AGRICULTURAL DEVELOPMENT SYSTEMS
EGYPT PROJECT

UNIVERSITY OF CALIFORNIA, DAVIS

TOWARDS A MODEL FOR FORECASTING COTTON PRICES

By

Donald G. Heckerman
University of Arizona

WORKING PAPER

ADS/EGYPT

TOWARDS A MODEL FOR FORECASTING COTTON PRICES

By

Donald G. Heckerman
University of Arizona

Economics
Working Paper Series
No. 52

Note: The Research Reports of the Agricultural Development Systems: Egypt Project, University of California, Davis, are preliminary materials circulated to invite discussion and critical comment. These papers may be freely circulated but to protect their tentative character, they are not to be quoted without the permission of the author(s).

December, 1981

Agricultural Development Systems:
Egypt Project
University of California
Davis, Ca 95616

Uses of International Cotton Market Price Forecasts

It is important to identify the potential applications of cotton price forecasts, for the nature of the applications will determine the relative importance of producing different types of forecasts. Some applications require forecasts for periods which extend to 10, 20 or even 30 years into the future. Other applications require no more than a one year horizon, while still others have horizons as short as one to three months. Some applications require a forecast for only one date in the future, while other applications require forecasts not only for the horizon date but also for periodic intermediate dates. Different applications also have different requirements for the frequency with which forecasts are revised in response to the arrival of new information. In the face of limited resources available for the forecasting activity, it is important to identify priority uses of forecasts so that price forecasting activity may be of maximum usefulness.

Among the possible uses of cotton price forecasts are as inputs to decisions about the number of feddans to plant, about the amounts of variable factors such as fertilizer and pesticides to apply to cotton plants already in the ground, about optional timing for selling in the international market the harvested crop and/or whether to use the cotton futures market as part of that marketing effort. In addition to these familiar uses of forecasting information, planners may use cotton price forecasts to determine the potential benefits of land reclamation projects whose principal crop will be cotton or to make decisions about other capital improvements which are specific to cotton production. Planners may also use cotton price forecasts to make the projections of foreign exchange revenues which underlie short- and medium-term economic plans.

Each of these possible uses require a different forecasting horizon as well as having different requirements for frequency of intermediate estimates and for revisions. If planting decisions are to be responsive to expected future international cotton market conditions, a forecast horizon of at least a year is required. If decisionmakers are able to implement these planting decisions with minimal lead time, the forecasting strategy which maximizes forecast accuracy is to make the forecast at the last possible moment at which planting decisions can be altered. The object of such a forecast could be the average spot price during the next crop year or, alternatively, the spot price on a particular day (or average of days for some period less than a year) on which the crop is to be sold. If decisionmakers find it more difficult to alter (tentative) planting decisions as the date of planting approaches, it may be optimal to make available on the first date at which planting decisions become inflexible an initial forecast of selling prices for the new crop and then periodically revise that forecast up to the last date at which planting decisions can be altered. If such forecast revisions are substantial, decisionmakers may find it worthwhile to revise ideal planting dates. Once the cotton crop is in the ground, periodic forecast revisions may still be useful in decisionmaking on the optimal use of such post-planting inputs as applications of fertilizers and pesticides. Decisions on the optimal timing of international sales require month-by-month forecasts for as long a period as marketing authorities are willing to consider. This may include month-by-month forecasts for the crop year with the first forecast made before harvest if authorities are willing to consider transactions in the futures market and which extends into the next crop year if authorities are willing to contemplate a carryover of cotton stocks from one year to another. These month-by-month forecasts should also be regularly revised. The forecast requirements of planners extend the forecast horizon to 2-5 years for medium term revenue forecasts. Such forecasts, which should be revised as often as planners are willing to review their plans, should include forecasts for each year out to the horizon limit. Forecasts which are used in the allocation of capital improvements specific to cotton production ideally require annual forecasts for each year out to horizons of 10, 20 or even 30 years. They need be revised only when new or revised investment plans are considered.

A Supply-Demand Approach to the Determination of Cotton Prices

A model of cotton prices will be developed which is based on supply and demand characteristics. The model will include in addition to current cotton production and current processor demand for cotton fibers the effects of expected future production and demand via the supply of storage. Cotton market transactions are often made with a view toward resale at a later date or, in what amounts to the same thing, stocks of cotton are withheld from the market in anticipation of higher future prices. As a result of these speculative transactions, today's price is in large part determined by what tomorrow's price is expected to be. To capture these speculative effects, a cotton pricing equation must reflect the influence of stocks of raw cotton on hand and the effects of future demand and production.

The following set of equations characterize the model of the cotton market.

(1) $Q_t + S_{t-1} = D_t + S_t$

(2) $Q_t = Q(E_{t-1}[P_t], E_{t-1}[A_t], C_{t-1})$

(3) $D_t = D(P_t, Y_t, B_t, F_t)$

(4) $E_t[P_{t+1}] = P_t\ (1 + r_t + s_t - c_t)$

(5) $E_t[Q_{t+1}] + S_t = E_t[D_{t+1}] + E_t[S_{t+1}]$

where Q_t = production available to the market during period t

D_t = processors' demand for cotton fibers during period t

C_{t-1} = cost of producing cotton fibers marketed during period t

S_t = the stock of inventories at the end of period t

P_t = price of cotton fibers during period t

A_t = price of other agricultural commodities during period t

r_t = cost of capital applicable to speculators in cotton fibers during period t

c_t = "convenience yield" associated with holding inventories of cotton fibers during t

B_t = the prices of other consumer goods during period t

F_t = the prices of substitutes for cotton fibers during period t

s_t = cost of storage during period t

$E_t[\quad]$ = expectation operator for expectations formed during period t.

Y_t = World income

Equation (1) is the equilibrium condition for the cotton fibers market during period t. Total supply available to the market consists of current production, Q_t, and the stock of inventories, S_{t-1}, carried over from the preceding period. Total demand during period t includes D_t, the demand by processors for fiber to be converted into other products, and S_t, the demand for stocks to be carried over to the next period. Equation (2) indicates that production decisions, made during the preceding period, depend on the expected price of cotton fibers at the time they come to market, the expected prices of other agricultural commodities which producers could offer to the market, and the costs of production. End use demand for cotton fibers (see equation (3)) depends on the price of cotton fibers, the income of consumers, the prices of other consumer goods, and the prices of substitutes for cotton fibers. Equation (4) states the condition which underlies the demand for the stock of inventories to be carried over into the next period. This condition, better known as the decision rule for the "supply of storage," suggests that a speculator will accumulate stocks up to the point where the expected future price for the next period is equal to the current price plus appropriate rewards for risk capital, storage facilities and a convenience yield which compensates for the costs of a stock-out.[1] Equation (5) is the expected market-clearing condition for period t+1 which can be thought of as the condition which determines $E_t[P_{t+1}]$ and thus drives the demand for S_t during the current period.

---

[1]See Michael J. Brennan, "The Supply of Storage", American Economic Review, March 1958, pp. 50-72, for a derivation of this condition and a statement that r, s, and c are best thought of as schedules which depend on other variables, including S_t).

The equation set (1)-(5) reflects an underdetermined system. Even if S_{t-1}, $E_{t-1}[A_t]$, C_{t-1}, Y_t, B_t, F_t, r_t, s_t and c_t are taken to be predetermined variables or as variables to be determined outside the cotton market, there remain eight variables-- Q_t, D_t, S_t, P_t, $E_t[P_{t+1}]$, $E_t[Q_{t+1}]$, $E_t[D_{t+1}]$ and $E_t[S_{t+1}]$-- and only five equations. The problem, of course, is that in a speculative market today's price depends not only on current production and final demand but upon all future expected production and final demands. In the model above, which is an example of a "rational expectations" approach to price determination in speculative markets, this is reflected by the inclusion of $E_t[Q_{t+1}]$ and $E_t[D_{t+1}]$, which represent production and final demand in the next period, and $E_t[S_{t+1}]$, which in turn is determined by expected production and final demand in all periods beyond t+1.

If this model of cotton pricing is to be applicable to the real world, a feasible solution to the problem of dealing with production and final demand over many periods must be found. Feasibility usually requires major simplifications. One approach is to ignore entirely the stockbuilding dimension of the problem. In its simplest form this approach restates (1) as $Q_t = D_t$, i.e., stocks are ignored, or, in effect, S_t is assumed to be equal to S_{t-1}. With this assumption, equations (4) and (5) become irrelevant and the system of equations is found to be determined. A second possible approach is to assume that future production, final demands, and end of period inventories are all "normal." In that case S_t becomes a "normal carryover," s_t and c_t are normal, and $E_t[P_{t+1}]$ is related to P_t by the cost of capital, which presumably depends on the expected inflation rate. If the expected inflation rate is assumed to be independent of the cotton market, the system is determinate, making a solution possible. A still more general approach is to assume normality of production, final demand and end-of-period inventories in all periods beyond t+1 while allowing for other patterns for Q_t, $E_t[Q_{t+1}]$. D_t, $E_t[D_{t+1}]$, and S_t. This allows for more general patterns between P_t and $E_t[P_{t+1}]$ while constraining $E_t[P_{t+1+i}]$, $i > 0$, to be consistent with the expected rate of inflation. Of course, other possible solutions to the problem of an underdetermined system exist. It is useful, however, to note that most of the solutions are likely to take the form of simplified assumptions about how price expectations for future periods are formed.

Quantifiable Versions of the Model of a Speculative Market

This section identifies alternative approaches to obtaining quantifiable versions of the model described above. As an aid to specifying quantifiable equations, an explicit forecasting horizon relevant to the planting decision is chosen. Alternative approaches to quantifying the forecasting model are presented and their advantages and disadvantages are considered. The results presented here are compared to forecasting methodologies offered by others.

Consider the forecasting requirements of planners who must make planting decisions. Since cotton planting begins in February, planners require a forecast by no later than early January if they are to have sufficient time to make planting decisions and to implement those decisions in time to get the new crop into the ground on schedule. Since the new crop will be available for marketing in November or December of the next year, forecasts used to make planting decisions require a horizon of at least a year. That is, a forecast for the December price for the following year might be made each December. While planners might prefer a forecast which becomes available before December, efforts to forecast at an earlier date are likely to be considerably less reliable, since accurate data on the recent harvest as well as data on carryover stocks from the preceding crop year are unlikely to be available much before December. If planting is the only decision likely to be influenced by price forecasts, only one forecast, made in December for the following December, is necessary.

This discussion reflects the assumption that there exists only one grade of cotton. Hence one would forecast "the" cotton price. In a world of many grades, there is a need for forecasts for at least one, and maybe more, specific grades. At this point, it seems wise to develop first a model for forecasting an average price for all grades or a price for the most frequently traded grade and then estimate a separate relationship for predicting the size of the premium or discount for particular grades relative to the average price or price of the base grade. Specification of the model discussed below will be based on these assumptions about horizon, absence of need for forecast revision, and need to forecast prices for particular grades of cotton.

Consider how one might use the model described by equations (1) - (5) in order to generate a forecast of $E_t[P_{t+1}]$. It will be assumed that the forecaster possesses data for S_{t-1}, Q_t, and P_t. The unknowns, in addition to $E_t[P_{t+1}]$ are D_t, S_t, $E_t[Q_{t+1}]$, $E_t[D_{t+1}]$, $E_t[S_{t+1}]$, etc. The first step to be taken is to assume that beyond some point Q, D, and S assume the kinds of normal patterns identified above. We shall assume that $E_t[S_{t+1}]$ $E_t[Q_{t+2}]$, $E_t[D_{t+2}]$, and all following values are normal and therefore quantifiable based on observable patterns in the past. As a result, there remain five unknowns, -- $E_t[P_{t+1}]$, D_t, S_t, $E_t[Q_{t+1}]$, and $E_t[D_{t+1}]$ -- and six equations -- (1), (3), (4), (5) as well as versions of (2) and (3) appropriate for Q_{t+1} and D_{t+1}. As we shall see somewhat later, this overdetermined system, which arises from (at least partial) knowledge of P_t, gives rise to alternative forecasting approaches.

One forecasting alternative is to specify empirical versions of behaviorial equations (2) and (3) in terms of current and past variables, estimate the coefficients of these equations, and then use observable values together with (1), (5), and the assumed normal value of $E_t[S_{t+1}]$ to solve sequentially for a forecast of $E_t[P_{t+1}]$. As an example of this one could estimate equation (3) for D_t and substitute into it forecasts of world income, world price levels, and the prices of fiber substitutes for the current year in order to generate a forecast for D_t. This forecast for D_t together with known values of Q_t and S_{t-1} will from equation (1) permit a forecast of S_t, the expected carryover into the next year. In order to forecast Q_{t+1}, one could estimate some version of equation (2) and, upon plugging in observed values of the right hand side variables, generate a forecast for Q_{t+1}. With S_t, Q_{t+1} and S_{t+1} thus qualified one can solve equation (5) for the price which makes D_{t+1} from a version of (3) compatible with the determined values of S_t, Q_{t+1} and S_{t+1}. It should be noted that this solution technique does not make use of equation (4). This occurs because the available December price is taken to be a proxy for P_t. While such an approach has the merit of using all the price information available, it has the disadvantage of making D_t and $E_t[Q_{t+1}]$ functions only of the December price when in fact those decisions surely reflect prices at other times during the year.

There are at least two ways in which the supply of storage relationship could be used in principle to improve the information built into the forecasting procedure. One way is to use a monthly or quarterly version of (4) to generate expected prices at other points during the year. These prices would allow calculation of an average price for the year, which is the more appropriate variable to enter into the demand equation. An additional use of (4) is to produce a forecast of P_{t+1} which does not utilize forecasts of Q_{t+1} and D_{t+1} nor the assumption about a "normal" S_{t+1}. The centerpiece of this approach is equation (4) together with the assumption that r_t, s_t, and c_t are functions of S_t. The first step is to forecast S_t as the residual from (1) after quantifying D_t with the help of (3). With a value of S_t available, one calculates values for r_t, s_t, and c_t as functions of S_t and substitutes these values into equation (4) to produce a forecast of P_{t+1}.

Each of the approaches described above required the estimation of a set of structural equations which are then solved recursively to forecast one variable at a time with the forecasted variable becoming an input into the process of forecasting the next dependent variable. Another approach is to solve the system for $E_t[P_{t+1}]$ in terms of observable variables, Q_t, P_t, S_{t-1} and forecasted exogenous variables Y_t, B_t, F_t, and A_t. One can estimate this reduced form equation. By substituting observed and forecasted values into the estimated reduced form equation, P_{t+1} can be forecasted without calculating values for the other endogenous variables.

It is instructive to compare the alternative approaches set out above to approaches offered by others. Two other studies will be summarized. One is the work presented in F.G. Adams and J.R. Behrman, Econometric Models of World Agricultural Commodity Markets: Cocoa, Coffee, Tea, Wool, Cotton, Sugar, Wheat, Rice (Ballinger Press, 1976). This volume presents the results of the efforts of a group of scholars working under the auspices of the Wharton Econometric Forecasting Unit and financed by UNCTAD and the Rockefeller Foundation. Forecasts based on these equations are offered periodically by Wharton Econometric Forecasting Associates. The other study to be summarized is that by E.C. Hwa, "Price Determination in Several International Primary Commodity Markets: A Structural Analysis," International Monetary Fund Staff Papers, March 1979, pp. 157-192 (Vol. 25, No. 1). At the time the study was done, Mr. Hwa was an economist in the Commodities Division of the Research Department of the International Monetary Fund.

The Adams-Behrman (AB) model of commodities markets consists of four equations. The first three equations are similar to equations (1), (2), and (3) above. The fourth equation is identified as a "price determination mechanism (which) takes into account the relationship between available inventory stocks and the level of demand."[2] The price determination equation is specified as

$$(6) \quad \frac{P_t}{B_t} = f\left(\frac{S_t}{D_t}, \frac{S_{t-1}}{D_{t-1}}, \ldots P_{t-1}\right).$$

---

[2]AB, p. 10.

AB indicate that "this relationship can be rewritten as a demand for inventories relationship which implies that inventories are proportional to the level of demand."[3]

Accepting AB's interpretation of (6) as a stock demand equation allows us to understand how AB solve the problem of today's price as dependent on expected future production and demand. They simply assume away the role of future production and demand. They simply assume away the role of future production and demand by hypothesizing that the stock of inventories does not depend upon $E_t[P_{t+1}]$. Thus $E_t[P_{t+1}]$ does not enter their model. While they recognize that "the costs of holding inventories are not well represented," they express the hope that "to the extent that the distributed lags in the model are consistent with the formation of expectations in relative price changes, some aspects of speculative inventory behavior may be captured."[4]

The discussion in AB does not indicate what forecasting approach they use in applying the model to P_{t+1} at a time when at least some information about P_t is already available from the market. It is important to note, however, that when they estimate a version of (6) which contains the lagged cotton price P_{t-1}, they discover that the coefficient on $\frac{S_t}{D_t}$ is statistically insignificant.[5] The implication of this for forecasting P_{t+1} when P_t is at least partially revealed by the market is that only information on D_t and S_t are needed to generate a forecast of P_{t+1}.

---

[3] AB, p. 11.

[4] AB, p. 11.

[5] AB, p. 38.

Hwa's primary interest is in testing partial adjustment pricing mechanisms and in the role of beginning-of-period inventories on price rather than in forecasting. Nevertheless, it is useful to consider how his model lends itself to forecasting.

Hwa's model is virtually identical to equations (1) - (4) except that he states (4) as a desired demand for inventories which may not be equal to the actual end-of-period stock of inventories. Thus Hwa allows for the possibility of market disequilibrium. This leads him to an equation for P_t which, in addition to the variables in equations (1) - (4), includes P_{t-1} as a reflection of the partial adjustment mechanism. Hwa hypotheses that expectations about future prices are given by

$$(7)\ E_t[P_{t+1}] = \alpha_0 + \alpha_1 B_t + \alpha_2 Y_t + \alpha_3 P_{t-1}.$$

(While he claims that this is a "rational" hypothesis about the formation of price expectations, he does not demonstrate that such an expected price flows from his pricing model, which it surely does not.) After substituting (7) into his model, he derives a partial adjustment theory of current price which takes the form,

$$(8)\ P_t = b_0 + (b_1 + b_4)D_t - b_4(S_{t-1} + Q_t) + b_2 B_t - b_3\ \text{TREND} + b_5 P_{t-1}.$$

Consider how Hwa's approach could be used for forecasting. One possibility is simply to accept his rational expectations hypothesis at face value, estimate an empirical version of (7), and use it to forecast. (Such an equation is not unlike the reduced form equation discussed above.) The more interesting alternative, however, is to assume that (8), after estimation of the b_i's, is also appropriate for forecasting $E_t[P_{t+1}]$ when $E_t[D_{t+1}]$, $E_t[S_t + Q_{t+1}]$, $E_t[B_{t+1}]$ and P_t are used as right hand side variables. Values for $E_t[D_{t+1}]$, $E[S_t]$ and $E[Q_{t+1}]$ could be estimated using appropriate versions of equations (1), (2), (3), and (5).

Forecasting Prices for Premium Staple Lengths

The models discussed above are appropriate for forecasting the price of a high volume grade such as medium staple or perhaps an average of prices for different qualities. The models are not appropriate, however, for forecasting prices for premium staple lengths. This section contains a discussion of alternative methodologies for forecasting prices for premium staple lengths. These methodologies will use forecasted prices for some base staple length or average price to produce a forecast for premium staple lengths. Ideally, forecasts for both long staple and extra long staple cotton prices should be available so that planners may make appropriate planting decisions where the alternatives include extra long staple cotton, long staple cotton, and other crops.

The appropriate approach to price forecasting for premium staple lengths depends on whether there exist very close substitutes for any particular staple length. If substantial numbers of users of extra long staple cotton find it possible because of the emergence of synthetic fibers and advances in spinning technologies to substitute other fibers for extra long staple cotton, the price of premium grades purchased by these users must depend entirely on the prices of the substitutes. In this case producers of premium grades are price takers. They have no choice but to sell at or below the price determined by the market for substitutes, and the forecast they require is a forecast of that market-determined premium. If, on the other hand, the number of users of premium staple lengths who find it expensive to substitute for particular qualities bulk large in the total market, a large producer may have some market power. In this case, the producer has the choice of selling large amounts at a relatively small premium to users who have the option of substituting other fibers or of keeping prices high and supplying that portion of fiber users who are unwilling even at high prices to substitute other fibers for premium staple lengths. Such a producer, who is a potential price maker, requires a forecast of the entire demand schedule (i.e., a forecast of demand at different premium levels) in order to determine the most profitable production level.

For the case in which the producer is a price taker, an appropriate approach to forecasting a price premium is to estimate a hedonic price index similar to that developed by Petzel and Monke in "Integration of the International Market for Cotton" Working Paper No. 4 for the UCD-ARE-AID Project, The Interaction of Demand, Supply and Government Policy: A Case Study of Egyptian Cotton. Quoting from page 5 of that paper,

> "Hedonic analysis involves the association of a set of quality characteristics (entered as integer-valued dummy variables) which correspond to each observed price. The estimated equation is of the form
>
> (9) $P = \beta_0 + \beta_1 D_1 + \beta_2 D_2 + \beta_3 D_3 + \ldots.$
>
> where
>
> P = national logarithm of average price
>
> $D_1, D_2, D_3 \ldots$ = different quality characteristics.

Equation (9) provides the basis for the calculation of the premia and discounts awarded to particular quality characteristics."

Price forecasting for a potential price maker requires a forecast of the quantity demanded at different price levels. Such a demand schedule might take the form:

(10) $D_{ELS} = \gamma_0 + \gamma_1 P_{ELS} + \gamma_2 P_{MS} + \gamma_3 P_S + \gamma_4 Y$

where

D_{ELS} = Processor's demand for extra long staple cotton fibers during the period

P_{ELS} = Price of extra long staple cotton

P_{MS} = Price of medium staple cotton (or alternatively the average price).

P_S = Price of synthetic fibers.

One would need first to estimate such an equation using existing data and then forecast a $D_{ELS} - P_{ELS}$ schedule, given forecasts for P_{MS}, P_S, and Y. This schedule, together with forecasts for quantities of ELS fibers to be offered by other producers would serve as an input into optimal planting decisions.