suggesting that the error terms are independently distributed. A test for whether the difference between the two RSS is statistically significant can be derived from the theories of Davies and Andrews-Ploberger. Specifically, it is known that if the random error term (ν_t) is iid, a test with near-optimal power against alternatives distant from the null hypothesis is the pointwise F-statistic:

$$(4) \qquad F_T = T(\mathrm{RSS_{AR}} - \mathrm{RSS_{TAR}})/\mathrm{RSS_{TAR}}$$

The problem, however, is that since TR is not identified under the null hypothesis, the asymptotic distribution of F_T is not χ^2. Hansen (1996, p. 4) proposes a bootstrapping procedure to approximate the asymptotic distribution of F_T in the case of standard TAR models. Since the models in this paper have systematic components, this procedure is applied on the basis of the AR and TARSC model residuals to obtain the critical values for this test statistic for an α of 0.001. In the case of the coffee price model, the value is 48.01 versus a computed test statistic of $F_T = 92.93$. This implies that the observed RSS difference is significant at an α of 0.001, that is, that the LS-TARSC model is statistically superior to the standard AR.

The fifth-order autoregressive coefficient in the initial AR(5) model is statistically insignificant ($\alpha = 0.10$). Exclusion of this parameter leads to the final AR(4) model presented in Table 1. A Box-Pierce statistic of 11.02 does not reject the null hypothesis of no error term autocorrelation either ($\alpha = 0.25$). The initial LS-TARSC(5) model of Brazilian coffee prices also includes several statistically insignificant autoregressive coefficients ($\alpha = 0.10$). Exclusion of those parameters leads to the final LS-TARSC model (Table 1). As in the AR, a Box-Pierce statistic of 17.79 fails to reject the no autocorrelation hypothesis ($\alpha = 0.25$). The difference between these two models' RSS is significant ($\alpha = 0.001$) as well.

As expected, the R^2 of the LS-TARSC model (0.905) is noticeably higher than the AR's 0.847. The root mean square of the within sample forecast error (RMSFE) is 17.74 cents/lb under the LS-TARSC versus 22.42 cents/lb, or 26.4% higher, under the AR (Table 1). In order to further evaluate forecasting performance differences, 40 additional sets of models are estimated on the basis of samples ending from the first quarter of 1998 to the last quarter of 2007 and one-, two-, and three-period-ahead out-of-sample forecasts for the last 40 quarters are obtained using those models. The RMSE of these out-of-sample forecasts from the LS-TARSC models are 9.99, 17.03, and 23.73 cents/lb, respectively, versus 12.44, 20.86, and 25.97 cents/lb in the case of the AR models, for an average percentage difference of 18.81%. By any standards, these substantial differences in forecasting precision would justify using the more sophisticated LS-TARSC modeling technique.

In addition, the TARSC model provides useful insights into the dynamics of Brazilian coffee price cycles that are quite different from those implied by the statistically inferior AR model. Specifically, 39.6% of the residuals are

Table 1. Key Statistics of Models for Quarterly Brazilian Coffee Spot and U.S. Soybean Future Prices

PR	β_0	β_1	ϕ_{1p}	ϕ_{2p}	ϕ_{3p}	ϕ_{4p}	ϕ_{5p}	ϕ_{1n}	ϕ_{2n}	ϕ_{3n}	ϕ_{4n}	ϕ_{5n}
Final AR Model of Coffee Prices												
PE	135.10	–0.27	1.26	–0.63	0.40	–0.16	0.00	1.26	–0.63	0.40	–0.16	0.00
SE	30.32	0.34	0.08	0.13	0.13	0.08	—	0.08	0.13	0.13	0.08	—
PV	0.00	0.43	0.00	0.00	0.00	0.06	—	0.00	0.00	0.00	0.06	—
	RSS = 72,352.24				WFE = 22.42					$R^2 = 0.847$		
Final LS-TARSC Model of Coffee Prices												
PE	240.32	–0.99	0.81	–0.76	0.45	–0.48	0.00	1.12	–0.15	0.00	0.00	0.00
SE	—	—	0.14	0.17	0.18	0.12	—	0.09	0.09	—	—	—
PV	—	—	0.00	0.00	0.01	0.00	—	0.00	0.09	—	—	—
	RSS = 45,303.04				WFE = 17.74					$R^2 = 0.905$		
Final OLS-TARSC Model of Coffee Prices												
PE	135.10	–0.26	1.35	–0.91	0.75	–0.46	0.00	1.27	–0.31	0.00	0.00	0.00
SE	—	—	0.20	0.33	0.36	0.25	—	0.09	0.09	—	—	—
PV	—	—	0.00	0.01	0.04	0.07	—	0.00	0.00	—	—	—
	RSS = 66,590.47				WFE = 21.21					$R^2 = 0.862$		
Final RLS-TARSC Model of Coffee Prices												
PE	214.22	–0.26	0.91	–1.02	0.89	–0.94	0.38	1.29	–0.32	0.00	0.00	0.00
SE	—	—	0.25	0.31	0.37	0.36	0.23	0.09	0.09	—	—	—
PV	—	—	0.00	0.00	0.02	0.01	0.10	0.00	0.00	—	—	—
	RSS = 56,567.02				WFE = 19.55					$R^2 = 0.881$		
Final AR Model of Soybean Prices												
PE	254.83	2.46	0.91	0.00	0.00	0.00	0.00	0.91	0.00	0.00	0.00	0.00
SE	81.69	0.67	0.03	—	—	—	—	0.03	—	—	—	—
PV	0.00	0.00	0.00	—	—	—	—	0.00	—	—	—	—
	RSS = 587,181.41				WFE = 54.32					$R^2 = 0.911$		
Final LS-TARSC Model of Soybean Prices												
PE	459.05	2.44	0.45	0.00	0.00	–0.29	0.28	0.87	0.00	0.30	0.00	–0.18
SE	—	—	0.16	—	—	0.15	0.14	0.07	—	0.09	—	0.07
PV	—	—	0.01	—	—	0.06	0.04	0.00	—	0.00	—	0.01
	RSS = 504,981.65				WFE = 50.37					$R^2 = 0.923$		
Final OLS-TARSC Model of Soybean Prices												
PE	276.78	2.19	0.78	0.00	0.29	–0.50	0.44	0.92	0.00	0.00	0.00	0.00
SE	—	—	0.10	—	0.17	0.21	0.17	0.04	—	—	—	—
PV	—	—	0.00	—	0.10	0.02	0.01	0.00	—	—	—	—
	RSS = 541,556.44				WFE = 52.17					$R^2 = 0.918$		

Table 1. Continued.

Final RLS-TARSC Model of Soybean Prices											
PR	β_0	β_1	ϕ_{1p}	ϕ_{2p}	ϕ_{3p}	ϕ_{4p}	ϕ_{5p}	ϕ_{1n}	ϕ_{2n}	ϕ_{3n}	ϕ_{4n}	ϕ_{5n}
PE	488.98	2.19	0.28	0.00	0.00	0.00	0.00	0.93	0.00	0.19	0.00	–0.15
SE	—	—	0.18	—	—	—	—	0.07	—	0.08	—	0.07
PV	—	—	0.11	—	—	—	—	0.00	—	0.02	—	0.02
	RSS = 515,245.31				WFE = 50.88					R^2 = 0.922		

Notes: PR, PE, SE, and PV stand for parameter, parameter estimate, standard error estimate, and *p*-value; RSS is the minimum value reached by the residual sum of squares; WFE is the within sample root mean square error of the one-period-ahead forecasts; and the R^2 is computed as the square of the correlation coefficient between the within sample one-period-ahead autoregressive predictions and the actual dependent variable values. Autoregressive parameters that are statistically insignificant at the 10% level have been set equal to zero.

expected to be above the estimated error term threshold (TR = –30.59), that is, 39.6% of the price realizations are anticipated to be above and 60.4% below the estimated price (y_t) threshold PTR = $\mathbf{x_t\beta}$ + TR = 240.32 –0.987t – 30.59 = 209.73 – 0.987t. The dynamics of the upward cycles are found to be very different from those of the downward cycles (Table 2). Only 5.66% of the prices crossing PTR from below are expected not to be followed by at least one more price realization above that threshold, while 13.01% of the prices crossing PTR from above will not be followed by additional price occurrences below PTR. Interestingly, the AR implies that 20.27% of the prices crossing over or under this model's estimated long-term trend equation will go back across the next quarter.

On the other hand, the LS-TARSC model suggests that only 5.78% of the upward cycles, versus 18.78% of the downward cycles, will last just two or three quarters; while the AR model implies that 26.30% of the cycles will be of such length. In contrast, nearly 61% of the upward cycles, but only 25.90% of the downward cycles (and 25.20% of the AR cycles), are predicted to last between four and eight quarters. And, while almost 36% of the downward cycles may last over 10 quarters, less than 19% of the upward cycles (and about 22% of the AR cycles) are expected to be that long.

Another practical advantage of TARSC models is that they are useful to examine possible differences in error term variability and, thus, forecasting precision, in upward versus downward cycles. This can be accomplished by computing separate RMSFE statistics for observations above and below PTR. The LS-TARSC model of coffee prices suggests an expected prediction error of RMSFE_p = 24.35 cents/lb in the upward cycles and of RMSFE_n = 13.08 cents/lb in the downward cycles. This indicates that the level of unpredictable variation in the upward price cycles is nearly twice as high as in the downward cycles; which is evident in the observed Brazilian coffee price data (Figure 1).

In short, the coffee price cycle dynamics implied by the estimated TARSC model are markedly asymmetric and, therefore, vastly different from what can be accommodated by a standard AR model. The LS-TARSC suggests that most of the upward cycles will quickly reach high price levels but be moderately-lived (four to eight quarters), with only about 1 in 10 lasting more than 3 years. In contrast, it predicts that nearly one-third of the downward cycles will be more than 3 years long and close to one-fifth will last in excess of 5 years. Such long downward cycles can take coffee prices to very low levels, but in a gradual manner. Note that this cycling dynamics implied by the estimated TARSC model are consistent with the observed behavior of Brazilian coffee prices (Figure 1).

As previously discussed, the simulation analysis suggests that both AR and OLS yield biased intercept and unbiased but inefficient slope parameter estimates when the error term process is TAR. In this case, the OLS estimates for those parameters (135.10 and –0.26) are noticeably different from the LS-TARSC's (240.32 and –0.987). Because the OLS-TARSC

Table 2. Relative Frequencies of the Cycle Durations Implied by the Estimated LS-TARSC and AR Models for Brazilian Coffee and U.S. Soybean Prices

Cycle Length (quarters)	LS-TARSC–Coffee Prices Upward	LS-TARSC–Coffee Prices Downward	AR–Coffee Prices	LS-TARSC–Soybean Prices Upward	LS-TARSC–Soybean Prices Downward	AR–Soybean Prices
1	5.66%	13.01%	20.27%	21.51%	34.05%	30.12%
2	2.25%	10.41%	16.55%	9.39%	16.20%	14.16%
3	3.53%	8.37%	9.75%	6.52%	4.89%	8.91%
4	7.10%	7.14%	6.55%	9.50%	2.49%	6.36%
5	13.15%	5.80%	5.70%	8.87%	2.86%	4.76%
6	19.22%	4.81%	4.96%	6.41%	2.41%	3.82%
7	13.65%	4.28%	4.25%	4.98%	2.00%	3.18%
8	7.82%	3.87%	3.73%	3.69%	1.81%	2.70%
9	5.24%	3.34%	3.25%	3.00%	1.80%	2.36%
10	3.77%	3.01%	3.03%	3.01%	1.52%	1.96%
11	2.91%	2.73%	2.62%	2.60%	1.37%	1.88%
12	2.55%	2.40%	2.27%	2.39%	1.20%	1.62%
13	2.22%	2.22%	1.99%	2.05%	1.20%	1.48%
14	2.02%	1.96%	1.78%	1.72%	1.08%	1.34%
15	1.65%	1.98%	1.54%	1.66%	0.99%	1.25%
16	1.34%	1.64%	1.35%	1.37%	1.01%	1.10%
17	1.12%	1.56%	1.29%	1.31%	0.94%	0.95%
18	0.87%	1.47%	1.03%	1.10%	0.79%	0.95%
19	0.75%	1.35%	0.95%	0.98%	0.84%	0.82%
20	0.59%	1.27%	0.84%	0.93%	0.74%	0.78%
21	0.44%	1.16%	0.72%	0.77%	0.74%	0.72%
22	0.39%	1.04%	0.67%	0.72%	0.71%	0.63%
23	0.35%	1.03%	0.57%	0.62%	0.66%	0.58%
24	0.25%	0.88%	0.55%	0.53%	0.59%	0.59%
25	0.21%	0.89%	0.47%	0.47%	0.60%	0.50%
26	0.16%	0.81%	0.38%	0.42%	0.53%	0.50%
27	0.14%	0.67%	0.36%	0.40%	0.54%	0.46%
28	0.13%	0.72%	0.31%	0.35%	0.51%	0.41%
29	0.11%	0.68%	0.25%	0.27%	0.49%	0.38%
30	0.08%	0.60%	0.26%	0.31%	0.44%	0.33%
31	0.07%	0.60%	0.22%	0.21%	0.48%	0.33%
32	0.04%	0.54%	0.19%	0.22%	0.45%	0.32%
33	0.05%	0.51%	0.16%	0.19%	0.42%	0.29%
34	0.04%	0.41%	0.14%	0.17%	0.40%	0.25%
35	0.03%	0.43%	0.11%	0.17%	0.39%	0.23%
36	0.02%	0.40%	0.11%	0.13%	0.34%	0.24%

Note: The Gauss program used to compute these frequencies will be made available upon request.

method is based on the OLS estimates, with a R^2 of 0.862, WFE of 21.21 cents/lb and RSS of 66,590.47, the performance of the OLS-TARSC model in this application is closer to the AR's (R^2 = 0.847, WFE = 22.42 cents/lb, RSS = 72,352.24) than to the LS-TARSC's (R^2 = 0.905, WFE = 17.74 cents/lb, RSS = 45,303.04) (Table 1). In contrast, as expected from the simulation results, the RLS-TARSC model performance statistics (R^2 = 0.881, WFE = 19.55 cents/lb, RSS = 56,567.02) are relatively close to those of the LS-TARSC. Therefore, this application supports the contention that the RLS method for estimating TARSC models is a suitable albeit not perfect alternative to the LS procedure.

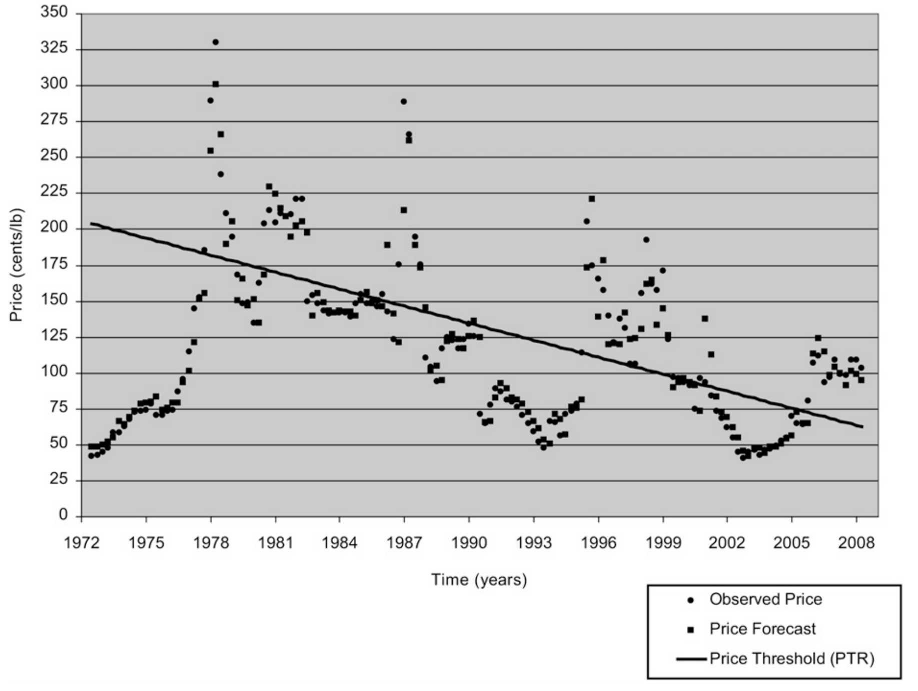

Figure 1. Observed versus Predicted Prices and Price Threshold from the LS-TARSC Model of Brazilian Coffee Prices

In the case of U.S. soybean future prices, the AR(5) and LS-TARSC(5) models reach minimum RSS values of 569,146.27 and 496,065.37, respectively. Box-Pierce statistics of 4.11 and 6.54 do not reject the null hypothesis that the population autocorrelation coefficients between these models' errors and their first 20 lags are jointly equal to zero (α = 0.25). The critical value of the test statistic for the significance of the difference between these two RSS is 30.05, versus a computed value of F_T = 28.63. That is, the observed RSS difference is significant at an α of 0.005, that is, the LS-TARSC model for U.S. soybean future prices is statistically superior to the AR as well.

In the case of the AR(5) model, only the first-order autoregressive coefficient is statistically insignificant (α = 0.10). Exclusion of the nonsignificant parameters leads to the final AR(1) model presented in Table 1. A Box-Pierce statistic of 12.31 does not reject the null hypothesis of no error term autocorrelation either (α = 0.25). The initial LS-TARSC(5) model of U.S. soybean future prices also contains several statistically insignificant autoregressive coefficients (α = 0.10). Exclusion of those parameters leads to the final LS-TARSC model (Table 1). As in the AR, a Box-Pierce statistic of 4.48 fails to reject the no autocorrelation hypothesis (α = 0.25). The null hypothesis of equality between the RSS of these two final models is easily rejected (α = 0.005) as well.

As expected, the LS-TARSC R^2 of 0.923 is higher than the AR's 0.911. Within sample, the one-period-ahead RMSFE is 50.37 cents/bushel under the LS-TARSC versus 54.32 cents/bushel, or 7.85% higher, under the AR. As with coffee prices, 40 additional sets of models are estimated on the basis of samples ending from the first quarter of 1998 to the last quarter of 2007 and one-, two-, and three-period-ahead out-of-sample forecasts for the last 40 quarters are obtained using those models. The RMSE of these out-of-sample forecasts from the LS-TARSC models are 59.74, 96.73, and 139.52, respectively, versus 67.27, 108.75, and 154.54 in the case of the AR models, for an average difference of 11.93%. Although these differences in forecasting precision are not as striking as in the case of coffee prices, they amply justify using the more sophisticated LS-TARSC modeling technique in this application as well.

The TARSC model also provides valuable insights into the dynamics of soybean price cycles. Specifically, 35% of the price realizations are anticipated to be above and 65% below the price threshold PTR = $\mathbf{x_t\beta}$ + TR = 459.05 + 2.44t – 80.38 = 378.66 + 2.44t. As in the case of coffee prices, the dynamics of the upward cycles are very different from those of the downward cycles (Table 2). For instance, only 21.51% of the prices crossing PTR from below are expected not to be followed by at least one more price realization above that threshold equation, while 34.05% of the prices crossing PTR from above will not be followed by additional price occurrences below the threshold. Interestingly, the AR implies that 30.12% of the prices crossing over or under this model's estimated long-term trend equation will go back across the following quarter.

In contrast, according to the estimated TARSC model, over 36% of the high-price cycles will last between three and seven quarters while less than 15% of the downward cycles (and 27.02% of the AR cycles) are expected to be of such duration. In addition, 56.40% of the upward cycles are predicted to span between 3 and 15 quarters while only 25.62% of the low-price cycles (and 41.61% of the AR cycles) are expected to be within that range. On the other hand, only 12.71% of the high-price cycles, versus 24.13% of the downward cycles, are likely to last more than 5 years. As in the case of coffee prices, the cycling behavior implied by the LS-TARSC model appears to match the oscillations of the observed soybean price data (Figure 2).

The LS-TARSC forecast errors are estimated to be RMSFE$_p$ = 67.44 for the upward and RMSFE$_n$ = 44.89 cents/bushel for the downward cycles. That is, as in the coffee price

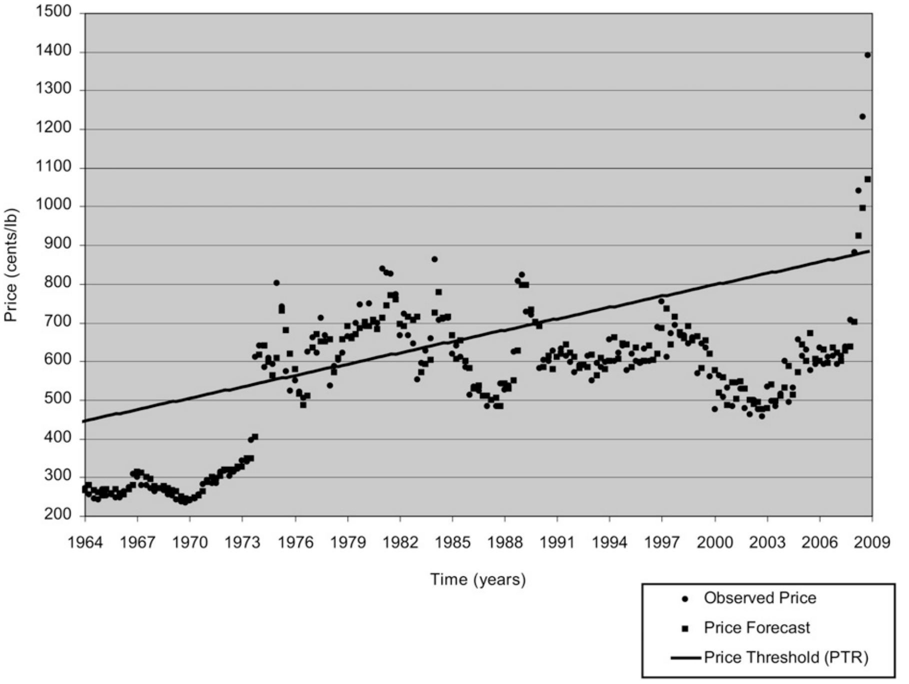

Figure 2. Observed versus Predicted Prices and Price Threshold from the LS-TARSC Model of U.S. Soybean Prices

model, unpredictable variation in the upward price cycles is markedly higher than in the downward cycles. In short, the estimated LS-TARSC model suggests that, while they tend to be more frequent, high price cycles are of moderate duration, with nearly half of them lasting between 3 and 10 quarters. Downward cycles are less frequent, but when they occur, they tend to last longer. Specifically, out of the two-thirds of the low-price observations that extend for more than two consecutive periods, nearly 40% are likely to turn into cycles of more than 10 quarters.

Finally, in this application, the OLS-TARSC model R^2 (0.918), RMSFE (52.17 cents/bushel), and RSS (541,556.44) are more or less in between the AR's (0.911, 54.32, and 587,181.41) and the LS-TARSC's (0.923, 50.37, and 504,981.65) (Table 1). As in the case of coffee prices, the RLS-TARSC model performance statistics (R^2 = 0.922, RMSFE = 50.88 cents/bushel, and RSS = 515,245.31) are quite close to those of the LS-TARSC.

Concluding Remarks

This paper estimates threshold autoregressive (TAR) models for U.S. soybean and Brazilian coffee prices and evaluates their ability to explain the time series behavior of these two price series in comparison to that of standard AR models. The TAR models estimated using the proposed LS-TARSC method provide for substantially more precise forecasts and a much improved understanding of the cycle dynamics of these two price series in comparison to standard autoregressive (AR) models. In general, it is hoped that similar benefits may be reaped by using TARSC models for the analysis of other agricultural time series variables.

Simulation results presented in this article confirm the two main conclusions derived from the previously discussed applications: (a) substantial gains in forecasting precision in relation to the standard AR models can be achieved by using the proposed (LS- or RLS-TARSC) estimation method when the dependent variable is characterized by both a systematic and a random component and the random component follows a TAR rather than an AR process, (b) in such cases, the TARSC models will also provide empirically valuable insights about the differential dynamics of the upward versus the downward cycles of the dependent variable, which are not afforded by standard AR models. In short, researchers interested in thoroughly understanding the cycling behavior and obtaining more reliable forecasts for time series variables should consider using of the proposed procedures to ascertain if a TARSC model is more suitable than a standard AR model in a particular application.

[Received March 2008; Accepted January 2009.]

References

Andrews, D.W.K., and W. Ploberger. “Optimal Tests when a Nuisance Parameter is Present Only Under the Alternative.” *Econometrica* 62(1994):1383–1414.

Balke, N.S., and T.B. Fomby. “Threshold Cointegration.” *International Economic Review* 38,3(1997):627–45.

Bradley, M., and D. Jansen. “Non-Linear Business Cycle Dynamics: Cross-Country Evidence on the Persistence of Aggregate Shocks.” *Economic Inquiry* 35(1997):495–509.

Brockwell, P.J., J. Liu, and R. Tweedie. “On the Existence of Stationary Threshold Autoregressive Moving Average Processes.” *Journal of Time Series Analysis* 13(1992):95–107.

Chang, K.S. “Consistency and the Limiting Distribution of the Least Squares Estimator of a Threshold Autoregressive Model.” *The Annals of Statistics* 21(1993):520–33.

Davies, R.B. “Hypothesis Testing when a Nuisance Parameter is Present Only Under the Alternative.” *Biometrika* 74(1987):33–43.

Enders, W., and C.W.J. Granger. “Unit-Root Tests and Asymmetric Adjustment with an Example Using the Term Structure of Interest Rates.” *Journal of Business & Economic Statistics* 16,3(1998):304–11.

Enders, W., and P.L. Siklos. “Cointegration and Threshold Adjustment.” *Journal of Business & Economic Statistics* 19,2(2001):166–76.

Goodwin, B.K., and N.E. Piggott. “Spatial Market Integration in the Presence of Threshold Effects.” American *Journal of Agricultural Economics* 83(2001):302–17.

Granger, C.W.J., and T.H. Lee. “Investigation of Production, Sales and Inventory Relationships

Using Multicointegration and Non-Symmetric Error-Correction Models." *Journal of Applied Econometrics* 4(1989):S145–59.

Hansen, B.E. "Inference in TAR Models." *Studies in Nonlinear Dynamics and Econometrics, Berkeley Electronic Press* 2,1(1997):1–14.

———. "Inference When a Nuisance Parameter is not Identified Under the Null Hypothesis." *Econometrica* 64(1996):413–30.

Judge, G.G., W.E. Griffiths, R. Carter Hill, H. Lutkepohl, and T.-C. Lee. *The Theory and Practice of Econometrics*. 2nd ed. New York: John Wiley & Sons, Inc., 1985.

Ker, A.P., and K. Coble. "Modeling Conditional Yield Densities." *American Journal of Agricultural Economics* 85,2(2003):291–304.

Moss, C.B., and J.S. Shonkwiler. "Estimating Yield Distributions: Using a Stochastic Trend Model and Nonnormal Errors." *American Journal of Agricultural Economics* 75,5 (1993):1056–62.

Nelson, C.H., and P.V. Preckel. 1989. "The Conditional Beta Distribution as a Stochastic Production Function." *American Journal of Agricultural Economics* 71,2(1989):370–78.

Obstfeld, M., and A.M. Taylor. "Nonlinear Aspects of Goods-Market Arbitrage and Adjustment; Heckscher's Commodity Points Revisited." *Journal of the Japanese and International Economies* 11(1997):441–79.

Petrucelli, J.D., and S. Woolford. "A Threshold AR(1) Model." *Journal of Applied Probability* 21(1984):270–86.

Potter, S. "A Non-Linear Approach to U.S. GNP." *Journal of Applied Econometrics* 10(1995):109–25.

Ramirez, O.A. "Estimation and Use of a Multivariate Parametric Model for Simulating Heteroskedastic, Correlated, Non-Normal Random Variables: The Case of Corn-Belt Corn, Soybeans and Wheat Yields." *American Journal of Agricultural Economics* 79,1(1997):191–205.

Ramirez, O.A., and E. Somarriba. "Risk and Returns of Diversified Cropping Systems under Non-Normal, Cross and Autocorrelated Commodity Price Structures." *Journal of Agricultural and Resource Economics* 25,2(2000):653–68.

Taylor, C.R. "Two Practical Procedures for Estimating Multivariate Non-Normal Probability Density Functions." *American Journal of Agricultural Economics* 72,1(1990):210–17.

Tong, H. *Threshold Models in Non-Linear Time Series Analysis. Lecture Notes in Statistics 21*. Berlin: Springer, 1983.

Tsay, R.S. "Testing and Modeling Threshold Autoregressive Processes." *Journal of the American Statistical Association* 84(1989): 231–40.

Appendix A. Simulation Methods and Results

In the simulation experiments $\mathbf{x_t \beta} = \beta_0 + \beta_1 x_t = -1 + 1x_t$, were x_t is a Bernoulli random variable with P = 0.5. Two TAR(1) [$\phi_{1p} = 0.9$, $\phi_{1n} = 0.0$ and $\phi_{1p} = 0.9$, $\phi_{1n} = -0.8$] and four TAR(2) [$\phi_{1p} = 1.2$, $\phi_{2p} = -0.8$, $\phi_{1n} = 0.8$, $\phi_{2n} = 0$; $\phi_{1p} = 1.5$, $\phi_{2p} = -0.8$, $\phi_{1n} = 0$, $\phi_{2n} = 0$; $\phi_{1p} = 1.3$, $\phi_{2p} = -0.6$, $\phi_{1n} = 0.5$, $\phi_{2n} = 0.4$; and $\phi_{1p} = 1.5$, $\phi_{2p} = -0.8$, $\phi_{1n} = -0.9$, $\phi_{2n} = 0$] processes are assumed in conjunction with this systematic component. Their white noise term variance (σ) is set equal to one. The table below shows select Monte Carlo simulation statistics about LS-, OLS-, and RLS-estimated TARSC models and of standard AR models under these six TAR error term processes and two sample sizes.

Appendix A.

TAR(1) Process with Parameters $\phi_{1p} = 0.9$, $\phi_{2p} = 0.0$, $\phi_{1n} = 0.0$, $\phi_{2n} = 0.0$

EM	T	RSS	β_0	SE_1	ϕ_{1p}	ϕ_{2p}	ϕ_{1n}	ϕ_{2n}	R^2	WFE	FE1	FE2	FE3
AR	L	1.06	0.35	0.076	0.76	—	0.76	—	0.62	1.03	1.04	1.28	1.42
LS	L	1.00	–0.98	0.071	0.89	—	–0.02	—	0.65	1.00	1.00	1.24	1.39
OLS	L	1.06	0.35	0.076	0.83	—	0.68	—	0.63	1.03	1.03	1.27	1.41
RLS	L	1.00	–0.99	0.076	0.89	—	–0.02	—	0.65	1.00	1.00	1.24	1.39
AR	S	1.04	0.35	0.174	0.73	—	0.73	—	0.59	1.02	1.06	1.31	1.45
LS	S	0.96	–0.89	0.166	0.86	—	–0.12	—	0.62	0.98	1.03	1.28	1.42
OLS	S	1.02	0.35	0.174	0.79	—	0.65	—	0.59	1.01	1.05	1.30	1.44
RLS	S	0.98	–0.88	0.174	0.86	—	–0.10	—	0.62	0.98	1.03	1.28	1.42

TAR(1) Process with Parameters $\phi_{1p} = 0.9$, $\phi_{2p} = 0.0$, $\phi_{1n} = -0.8$, $\phi_{2n} = 0.0$

EM	T	RSS	β_0	SE_1	ϕ_{1p}	ϕ_{2p}	ϕ_{1n}	ϕ_{2n}	R^2	WFE	FE1	FE2	FE3
AR	L	1.16	0.62	0.082	0.72	—	0.72	—	0.56	1.09	1.10	1.32	1.43
LS	L	1.00	–0.99	0.069	0.90	—	–0.80	—	0.64	1.00	1.00	1.26	1.39
OLS	L	1.14	0.62	0.082	0.82	—	0.59	—	0.57	1.07	1.08	1.29	1.40
RLS	L	1.00	–0.99	0.082	0.90	—	–0.79	—	0.64	1.00	1.01	1.27	1.39
AR	S	1.12	0.61	0.188	0.68	—	0.68	—	0.53	1.06	1.11	1.33	1.45
LS	S	0.98	–0.92	0.163	0.87	—	–0.69	—	0.62	0.98	1.03	1.28	1.42
OLS	S	1.10	0.61	0.188	0.77	—	0.56	—	0.54	1.05	1.09	1.30	1.43
RLS	S	0.98	–0.91	0.188	0.87	—	–0.67	—	0.62	0.98	1.02	1.28	1.41

TAR(2) Process with Parameters $\phi_{1p} = 1.2$, $\phi_{2p} = -0.8$, $\phi_{1n} = 0.8$, $\phi_{2n} = 0.0$

EM	T	RSS	β_0	SE_1	ϕ_{1p}	ϕ_{2p}	ϕ_{1n}	ϕ_{2n}	R^2	WFE	FE1	FE2	FE3
AR	L	1.18	–1.35	0.068	0.97	–0.33	0.97	–0.33	0.62	1.10	1.13	1.58	1.72
LS	L	0.98	–1.00	0.057	1.20	–0.80	0.79	0.00	0.69	0.99	1.03	1.47	1.64
OLS	L	1.04	–1.35	0.068	1.19	–0.67	0.76	–0.01	0.67	1.02	1.05	1.49	1.65
RLS	L	0.98	–1.00	0.068	1.20	–0.80	0.80	0.00	0.69	0.99	1.04	1.48	1.63
AR	S	1.14	–1.35	0.158	0.96	–0.34	0.96	–0.34	0.61	1.07	1.12	1.60	1.76
LS	S	0.94	–1.01	0.142	1.17	–0.80	0.77	0.00	0.69	0.96	1.04	1.50	1.67
OLS	S	1.00	–1.35	0.158	1.17	–0.66	0.76	–0.04	0.66	1.00	1.06	1.52	1.70
RLS	S	0.94	–1.01	0.158	1.17	–0.79	0.77	–0.01	0.68	0.96	1.04	1.50	1.68

TAR(2) Process with Parameters $\phi_{1p} = 1.5$, $\phi_{2p} = -0.8$, $\phi_{1n} = 0.0$, $\phi_{2n} = 0.0$

EM	T	RSS	β_0	SE_1	ϕ_{1p}	ϕ_{2p}	ϕ_{1n}	ϕ_{2n}	R^2	WFE	FE1	FE2	FE3
AR	L	1.42	0.28	0.075	1.03	–0.52	1.03	–0.52	0.63	1.24	1.25	1.80	1.96
LS	L	1.00	–1.00	0.049	1.50	–0.80	0.00	0.00	0.76	0.99	1.01	1.60	1.88
OLS	L	1.30	0.28	0.075	1.35	–0.78	0.76	–0.33	0.67	1.16	1.16	1.71	1.94
RLS	L	1.00	–1.00	0.075	1.50	–0.80	0.00	0.00	0.76	0.99	1.01	1.61	1.89
AR	S	1.38	0.28	0.173	1.02	–0.52	1.02	–0.52	0.63	1.21	1.29	1.84	1.98
LS	S	0.96	–0.98	0.121	1.48	–0.79	–0.02	0.01	0.76	0.96	1.07	1.64	1.91

Appendix A. Continued.

TAR(2) Process with Parameters $\phi_{1p} = 1.3$, $\phi_{2p} = -0.6$, $\phi_{1n} = 0.5$, $\phi_{2n} = 0.4$													
EM	T	RSS	β_0	SE_1	ϕ_{1p}	ϕ_{2p}	ϕ_{1n}	ϕ_{2n}	R^2	WFE	FE1	FE2	FE3
OLS	S	1.26	0.28	0.173	1.32	–0.77	0.77	–0.34	0.67	1.13	1.19	1.76	1.98
RLS	S	0.98	–0.98	0.173	1.48	–0.79	–0.01	–0.01	0.76	0.97	1.07	1.65	1.90

TAR(2) Process with Parameters $\phi_{1p} = 1.3$, $\phi_{2p} = -0.6$, $\phi_{1n} = 0.5$, $\phi_{2n} = 0.4$													
EM	T	RSS	β_0	SE_1	ϕ_{1p}	ϕ_{2p}	ϕ_{1n}	ϕ_{2n}	R^2	WFE	FE1	FE2	FE3
AR	L	1.28	–1.09	0.077	0.89	–0.09	0.89	–0.09	0.69	1.15	1.17	1.57	1.78
LS	L	0.98	–1.00	0.056	1.30	–0.60	0.49	0.40	0.77	0.99	1.02	1.42	1.64
OLS	L	1.04	–1.09	0.077	1.25	–0.54	0.52	0.36	0.75	1.02	1.04	1.45	1.67
RLS	L	1.00	–1.00	0.077	1.30	–0.60	0.50	0.39	0.77	0.99	1.02	1.43	1.65
AR	S	1.22	–1.07	0.176	0.87	–0.11	0.87	–0.11	0.66	1.11	1.18	1.59	1.82
LS	S	0.94	–0.99	0.144	1.29	–0.62	0.46	0.38	0.76	0.96	1.05	1.47	1.70
OLS	S	1.04	–1.07	0.176	1.18	–0.49	0.56	0.25	0.72	1.02	1.09	1.50	1.73
RLS	S	0.94	–1.00	0.176	1.28	–0.61	0.47	0.37	0.75	0.97	1.05	1.47	1.70

TAR(2) Process with Parameters $\phi_{1p} = 1.5$, $\phi_{2p} = -0.8$, $\phi_{1n} = -0.9$, $\phi_{2n} = 0$													
EM	T	RSS	β_0	SE_1	ϕ_{1p}	ϕ_{2p}	ϕ_{1n}	ϕ_{2n}	R^2	WFE	FE1	FE2	FE3
AR	L	1.96	1.06	0.099	0.99	–0.56	0.99	–0.56	0.61	1.62	1.65	2.33	2.45
LS	L	1.00	–1.00	0.048	1.50	–0.80	–0.89	0.01	0.85	0.99	1.01	1.66	1.99
OLS	L	1.78	1.06	0.099	1.32	–0.77	0.64	–0.38	0.68	1.47	1.43	2.06	2.34
RLS	L	1.00	–1.00	0.099	1.50	–0.80	–0.90	0.00	0.85	1.00	1.02	1.63	1.98
AR	S	1.88	1.06	0.229	0.98	–0.57	0.98	–0.57	0.61	1.57	1.69	2.41	2.50
LS	S	0.96	–0.98	0.115	1.49	–0.79	–0.85	–0.01	0.84	0.97	1.06	1.69	2.04
OLS	S	1.72	1.06	0.229	1.29	–0.76	0.66	–0.41	0.67	1.43	1.49	2.16	2.43
RLS	S	1.00	–0.97	0.229	1.48	–0.79	–0.84	–0.03	0.84	0.99	1.06	1.70	2.05

Notes: The statistics are over 10,000 models estimated on the basis of a similar number of simulated samples. EM refers to the type of model being estimated: AR is the standard autoregressive model; LS is a TARSC model estimated using the proposed least squares-based method; OLS is a TARSC model estimated on the basis of the OLS residuals; and RLS is a TARSC model estimated using the proposed restricted least squares procedure, that is, using a grid search over the intercept with the slope parameter set equal to its OLS estimate, combined with OLS estimation of $\boldsymbol{\phi}_\mathbf{p}$ and $\boldsymbol{\phi}_\mathbf{n}$. T is the sample size [L = large ($T = 500$), and S = small ($T = 100$)]. RSS indicates the average minimum value reached by the corresponding residual sums of squares. β_0, ϕ_{1p}, ϕ_{2p}, ϕ_{1n}, and ϕ_{2n} refer to the averages of the estimates for the intercept and the four autocorrelation process parameters, respectively. $\mathbf{SE_1}$ stands for the standard deviation of the 10,000 slope parameter estimates. The R^2 is computed as described in Table 1. WFE stands for the average of the within sample root mean square error of the one-period ahead forecasts; and FE1, FE2, and FE3 refer to the root mean square errors of the one-, two-, and three-period-ahead out-of-sample forecasts.

Appendix B. Multi-Period TARSC Model Forecasts

These formulas needed to compute multi-period TARSC model forecasts are obtained by computing the expected values of the future errors conditional on the previous residuals. The forecasting formulas for a TARSC(2) are derived next since the TARSC(1) is a trivial case and the formulas corresponding to higher order processes are a logical extension of the TARSC(2)'s. As in standard AR(2) models, the one-period-ahead forecast from a TARSC(2) model is:

$$
\begin{aligned}
y_{\mathrm{FT}+1} &= \mathrm{E}[y_{\mathrm{T}+1}|e_{\mathrm{T}},e_{\mathrm{T}-1}] = \mathrm{E}[\mathrm{X}_{\mathrm{T}+1}\beta] \\
&\quad + \mathrm{E}[e_{\mathrm{T}+1}|e_{\mathrm{T}},e_{\mathrm{T}-1}] = x_{T+1}\beta \\
&\quad + E[I_{Tp}\phi_{1\mathrm{p}}e_{\mathrm{T}} + I_{\mathrm{Tn}}\phi_{1\mathrm{n}}e_{\mathrm{T}}] \\
&\quad + \mathrm{E}[I_{\mathrm{Tp}}\phi_{2\mathrm{p}}e_{\mathrm{T}-1} + I_{\mathrm{Tn}}\phi_{2\mathrm{n}}e_{\mathrm{T}-1}];
\end{aligned} \tag{1}
$$

where the subscript F indicates forecast, T refers to the time period corresponding to the last available observation, I_{Tp} and I_{Tn} are indicator variables such as I_{Tp} equals one if $e_{T-1} \geq \mathrm{TR}$ and zero otherwise and $I_{Tn} = (1 - I_{Tp})$, and everything else is as defined before. Note that (1) is easily computed since e_T, e_{T-1}, I_{Tp}, and I_{Tn} are known at time $T + 1$ and β, ϕ_{1p},

ϕ_{1n}, ϕ_{2p}, and ϕ_{2n} can be replaced by their LS estimates. The two-period-ahead forecast from a TAR(2) model is:

$$
\begin{aligned}
y_{FT+2} &= E[y_{T+2}|e_T, e_{T-1}] = E[\mathbf{x_{T+2}\beta}] \\
&\quad + E[e_{T+2}] = \mathbf{x_{T+2}\beta} \\
&\quad + E[I_{T+1p}\phi_{1p}e_{T+1} \\
&\quad + I_{T+1n}\phi_{1n}e_{T+1}] + E[I_{T+1p}\phi_{2p}e_T \\
&\quad + I_{T+1n}\phi_{2n}e_T]
\end{aligned} \tag{2}
$$

where (from here on) all expectations are conditional on the known e_T and e_{T-1} values and $\boldsymbol{\beta}$, ϕ_{1p}, ϕ_{1n}, ϕ_{2p}, and ϕ_{2n} denote the LS estimates for these parameters, which are independent of all other random variables in the following equations. The last term of (2) is:

$$
\begin{aligned}
&E[I_{T+1p}\phi_{2p}e_T + I_{T+1n}\phi_{2n}e_T] \\
&\quad = E[I_{T+1p}]\phi_{2p}e_T + E[I_{T+1n}]\phi_{2n}e_T,
\end{aligned} \tag{3}
$$

where the expected values of the indicator variables at T+1 are computed as follows:

$$
\begin{aligned}
E[I_{T+1p}] &= \text{Prob}[e_{T+1} > 0] \\
&= \text{Prob}[E[e_{T+1}] + \nu_{T+1} > 0] \\
&= \text{Prob}[\nu_{T+1} > -E[e_{T+1}]] \\
&= \text{Prob}[\nu_{T+1} > -(I_{Tp}\phi_{1p}e_T \\
&\quad + I_{Tn}\phi_{1n}e_T + I_{Tp}\phi_{2p}e_{T-1} \\
&\quad + I_{Tn}\phi_{2n}e_{T-1})] = \int_C^\infty f_V;
\end{aligned} \tag{4}
$$

where $C = -(I_{Tp}\phi_{1p}e_T + I_{Tn}\phi_{1n}e_T + I_{Tp}\phi_{2p}e_{T-1} + I_{Tn}\phi_{2n}e_{T-1})$ is known at $T + 2$ and fv is a normal density with mean zero and variance σ^2; and $E[I_{T+1n}] = \text{Prob}[e_{T+1} < 0] = 1 - \text{Prob}[e_{T+1} > 0] = 1 - E[I_{T+1p}]$.

The second term of (2) ($E[I_{T+1p}\phi_{1p}e_{T+1} + I_{T+1n}\phi_{1n}e_{T+1}]$) is more complicated to compute because it involves e_{T+1}. Specifically, substituting $E[e_{T+1}] + v_{T+1}$ for e_{T+1} yields:

$$
\begin{aligned}
&E[I_{T+1p}\phi_{1p}e_{T+1} + I_{T+1n}\phi_{1n}e_{T+1}] \\
&= E[I_{T+1p}\phi_{1p}(E[e_{T+1}] + \nu_{T+1}) \\
&\quad + I_{T+1n}\phi_{1n}(E[e_{T+1}] + \nu_{T+1})] \\
&= E[I_{T+1p}\phi_{1p}E[e_{T+1}] \\
&\quad + I_{T+1n}\phi_{1n}E[e_{T+1}]] \\
&\quad + E[I_{T+1p}\phi_{1p}\nu_{T+1} + I_{T+1n}\phi_{1n}\nu_{T+1}].
\end{aligned} \tag{5}
$$

Since $E[e_{T+1}]$ is a known constant (defined as $-C$ above), $E[I_{T+1p}\phi_{1p}E[e_{T+1}] + I_{T+1n}\phi_{1n}E[e_{T+1}]] = -C(\phi_{1p}E[I_{T+1p}] + \phi_{1n}E[I_{T+1n}])$, where $E[I_{T+1p}]$ and $E[I_{T+1n}]$ are computed as described in equation (4). Calculation of $E[I_{T+1p}\phi_{1p}\nu_{T+1} + I_{T+1n}\phi_{1n}\nu_{T+1}]$, on the other hand, requires knowledge of $E[I_{T+1p}\nu_{T+1}]$ and $E[I_{T+1n}\nu_{T+1}]$, which are obtained as follows:

$$
\mathrm{E}\left[I_{\mathrm{T+1p}}\nu_{\mathrm{T+1}}\right] = \int_C^\infty \nu\, f_\nu, \tag{6}
$$

where C and fv are as defined above; and, since $E[I_{T+1n}\nu_{T+1}] + E[I_{T+1p}\nu_{T+1}] = E[\nu_{T+1}] = 0$, $E[I_{T+1n}\nu_{T+1}] = -\mathrm{E}[I_{\mathrm{T+1p}}\nu_{\mathrm{T+1}}]$. Finally, the first term of (2) ($\mathbf{x_{T+2}\beta}$) is obtained by replacing $\boldsymbol{\beta}$ with its LS estimate. The three-period-ahead forecast involves the following computations:

$$
\begin{aligned}
y_{FT+3} &= E[y_{T+3}|e_T, e_{T-1}] \\
&= E[\mathbf{x_{T+3}\beta}] + E[e_{T+3}] \\
&= \mathbf{x_{T+3}\beta} + E[I_{T+2p}\phi_{1p}e_{T+2} \\
&\quad + I_{T+2n}\phi_{1n}e_{T+2}] \\
&\quad + E[I_{T+2p}\phi_{2p}e_{T+1} \\
&\quad + I_{T+2n}\phi_{2n}e_{T+1}]
\end{aligned} \tag{7}
$$

The first term in (7) is obtained by replacing $\boldsymbol{\beta}$ with its LS estimate. Then, after substituting $E[e_{T+2}] + \nu_{T+2}$ and $E[e_{T+1}] + \nu_{T+1}$ for e_{T+2} and e_{T+1}, the second and third terms become:

$$
\begin{aligned}
&E\left[I_{T+2p}\phi_{1p}e_{T+2} + I_{T+2n}\phi_{1n}e_{T+2}\right] \\
&\quad = E[I_{T+2p}\phi_{1p}E[e_{T+2}]] \\
&\qquad + E[I_{T+2n}\phi_{1n}E[e_{T+2}]] \\
&\qquad + E\left[I_{T+2p}\phi_{1p}\nu_{T+2}\right] \\
&\qquad + E[I_{T+2n}\phi_{1n}\nu_{T+2}],
\end{aligned} \tag{8}
$$

and

$$
\begin{aligned}
&E\left[I_{T+2p}\phi_{2p}e_{T+1} + I_{T+2n}\phi_{2n}e_{T+1}\right] \\
&\quad = E[I_{T+2p}\phi_{2p}E[e_{T+1}]] \\
&\qquad + E[I_{T+2n}\phi_{2n}E[e_{T+1}]] \\
&\qquad + E\left[I_{T+2p}\phi_{2p}\nu_{T+1}\right] \\
&\qquad + E[I_{T+2n}\phi_{2n}\nu_{T+1}].
\end{aligned} \tag{9}
$$

Since $E[e_{T+1}]$ is known at $T + 2$ ($-C$ above), the terms in (9) are computed as follows:

$$
E\left[I_{T+2p}\phi_{2p}E[e_{T+1}]\right] = -C\phi_{2p}E\left[I_{T+2p}\right]; \tag{10}
$$

where $E[I_{T+2p}] = \text{Prob}[e_{T+2} > 0] = \text{Prob}[E[e_{T+2}] + \nu_{T+2} > 0] = \text{Prob}[\nu_{T+2} > -$

$E[e_{T+2}]] = \text{Prob}[\nu_{T+2} > -(I_{T+1p}\phi_{1p}e_{T+1} + I_{T+1n}\phi_{1n}e_{T+1} + I_{T+1p}\phi_{2p}e_T + I_{T+1n}\phi_{2n}e_T)] = \text{Prob}[\nu_{T+2} > -(I_{T+1p}\phi_{1p}(E[e_{T+1}] + \nu_{T+1}) + I_{T+1n}\phi_{1n}(E[e_{T+1}] + \nu_{T+1}) + I_{T+1p}\phi_{2p}e_T + I_{T+1n}\phi_{2n}e_T)] = \int_D^\infty \int_{-\infty}^\infty f\nu 1\nu 2$; where $D = -(I_{T+1p}\phi_{1p}(E[e_{T+1}] + \nu_{T+1}) + I_{T+1n}\phi_{1n}(E[e_{T+1}] + \nu_{T+1}) + I_{T+1p}\phi_{2p}e_T + I_{T+1n}\phi_{2n}e_T)$, $E[e_{T+1}] = -C$, $I_{T+1p} = 0.5(1 + \{e_{T+1}/|e_{T+1}|\}) = 0.5(1 + \{(E[e_{T+1}] + \nu_{T+1})/|E[e_{T+1}] + \nu_{T+1}|\}) = 0.5(1 + \{(-C + \nu_{T+1})/|-C + \nu_{T+1}|\})$, $I_{T+1n} = 1 - I_{T+1p}$, and $fv1v2$ is a bivariate normal density with means [0,0] and variances $[\sigma^2,\sigma^2]$.

$$(11) \qquad E[I_{T+2n}\phi_{2n}E[e_{T+1}]] = -C\phi_{2n}E[I_{T+2n}];$$

where $E[I_{T+2n}] = \text{Prob}[e_{T+2} < 0] = 1 - \text{Prob}[e_{T+2} > 0] = 1 - E[I_{T+2p}]$; and $E[I_{T+2p}]$ is computed as in (7).

$$(12) \qquad E[I_{T+2p}\phi_{2p}\nu_{T+1}] = \phi_{2p}E[I_{T+2p}\nu_{T+1}] = \phi_{2p}\int_D^\infty \int_{-\infty}^\infty \nu 1 f\nu 1\nu 2;$$

$$(13) \qquad E[I_{T+2n}\phi_{2n}\nu_{T+1}] = \phi_{2n}E[I_{T+2n}\nu_{T+1}] = -\phi_{2n}E[I_{T+2p}\nu_{T+1}].$$

The last two terms in (8) ($E[I_{T+2p}\phi_{1p}\nu_{T+2}]$ and $E[I_{T+2n}\phi_{1n}\nu_{T+2}]$) are analogously computed:

$$(14) \qquad E[I_{T+2p}\phi_{1p}\nu_{T+2}] = \phi_{1p}E[I_{T+2p}\nu_{T+2}]; \text{ where } E[I_{T+2p}\nu_{T+2}] = \int_D^\infty \int_{-\infty}^\infty \nu 2 f\nu 1\nu 2;$$

where $E[I_{T+2n}\phi_{1n}\nu_{T+2}] = \phi_{1n}E[I_{T+2n}\nu_{T+2}] = -\phi_{1n}E[I_{T+2p}\nu_{T+2}]$ and $E[I_{T+2p}\nu_{T+2}]$ is computed as above. Lastly, computation of the first two terms in (7) is carried out as follows:

$$(15) \qquad E[I_{T+2p}\phi_{1p}E[e_{T+2}]] = \phi_{1p}E[I_{T+2p}\{I_{T+1p}\phi_{1p}(E[e_{T+1}] + \nu_{T+1}) + I_{T+1n}\phi_{1n}(E[e_{T+1}] + \nu_{T+1}) + I_{T+1p}\phi_{2p}e_T + I_{T+1n}\phi_{2n}e_T\}],$$

and

$$(16) \qquad E[I_{T+2n}\phi_{1n}E[e_{T+2}]] = \phi_{1n}E[I_{T+2n}\{I_{T+1p}(E[e_{T+1}] + \nu_{T+1}) + I_{T+1n}\phi_{1n}(E[e_{T+1}] + \nu_{T+1}) + I_{T+1p}\phi_{2p}e_T + I_{T+1n}\phi_{2n}e_T\}].$$

Noting again that $E[e_{T+1}] = -C$ (defined above), computation of (15) and (16) requires finding the expected value of products random variables such as:

$$(17) \qquad E[I_{T+2p}I_{T+1p}] = \text{Prob}[e_{T+2} > 0 \text{ and } e_{T+1} > 0] = \text{Prob}[E[e_{T+2}] + \nu_{T+2} > 0 \text{ and } E[e_{T+1}] + \nu_{T+1} > 0] = \text{Prob}[\nu_{T+2} > -E[e_{T+2}] \text{ and } \nu_{T+1} > -E[e_{T+1}]] = \int_D^\infty \int_C^\infty f\nu 1\nu 2;$$

$$(18) \qquad E[I_{T+2p}I_{T+1p}\nu_{T+1}] = E[\nu_{T+1}|e_{T+2} > 0, e_{T+1} > 0] = E[\nu_{T+1}|E[e_{T+2}] + \nu_{T+2} > 0, \times\ E[e_{T+1}] + \nu_{T+1} > 0] = E[\nu_{T+1}|\nu_{T+2} > -E[e_{T+2}], \nu_{T+1} > -E[e_{T+1}]] = \int_D^\infty \int_C^\infty \nu 1 f\nu 1\nu 2.$$

The expected values of the remaining products are computed analogously. The Gauss programs needed to compute these one-, two-, and three-period ahead TARSC model forecasts will be made available upon request.