Stationarity Assumptions and Technical Change in Supply Response Analysis

John W. McClelland and Harry Vroomen

Abstract. *Proper stationarity assumptions (trend stationarity or difference stationarity) are important for modeling agricultural supply response in the context of time series analysis. Test results show that the assumption of trend stationarity should be a tested rather than a maintained hypothesis. We discuss implications of model misspecification in the interpretation of trend line regression coefficients as a proxy for technical change. The analysis suggests a more careful consideration of stationarity assumptions when this method is employed in the future.*

Keywords. *Stationarity, time series, supply response*

Analysts of supply response for major agricultural commodities often rely on time series data to estimate behavioral relationships econometrically. Forecast results are important to policymakers who must decide the direction of U.S. agricultural policy.

Researchers commonly decompose real variables, such as output or acres planted, into a growth component and a cyclical component. The growth component results from changes in factors such as capital stock, population, or technology, whereas the stationary cyclical component is the result of monetary or price factors. The econometric procedure relied on for the decomposition into growth and cyclical factors is often a regression with time as an independent variable. Residuals resulting from this detrending procedure are then treated as a stationary series.

Several papers have appeared in the economics literature discussing the problem of inappropriately detrending macroeconomic time series *(14, 15, 16)*.[1] We extend the investigation by applying current time series methods to time series data frequently used in agricultural supply analysis where detrending, by including time as an independent variable in supply response equations, is common practice. We question both the use of time as a proxy for technological change and time-associated coefficients as a measure of technical change or as an indicator of dynamic movements in the production system.

We discuss specification of time as an independent regressor in supply response equations, statistical analysis of time series, and different methods for decomposing time series data. The procedure is to make alternative stationarity assumptions (trend stationarity or difference stationarity) on the statistical structure of regression residuals. Our purpose is to develop an analytical framework for assessing the validity of *a priori* stationarity assumptions. We apply a test proposed by Dickey and Fuller *(3)* to time series data for yield response changes and acres planted for three major crops: corn, wheat, and soybeans. The results of this test suggest a lack of adequate diagnostic analysis of time series data used in studies of supply response. Trend stationarity, as a maintained hypothesis, is tenuous, producing statistical results that may be misleading. We consider the implications of structural misspecification and the possibility of spurious results from inappropriate stationarity assumptions.

Supply Response Analysis

Many models that analyze agricultural supply response contain a linear trend term as an independent regressor. The justification often given for including trend terms is their perceived ability to capture the effects of omitted or unmeasurable variables, which are thought to have an effect over time. The omitted variable is frequently assumed to be technology, suggesting smooth deterministic changes in technology and bounded uncertainty as opposed to irregular stochastic changes with unbounded uncertainty *(14)*. Specification of a functional dependence on time implies an assumption by the investigator of trend stationarity. We suggest, however, that this assumed functional dependence is an empirical question and that *a priori* assumptions about the stationarity of any particular time series and the nature of technical change and future uncertainty are ad hoc. Incorrect stationarity assumptions have serious consequences. We show how they lead to spurious regression results and erroneous conclusions about the nature and magnitude of uncertainty and technical change.

Analysts of agricultural supply response generally separate crop production into two categories: yield response

The authors are agricultural economists with the Resources and Technology Division, ERS.

[1] Italicized numbers in parentheses refer to items in the References at the end of this article.

and acreage response Examples of studies that consider yield response include Menz and Pardey *(12)*, Houck and Gallagher *(8)*, Reed and Riggins *(17)*, Butell and Naive *(1)*, and Lin and Davenport *(11)* These studies use models that specify yield per acre as a deterministic function of time The trend variable is assumed to measure technical change *(10)* Reed and Riggins *(17)* also employ a difference specification after discovering that the trend term explains most of the variation in corn yields

Models of supply response that analyze acres planted as the dependent variable in the regression include, Houck and others *(7)*, Gardner *(5)*, Houck and Ryan *(9)*, Morzuch and others *(13)*, and Ryan and Abel *(18)* Again, time is included as an independent regressor in the acreage response equations, because "inclusion of the trend variable (T) had the effect of increasing the t-values for the individual variables and improving the overall fit of the equation as compared with specifications not including T" (*7*, p 17) Trend is also included to "account for changes occurring through time which are not reflected by other variables" (*9*, p 190)

Statistical Background

Modeling time series data is fundamentally a choice between two hypotheses about the data-generation process We specify technical aspects of model specification without regard to any particular time series and show with a simple example that improper assumptions about the stationarity of a time series can have serious consequences, including unbounded forecast errors and uncertainty We specify the most elementary representations of statistical time series, namely, first-order trend stationary (TS) processes and first-order difference stationary (DS) processes *(14)* Extensions to higher order cases are discussed by Nelson and Plosser *(16)*, but Dickey and Fuller's tests *(2, 3)* are applicable only to the first-order cases presented here

Consider the sequence $\{y_t\}$ of an observed nonstationary time series If the nonstationarity in $\{y_t\}$ is assumed to be a linear dependence on time, then a model explaining the variation in y is properly specified as

$$y_t = \alpha + \beta t + u_t \tag{1}$$

where $\{u_t\}$ is the stationary cyclical component of the variation in equation 1 and is assumed to be independently and identically distributed with zero mean and constant variance, and α and β are fixed paramenters An alternative to equation 1 is to assume that y is stationary in first differences, for which the correct specification is given by

$$y_t - y_{t-1} = \beta + e_t \tag{2}$$

where $\{e_t\}$ is a stationary series of independently and identically distributed random disturbances with zero mean and constant variance, and β is a fixed parameter Equations 1 and 2 are alternative versions of a first-order transformation of nonstationary time series from which a stationary sequence of residuals is obtained Equation 1 is a linear TS specification, and equation 2 is a first-order DS specification

We can illustrate the fundamental difference between 1 and 2 by rewriting equation 2 as a recursive system

$$\begin{aligned} y_t &= y_{t-1} + \beta + e_t \\ y_{t-1} &= y_{t-2} + \beta + e_{t-1} \\ y_{t-2} &= y_{t-3} + \beta + e_{t-2} \\ &\vdots \end{aligned} \tag{3}$$

Successive substitution to some point in time, say y_0, yields

$$y_t = y_0 + \beta t + \sum_{i=1}^{t} e_i \tag{4}$$

which is the result of expressing a first-order DS process as a linear function of time Although equations 1 and 4 are similar in appearance, they are fundamentally different One difference is in the intercept term, the intercept in equation 1 is a fixed parameter, whereas the intercept in equation 4 depends on the arbitrary determination of y_0 *(14)* The error structure of the two equations is also different, equation 1 has a stationary error structure, but equation 4 has a nonstationary error structure because it is dependent on time We can easily show this nonstationarity by computing the variance of the residuals in 4 as

$$\begin{aligned} V(e_t) &= E[e_t - E(e_t)]^2 \\ &= E[\Sigma\, e_i^2] = e_1^2 + e_2^2 + \cdots + e_t^2 \\ &= t\sigma_e^2 \end{aligned} \tag{5}$$

Equation 5 is an important result because it shows that a first-order DS process expressed as a linear function of time will yield confidence intervals that increase without bound *(14)* The problem, however, is far more serious than unbounded confidence intervals Nelson and Kang *(15)* investigate the problem of inappropriate detrending of time series and find it "to produce evidence of periodicity which is not in any meaningful sense a property of the underlying system" (*15*, p 742) Their results "further suggest that the dynamics of econometric models estimated from such data may well be wholly or in part an artifact of the trend removal process" (*15*, p 742) They later show, through a decomposition of R^2, that the significance of coefficients from regressions of a random walk on time will be overstated and that R^2 "will exaggerate the extent to which movement in the data is accounted for by time" (*14*, p 74) The reported t-statistics for the OLS coefficients on time for data generated by our equation 2 and modeled as equation 1 are striking results of their Monte Carlo experiments Nelson and Kang's results reject the hypothesis of no functional dependence on time in 87 percent of the cases for samples of 100 observations at a 5-percent level of significance The hypothesis of no functional dependence on time is rejected in spite of the fact that no such time dependence actually exists It is similar to the spurious regression phenomena discussed by Granger and Newbold (*6*, sec 6 4) Granger and Newbold show that conventional t-statistics can indicate a high degree of fit when one independent random walk is regressed on another

Spurious regression results are a danger when time series data are detrended because the detrending procedure tends to remove much of the variation from the data (see *16*) Because random walk data often have the appearance of movement around a trend, it may seem reasonable to apply detrending procedures to achieve stationarity in the residuals The result, however, is not a stationary sequence of residuals, but the removal of about 86 percent of the stochastic variation in the data (*14*), and the attribution of that variation to assumed deterministic phenomena such as technical change

Dickey and Fuller have developed formal procedures for testing time series specifications (*2*, *3*) Each specification, TS and DS, is treated as one side of a mutually exclusive hypothesis and is combined into a single model One can write a model for testing the TS vs the DS hypothesis for our simple example by combining equations 1 and 2 as

$$Y_t = \alpha + \beta t + \phi Y_{t-1} + e_t \qquad (6)$$

and testing the null hypothesis, $\phi = 1$, $\beta = 0$ (*16*, p 144) Failure to reject the null hypothesis indicates an underlying DS process, whereas rejecting the null hypothesis implies an underlying TS process (*16*) Dickey and Fuller (*2*) represent the limiting distribution of $\hat{\phi}$, and they derive a test statistic $t(\hat{\phi})$ for testing this hypothesis Critical values are tabulated and presented in Fuller (*4*) for the one parameter test, and in Dickey and Fuller (*3*) for the likelihood ratio test on the entire parameter space where the null hypothesis is $(\hat{\alpha},\hat{\beta},\hat{\phi}) = (\alpha,0,1)$

The Dickey-Fuller test indicates model stationarity under the alternative hypothesis presented because it determines statistically the probability of a unit root in the characteristic equation of the model In our simple model, the value of ϕ in equation 6 must be estimated and compared with the hypothesized value If ϕ in equation 6 is significantly different from 1, then $\{y\}$ is a sequence that is stationary in trend, while also exhibiting autoregressive behavior However, if ϕ is equal to 1, the indication is nonstationary behavior characterized by a unit root in the characteristic equation, and equation 6 reduces to a random walk with drift under the null hypothesis Test statistics for the OLS estimator for ϕ do not conform to standard statistical distributions because the distribution centers about 1 and not zero

Test Results

We apply the Dickey-Fuller test of the TS vs the DS hypothesis to aggregate U S Department of Agriculture data for total crop yield and acres planted for corn, soybeans, and wheat Data are annual and observations are continuous for 1930–86 The table shows results of the Dickey-Fuller test for both yields and acres planted For the yield data, the null hypothesis HO $\phi = 1$ was rejected at the nominal 0 05 level[2] in all cases The test statistic for samples of 50 is −3 50, but for samples of 100 is −3 45 Thus, the true test statistic for our sample is between these two values The yield data imply an underlying TS-generating process and suggest that the appropriate specification is one that involves a deterministic function of time Results for acreage-planted data were the opposite We were unable to reject the null hypothesis in any case at the nominal 0 05 level The disagreement in these results requires further scrutiny

Results of testing for autoregressive unit roots

Item	Parameter estimate	Standard error	Dickey Fuller test statistic
Crop yield			
Corn	0 350	0 129	−5 04
Soybeans	015	138	−7 14
Wheat	528	113	−4 18
Acres planted			
Corn	0 783	0 086	−2 52
Soybeans	860	066	−2 12
Wheat	807	081	−2 38

Researchers who chose to model yields as a deterministic function of time make the correct *a priori* assumption, and those who chose to model acreage planted as

[2] In this case, the 0 05 level of significance is a stronger condition because the greater the significance level, the smaller the test statistic must be to maintain the null hypothesis

a deterministic function of time do not However, these results can be explained in terms of the underlying assumptions of the TS and DS specifications The critical assumption involves the nature of the technological change that the dynamic model is postulated to capture If technology does in fact change in a relatively smooth way, it is reasonable to assume a TS process

Technological change in agriculture can be characterized as a TS process because of active and independent research and innovation related to output-enhancing inputs Unlike many other types of production technology, agricultural technology is funded by both the private and public sector Many assets in agriculture also have a relatively short span of productivity 1 week, 6 months, a few years Therefore the turnover in assets is rapid in contrast to heavy industry where plants and equipment may have an economic life of 25 years or more Furthermore, complements of inputs in agricultural production are constantly changing, component by component, giving the effect of smooth changes in output For example, a major breakthrough in seed technology may be followed by an improvement in fertilizer, which in turn is followed by an advance in herbicides, and so forth Therefore, one could argue that aggregate yields for corn, soybeans, and wheat have increased along a deterministic trend

Acreage planted, in contrast, is more a function of uncertain policy changes from one farm bill to the next and of prices and price expectations These effects are likely to be random Therefore, data for acres planted would likely follow a DS specification This observation is particularly disturbing in light of the discussion of spurious regression phenomena provided by Nelson and Kang *(14, 15)* because several studies *(5, 7, 9, 13, 18)* employ a TS specification when analyzing acreage response Nelson and Kang's argument leads to the conclusion that results from an inappropriately detrended series can provide seriously misleading information about the relationship between changes in farmers' decisions and changes in policies These behavioral changes may be wholly unrelated to technical change, but largely attributed to it Thus, technical change is given a role in policymaking that is unwarranted and unwise

Conclusions

From the simple diagnostic example presented here, we have shown that the analysis of policy decisions over time and the evaluation of technical change and uncertainty, are closely linked to the assumptions and methodologies employed This linkage is particularly true when longrun projections are being considered, because model misspecification implies unbounded uncertainty in future time Researchers concerned with the accuracy of information generated for policy analysis must consider the consequences of *a priori* assumptions about data-generating processes Far more work is needed to resolve these methodological issues so policy analysis can be improved

References

1 Butell, R , and J J Naive "Factors Affecting Corn Yields " *Feed Situation* FDS-269 U S Dept of Agr , Econ Stat Coop Serv , May 1978, pp 14-16

2 Dickey, D A , and W A Fuller "Distribution of the Estimators for Autoregressive Time Series With a Unit Root," *Journal of the American Statistical Association,* Vol 74, 1979, pp 427-31

3 ———————— "Likelihood Ratio Statistics for Autoregressive Time Series With a Unit Root," *Econometrica* Vol 49, 1981, pp 1057-72

4 Fuller, W A *Introduction to Time Series* New York John Wiley and Sons, 1976

5 Gardner, B L "Futures Prices in Supply Analysis," *American Journal of Agricultural Economics,* Vol 58, 1976, pp 81-84

6 Granger, C W J , and P Newbold *Forecasting Economic Time Series* Orlando, FL Academic Press, 1986

7 Houck, J P , M E Abel, M E Ryan, P W Gallagher, R G Hoffman, and J B Penn *Analyzing the Impact of Government Programs on Crop Acreage* TB-1548 U S Dept of Agr , Econ Res Serv , Aug 1976

8 Houck, J P , and P W Gallagher "The Price Responsiveness of U S Corn Yields," *American Journal of Agricultural Economics,* Vol 58, 1976, pp 731-34

9 Houck, J P , and M E Ryan "Supply Analysis for Corn in the United States The Impact of Changing Government Programs," *American Journal of Agricultural Economics,* Vol 54, 1972, pp. 184-91

10 LaFrance, J T , and O R Burt "Modified Partial Adjustment Model of Aggregate U S Agricultural Supply," *Western Journal of Agricultural Economics,* Vol 8, 1983, pp 1-12

11 Lin, W , and G Davenport "Analysis of Factors Affecting Corn Yields Projections to 1985," *Feed*

Outlook and Situation, FDS-285, USDA, ERS, May 1982, pp 9-14

12 Menz, K.M., and P Pardey "Technology and U S Corn Yields Plateaus and Price Responsiveness," *American Journal of Agricultural Economics*, Vol 65, 1983, pp 558-62

13 Morzuch, B J , R D Weaver, and P.G Helmberger "Wheat Acreage Supply Response Under Changing Farm Programs," *American Journal of Agricultural Economics*, Vol 62, 1980, pp 29-37

14 Nelson, C R , and H Kang "Pitfalls in the Use of Time as an Explanatory Variable in Regression," *Journal of Business and Economics Statistics*, Vol 2, 1984, pp 73-82

15. __________ "Spacious Periodicity in Inappropriately Detrended Time Series," *Econometrica*, Vol 49, 1981, pp 741-751

16 Nelson, C R , and C I Plosser "Trends and Random Walks in Macroeconomic Time Series," *Journal of Monetary Economics*, Vol 10, 1982, pp 139-62

17 Reed, M.R , and S K Riggins "Corn Yield Response A Micro-Analysis," *North Central Journal of Agricultural Economics*, Vol 4, 1982, pp 91-94

18 Ryan, M E , and M E Abel "Corn Acreage Response and the Set-Aside Program," *Agricultural Economics Research*, Vol 24, 1972, pp 102-112