FS 98-02 January, 1998

LEADING INDICATORS FOR REGIONAL COTTON RESPONSE: STRUCTURAL AND TIME SERIES MODELING RESULTS

Paul A. Stavriotis, Jack E. Houston,
Christopher S. McIntosh, and Steven C. Turner

FS98-02 January, 1998

LEADING INDICATORS FOR REGIONAL COTTON RESPONSE: STRUCTURAL AND TIME SERIES MODELING RESULTS

Paul A. Stavriotis, Jack E. Houston,
Christopher S. McIntosh, and Steven C. Turner

Paul A. Stavriotis is a former graduate student, and Jack E. Houston, Christopher S. McIntosh, and Steven C. Turner are associate professors in the Department of Agricultural & Applied Economics, University of Georgia, Athens, GA.

Dept. of Agricultural & Applied Economics
College of Agricultural & Environmental Sciences
University of Georgia

LEADING INDICATORS FOR REGIONAL COTTON RESPONSE: STRUCTURAL AND TIME SERIES MODELING RESULTS

Paul A. Stavriotis, Jack E. Houston,
Christopher S. McIntosh, and Steven C. Turner

Department of Agricultural and Applied Economics

University of Georgia

Athens, GA 30602-7509

jhouston@agecon.uga.edu

ABSTRACT---

Resurging southeastern cotton production compels better cotton acreage forecasts for planning seed, chemical, and other input requirements. Structural models describe leading acreage response indicators, and forecasts are compared time-series models. Cotton price, loan rate, deficiency payments, lagged corn acreage, the PIK program, and previous cotton yield significantly influence response.

Leading Indicators for Regional Cotton Response:

Econometric and Time Series Modeling Results

Background

In the past decade, the role of cotton as one of the Southeast's most important field crops has been revived (Table 1). Acreage planted to cotton has increased almost 300 percent in the Southeast in a matter of six years. This resurgence compels better estimation of future cotton acreage in order for businesses to respond to seed, chemical, and other input requirements and for processing/ginning facilities planning. As the amount of cotton planted by farmers increases, so does the level of capital and other resources required. Accurate leading indicators are then more vital as cotton becomes more prominent in the crop-mix and decision-making process for producers. For example, acres planted to cotton in Georgia alone doubled from 1994 to 1995, reaching 1.5 million acres and the largest cotton crop in Georgia in more than 50 years.

Identifying leading indicators and forecasting responses to changing conditions in agricultural crop production have long been important factors in risk-management strategies. Acreage response is an important indicator when producers are making spring planting decisions for competing crops. Integrating these management components have become more important with the enactment of the Federal Agricultural Improvement and Reform (FAIR) Act in 1996. Under FAIR, producers must be more responsive to available market and production information.

Cotton was selected for this study because of its recent and dramatic increase in popularity among Southeastern producers. Can cotton acreage response be reliably forecasted, and what leading indicators are needed to do this? Acreage response enters producers' spring planting decisions. The purpose of this paper is to present a model which identifies leading indicators with respect to Southeastern planted cotton acres, so that farmers and supporting agribusiness suppliers can respond more quickly to changes in the cotton market's environment.

We will first review and summarize government programs and policy background and relevant work on acreage response. Theory and methods will be discussed briefly for some of the articles that influenced the development of these models. Development of the econometric and time-series models, a description of the data, and the forecasting results for the Southeastern models follow. Important conclusions and a discussion of the implications of the findings for the industry and for government policy makers complete the paper.

Government Programs

Price theory suggests that prices convey information that provides incentives for increasing or decreasing the supply of one commodity or another. In the 1950's and 1960's, the aim of U.S. government policy was to reduce cotton supply and stocks, as increased production following WWII and Korea was no longer necessary. This intention was reflected in the Agricultural Adjustment Act of 1954. Looking at the number of acres planted to cotton during these periods, we can see the great drop in acreage planted today compared to 30 or 40 years ago, as seen in Table 1.

As a result of these government programs, as well as increasing real costs and boll weevil damage and control costs, acreage in cotton declined and cotton prices were low. For example, the Cotton - Wheat Act of 1964 was aimed at beginning voluntary programs to reduce cotton production. After reviewing previous work on cotton acreage response (Duffy *et al*. 1987; Duffy *et al.* 1994; Parrot and McIntosh 1996), and acreage response in general (Gardner 1976; Pope 1981; Shideed *et al.* 1986; Shideed *et al*. 1987; Shideed and White 1989; McIntosh and Shideed 1989; McIntosh and Shumway 1994), the primary objective of this study was to develop an economically sound model for forecasting cotton acreage and resulting impacts on planting and marketing decisions. This goal has eluded accurate estimation in recent cropping years.

Model and Data

In this study, two different approaches are used to estimate acreage response: mulitvariate linear regression and Box-Jenkins Autoregressive Integrated Moving Average, or ARIMA, time-series models (Box and Jenkins, 1976). Yield and competitive crop prices are tested as factors in cotton acreage response decisions, and, following the opinion of Chavas *et al.*, include any and all rational cotton data in the testing. In particular, production and yield are important not only for cotton but for the other major field crops that compete with cotton acreage, because these variables are important in the process that influences farmers' planting decisions.

Previous studies have used relatively short periods of time and thus missed capturing longer trends and policy impacts in cotton production responses. Government program variables must be included in the development of such a model, according to Duffy *et al*., due to their ability to alter producers' planting decisions. The loan rate, target price and support payments also are included in the analysis to examine their influence on acreage response. A general economic model is specified as follows:

$$CTA=f(SBA_{t-2},SBP_{t-1},SBY_{t-2},CNP_{t-1},CNY_{t-2},CNA_{t-1},CTH_{t-2},$$
$$CTL_{t},CTP_{t-1},CTPD_{t-1},TGT_{t},D1_{t-1},D2_{t-1},D3_{t-1},D4_{t-1},CTY_{t-2}),$$

where the variable descriptions are presented in Table 2. The model describes planted cotton acreage as a relationship to 16 independent/pre-determined variables.

The Southeast region was selected for this study because of its high response elasticities, according to Duffy *et al.* (1994), and the dramatic increase in acreage in this area in the last several years. A combination of data directly related to cotton, as well as its major production substitutes, corn and soybeans, was gathered which encompassed 53 years of observations, 1944 to 1996. Most previous studies have used 30 years or less. Six observations, 1991 to 1996, were withheld for out-of-sample forecasting validation. Cotton data were obtained for planted acreage, harvested acreage, harvested yield, total production, and season average price per pound received by farmers for cotton in each of the Southeastern states. Observations were collected for corn and soybeans to include planted acreage, harvested yield, and season average price per bushel received by farmers. Government program data were collected for the national price support loan rate for cotton, target price, deficiency payments, disaster payments, diversion payments and payments-in-kind (PIK).

Data were collected for each of the Southeast region states, and the information was then processed into weighted Southeastern numbers, as described in the following. Planted and harvested acreage and season average price per bushel received by farmers were weighted by production in each state, as in the Duffy study, while harvested yield was weighted by acreage. Futures prices were not used in this study, because they are reported for the whole country and would reflect a bias towards other cotton producing regions of the country.

Model Results

The estimated structural model explains 92 percent of the variation in planted cotton acreage in the Southeast (Table 3). The model’s U-statistic of 0.78, its mean absolute percentage error (MAPE), and its mean squared error (MSE) in Table 4 indicate the strength of its forecasts as compared to several ARIMA forecasting models and to a composite forecast. A U-statistic between 0 and 1 indicates that the forecasts perform better than a naive, or random walk, model, while U-statistics greater than one indicate forecasting ability worse than a naive model.

Southeastern cotton acreage was found to be responsive at a significance level of 0.10 or better to three leading indicators: the cotton loan rate, cotton price, and the deficiency payment. Thus, in a year where the loan rate drops, one could expect planted acreage to rise. This is due to the change in the basis between the loan rate and the target price, which, depending on market prices, allowed eligible producers a chance at a larger deficiency payment. However, producers should be aware that without the target price in effect in the future, this response will be redundant. When cotton price is high one period removed, producers may expect increases in acreage the next year. The same can be said of the deficiency payment. This study used dummy variables for the deficiency payment, and results indicate that only having a payment is sufficient for producers to expect to increase acreage in the following year.

The forecasting capability of the structural model demonstrates that several of the variables in the model serve as significant leading indicators for planted cotton acreage. Based on statistical significance at the 10 percent level or greater, the following seven explanatory variables appear to be important leading indicators for producers and agribusiness firms in projecting cotton acres planted and identifying appropriate risk management strategies. These variables are categorized according to lag length and are presented along with impact elasticities based on mean values for the 1991-1996 forecast period (Table 5). In particular, a 10 percent increase in cotton price in one year would induce, on average, a 12.6 percent increase in plantings in the following year.

Box-Jenkins (ARIMA) time-series models are also estimated for comparison of the structural model's forecast statistics with a time-series approach on the same data. ARIMA models were run on actual acreage planted and on the first differences (annual changes) of acreage (Table 4). As determined by the U-statistic of the ARIMA forecasts, the ARIMA (0,2,1) model more successfully predicts Southeastern cotton acreage response than other ARIMA specifications. The ARIMA (0,2,1) exhibited a U-statistic of 0.66, indicating that it is superior to a naive model and out-performs the structural model in predicting annual changes in acreage planted. This suggests that ARIMA analysis is an effective complement to any forecasting venture that uses econometric tools. In fact, a composite forecasting model using a simple average of the structural and ARIMA (0,2,1) forecasts demonstrated superior u-statistics and MSE, with a comparable MAPE value.

Previous literature has focused on structural models of acreage response at the regional level, as in the Duffy *et al*. study (1994). Box-Jenkins (ARIMA) analysis can out-perform the forecast results of such structural models. However, the importance of the structural models is in their value as explanatory models. Next, we will compare the regional results of this study with those of the Duffy et al. study.

Structural Model Comparison to Previous Literature

The Duffy *et al*. model, the only previous regional cotton acreage model, was re-constructed on their original time period, 1959-1983, using the data and variables found in their study (Table 6). Our Southeastern structural model was estimated on the same time period for comparison. The Duffy model was also reconstructed for the extended time period of this study, 1945-1990. The methods for constructing these prices are complex and change whenever government policy changes with respect to acreage restriction, allotments, etc.

The Southeast model results indicate similar explanatory values in both time periods and both models pass the Durbin-Watson test, but the Duffy model has a higher F-statistics. The mean absolute errors are comparable, and the model from this study out-performs the Duffy *et al*. model in terms of the Theil u-statistic and forecasting values, although neither model is a valid forecasting tool on the 1959-1983 time period. This may indicate that the longer time period used in this study improves forecasting accuracy.

In the extended time period, the explanatory power of the Duffy model dropped somewhat. The model is still not a valid forecasting tool, as seen by its u-statistic, which is 1.38. The error values are also higher than the structural model postulated in this study. The results of this comparison indicate that the Southeastern structural model posited here is an improvement over the Duffy study's model in most statistics. The models in this study, however, provide more explanatory value, largely due to the inclusion of additional relevant decision variables and a longer time period.

Conclusions and Implications

The objective of this study was to develop a structural model that could identify leading indicators and accurately forecast cotton acreage plantings in the Southeast. Several variables in the current structural model serve as significant leading indicators for planted cotton acreage. Southeastern cotton acreage was found to be responsive to the cotton loan rate, cotton price and the deficiency payment, as well as to lagged corn acreage, the PIK program, and previous cotton yield. The use of such indicators and forecasts enables producers and agribusiness firms in this region to respond to changes in supply with a better understanding of the changes in the cotton market in the Southeast.

The structural model tested was found to be superior to a naive forecast in its ability to project cotton acres planted. The results of Box-Jenkins analysis confirm that some ARIMA models can out-perform the forecasting capabilities of structural models used thus far. Comparable models were tested from the literature review to evaluate the success of this study's model. The results of this comparison have shown that the structural model is an improvement over previous work in cotton acreage response, both in the simplicity of the model and in its forecasting ability. Time series analysis also was performed on cotton acreage and the annual differences of cotton acreage plantings--effectively turning point analysis. The results from ARIMA analysis show promise for this type of forecasting tool in decision analysis. In some respects, ARIMA out-performs the structural model in forecast accuracy. However, the econometric model better identifies leading indicators for cotton acreage.

Results of the structural and time series forecast modeling provide evidence of tools which may prove useful to current and potential cotton producers, their input suppliers, and those considering investing in ancillary services in Southeastern regional cotton and oilseed industries. Removing some uncertainty as to planting responses enables better decisions on following season planting decisions for cotton and its substitutes, for agribusinesses ordering seed and chemical inputs for the area, and for scaling prospective services necessary to handle the plantings and post-harvest handling. In particular, turning points and rates of change were much more accurately forecast than previously.

References

Box, G.E.P., and G.M. Jenkins. *Time Series Analysis, Forecasting and Control*. Revised Edition. Oakland, CA:Holden-Day, 1976.

Chavas, J.P., R.D. Pope, and R.S. Kao. "An Analysis of the Role of Futures Prices, Cash Prices, and Government Programs in Acreage Response." *Western Journal of Agricultural Economics* 8(1983):27-33.

Duffy, P.A., J.W. Richardson, and M.K. Wohlgenant. "Regional Cotton Acreage Response." *Southern Journal of Agricultural Economics* 19(1987):99-109.

Duffy, P.A., K. Shalishali, H.W. Kinnucan. "Acreage Response under Farm Programs for Major Southeastern Field Crops." *Journal of Agricultural and Applied Economics* 26(1994):367-378.

Gardner, B.L. "Futures Prices in Supply Analysis." *American Journal of Agricultural Economics* 58(1976):81-84.

McIntosh, C.S., and C.R. Shumway. "Evaluating Alternative Price Expectations Models for Multiproduct Supply Analysis." *Agricultural Economics* 10(1994):1-11.

McIntosh, C.S., and K.H. Shideed. "The Effect of Government Programs on Acreage Response over Time: The Case of Corn Production in Iowa." *Western Journal of Agricultural Economics* 14(1989):38-44.

Parrot, S.D., and C.S. McIntosh. "Nonconstant Price Expectations and Acreage Response: the Case of Cotton Production in Georgia." *Journal of Agricultural and Applied Economics* 28,1(1996):203-210.

Pope, R.D. "Supply Response and the Dispersion of Price Expectations." *American Journal of*

Agricultural Economics 63(1981):161-163.

Shideed, K.H., F.C. White, S.J. Brannen, and R.S. Glover. "Structural Changes of Corn Supply Response in Georgia." Georgia Agricultural Experiment Station Bulletin No. 367, 1987.

Shideed, K.H., S. J. Brannen, and R.S. Glover. "Econometric Estimation of Soybean Supply Response under the Rational Expectations Hypothesis." Georgia Agricultural Experiment Station Bulletin No. 352, 1986.

Shideed, K.H., and F.C. White. "Alternative Forms of Price Expectations in Supply Analysis for U.S. Corn and Soybean Acreages." *Western Journal of Agricultural Economics* 14(1989):281-292.

Tomek, W.G. and K.L. Robinson. *Agricultural Product Prices*. 3rd Edition. Ithaca, NY: Cornell University Press, 1990.

U.S. Department of Agriculture, Economic Research Service. "The Cotton Industry in the United States." Agricultural Economic Report No. 739: Washington, DC, 1996.

_______________. "Oil Crops Yearbook." Washington, D.C., 1996.

U.S. Department of Agriculture, National Agricultural Statistics Service. "Crop Production Annual Summary." Washington, DC, 1997.

_____________. "1996 Crop Values Summary." Washington, DC, 1996.

U.S. Department of Agriculture, Statistics Reporting Service. "Agricultural Statistics." Washington, DC: Government Printing Office. Various issues, 1945-1994.

______________. "Georgia Agricultural Facts." Georgia Crop Reporting Service, 1993.

Table 1. U.S. and Southeastern Region Cotton Acreage and Yields, 1945-1996.

	U.S.		Southeastern	
Years	Planted Acres 1,000 acres	Yield lb/acre	Planted Acres 1,000 acres	Yield lb/acre
1945-49	22,075	269	4,725	299
1950-54	24,641	296	4,494	286
1955-59	15,518	427	2,534	388
1960-64	15,728	475	2,583	403
1965-69	11,448	480	1,680	378
1970-74	12,892	469	1,545	444
1975-79	12,429	481	757	423
1980-84	11,856	528	656	559
1985-89	10,845	624	863	583
1990-94	13,359	660	1,626	662
1995	16,931	537	3,460	539
1996	14,240	703	3,164	724

Source: USDA, 1945-1994, 1997.

Table 2. Descriptions of Variables for Southeastern Cotton Structural Models.

Variable	Description
CTA	Cotton acreage (planted), 1,000 acres
SBA	Soybean acreage (planted), 1,000 acres
SBP	Soybean price, $/bu, season avg. price
SBY	Soybean yield, bu/acre
CNP	Corn price, $/bu, season avg. price
CNY	Corn yield, bu/acre
CNA	Corn acreage (planted), 1,000 acres
CTH	Corn acreage (harvest), 1,000 acres
CTL	Cotton loan rate, cents/lb
CTP	Cotton price, cents/lb, season avg. Price
CTPD	Cotton production, 1,000 bales
CTY	Cotton yield, lb/acre
TGT	Target cotton price, cents/lb
D1	Deficiency payment; binary, 1 yes, 0 otherwise.
D2	Disaster payment; binary, 1 yes, 0 otherwise
D3	Diversion payment; binary, 1 yes, 0 otherwise
D4	Payment-in-kind; binary, 1 yes, 0 otherwise

Table 3. Estimated Variable Parameters for the Southeast Structural Model, 1945-1990.

Variable	Coefficient	t statistic	p value
Constant	-1424.17	-0.9864	0.3324
SBA $_{t-2}$	-0.0166	-0.2325	0.8179
SBP $_{t-1}$	-216.545	-1.2888	0.2080
SBY $_{t-2}$	87.1832	1.6103	0.1186
CNP $_{t-1}$	198.587	0.6515	0.5200
CNY $_{t-2}$	4.9755	0.3571	0.7237
CNA $_{t-1}$	0.3639	3.9702	0.0005
CTH $_{t-2}$	0.1599	1.2900	0.2076
CTL $_{t}$	-70.5512	-3.4300	0.0019
CTP $_{t-1}$	38.1069	2.2771	0.0306
CTProd $_{t-1}$	0.8146	4.8349	0.0000
CTY $_{t-1}$	-4.348	-3.2861	0.0027
TGT $_{t}$	10.3222	0.6807	0.5016
D1 $_{t-1}$	1323.47	2.6622	0.0127
D2 $_{t-1}$	-3.5361	-0.0134	0.9894
D3 $_{t-1}$	327.809	1.0650	0.2960
D4 $_{t-1}$	924.796	2.3873	0.0240

N	46
Adjusted R-square	0.92
F-value	34.20
Mean absolute error	270.215
DW statistic	1.80

Table 4. Out-of-Sample Structural and ARIMA Model Forecasts of Southeast Cotton Acreage Response, 1991-1996.

Year	Actual	Structural Model	ARIMA 1,0,0	ARIMA 1,1,0	ARIMA 0,1,1	ARIMA 1,2,0	ARIMA 0,2,1	Composite Structural/ ARIMA 0,2,1
	1,000 acres							
1991	1579	1254	1140	1088	1086	1106	1342	1298
1992	1524	2440	1564	1508	1488	1536	1802	2121
1993	1727	1828	1512	1532	1517	1529	1749	1789
1994	2170	2168	1705	1694	1688	1707	1959	2064
1995	3460	2874	2126	2099	2081	2127	2415	2645
1996	3164	2944	3352	3255	3205	3333	3738	3341
U Statistic		0.78	1.02	1.04	1.05	0.79	0.66	0.56
MSE		224447	378612	394423	404015	380483	266660	191144
MAPE		18.42	18.13	17.93	18.18	17.90	15.43	15.77

Note: MSE and MAPE are for the out-of-sample validation period. (0,1,1) and (1,1,0) models have no constant terms; (1,0,0) and (0,2,1) models have constant terms; (0,2,1) and (1+5+6,2,0) models have no constant terms.

Table 5. Southeastern Cotton Acreage Structural Model Impact Response Elasticities for Leading Indicators, 1991-1996.

	Elasticity by lag length		
Variable	0	1	2
Cotton loan rate	-1.61		
Corn acreage		0.49	
Cotton price		1.26	
Deficiency		----	
PIK		----	
Cotton production		0.87	
Cotton yield			-1.11

Note: Elasticities are based on mean values for the out-of-sample forecast period.

Table 6. Duffy et al. (1987) Model Comparisons for Southeast cotton acreage, 1959-1983 and 1945-1996.

Variable	Duffy 59-83	Current 59-83	Duffy 45-96	Current 45-96
Constant	2159.79*** (8.5493)	1137.29 (0.4477)	1419.59** (2.6714)	-1424.17 (-0.9864)
CTSIP	16.5847*** (3.4683)	-	30.0792** (2.6422)	-
CNSIP	-408.562*** (-3.1721)	-	-399.444 (-1.5721)	-
CTA t-1	0.2641*** (3.1041)	-	0.6427*** (5.5591)	-
CNA t-1	-	0.1186 (0.4998)	-	0.3639 (3.9702)
CNP t-1	-	-80.2036 (-0.1615)	-	198.587 (0.6515)
CNY t-2	-	3.9668 (0.2747)	-	4.9755 (0.3571)
CTH t-2	-	0.1955 (0.4063)	-	0.1599 (1.2900)
CTL t	-	3.7616 (0.1269)	-	-70.5512 (3.4300)
CTP t-1	-	34.4902* (2.2624)	-	38.1069 (2.2771)
CTPD t-1	-	0.2971 (1.4849)	-	-
CTY t-2	-	-2.5184* (-1.8381)	-	-4.348 (-3.2861)
SBA t-2	-	0.1160 (0.3771)	-	-0.0166 (-0.2325)
SBP t-1	-	-177.865 (-0.8862)	-	-216.545 (-1.2888)
SBY t-2	-	-43.1236 (-0.7713)	-	87.1832 (1.6103)
TGT t	-	-28.8088 (-1.5431)	-	10.3222 (0.6807)
CTEDP	-178.324*** (-4.7783)	-	-27.1379 (-0.8230)	-
D1 t-1	-	-295.283 (-0.3787)	-	1323.47 (2.6622)
D2 t-1	-	-217.024 (-0.4506)	-	-3.5361 (-0.0134)
T	-73.173*** (-7.0931)	-	-49.4762*** (-2.9902)	-
D3 t-1	-	-213.823 (-0.3684)	-	327.809 (1.0650)
Adj. Rsquare	0.96	0.91	0.91	0.92
Durbin-Watson	2.12	1.73	1.88	
F	134.03	16.59	95.43	34.20
MAE	97.60	98.97	312.747	270.215
U statistic	3.99	3.19	1.38	

Note: *, **, and *** indicate significance levels of 0.1, 0.05, and 0.01, respectively. Numbers in parentheses represent t-values.