Forecasting the U.S. Cotton Industry: Structural and Time Series Approaches

by

Dean T. Chen and David A. Bessler

FORECASTING THE U.S. COTTON INDUSTRY: STRUCTURAL AND TIME SERIES APPROACHES

Dean T. Chen and David A. Bessler

I. INTRODUCTION

Structural econometric models and vector autoregressions have been widely adopted by agricultural economists for forecasting, policy analysis, and behavioral hypothesis-testing research. Traditionally these two approaches have been viewed as distinct and competitive. The structural model emphasizes the theoretical description of behavioral relations which impose identifying restrictions on model specification, while the vector autoregressive model, on the other hand, focuses upon reduced form estimation with few parameter restrictions and does not attempt structural interpretations of data.

In the past, the reseachers of these two approaches have generally maintained a self-contained attitude, claiming the superior predictive performance of their approach over the other. Despite this competitive state of affairs, some model-builders have begun to explore the complementary nature of these two approaches, in particular, the investigation of combining vector autoregressive models with structural models to determine the future values of exogenous variables (Fair) or to adjust the constant terms of stochastic equations of macroeconomic models (Klein, 1984).

This study attempts to combine the structural and vector autoregressive approaches of econometric modeling work in agriculture, using the applied setting of the U.S. cotton subsector. The cotton market is particularly interesting because it has recently undergone a substantial policy change which may be difficult for any model to capture in ex-ante forecasts. Due to the 1985 Farm Bill provisions for a marketing loan, U.S. cotton price fell from $.67 per pound to $.27 per pound within a one-week period. The drop created a challenging period for testing the model's capability in forecasting its future path. Under such circumstances, time series models are not expected to perform well. They should provide improvement, however, in forecasting performance in a relatively stable environment and may be useful to reduce uncertainty of exogenous variables as input to the structural model.

Another important motivation for this study comes from the methodological issues of forecasting accuracy analysis. Most of the methods that have been used in past model evaluation studies can only be considered oversimplified with the use of an unrealistic set of performance criteria. The procedures are in need of major modification in order to reflect actual forecasting situations. This study attempts a multi-dimensional approach for evaluating the forecasting performance of models, especially the important differences in forecasts between ex-post and ex-ante, static and dynamic, within sample and outside sample, parameter updates and no updates, single equation and complete model, and low and high frequency forecasting conditions. The first four areas are considered in the design of simulation experiements in this study.

The authors are professors of agricultural economics at Texas A&M University, College Station, Texas. The computational assistance by George T. C. Chiou, a graduate student at Texas A&M University, is gratefully acknowledged.

The paper is presented in five sections. First we discuss the theoretical aspects of model performance analysis. We then relate these to the issues of combining structural and time series models. A comprehensive set of the multi-dimensional forecast evaluation criteria is proposed. In section III an overview of the structural econometric model of the U.S. cotton industry is presented. This is followed by a brief discussion of the time series method employed in this simulation study. The predictive testing results of the structural and vector autoregressive models, both singly and in combination, are presented in section IV. The final section of this paper contains an overall summary and some suggestions for further research.

II. EVALUATION AND INTEGRATION OF FORECASTING MODELS

In evaluating econometric model performance, there are four major areas of concern: (1) the stochastic disturbance terms associated with the model, (2) the parameters of the model, (3) the assumed input values of exogenous variables and (4) the specification of the model. A structural econometric model in its forecasting application forms (pure and adjusted), can be written as (Chen, 1981):

$$(1) \quad \hat{Y}_t = g(\ 0,\ \hat{x}_t;\ \hat{\theta}),\ \text{or}$$

$$(1a) \quad \hat{Y}_t = g(\ 0,\ \hat{x}_t;\ \tilde{\theta}).$$

In an actual forecasting exercise, the model is used to generate the predicted values of the endogenous variable $\hat{Y}_t$ based on the estimated values of parameter $\hat{\theta}$, the expected values of the structural disturbances of $E(e_t)$ at zero, and the assumed input values of exogenous variables $\hat{x}_t$ over the prediction period. Largely due to data revision and the availability of non-sample and non-model information, the econometric forecasters must adjust the constant terms of the models in the preparation of forecast. Therefore, as shown by Equation (1a), a vector of the adjusted values of the parameters, $\tilde{\theta}$ instead of $\hat{\theta}$ is actually used in the model.

A Comprehensive Set of Evaluation Criteria

Model evaluation is a multiple dimension problem. We look at the model's forecast performance from several perspectives. A model can be evaluated as to its performance over the historical period of observation (within sample evaluation) for which the parameters ($\hat{\theta}$) of the model were estimated or over a different time period (outside sample evaluation). A model can be evaluated under an environment of no parameter updating or parameter updating can be allowed. The model's forecasts should take into account exogenous variable uncertainty and the fact that forecast error variances vary across time. If the actual values of the exogenous variables (x_t) are used, the evaluation is based on ex-post forecast, while the ex-ante forecast uses the predicted values of exogenous variables ($\hat{x}_t$). The model's forecast can be dynamic or static. The static forecast means the use of actual values of lagged endogenous variables (Y_{t-i}), while the dynamic forecast is based on the predicted value of lagged endogenous variables ($\hat{Y}_{t-i}$). In addition, a model can be evaluated on its own or as it contributes to forecasts from a combination with one or more additional forecasting models.

The design of the simulation experiments and the particular aspects of forecast evaluation which we considered in this paper are summarized in table 1. To highlight these forecast evaluations, a total of 30 simulation experiments are listed in the table.

Entries in Table 1 are of six general types. SIMi entries refer to forecasts from structural models for the time period of policy shocks due to the implementation of the marketing loan program in 1986. Here the index i runs from 1 to 7, as seven different aspects of forecast evaluation are considered. The VARi entries refer to forecasts from a vector autoregression. Finally, the SAVi entries refer to evaluations of the combined structural and autoregressive models. The same types of simulations were performed with another sample period of observations to represent an ordinary time period in 1984. Entries of TESi, VESi, and TAVi, are 1984 simulation runs of the structural model, vector autoregression, and their combinations, respectively.

Table 1. SIMULATION EXPERIMENT DESIGN:
Multi-Dimensional Predictive Performance Evaluation

	Within Sample	Outside Sample					
	No Parameter Updates	Parameter Updates			No Parameter Updates		
	Ex-Post	Ex-Post	Ex-Ante		Ex-Post	Ex-Ante	
	Static (1 pd)	Static (1 pd)	Static (1 pd)	Dynamic (5 pds)	Static (1 pd)	Static (1 pd)	Dynamic (5 pds)
1986 Policy Shock Period:							
Structure Model	SIM1	SIM2	SIM3	SIM4	SIM5	SIM6	SIM7
Vector Autoregr.			VAR3	VAR4		VAR6	VAR7
S/V Combined			SAV3	SAV4		SAV6	SAV7
1984 Ordinary Time Period:							
Structure Model	TES1	TES2	TES3	TES4	TES5	TES6	TES7
Vector Autoregr.			VES3	VES4		VES6	VES7
S/V Combined			TAV3	TAV4		TAV6	TAV7

A common approach in model evaluation is to track the model's performance within the sample period for which the parameters were estimated. In this context, historical curve fitting have become a dominant factor in model research. Te deviate from this approach, we extend our model evaluation to outside sample, and allow parameter updates for both the static (1 step ahead) and dynamic (5 steps ahead) conditions. Successive reestimation is assumed to improve parameter precision and model performance over its simulation path.

In a realistic forecasting environment, the input values of exogenous variables are also unknown. In this study, both the ex-post and ex-ante forecasting performance are evaluated. Although a good deal of intelligence is available to structural model forecasters on the future values of the exogenous variables, a simple extrapolation of the exogenous variable is commonly used. For our simulation research, a naive no-change extrapolation was used to generate the ex-ante values of exogenous variables under static and dynamic assumptions. In actual forecasting situations in which the future is truly unknown, a realistic model evaluation must be outside sample, ex-ante and dynamic -- the 7^{th} simulation entry presented in table 1. This study attempts a comprehensive set of model evaluation criteria. The procedure and a theoretical discussion of the various methods which we have considered in this study can be found in Fair (1986).

Integration of Structural and Time Series Models

Integration of alternative econometric models has been considered by several authors. Early work of Granger and his coworkers suggests integration of models through linear combinations of several forecasts. While this work continues to generate interest among applied researchers, it is not the path followed in this paper. An alternative approach of integration is outlined in a paper by Ashley. Here the author use time series analysis to model the residuals from a structural ecoometric model. Forecasts of the relevant dependent variable are then generated as the sum of the forecast of the econometric model and the forecasts of the residuals. A variant of this procedure is suggested by Klein and Sojo - where they use time series analysis to extrapolate the high-frequency (monthly) data to improve the ex-ante forecast of the related low frequency (quarterly) endogenous variables in the structural model. An adjustment is added to the right-hand side of each econometric equation so that the equation generates the same value for the left-hand side as that given by the time series model. This procedure is an alternative to the usual subjective adjustment factor, which is often used in real time econometric forecasting. We do not follow this approach in this paper either.

Here we follow the idea given in Fair, where we use time series analysis to generate forecasts of the relevant exogenous variables in an econometric equation. These forecasts are generated in the usual way from the historical regularities in the data.

Consider the structural equation given in equation (1), where y_t represents an endogenous variable of the forecasting model; $g(\bullet)$ represents a functional relationship which transforms the contents of the parenthesis into y_t; x_t represents a vector of observed exogenous variables; and θ represents a matrix of unknown parameters, which are to be estimated with an appropriate GNLS (generalized nonlinear least squares) estimator.

In many cases the form of equation 1 is simplified to be linear. This permits a simple representation as in equation (2):

(2) $Y_t = A_0 + A_1 X_t + \epsilon_t$,

where X_t is again an exogenous variable, ϵ_t a white noise residual and A_0, A_1

appropriately estimated GLS parameters. A common procedure would be to substitute lagged X_t (X_{t-1}) into equation (2) to generate a structural forecast of Y_{t+k} at all future dates.

If, however, through time series techniques, one can identify the process which generates X_t as a $k^{\underline{th}}$-order AR, then an alternative to the above procedure would be to substitute the time series forecasts from this autoregression into equation (2). Suppose that AR is given as equation (3):

$$(3) \quad X_t = B_0 + B_1 X_{t-1} - \ldots - B_k X_{t-k} + v_t,$$

where X_t is the exogenous variable in equation (2) and v_t is a white noise innovation. Forecasts from equation (3) can be generated by application of the chain-rule of forecasting -- where forecasts at horizons h > 1 are found by substituting earlier forecasts into equation (3) and treating them as if they were the actual values. Equation (4) is a representation of this rule for the h-step-ahead forecast:

$$(4) \quad \hat{X}(t+h) = \hat{B}_0 + \hat{B}_1 \hat{X}(t+h-1) + \hat{B}_2 \hat{X}(t+h-2) + \ldots$$

$$+ \hat{B}_h X(t) + \ldots + \hat{B}_k X(t+h-k).$$

Here $\hat{B}_0$, $\hat{B}_1$, ..., $\hat{B}_k$ are appropriately estimated GLS parameters. Note here that $\hat{X}(t+h-k)$ represents the earlier k<u>th</u>-step-ahead forecast from the recursion. The entries in table 1 labeled SAVi and TAVi represent forecasts from combinations of this type.

III. <u>AN OVERVIEW OF THE STRUCTURAL MODEL</u>

The structural model of the U.S. cotton industry presented in this paper represents an ambitious attempt to model the farm commodity sector. Guiding this model-building effort is the priority consideration given to theoretical structure, information content, and practical operation.

A unique feature of this model is the implicit revenue function, a new theoretical framework for the derivation of structural relations linking government farm programs with almost every stage of farm decision process. This model also emphasizes its information processing role, seeking maximum points of contact with regular government reports and industry data. For operational efficiency, this microcomputer-based econometric model is run on an IBM-PC(AT), utilizing Advanced Retrieval Econometric Modeling System (AREMOS) software for database management, model estimation, forecast solution, and impact simulation.

<u>Implicit Revenue Function and Farm Program Simulator</u>

In modeling farm commodity sectors for policy analysis, numerous alternative specifications have been explored in the past. Previous model work has concentrated on either supply response studies which elaborate farm program analysis or the construction of complete sector models which include a few ad hoc government policy variables. To provide comprehensive treatment of the impact of government policy actions, a theoretical specification of implicit revenue function is introduced.

Based upon microeconomic theory, econometric relations of firm take several different forms, e.g., the production function and efficiency conditions, production and factor demand functions, supply functions, and cost and revenue functions. Although they have equal standing in economic theory, the specifications are different in terms of parametric information and data requirements (Klein 1982). This cotton model follows the cost and revenue function approach, for which the producer is assumed to maximize the expected net returns subject to the constraints imposed by government programs.

This specification takes into account the implicit nature of producers' revenue in terms of direct and indirect government program benefits. A simplified representation of the producers' net operating return (NOR), is expressed as the difference between total operating return (OR) and cost (OC):

$$(5)\ NOR = OR - OC$$

$$= \Sigma\ R_i * SY + \Sigma\ G_j * SYG - \Sigma\ C_k * SY$$

where $i = 1,2;\ j = 1,\ldots,4;$ and $k = 1,\ldots,k.$

In equation (5), SY and SYG are actual yield per acre and program payment yield per acre, respectively; while R_1 multiplied by SY represents cash receipts from marketing per acre, R_2 multiplied by SY represents net loan receipts, G_j multiplied by SY represents the four categories of government's direct payment considered in this model, and C_k represents the cost items.

Based upon the implicit revenue function specification, the interactions of program instruments with cotton market variables are summarized in the following equation (Chen, 1987):

$$(6)\ OR = CR + NLR + DFG + LDFG + DVG + DAG$$

$$= [\beta_o * PF * SY * SA] + [(1-\beta_o) * PNL * SY * SA]$$

$$+[(PT - Max(PL, PF)]*SYG*[1 - (\gamma_1*SARP+\gamma_2*SPLD)]*SALO$$

$$+[(PL - PLR) * SYG * [1 - (\gamma_1*SARP + \gamma_2*SPLD)]*SALO$$

$$+[PDVG * SYG * (\gamma_2 * SPLD * SALO)] + [0.75 * SYG * (0.33 * PT)].$$

Here, producer's operating return (OR) is a sum of cash receipts (CR), net loan receipts (NLR), deficiency payment (DFG), loan deficiency payment (LDFG), diversion payment (DVG), and disaster payment (DAG). This equation also contains the various program instruments: loan rate (PL), loan repayment rate (PLR), net loan rate (PNL), target price (PT), program payment yields (SYG), percent of ARP acreage reduction (SARP), percent of paid land diversion (SPLD), base acreage (SALO), and diversion payment rate (PDVG), in addition to the cotton market variables of price received by farmer (PF), yield per acre (SY), and planted acres (SA). There are two sets of behavior response parameters, β_i, and γ_i, in equation (6). The former, β_i, describes producers' decision in allocating cotton output to be sold on the spot market and in determination of CCC loan entry or redemption. The latter, γ_i, represents producers' decision for participating in the acreage reduction program, either the mandatory ARP (acreage reduction program) or the voluntary PLD (paid land diversion).

In view of the complexity of the current farm program, a separate Farm Program Simulator is developed. Based upon the provisions of the 1985 Food Security Act, six important program parameters, α_i (i=0,...,5) are used in the simulator. These include one parameter for target price determination (α_o), four for loan rate determination -- spot market calculation α_1, Northern European calculation α_2, percent adjustment from preceding year α_3, low limit permitted over program year α_4, and one for loan repayment rate determination (α_5), under either Plan A or Plan B of the marketing loan provisions.

Development of the Farm Program Simulator helps provide the transmission mechanism in the model, tracing the effect of farm program changes on acreage response, market price determination, CCC loan activity, inventory stock adjustment, farm income and government payment.

<u>Structural Characteristics of the Model</u>

The current version of this cotton model is a 67-equation system with 15 behavioral equations and 52 identities. From the specification viewpoint, it is a fully integrated model linking the domestic market block with a Farm Program Simulator and the world market block. The Farm Program Simulator is by far the largest block in the model with 58 variables describing policy instruments and parameters, and producers' operating returns and costs in detail. Gross and net operating returns are determined either on per acre or per pound basis. To analyze the effect of intercrop competition, operating returns and cost variables for five major crops including barley, corn, sorghum grain, soybean, and wheat, are used in the model.

The domestic market block contains monthly equations of domestic mill consumption, ginning, and export sales. Memphis prices, average price received by farmers, cash receipts, and other income components are also determined monthly. The annual crop production equations include planted acreage for four major crop regions of the Southeast, Southwest, Delta, West, and the State of Texas. Yield per acre equation for the U.S. is endogenously determined. This allows crop production estimates for the U.S. as a whole, providing acreage detail on a regional basis. The acreage response equation reflects profit maximization behavior derived from the implicit revenue function. Producers' net operating return, diversion payment, and intercrop competition are the key exogenous variables. Soybean was found to be a significant competing crop in the Southeast and Delta areas, as sorghum grain was in the Southwest and Texas.

In developing world market equations, the theoretical specification of trade flow and market share, particularly the two-stage decision process model, was adopted. The model contains annual equations for total world cotton import demand and U.S. export market share, and monthly equations for U.S. cotton exports. The key variables in export equations are U.S. cotton prices at the Memphis market and world market prices at Liverpool and the weighted average exchange rates of six major countries. Total mill consumption, harvest acreage, and production for the Rest-of-World totals are also determined endogenously in the model. The basic identity for achieving the supply-utilization balance of this two-region model is also included.

This cotton model is constructed to combine data of multi-frequencies, annual and monthly. This type of modeling work requires a proper linkage mechanism for data conversion and the generation of multi-frequency simultaneous equation model solution. An important consideration in modeling the cotton sector is the development of market expectation variables in the price equations for U.S. (Memphis) and world (Liverpool) markets. This model emphasizes "forward-looking", rather than "backward-looking" expectation formulations. Based upon the monthly data released by the World Agricultural Outlook and Situation Board of USDA, a dynamic market expectation model is constructed. Therefore, the model is useful for impact simulation analysis of the USDA supply-utilization outlook estimates.

Price Determination Equations for Policy Impact Simulation

Price and income equations are the heart of commodity sector models for forecasting and policy analysis. These equations are subject to critical evaluation through an ordinary time period of stable price movement and a period of substantial policy changes. In the past eighteen months, cotton market price has been dominated by the influence of the 1985 Farm Bill, because of the implementation of the marketing loan program. This policy action was designed to boost U.S. export, to reduce stock levels, and to ensure competitive U.S. prices on the international market. The implication from this action was a shift of the effective price floor from the domestic loan rate to adjusted world prices, the former being substantially higher. Developments of this type had never occurred in the historical period. One would suspect that the structural parameters of the model would be unstable. Consequently the model would not be useful for forecasting the future path.

However, the theoretical structure of the price equation selected for simulation experiments has properties suitable for forecasting these type of policy changes. The theory underlying this price equation can be sketched as follows: First, the behavior equation for U.S. monthly cotton price is estimated by a deviation term relating Memphis price to the effective price floor. Second, through an identity relation, as shown in equation (7), we can change the effective price floor from the original specification of effective domestic loan rate to adjusted world prices. This mechanism is particularly useful for impact simulation in the marketing loan program. Third, these price equations are constructed to reflect the theoretical framework of stock-demand functions. Following the conditional expectation hypothesis, three expectation terms are also used in the model (equation (8)). This specification has proven to be particularly valuable in tracking developments in the domestic and international markets and in reflecting the dynamic process of market equilibrium.

(7) Identity for Memphis Cotton Price

COLPMME116 = COLPMDPLL + (COLPFLLD1 * COLPLE + COLPFLLD2 * COLAWP)

where COLPMME116 is the cotton market price, c/lb, Memphis slm 1 1/16 inch; COLPMDPLL is the deviation of Memphis price from effective price floor; COLPLE is the effective loan rate, c/lb, using base loan rate adjusted by interest charge and storage costs through the crop season; COLAWP is the adjusted world prices, c/lb, Liverpool market, the A index series, adjusted by transportation costs and quality differences between the U.S. and Liverpool markets; COLPFLLD1 and COLPFLLD2 are two dummy variables used to represent policy changes, implementation of the 1985 Farm Bill provision of marketing loan August 1, 1986; COLPFLLD1 equals one prior to August 1986 and zero otherwise; COLPFLLD2 equals one after August 1986 and zero otherwise.

(8) Behavioral Equation for Monthly Cotton Price, Deviation from Effective Price Floor

$$\begin{aligned} COLPMDPLL = & - \underset{(2.2820)}{0.3899} * COLPMME116F - \underset{(1.4744)}{0.2234} * COLHTDT - \underset{(2.2011)}{0.0734} * COLHTDTX \\ & - \underset{(3.5682)}{0.0009} * COLDA2 - \underset{(4.2167)}{21.1698} * COLHTDTR + \underset{(12.2323)}{17.0763} * USMXPRC + \underset{(4.1296)}{73.7951} \end{aligned}$$

Sum Sq	2104.95	Std Err	4.5652	LHS Mean	9.9990
R Sq	0.8063	R Bar Sq	0.7947	F 6,101	70.051
D.W.(1)	0.9062	D.W.(12)	1.9932		

where COLPMDPLL is cotton price deviation from effective price floor, c/lb, Memphis market; COLPMME116F is the seasonal adjustment factor of Memphis cotton price estimated by Census Bureau X11M method; COLHTDT is the U.S. monthly-ending stock-to-demand ratio; COLHTDTX is the expected U.S. stock-to-demand ratio at the end of the current crop year; COLDA2 is the expected U.S. total supply for the second crop year ahead, including expected ending carryover stock and the new crop; COLHTDTR is the Rest-of-World ending stock-to-demand ratio; and USMXPRC is the dummy variable for measuring the entry of P.R.C. into U.S. export market in the early 1980s.

IV. THE TIME SERIES MODEL

The method used to summarize the time series properties of the data is a Bayesian vector autoregression (see Litterman). This prior treats each variable as a random walk, with varying degrees of tightness to permit differential degrees of series interactions. We include in this general specification the seasonal dummy variables, for which we provide no prior.

The essential features of this model are that the researcher specifies the degree of interaction among the variables of a multiple time series. While the prior is centered on a random walk for each variable, by specifying differential levels of tightness on each variable in each VAR equation, the researcher can allow the data to have some influence on the resulting forecast.

Three types of information must be specified in the "Litterman prior." Overall tightness reflects the prior standard deviation on the coefficient on the first lag of the dependent variable. This was set at .25 to be consistent with our earlier study of empirical data (see Bessler and Kling). The rate of decay on tightness of coefficients of lagged variables (beyond one period) was set at 1.0. Finally, the interseries tightness parameters were set following some initial pretesting. Table 2 summarizes that information. Based on work of the second author, the model was specified by application of Hsiao's recursive procedure. This provided a guide on where to place strong or loose restrictions on the data. When the FPE model included a variable in a particular equation, that variable in that equation was assigned a prior tightness value of .8 -- indicating that the data were given considerable influence on the resulting forecast.

We deviated slightly from this procedure by assigning a prior tightness value of 1.0 to own lags of a particular variable (see diagonal elements of table 2). When the FPE-specified model did not include a variable in a particular equation, that variable was given a prior tightness value of .1 in that particular equation - - indicating that this variable was not given much influence on the resulting forecast of the particular equation.

TABLE 2. Prior Tightness Levels of the Vector Autoregressive Model

	Equation				
lagged variables	World price	Expected supply	U.S. stocks	World stocks	Memphis price
World price	1.0	.1	.1	.8	.1
Expected supply	.8	1.0	.8	.8	.1
US stocks	.8	.1	1.0	.1	.1
World stocks	.1	.8	.1	1.0	.1
Memphis price	.1	.1	.1	.1	1.0

Priors were specified by pretesting the data with FPE criteria.

Clearly our procedure is an ad hoc way of providing priors (indeed, some may not call what results a prior at all); however, it does provide a quick method of reducing the variable interactions. The alternative procedure of searching for optimal tightness settings over earlier periods (see Doan, Litterman, and Sims) was viewed as too costly and not followed. A third procedure of actually eliciting priors of real world experts was not considered either. The forecasts from this specification are discussed in the results section which follows.

V. THE EMPIRICAL SIMULATION RESULTS

Results from forecasting over two periods are given in tables 3 and 4. Table 3 is associated with forecasts over the period August through December 1986. This period is significant because the government program for cotton changed such that prices in August were substantially different from those in previous periods. Forecasts from the structural model are labeled "Structure" in the table; those from the times series model are labeled "Vector Auto"; and those from the combined model are labeled "Combined". The forecasts are evaluated under seven dimensions of the forecasting problem -- within sample with no parameter updating, etc. The VAR models and the Combined Structural/VAR forecasts are presented for just four model types -- outside sample with parameter updating and static; outside sample with parameter updating and dynamic; outside sample with no parameter updating and static; and outside sample with no parameter updating and dynamic.

From table 3 note that the structural model outperforms the VAR and the combined forecast at all horizons and over all model types. The VAR failed to forecast the drop in price in August. While the VAR adjusts quickly to the new level of prices in the static model (in Sept. the VAR begins to forecast in the low twenty cents per pound range), it shows no evidence of adjusting in the dynamic forecasts.

Structure, Vector Autoregressive, Combined

Table 3 COTTON PRICE | Table 4 COTTON PRICE

	1986 AUG	1986 SEP	1986 OCT	1986 NOV	1986 DEC	1984 AUG	1984 SEP	1984 OCT	1984 NOV	1984 DEC
	POLICY SHOCK:MARKETING LOAN PERIOD.........					ORDINARY TIME PERIOD SIMULATION..........				
ACTUAL PRICE, COTTON.. Memphis 1-1/16 c/lb.	26.60	33.58	41.93	44.09	51.09	63.05	60.67	60.83	60.44	60.83
WITHIN SAMPLE NOPARA. UPDATE EX-POST STATIC ...Structure SIM1...	25.13	33.76	42.95	44.04	51.85	61.48	63.49	62.12	61.54	62.08
OUTSIDE SAMPLE PARA. UPDATE EX-POST STATIC ...Structure SIM2...	25.27	33.74	42.97	44.18	51.99	68.91	68.49	63.28	62.19	62.36
OUTSIDE SAMPLE PARA. UPDATE EX-ANTE STATIC ...Combined SAV3....	27.23	25.49	35.37	47.57	44.86	76.56	71.59	65.13	67.84	62.25
...Structure SIM3...	28.54	29.69	36.00	46.20	47.19	67.32	68.96	64.33	63.92	62.30
...Vector Auto VAR3.	64.16	20.67	35.03	44.33	46.09	65.72	60.40	59.83	60.45	60.14
OUTSIDE SAMPLE PARA. UPDA EX-ANTE DYNAMIC. ...Combined SAV4....	24.09	25.68	25.97	26.39	26.29	78.15	77.69	73.60	73.52	72.96
...Structure SIM4...	24.60	42.88	43.45	44.78	44.93	68.91	68.96	64.81	65.00	65.25
...Vector Auto VAR4.	64.16	55.73	55.97	55.98	55.98	65.72	65.15	65.17	65.20	65.20
OUTSIDE SAMPLE PARA. UPDATE EX-POST STATIC ...Structure SIM5...	25.14	33.82	43.02	44.09	51.91	68.52	70.68	69.24	68.68	69.14
OUTSIDE SAMPLE NOPARA. UPDATE EX-ANTE STATIC ...Combined SAV6....	27.23	25.61	34.78	46.29	44.62	76.08	72.03	68.98	70.87	66.52
...Structure SIM6...	28.54	29.90	36.17	46.25	47.23	66.84	68.96	68.30	68.46	67.19
...Vector Auto VAR6.	64.16	22.28	35.06	44.60	46.39	65.72	60.64	59.95	60.32	60.04
OUTSIDE SAMPLE NOPARA. UPDA EX-ANTE DYNAMIC. ...Combined SAV7....	24.09	23.63	22.04	21.99	21.82	75.54	75.84	73.92	74.09	73.72
...Structure SIM7...	25.41	43.09	43.62	44.81	44.97	66.30	68.96	68.77	69.55	70.01
...Vector Auto VAR7.	64.16	62.86	62.72	62.31	62.03	65.72	63.93	63.77	63.69	63.76

Multi-Dimensional Evaluation of Cotton Price Forecast Simulations

Quite understandably, the VAR did not have the structural change information in the latter (dynamic) models and thus continued to forecast business as usual (the dynamic model is a five step ahead forecast while the static model is a one step ahead forecast). Parameter updating is not all that helpful in improving the static model, but does show considerable improvement in the dynamic model. In fact, under the static version of the forecast the MSE of the VAR is actually higher under the parameter updating scenario; while the MSE falls by about 15% when parameter updating is allowed in the dynamic version of model evaluation.

The structural model tracks the 1986 cotton prices very well. Under both static (one step horizons) and dynamic (five step ahead horizons) simulations, the forecasts are quite close to the actual prices realized in the period.

Table 4 presents the forecasts from the period August 1984 through December 1984. The same forecast (simulation) types are presented for this early period. Our reasons for considering this period is that it represents a more business as usual period (no structural change), even though we found it was a period of PIK when the farm commodity market was under substantially downward price pressures due to macroeconomic policy and world-wide surplus conditions of grains and oil crops.

Table 4 contains the forecasts of all models over the 1984 period. Here the results are quite the opposite from those presented for the 1986 period. The MSE calculations for this early period are given in table 6. The VAR model outperforms both the structural model and the combined forecasts according to the MSE metrics in all versions of the forecast simulations. Parameter updating of the VAR does not seem to be benefical; in fact it results in higher MSE in both the static and dynamic specifications. The static model outperforms the dynamic model, as we would naturally expect. The structural model performs better under the static version of the simulation experiment in the early time period (again as we would expect). Parameter updating is quite helpful in both the static and dynamic specifications (it reduces MSE by about 31% in the former and about 28% in the latter). This is quite different than the result found in the VAR simulations (where parameter updating was apparently harmful).

The combined forecasts for the early period (1984) did not perform well. They were actually worse than both the VAR forecasts and the structural forecasts. This is apparently due to either misspecification of the time series process, which generated the VAR variables, or misspecification of the econometric model. Under the former hypothesis we probably would expect to see poor forecasts of the Memphis price from the VAR as well; which of course we do not observe. Of course, the structural model was not constructed with the VAR forecasts. If the time series model and the structural model were constructed as one system, the combined model may have shown better results.

The results of all simulation experiments for the time periods of 1986 and 1984 are summarized by the Root Mean Squared Error (RMSE) statistics in Tables 5 and 6.

From table 5 note that parameter updating is not all that useful in improving the forecast performance of the structural model. Updating the static forecasts results in a slightly higher RMSE, while updating the dynamic forecasts results in a .5% reduction in RMSE. As one would expect, the

Tab 5 ERROR ANALYSIS

Policy Shock-- 1986 Marketing Loan Period

	WITHIN SAMPLE EXPOST NOPARA UPDATE STATIC	OUTSIDE SAMPLE EX-ANTE PARAMETER UPDATE STATIC			OUTSIDE SAMPLE EX-ANTE PARAMETER UPDATE DYNAMIC			OUTSIDE SAMPLE EX-ANTE NOPARAMETER UPDATE STATIC			OUTSIDE SAMPLE EX-ANTE NOPARAMETER UPDATE DYNAMIC		
	Struc. Model	Str/Var Combined	Struc. Model	Vector Autoregressive	Str/Var Combined	Struc. Model	Vector Autoregressive	Str/Var Combined	Struc. Model	Vector Autoregressive	Str/Var Combined	Struc. Model	Vector Autoregressive
1986 DEC	0.88	5.65	3.84	18.17	15.82	5.12	21.28	5.69	3.74	17.93	19.21	5.15	25.11

Tab 6 ERROR ANALYSIS

Ordinary Time Period--1984 Simulations

1984 DEC	1.72	8.68	4.76	1.32	14.14	5.64	4.19	10.06	6.98	1.31	13.49	7.87	3.02

Root Mean Squared Error, For 5 periods ending Dec.

static version of the forecast (the one step ahead forecast) is better than the dynamic version of the forecast. Under the parameter updating simulations, the static model represents an approximate 25% improvement over the dynamic forecast; while the static model represents a 27% improvement in MSE over the dynamic model when no parameter updating is allowed. The combined forecasts do not perform particularly well in the 1986 period. In fact, in terms of the MSE metric these forecasts fall between the structural model and the VAR. Generally, the combined forecasts seem to get closer to the structural model's performance in the static simulations, than in the dynamic simulations.

As shown in table 6, the combination of VAR with structural model failed to improve the 1984 forecasting performance in virtually all cases. The results seem to reject our hypotheses that the VAR forecast of the exogenous variables would help improve the accuracy of structural model results. A careful evaluation of 1984 period cotton price forecasts indicates an upward forecasting bias of cotton price by the structural model due mainly to a general commodity price downturn in this period. This is a period in which the commodity market was heavily under the pressure of macroeconomic policies and the resultant implementation of PIK.

VI. SUMMARY AND SUGGESTIONS FOR FUTURE RESEARCH

In the early works on composite forecasting the authors argue that several alternative models can often be combined to yield mean squared errors which are lower than either of the individual forecasts (Granger and Ramaanathan, and many others). Rarely have the composites performed worse than all of the individual models. In fact, some researchers express disappointment when the composite does not show improvement over the best individual method! Here we have a situation where a combination does worse than the individual methods over one forecast evaluation period (1984) and shows no improvement over the best method over another forecast evaluation period. Thus, we must label our attempt to combine the two approaches to building forecasting as a failure. However, the results do suggest areas where we can look to improve future efforts.

The areas for future study seem to be in joint specification of the structural model and the time series properties of the exogenous variables. Here we treated the problems as separate and merely fed the process which generated the exogenous variables into the structural specification. Perhaps treating them as a system would improve forecast performance. Errors in the process which generated several of the exogenous variables could quite likely be correlated with those from the structural equations (perhaps a SUR estimation could be studied).

Other areas include more detailed analysis of the time series prior imposed on the data. Recall that we used a random walk prior with variable degrees of series interactions. This may prove to be an unreasonable prior under further study. We suggest a more formal data analysis (similar to that done in Doan, Litterman and Sims) be undertaken in future research with these data.

The 1984 experience of the structural model suggests that an important consideration should be given to the non-sample and non-model information such as the situation in the PIK period of 1984. A promising route which may be studied in the future is to use the VAR forecasts as "objective" adjustments to the structural equation intercept. Here we used the VAR forecasts of exogenous variables as input to the structural equation. An alternative procedure would be to use the VAR forecasts to adjust the structural forecast. Following Klein (1986), we can adjust the structural forecast equation with the VAR forecast - in normal time periods. This will be equivalent to adjusting the intercept on the structural equation in an objective - replicatable fashion. Heretofore the subjective adjustment procedure has been criticized as being not replicatable.

The procedure is suggested for normal - time forecasting and probably should be viewed cautiously in periods where the historical regularities are thought to be changing. In particular, we would not expect this procedure to work well in the Aug. - Dec. 1986 period. Here the model builder had prior information that a change was about to occur. He would not be well advised to ignore that by conditioning on the VAR forecast.

This structural model for the U.S. cotton industry has been tested for predictive performance on the basis of seven simulation experiments for two different time periods in 1984 and 1986. The results are generally quite promising for forecasting and policy analysis purposes. Further validation tests should be in the area of ongoing forecasting operation with adequate combination of the time series models.

References

Ashley, R. "Residual Analysis of Macroeconomic Models" Journal of Macroeconomics, 18(1983):73-89.

Bessler, D. and J. Kling. "Forecasting Vectorautoregressions with Bayesian Priors," American Journal of Agricultural Economics, 68(1986):144-151.

Chen, Dean T. "An Integrated Cotton Model with Interactions Between Domestic Program and International Market." Proceedings of the Beltwide Cotton Production Research Conferences, Dallas, Texas, January 7-8, 1987, forthcoming.

Chen, Dean T. "Error Decomposition in Commodity Price Forecast Evaluation." Proceedings of the Conference on Applied Commodity Price Analysis and Forecasting, October 12, 1981, Ames, Iowa.

Doan, T., R. Litterman, and C. Sims. "Forecasting and Conditional Projections Using Realistic Prior Distributions," Economic Reviews, 2(1984):1-48.

Fair, R. "Evaluating the Predictive Accuracy of Models", Handbook of Econometrics, Vol. III, eds., Z. Griliches and M.D. Intriligator, Elsevier Science Publishers, 1986, pp. 1980-1995.

Granger, C. and R. Ramanathan. "Improved Methods for Combining Forecasts", Journal of Forecasting, 3(1984):197-204.

Hsiao, C. "Autoregressive Modeling of Canadian Money and Income Data," Journal of the American Statistical Association. 74(1979):533-60.

Klein, L. "Economic Theoretic Restrictions in Econometrics," in Evaluating the Reliability of Macroeconomic Models, eds., Gregory C. Chow and Paolo Corsi, Chichester: John Wiley, 1982, pp. 23-28.

Klein, L. "The Importance of the Forecast", Journal Of Forecasting, 3(1984):1-9.

Klein, L. and E. Sojo. "Combinations of High and Low Frequency Data in Macroeconometric Models". Unpublished manuscript, Department of Economics, University of Pennsylvania, Philadelphia, PA., 1986.

Litterman, R. "Specifying Vector Autoregressions for Macroeconomic Forecasting", Bayesian Inference and Decision Techniques with Applications: Essays in Honor of Bruno de Finetti, eds., Goel, P. and Zellner, A., Amsterdam: North Holland Publishing Company, 1986.