An Automated Econometric Approach for Estimating and Reporting Daily Cotton Market Prices

Jeff E. Brown, Don E. Ethridge, Darren Hudson and Carlos Engels[*]

Abstract

An automated price reporting system developed through computer technology and hedonic price theory is used to estimate daily cotton market prices, premiums, and discounts. This objective approach for reporting cotton market prices was developed to complement the objective measures of high volume instrument grading of cotton. The computerized, econometric system is limited to the Texas and Oklahoma marketing regions where sales are readily available from electronic markets. The econometric based system has shown all the characteristics of an efficient price reporting system; it is accurate, reliable, consistent, and repeatable in its working process and price estimates.

Key words: computer automation, cotton marketing, econometrics, hedonic price analysis, price reporting

Introduction

Accurate and timely price information is a key to the existence and function of an efficient competitive market. In a commodity market, price information facilitates production, selling, and purchasing decisions for all market participants. Consequently, price reporting systems have been established to disseminate price information for many commodities, including prices of various qualities of some commodities. Price reporting systems are expected to provide accurate and reliable information to facilitate the trading of a commodity and aid in the price discovery process. To accomplish this, reported prices must be perceived as objective and accurate. This allows market participants to make decisions based on reliable information and to be equally informed about market conditions.

For price reporting on quality differentials to meet these goals, the initial measures of the quality of the product must be accurate (reliable) and timely. In cotton marketing, the industry recognizes the quality characteristics embodied within each bale of cotton as constituting reliable product information on which the market can function. The United States Department of Agriculture (USDA) administers and maintains a grading system to measure the quality characteristics of cotton fibers. The system of grading cotton by its quality attributes has evolved from a few subjective measures to more objective measures of a larger number of attributes. This evolution has

accelerated since the inception of high volume instrument (HVI) grading and its adoption industry-wide in the U.S. in 1991 (Ethridge, Engels, and Brown). The HVI grading technology has allowed cotton grading to become more rapid, accurate, consistent, and objective in the measuring of cotton quality attributes. For efficient market performance, price information concerning the various quality attributes must also be known.

The official cotton price report in the U.S. for the cash (spot) market is administered by the USDA's Agricultural Marketing Service (AMS), Cotton Division, in the form of the Daily Spot Cotton Quotations (DSCQ). The DSCQ provide estimates of prices and quality premiums and discounts for each of the seven designated markets in the U.S. for each trading day (USDA, daily issues). AMS reports a base price for each region for the "base" combination of quality attributes (the price for grade 41, staple 34, micronaire 3.5-4.9, strength 24 and 25). Premiums and discounts are determined for all other combinations of quality attributes to be added to or subtracted from the base price. The quality attributes in cotton are grade (a combination of leaf, bark, grass, and other foreign matter content and color), staple (fiber length), fiber strength, micronaire (a measure of fiber fineness and maturity), and length uniformity of the cotton fibers. The various components of "grade" were separated into color grade, leaf grade, and other foreign matter in the quality grading system in 1993. This separation added two more quality dimensions to price reporting. Both grading systems are covered in this paper.

The increased number of quality attributes recognized by the industry has made it increasingly difficult for the price reporting system used by AMS. The AMS market reporters collect samples (i.e., recaps of mixed-lot sales) of market transactions and conduct interviews with market participants (primarily merchants and marketing organizations) to obtain market information. The market reporter then arrives at an intuitive (subjective) determination of the market activity and the DSCQ are formulated (Hudson et al.). The grading system for cotton has become more precise with the use of HVI technology to measure nine different aspects of the fiber. Levels of these nine attributes make up more than 25,000 quality combinations that have different implications to

textile manufacturers and thus have different values in the market. The increased number of HVI quality attributes has made separating the effects of the quality attributes on price more difficult and increased the risk of error in the interpretation of the market information.

A new objective approach for estimating and reporting cotton prices has been developed and tested for the Texas and Oklahoma markets (West Texas and East Texas-Oklahoma). This approach, referred to as the Daily Price Estimation System (DPES), has been in development since 1988 to complement the more accurate and objective HVI grading system. The objectives of the DPES are to provide an objective and accurate tool to support reporting prices, premiums, and discounts. More specifically, the DPES is a computer-automated system for receiving and statistically analyzing sales data, estimating prices using an econometric model developed from hedonic price research, and transmitting the results to market participants. The objectives of this paper are to describe the development and current operation of the DPES and provide evidence for its accuracy, application, and inclusion in the daily reporting of cotton market prices.

Hedonic Price Analysis: Concepts and Requirements

The premise of hedonic price analysis is that a commodity is valued by the amount of utility-bearing characteristics associated with that commodity (Rosen). Therefore, a product is a collection of characteristics or attributes that have value. The explicit price of the product is observable in the market, and is composed of the aggregated values of these characteristics. The characteristics' values are termed (marginal) implicit prices. The implicit prices of the characteristics of products are not readily observable in the market because the attributes are not bought and sold independently; they are embodied in the product. However, to the extent that the characteristics are observable and/or measurable, they can be derived through a process of disaggregating the prices (see Ethridge and Morse for an overview of studies on these procedures).

These implicit prices may be determined by regressing the product's price against its attributes

and then differentiating the resulting equation with respect to each attribute. These procedures have been useful in determining the prices of attributes in cotton (Ethridge and Davis; Ethridge, Shaw, and Ross; Robinson et al.; Ethridge, Sudderth, and Moore; Ethridge and Mathews; Hembree, Ethridge, and Neeper; Ethridge and Neeper; Bowman and Ethridge; Ethridge, Engels, and Brown; Chen, Chiou, and Anderson). Statistically derived implicit prices indicate the average implicit price of each attribute at each attribute level, *ceterus paribus*, thus determining the average price of each attribute as observed in the specified market.

Hedonic Model For Daily Cotton Market Price Reporting

From 1988 through 1992, 30% of the total U.S. cotton production was produced in Texas and Oklahoma (USDA, 1993). The large number of market participants in the region meets the condition for a competitive market structure. The availability of marketed cotton provides the daily sales observations of lot prices and average lot values for each of the quality attributes for a daily hedonic pricing model for cotton. The data are obtained from two electronic spot marketing systems operating in the Texas and Oklahoma regions. The two electronic marketing systems are TELCOT, which merchandizes cotton throughout Texas and Oklahoma (Ethridge, 1978), and The Network, which operates primarily in the Texas Rolling Plains. The observed sales of cotton constitutes over 25% of all Texas and Oklahoma cotton and about 33% of all cotton in the spot market. The sample also represents purchases (buying behavior) of approximately 75% of the total merchants in the two regions. The daily sets of data contain the observed sale price of each mixed lot, the average quality attributes of classer's trash and leaf grade, reflectance, yellowness, and color grade, staple (fiber length), strength, micronaire, length uniformity, bark content, other extraneous matter content, and other relevant information about each sale, such as the location (region) and lot size.

Prior research on hedonic cotton models provides some evidence on the mathematical model structure used by the DPES. Such past works dealt with both aggregated data and individual lot sales (less aggregated) data. A linear functional form

performed well in cases where aggregated data were used (Hembree, Ethridge, and Neeper; Chen, Chiou, and Anderson; Bowman and Ethridge). As data become less aggregated--e.g., more toward individual sales and less toward regional or national averages--the non-linear functional forms have performed better (Ethridge and Mathews; Ethridge and Neeper).

Initial investigation into the development of the system for daily cotton market price reporting recognized that the most important ingredient would be the mathematical model structure. The initial research concentrated on identifying and evaluating possible mathematical model structures. The selection was narrowed, based on statistical evaluation, from 12 models down to 2 specific mathematical model structures (Ethridge, Engels, and Brown).

The two mathematical model structures were a "linear difference" model (Bowman and Ethridge) and a multiplicative or "double-log" model (Ethridge, Engels, and Brown). The "linear difference" model was determined to be inferior to the double-log model on the basis of conceptual and empirical criteria. As well, Ethridge, Engels, and Brown studied various weighting procedures for grouping days of data and for evaluating sales lot size impact on price. Both of these procedures failed to improve estimates based on estimation accuracy as the criteria for model selection. Therefore, an unweighted single-day sales model was concluded to produce the most accurate estimates.

An extensive search for the appropriate functional form using a single day of unweighted sales was performed (Brown and Ethridge). That study determined the "best" mathematical model structure for the pre-1993 grading system and the "expected best" for the 1993 grading system, which separated the grade code into its component parts of color, leaf content, and bark content. Each model used in estimating prices for the two grading systems was developed with consideration for both conceptual and empirical factors. The expected marginal productivity of each attribute in the production of yarn was one such conceptual consideration. Empirical considerations included partial regression analysis (Neter, Wasserman, and

Kutner) to determine the general relationships between price and each quality attribute and an extensive analysis of residual error terms from various statistically estimated functional forms. Both empirical considerations allowed the "best" mathematical model specifications to be determined. The "best" functional form daily model for the pre-1993 grading system was:

$$P_i = \beta_0 \ e^{\beta_1 G1_i^2 + \beta_2 BG_i (G1_i) + \beta_3 G2_i + \beta_4 G2_i^2 + \beta_5 STA_i + \beta_6 STA_i^2 + \beta_7 STR_i + \beta_8 M_i + \beta_9 M_i^2 + \beta_{10} R_i} \varepsilon_i$$

(1)

where:

P_i =   price of mixed lot i in cents per pound;

$G1_i$ =   average of the first digit of the grade code in mixed lot i;

BG_i =   binary indicator variable denoting below grade cotton ($BG = 0$ if $G1 \leq 7$ and $BG = 1$ if $G1 > 7$);

$G2_i$ =   average of the second digit of the grade code in mixed lot i;

STA_i =   average staple length of cotton in mixed lot i in 32^{nds} of an inch;

STR_i =   average strength reading of cotton in lot i in grams per tex;

M_i =   average micronaire index reading of mixed lot i;

R_i =   binary indicator variable denoting region ($R = 0$ if market reporting region is West Texas and $R = 1$ if market reporting region is East Texas or Oklahoma);

ε_i =   stochastic error term.

Table 1. Estimated Parameters of Model 1 from a Random Sample of 5 Days

Inde-pendent Variables	Coefficients[a]				
	Dates				
	03NOV92	01DEC92	23DEC92	14JAN93	21JAN93
Int.	-0.2904	0.1138	-3.4552	-0.7765	1.0485
	(-0.26)	(0.11)	(-1.75)	(-0.93)	(0.56)
$G1^2$	-0.0077	-0.0028	-0.0044	-0.0042	-0.0047
	(-16.70)	(-9.27)	(-7.90)	(-14.04)	(-10.31)
$BG(G1)$	0.0000	0.0000	-0.0293	-0.0193	-0.0152
	(0.00)	(0.00)	(-6.67)	(-8.09)	(-4.49)
$G2$	-0.0346	-0.0096	0.1543	-0.0442	-0.1123
	(-1.57)	(-0.99)	(3.75)	(-4.14)	(-5.28)
$G2^2$	-0.0056	-0.0063	-0.0542	-0.0021	0.0137
	(-1.12)	(-0.27)	(-5.27)	(-0.96)	(3.06)
STA	0.174	0.1305	0.3234	0.1646	0.0494
	(2.62)	(2.18)	(2.64)	(3.20)	(0.43)
STA^2	-0.0024	-0.0019	-0.0048	-0.0024	-0.0006
	(-2.45)	(-2.12)	(-2.58)	(-3.06)	(-0.38)
STR	0.0040	0.0024	0.0040	0.0020	0.0028
	(1.50)	(2.42)	(1.49)	(1.63)	(1.56)
M	0.5079	0.7594	0.8337	0.9640	1.0720
	(8.15)	(18.99)	(11.30)	(25.14)	(16.58)
M^2	-0.0613	-0.0938	-0.1006	-0.1187	-0.1350
	(-7.90)	(-17.93)	(-10.32)	(-22.54)	(-14.81)
R	0.0016	-0.0009	0.0051	-0.0074	-0.0180
	(0.20)	(-0.23)	(0.46)	(-1.34)	(-2.13)
CT	0.0000	0.0397	0.0000	0.0000	-0.0370
	(0.00)	(1.25)	(0.00)	(0.00)	(-2.37)
R^2	0.9275	0.7691	0.8894	0.8965	0.8764
F-Value	278.62	122.61	71.58	335.27	188.18
d.f.[b]	196	368	89	387	292
N.O.B.[c]	208	380	101	399	304

[a]Numbers below coefficients are t-values.

[b]degrees of freedom.

[c]number of observations.

With the implementation of the 1993 grading system, the mathematical model structure was further evaluated and revised to accommodate the form of the quality information available to market participants. The functional form daily model to accommodate the grading system which began in 1993[1] is:

$$P_i = \beta_0 \; e^{\beta_1 CTR_i^2 + \beta_2 C1_i^2 + \beta_3 C2_i^2 + \beta_4 STA_i + \beta_5 STA_i^2 + \beta_6 STR_i + \beta_7 M_i + \beta_8 M_i^2 + \beta_9 R_i + \beta_{10} LB_i + \beta_{11} HB_i + \beta_{12} LO_i + \beta_{13} HO_i} \varepsilon_i \qquad (2)$$

Table 2. Estimated Parameters of Model 2 from a Random Sample of 5 Days

Inde-pendent Variables	Coefficients[a]				
	Dates				
	18OCT93	03NOV93	19NOV93	07DEC93	29DEC93
Int	-2.5654	0.7324	1.2140	1.0795	3.0625
	(-3.36)	(1.42)	(4.39)	(6.00)	(9.67)
CTR^2	-0.0008	-0.0017	-0.0025	-0.0013	-0.0006
	(-1.18)	(-4.90)	(-19.51)	(-16.70)	(-4.33)
$C1^2$	-0.0002	-0.0004	-0.0002	-0.0021	-0.0017
	(-0.13)	(-0.47)	(-0.58)	(-10.89)	(-5.48)
$C2^2$	-0.0095	-0.0055	-0.0065	-0.0065	-0.0057
	(-4.42)	(-3.99)	(-11.79)	(-22.17)	(-14.07)
STA	0.2973	0.1230	0.1033	0.1176	0.0269
	(6.45)	(4.03)	(6.16)	(10.86)	(1.40)
STA^2	-0.0043	-0.0016	-0.0014	-0.0016	-0.0003
	(-6.15)	(-3.58)	(-5.35)	(-9.76)	(-1.10)
STR	0.0000	0.0019	0.0021	0.0016	0.0010
	(0.00)	(2.11)	(5.25)	(6.92)	(3.00)
M	0.6815	0.4203	0.3738	0.3683	0.2179
	(10.53)	(8.46)	(17.94)	(25.70)	(14.76)
M^2	-0.0835	-0.0516	-0.0451	-0.0446	-0.0264
	(-10.13)	(-8.48)	(-17.91)	(-25.86)	(-14.57)
R	0.0119	0.0006	-0.0030	0.0023	0.0097
	(1.62)	(0.14)	(-1.69)	(2.23)	(4.74)
LB	-0.0234	-0.0366	-0.0186	-0.0206	-0.0089
	(-2.93)	(-6.73)	(-7.36)	(-14.43)	(-4.68)
HB	-0.8873	0.0000	0.0000	0.0000	0.0000
	(-0.54)	(0.00)	(0.00)	(0.00)	(0.00)
LO	-0.0620	-0.0160	-0.0788	-0.0911	-0.0483
	(-1.61)	(-0.94)	(-7.36)	(-12.22)	(-2.62)
HO	0.0000	0.0000	-0.1238	-0.0781	-0.0399
	(0.00)	(0.00)	(-6.73)	(-7.41)	(-3.73)
R^2	0.8790	0.8341	0.7837	0.8543	0.7624
F-Value	55.45	74.02	285.31	663.87	138.27
$d.f.$[b]	84	162	945	1359	517
$N.O.B.$[c]	98	176	959	1373	531

[a]Numbers below coefficients are t-values.

[b]degrees of freedom.

[c]number of observations.

where:

CTR_i	=	average of the classer's trash designation in mixed lot i;
$C1_i$	=	average of the first digit of the color grade in mixed lot i;
$C2_i$	=	average of the second digit of the color grade in mixed lot i;
LB_i	=	percentage of level 1 bark in mixed lot i;
HB_i	=	percentage of level 2 bark in mixed lot i;
LO_i	=	percentage of level 1 other extraneous matter content in mixed lot i;
HO_i	=	percentage of level 2 other extraneous matter content in mixed lot i;

and the other attributes are the same as defined in model 1. A sample of daily model results (i.e., coefficients, t-values, R^2's, F-values, degrees of freedom, and number of observations) for both equations 1 and 2 are shown in tables 1 and 2. The variability in the magnitudes of the coefficients can generally be explained by the supply and demand of the specific levels of each attribute over the marketing period.

These models both have *a priori* conceptual appeal. The price-quality relationships develop premium and discount patterns that are curvilinear rather than linear, and the coefficient magnitudes suggest decreasing marginal productivity of the desirable attributes. Each functional form was selected on the basis of no patterns in the error terms with respect to each attribute and they had the greatest accuracy in their estimates. In addition, they can be easily estimated using the ordinary least squares regression technique and provide prices, premiums, and discounts for the DPES.

## Daily Price Estimation System Structure and Operation

The Daily Price Estimation System is made possible by the existence of computer technology. The centerpiece of the system is the econometric model (equation 1 or 2). However, computer hardware and software, including communication capabilities, give the system the capacity to function as a daily price information system. A schematic diagram of the DPES operations is shown in figure 1.

The DPES procedures are initiated by a data contributor at the end of each trading day. The data contributor uploads the sales data via telephone communication to the DPES computer, then the DPES computer dials and downloads the data from the other data contributor. The DPES system analyzes the lot sales data received from the contributors with the aid of a spreadsheet. The initial analysis of the data determines whether the values for each variable are realistic by determining if maximum and minimum values lie within an expected range (e.g., staple length between 28 and $38/32^{nds}$ of an inch). If the reported data are outside the specified range, the system ceases operation and produces an error message to all parties.

Data are up-loaded to a mainframe computer where Statistical Analysis System (SAS Institute, Inc.) is used to estimate the ordinary least squares regression coefficients for equation 1 or 2 and derive test statistics. The means, standard deviations, maximums, and minimums of that day's lot sales data are also obtained. The program also checks for conceptually reasonable model coefficients. If a coefficient has a wrong sign, the respective variable/attribute is excluded from the model and the model is re-estimated. This procedure of excluding variables from the model has occurred on only 1.74% of the days since initiating operation in February, 1989. To date, there has been no instance in which a wrong sign coefficient has been statistically significant. The results from the SAS runs are then down-loaded to the DPES computer where they are again evaluated (i.e., coefficients for wrong signs and maximums and minimums for relevant ranges) and permanently stored.

The end products of the DPES are a Daily Spot Cotton Price Estimates table for (a) the West Texas region (table 3) and (b) the East Texas and Oklahoma region. This table provides a base price (price for base quality--color grade 41, staple 34, micronaire 3.5-4.9, leaf grade 4, strength 24-25, and no bark or other extraneous matter). The base price in table 3 is 53.22 ¢/lb. Premiums and discounts

Figure 1. A Schematic Diagram of the Daily Price Estimation System
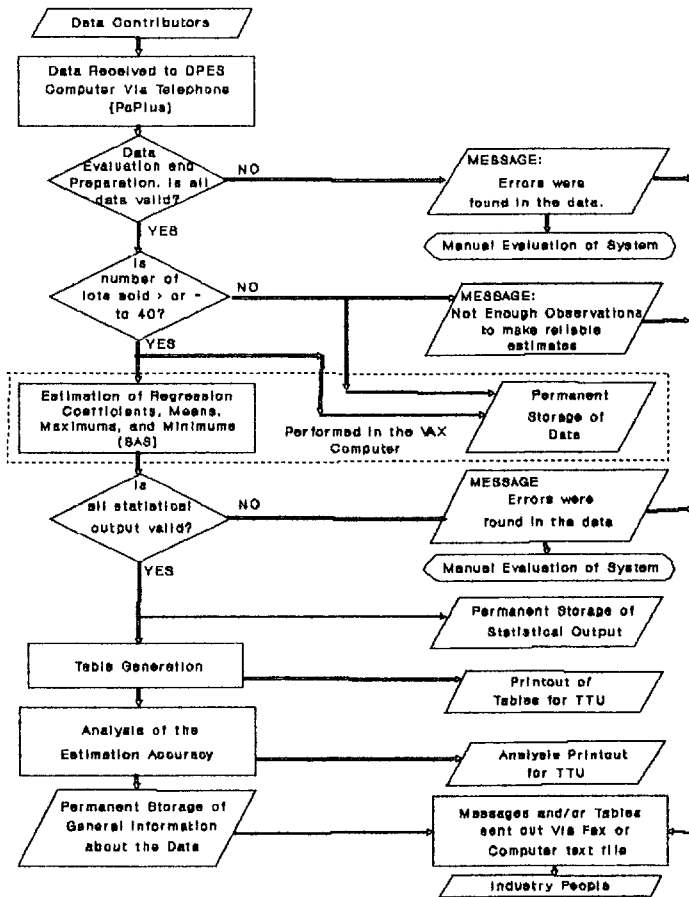
are shown for all other quality combinations in points/lb. (100 points = 1¢). For example, the discount for color grade 61, staple 33, all other attributes at base level, is 276 points (2.76 ¢/lb.), making the price for that quality 50.46 ¢/lb. The DPES, via a spreadsheet, uses the model and its estimated coefficients, inserts the various quality combinations into the equation, and calculates the prices and the various premiums and discounts in the tables. The system also measures the accuracy of the estimated prices by running regressions of the model's estimated prices against the actual prices. This provides a daily gauge of the performance of the model's estimates. Finally, the DPES transmits, via FAX or computer text file, a copy of both regions' tables to the National Cotton Council's COTNET (computer bulletin board) and several other outlets. The tables could be sent to other destinations. The entire working process of the DPES is usually performed during the night and is completed in a few hours.

Reported Price Comparison Analysis

Prior research has studied the relationship between the DPES and DSCQ reported prices. Hudson et al. found that the DSCQ reported prices were an average 5% (1.54¢/lb.) higher than the DPES for the period 1989-1992. That study also indicated that differences existed in structural patterns of the premiums/discounts across attribute levels between the DPES and DSCQ reported prices. Ethridge, Engels, and Brown found errors in prices reported by both the DPES and DSCQ against actual prices for the period of March 1989 through April 1990. DSCQ estimates systematically overestimated high priced cotton and underestimated low priced cotton. The DPES model used during that period also had a smaller, but significant, systematic error in the opposite direction. Subsequent analysis of the DPES model structure (Brown and Ethridge) resulted in changes that eliminated systematic error in the DPES estimates.

Table 3. Sample Table from the Daily Price Estimation System, West Texas Market, 1993 Grading System[a]

```
Daily Spot Cotton Price Estimates
Estimates by the Department of Agricultural Economics, Texas Tech Univ.
Date:   07DEC93;  Region:  West Texas                    # of Bales: 51,360
Color Grade and Staple Premiums and Discounts in Points/lb. # of Sales: 916
```

11	21	31	41	51	Staple 61	71	12	22	32	42
-418	-449	-500	-571	-660	28 -767	-891	-512	-542	-593	-662
-284	-316	-369	-441	-533	29 -643	-770	-381	-412	-464	-535
-163	-196	-250	-324	-419	30 -531	-661	-263	-295	-347	-421
-57	-90	-145	-221	-317	31 -432	-565	-158	-191	-244	-319
35	2	-54	-132	-230	32 -347	-482	-68	-101	-156	-232
112	77	21	-58	-157	33 -276	-413	7	-27	-82	-159
171	137	79	53.22	-100	34 -220	-359	66	32	-25	-102
214	179	121	53	-60	35 -181	-320	108	73	17	-62
240	205	146	66	-35	36 -157	-297	133	98	41	-38
247	212	154	74	-28	37 -150	-290	140	106	49	-30
247	212	154	74	-28	38 -150	250	140	106	49	-30

52	62	23	33	43	Staple 53	63	34	44	54
-750	-855	-695	-743	-811	28 -896	-1000	-946	-1011	-1092
-625	-733	-569	-619	-688	29 -775	-880	-827	-893	-977
-513	-623	-455	-506	-577	30 -666	-773	-719	-787	-873
-413	-526	-354	-406	-478	31 -570	-679	-624	-693	-780
-328	-442	-267	-320	-394	32 -487	-598	-542	-612	-701
-257	-373	-195	-249	-324	33 -418	-531	-474	-545	-635
-201	-318	-139	-193	-269	34 -364	-478	-421	-493	-584
-161	-279	-99	-153	-229	35 -325	-440	-383	-455	-547
-137	-256	-75	-130	-206	36 -302	-418	-360	-433	-525
-130	-249	-67	-122	-199	37 -295	-411	-353	-426	-518
-130	-249	-67	-122	-199	38 -295	-411	-353	-426	-518

Mike Differences Points/lb.   Trash Differences-Points/lb.   Bark Differences Points/lb.   Strength Differences Points/lb.

Mike Range	Discount	Trash Code	White	Light Spot	Spot	Tinged	Bark Code	Disc.	Level	Prem/ Disc
26 & below	-524	1	104	102	99	95	Level 1	-108	<18	-54
27-29	-400	2	83	82	79	76	Level 2	--	19	-45
30-32	-243	3	48	47	46	44			20	-37
33-34	-140	4	-0	0	-0	-0			21	-29
35-49	-0	5	-62	-60	-58	-56	Other		22	-21
50-52	-220	6	-136	-133	-129	-123	Differences		23	-12
53 & above	-317	7	-222	-218	-211	-202	Points/lb.		24&25	-0
									26	12
							Other		27	21
							Code Disc.		28	29
									29	37
							Level 1	-463	30	46
							Level 2	--	31+	54

Results are not for general distribution.

[a]This table represents a sample table produced by the DPES on a daily basis. It is intended to provide an example of the array of prices a producer has to make decisions from as well as provide a general picture of the information provided by the DPES.

Analysis of tests for accuracy of both DSCQ and DPES price estimates for the 1993/94 marketing year are presented here. Accuracy of the DPES estimates, both the tables and the formula, and the DSCQ estimates were examined. These comparisons were made against actual producer lot sales transactions. It is important to note that the statistical reliability of the DSCQ cannot be directly determined. That is, since the characteristics of the sample (e.g., size and representativeness) are not known, direct statistical measurements of the accuracy of the DSCQ in reporting prices, premiums, and discounts are not possible (Hudson et al.).

To perform this analysis, a random sample of 350 observations from 34,285 mixed lot sales during the 1993/94 cotton marketing year (August 1, 1993 to July 31, 1994) was drawn.[2] The sample was verified to accurately represent sales transactions during the marketing year by comparing the mean price and quality values of the sample

against the 34,285 mixed lot sales means. Mean difference tests indicated no significant differences between the sample and the population.

The DPES was evaluated in two ways. The daily model estimates based on the regression equations were calculated for each sale in the sample and DPES table prices were also used to calculate the prices for each sale. Since the tables identify only discrete quality categories, lot average price estimates were calculated as extrapolations between the discrete prices. These calculated prices are often not exactly the same as the values determined directly from the formula. The DSCQ prices were determined in a manner similar to that performed in determining the DPES table prices. Since some qualities in the random sample are excluded from DSCQ reporting, the sample for comparison was reduced from 350 to 326 mixed lot sales. The fewer number of observations in the sample had no significant effect on its representativeness of the population.

Two analyses were performed on the accuracy of each set of estimated prices in tracking actual prices: (1) Theil's inequality measure was calculated for both the DPES and the DSCQ against actual prices and (2) the price estimates from each were regressed against the actual prices.

Analysis using Theil's inequality coefficient (U) (Pindyk and Rubinfeld) revealed some differences between the DSCQ and DPES when compared to the actual prices. The DSCQ had a U = 0.0434 while the U values for the DPES were 0.0052 and 0.0018 for the table and formula estimates, respectively. None of the coefficients were close to one. Nevertheless, both of the DPES estimates appear to have performed better than the DSCQ. The primary disadvantage of Theil's inequality procedure is the inability to test for statistical differences.

Regression analysis was also used to examine the relationships between actual prices and DSCQ and DPES prices. A simple linear model of prices was:

$$P_{A,j} = \alpha + \beta P_{E,j} + \varepsilon, \qquad (3)$$

where $P_{A,j}$ is the actual price of mixed lot j, $P_{E,j}$ is either the DSCQ, DPES table, or DPES formula estimated price of lot j, α and β are parameters, and ε is the normally distributed, mean zero error term. The standard null hypothesis is that $\alpha = 0$ and $\beta = 1$. Any significant deviations from this result provides evidence of systematic error in the price estimates (Holden and Peel; Belongia).

Since these data are not a true time series (each day's estimate is from cross-section data and the random sample was drawn from the large set which varied over time) time series properties are not applicable.[3] This regression determines whether the estimated prices of a mixed lot (DPES or DSCQ) are equal to the actual prices of those lots, on average. Given the large sample size, regression parameters are expected to be unbiased (Elam and Dixon).

Results indicate the neither the DPES formula nor the DPES table prices showed a systematic bias with respect to actual mixed lot prices (table 4). However, the DSCQ showed a systematic bias ($\alpha > 0$ and $\beta < 1$; significant F). This indicates that the DPES estimated prices were closer to the actual than the DSCQ.

The F-values shown in table 4 were calculated following Harrison. The standard null hypothesis of $\alpha = 0$ and $\beta = 1$ from Equation 3 is tested jointly. That is, the F-test shows biasedness in the model without directly testing the individual parameters.   The limitation here is that no knowledge is obtained of the nature of the bias. Therefore, both the t and F-tests are presented.

Taken together, the results of the Theil's inequality and regression analysis support the conclusion that the DPES performs better at estimating market prices than the DSCQ. Both analyses indicated that the DPES prices were closer to the actual prices than the DSCQ. Additionally, the regression analysis indicated that the DPES estimated prices had no systematic bias.

Table 4. Regression Parameters of Estimated Lot Prices vs. Actual Sales Prices in Texas-Oklahoma Markets, Based on a Sample of 326 Mixed Lot Observations[a]

Source of Estimated Prices	Intercept	Slope	R^2	F-Value
DSCQ Tables	8.7601	0.7976	0.8430	1919.68
	(8.26)[c]	(10.58)[d]		
DPES Tables	0.3958	0.9873	0.9460	5.94[b]
	(0.57)	(0.97)		
DPES Models	0.8851	0.9815	0.9442	6.49[b]
	(1.26)	(1.40)		

[a] Numbers in parentheses are t-values.
[b] The null hypothesis of $\alpha = 0$ and $\beta = 1$ cannot be rejected at the $\alpha = .001$ level.
[c] Significantly different from 0 at the $\alpha = .01$ level.
[d] Significantly different from 1 at the $\alpha = .01$ level.

Implications for Price Reporting

The use of econometric procedures and computer technology such as those contained in the DPES has the potential to affect price reporting. The greatest potential impact may be in a procedural context. These procedures make the DPES objective; however, data requirements may limit the proportion of time that estimates are forthcoming when compared to a subjectively determined price report. This section addresses some of the impacts and limitations of the procedures used by DPES.

The DPES approach minimizes human error and/or subjective judgement in the process of analyzing prices to support price reporting. The system is automated, thus reducing the possibility of human error with the exception of the initial entry of price and quality data on the sales transactions. However, these types of errors are minimized when the computerized electronic sales systems are providing data. Another possibility of human error is the development of the daily model structure that is used by the system for the estimation of the prices. However, with the prior investigation into the appropriate functional form of the daily model (Brown and Ethridge), plus the diagnostics of accuracy on the daily estimates, human error of this type should be minimal.

The DPES price estimates cannot be accepted as truth. This was particularly evident in

the 1991/92 crop year in which a record number of "below grade" (very high trash content) cotton bales were produced in the two regions. The large proportion of below grade cotton, which sold for a minimal price, caused the magnitude of the trash premiums and discounts to deviate out of any reasonable proportion. The below grade cotton caused an increase in the slope of the premiums and discounts that would not have been discovered without a thorough inspection of the initial sales data. The model was adjusted using a dummy variable to identify and isolate the below grade cotton in the estimates. In the final analysis, ongoing evaluation, testing, and experimentation is the most effective safeguard against these sources of error.

An advantage of the DPES as opposed to subjectively estimated prices is its inability to adjust to events occurring within the market not shown by actual sales. Thus, if trade or harvest rumors arise, the DPES will not react unless market participants react in their buying behavior. This makes the DPES results more consistent with market participants' behavior without a preconceived idea of how prices should be progressing.

Since the system is automated, it allows prices to be reported on heavy trading days as easily and timely as when trading is less active, with the same effort on the part of an analyst. However, analysis has shown that days with less than 40 observed lot sales throughout Texas and Oklahoma,

the DPES does not produce statistically reliable estimates. Thus, the system is designed not to estimate or report prices if less than 40 lot sales are traded for any given trading day. The perspective that "no information is better than misinformation" has been adopted to date. This is a potential problem for users, but could be lessened with the addition of other participating merchandizing firms.

The computerized econometric approach is faster and more accurate than an "impressionistic analysis" of the market. The underlying premise of least squares procedure is to minimize the squared errors; therefore, the estimates are as accurate as statistical methods will allow. The system provides estimates that are easily repeatable. The data are easily stored, and the model can be re-estimated at any time and the same prices, premiums, and discounts obtained. The prices, premiums, and discounts currently provided by AMS are based on the market news reporter having a sample of cotton traded and market intelligence based on discussions with merchants to form a mental picture of the way trading has been that day.

An approach such as the DPES has qualities that warrant consideration in market reporting. The DPES is easily changed to accommodate changes in the large amount of information provided by the grading system. As grading adopts new measurement technology, the DPES can accommodate the new information.

Summary and Conclusions

The primary purpose of this paper is to describe an objective approach for daily cotton market price reporting. The Daily Price Estimation System has been developed, tested, and evaluated, and is operating on a daily routine. Because the market in which it operates is dynamic, the DPES continues to require and receive periodic testing and validation. However, the basic operational and conceptual structure of the system is reliable.

The Daily Price Estimation System is ready to be introduced to a broader spectrum of the market. It has been operational for three years, and price estimates have been generated and tested for almost five years. It is currently limited to the

producer sale pricing point in the Texas and Oklahoma markets because of the availability of electronic sales transactions data. However, expansion of the Daily Price Estimation System is feasible. The DPES could be adapted relatively easily to become a national system for cotton price reporting, but the data needs would necessitate broad-based participation of marketing firms-- marketing associations and merchants, and perhaps some textile processors. The sales information could be keyed into the system for regions that do not have operational computerized electronic markets. This process would be logistically demanding, yet realistic with existing technology.

The most relevant concern in expanding the system to other regional markets would be the functional form of the daily hedonic model. Cotton quality varies across regions. Therefore, the functional form used for the Texas and Oklahoma markets may not necessarily be the appropriate functional form for the other market regions. Consequently, it would require model development work in each market to determine the appropriate functional form for each market. Whether it could be one model for all regions or a specific model for each region would have to be ascertained. However, work has been conducted in the area of regional price differences (Bowman and Ethridge; Chen, Chiou, and Anderson; Chen and Ethridge), and other analysis is in progress.

The system might also be expanded to include all pricing points throughout the marketing channels. In an efficient market, the only change in prices throughout the marketing channels should reflect the marketing margins, assuming equivalent qualities (Kohls and Downey). The structure of the premium and discount patterns should not be substantially affected[4]. Therefore, the functional form of the model under these conditions would not change with respect to the implicit prices reflected by the model if the markets are efficient in terms of conveying price signals. Initial indications are that the marketing system is not operating efficiently in getting price signals through to cotton growers (Ethridge and Chen). Thus, the producer prices, premiums, and discounts resulting from this model may not reflect the quality premiums and discounts at other pricing points.

420

References

Alder, H.L., and E.B. Roessler. *Introduction to Probability and Statistics.* 6th ed. San Francisco: W.H. Freeman and Co., 1977, pp. 163-166.

Belongia, M.T. "Are Economic Forecasts by Government Agencies Biased? Accurate?" *Econ. Rev.*, Federal Reserve Bank of St. Louis, Nov./Dec. 1988, pp. 15-23.

Bowman, K. R. and D. E. Ethridge. "Characteristic Supplies and Demands in a Hedonic Framework: U.S. Market for Cotton Fiber Attributes," *Amer. J. of Agr. Econ.*, 74 (1992): 991-1002.

Brown, J.E., and D.E. Ethridge. "Functional Form Model Specification: An Application to Hedonic Pricing." *Agr. and Resource Econ. Res.* (1995) (forthcoming).

Chen, D. T., G. T. Chiou, and C. G. Anderson. "Inter Regional Pricing of Cotton Fiber Attributes: Does Quality Matter?" *1991 Beltwide Cotton Conference Proceedings*, Cotton Economics and Marketing Conference, National Cotton Council, pp. 367-373.

Chen, C., and D. Ethridge. "Premiums and Discounts Paid for Cotton Fiber Quality Attributes by U.S. Textile Manufacturers" *1995 Beltwide Cotton Conference Proceedings*, Cotton Economics and Marketing Conference National Cotton Council, Memphis, TN, (addendum), pp. 92-96.

Elam, E., and B. Dixon. "Examining the Validity of a Test of Futures Market Efficiency." *J. of Fut. Mkts.*, 8 (1988): 365-372.

Ethridge, D. E. "A Computerized Remote-Access Commodity Market: Telcot," *S. J. of Agr. Econ.*, 10 (1978): 177-182.

Ethridge D. E. and C. Chen. "Cotton Price Responsiveness to Quality in the U.S.: Textile Mill Prices Paid VS. Producer Prices," *1993 Beltwide Cotton Conference Proceedings*, Cotton Economics and Marketing Conference, National Cotton Council, pp. 441-444.

Ethridge D. E. and B. Davis. "Hedonic Price Estimation for Commodities: An Application to Cotton," *W. J. of Agr. Econ.*, 7 (1982): 156-163.

Ethridge, D. E., C. Engels, and J. E. Brown. "An Econometric Approach for Estimating Daily Market Prices," *1992 Beltwide Cotton Conference Proceedings*, Cotton Economics and Marketing Conference, National Cotton Council, pp. 399-402.

Ethridge, D. E. and K. H. Mathews. "Reliability of Spot Cotton Quotations for Price Discovery in the West Texas Cotton Market," Agricultural Economics Dept., Texas Tech University College of Ag. Sciences. Pub. No. T-1-212., Aug. 1983.

Ethridge, D. E. and S. Morse. "Hedonic/Characteristics Price Analysis in Agricultural Economics: Progress, Issues, and Needs," Agricultural Economics Dept., Texas Tech University College of Ag. Sciences. Pub. No. T-1-347. Aug. 1991. Abstract in *Amer. J. of Agr. Econ.*, 73 (1991): 1528-1529.

Ethridge, D. E. and J. T. Neeper. "Producer Returns From Cotton Strength and Uniformity: An Hedonic Price Approach," *S. J. of Agr. Econ.*, 19 (1987): 91-97.

Ethridge, D. E., D. L. Shaw, and J. E. Ross. "An Evaluation of the Impact of Instrument Test Line Values on Cotton Marketing and Use: Progress Report," CED Working Paper, Economic Research Service, USDA, Aug. 1977.

Ethridge, D. E., R. Sudderth, and M. Moore. "A Model for Cotton Variety Selection with Yield and Quality Considerations," *1983 Beltwide Cotton Research Conference Proceedings.*, Cotton Economics and Marketing Conference, National Cotton Council, pp. 251-253.

Harrison, S.R. "Regression of a Model on Real System Output: An Invalid Test of Model Validity." *Agr. Sys.*, 34(1990): 183-190.

Hembree, J. F., D. E. Ethridge, and J. T. Neeper. "Market Values of Fiber Properties in Southeastern Textile Mills," *Textile Res. J.*, 56 (1986): 140-141.

Holden, K., and D.A. Peel. "On Testing for Unbiasedness and Efficiency of Forecasts." *The Manchester School*, 58(1990): 120-127.

Hudson, D., D. Ethridge, and J. Brown. "Producer Prices in Cotton Markets: An Evaluation of Reported Price Information Accuracy." *Agribusiness: An Intl. J.*, (in press).

Kohls, R. L. and W. D. Downey. "Cotton Marketing," *Marketing of Agricultural Products*, 4th ed., New York, New York: The Macmillan Company, 1967, pp. 392-407.

Neter, J., W. Wasserman, and M. H. Kutner. *Applied Linear Regression Models.* 2nd ed. Homewood, IL.: Richard D. Irwin, Inc., 1989.

Pindyk, R., and D. Rubinfeld. *Econometric Models and Economic Forecasting.* 3rd ed. New York: The Dryden Press, 1991, pp. 216-217.

Robinson, J. A., D. E. Ethridge, D. L. Shaw, and C. D. Rogers. "Market Impact Evaluation of the Instrument Testing of Cotton: 1977 Crop Progress Report." CED Working Paper, Economic, Statistics, and Cooperative Service, USDA, May 1979.

Rosen, S. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *J. Polit. Econ.*, 82 (1974): 34-55.

SAS Institute, Inc. "SAS/STAT User's Guide, Release 6.03 Edition." Cary, NC, 1988.

U.S. Department of Agriculture, "Daily Spot Cotton Quotations." Market News Branch, Cotton Division, Agricultural Marketing Service, Memphis, TN, daily issues.

U.S. Department of Agriculture, "Cotton and Wool Situation and Outlook Report," Commodity Economics Division, Economic Research Service, CWS-73, August 1993.

Endnotes

1. The functional form is a modification of that found by Brown and Ethridge to facilitate the reporting of prices, premiums, and discounts. The difference is the relinquishing of a quadratic term on level 2 bark, the addition of the two other extraneous matter designations, and use of the color grade, which were not available in that study.

2. Using Alder and Roessler's method for determining survey sample size, $n = [t(s)/d]$; $t = t$-value at $\alpha = 0.9995$, $v = \infty$ (3.291); $s =$ standard deviation; $d =$ maximum allowable error (set at 1 ¢/lb.). A sample of prices from 4349 observed sales gave a mean price 53.778 and $s = 5.3859$. Thus, calculated n was 314.18.

3. Alternative measures such as the Dickey-Fuller unit root test were considered. However, given the cross-sectional nature of the data set, these approaches are not applicable.

4. The time value of money may affect the patterns of the premiums and discounts of various qualities through the marketing channel; however, the magnitude of these effects should be minimal.