Hedonic Price Estimation for Commodities: an Application to Cotton

Don E. Ethridge and Bob Davis

A model of hedonic prices — implicit prices of embodied quality attributes — was developed for cotton lint and the relative importance of various quality attributes were estimated with regression analysis from sample data on observed sales of cotton. Results indicated that producer prices were sensitive to variations in fiber length, micronaire, and trash content. Results also revealed differences in relative importance and sensitivity between years.

Hedonic prices are the implicit prices of attributes or characteristics embodied in a commodity as opposed to the price of the commodity itself. The underlying hypothesis is that goods are valued for their utility-bearing characteristics and that prices of goods vary with the specific amounts of those characteristics associated with them [Rosen]. Thus, markets generate observed product prices which are a composite of some (often undefined) set of embodied characteristics.

The general purpose of this paper is to present an approach to the estimation of hedonic prices for a semi-processed agricultural commodity, cotton lint. The specific purpose of the study was to determine the relative impacts of the various quality attributes of cotton lint on producer prices, i.e., to determine the value of fiber length (or micronaire, color, or trash content) as it conributes to the value of a bundle of cotton in which it is embodied.

Most empirical hedonic work has concentrated on hedonic price indexes — removing quality change from price indexes [Griliches, 1971]. In addition, the empirical studies have dealt predominantly with manufactured industrial products such as automobiles [Fisher, Griliches, and Kaysen; Griliches, 1961; Cagan; Dhrymes], tractors [Fettig], houses [Bailey, Muth, and Nourse; Musgrave], diesel engines [Kravis and Lipsey], refrigerators [Dhrymes], washing machines and carpets [Garett], and electric apparatus [Dean and DePodwin]. Although not labeled hedonic price estimation, Waugh's 1929 study of vegetable prices was an early attempt at such an endeavor. The empirical procedure for estimation of hedonic price indexes is useful wherever underlying product characteristics are measurable, but their impact is not necessarily obvious. For example, the early estimation by Griliches [1961] examined the impacts of horsepower, weight, length, body style, engine and transmission types, and other factors on U.S. passenger car prices.

It is important to note that estimated hedonic price functions typically identify neither demand nor supply functions [Rosen] although the literature contains attempts to treat product quality in consumer demand theory [Murphy; Officer] and component pricing of commodities [Perrin]. However, both observed prices and implicit prices of embodied attributes may be affected by market demand and/or supply considerations; for example, they may change as quantities of the product change. Because of market forces, the implied value of an embodied quality attribute may not be constant over

Don E. Ethridge and Bob Davis are Associate Professors in the Department of Agricultural Economics at Texas Tech University. The authors wish to acknowledge contributions made by Sujit Roy, David Pyles, Stanley Johnson, Dean Ethridge, Joel Hembree, Dale Shaw, the Editors, and three anonymous reviewers.

Texas Tech University College of Agricultural Sciences Publication No. T-1-204.

time and may vary with the specific market (regional, end-use, etc.). The hedonic estimation process may then have to adjust for effects of changes in market forces over time when time series data are used and provide a means for comparison of hedonic prices at different points in time when cross section data are employed.

This analysis was applied to a semi-processed agricultural product, cotton lint. Impact of quality attributes on price rather than a price index was the main focus of the study, and primary as opposed to secondary data were utilized. The approach also entailed use of generalized least squares estimation procedures to handle data problems, especially autocorrelation. The procedures may be applicable to many other types of commodities.

The paper is divided into three sections: (1) an explanation of the hedonic model including institutional attributes of the cotton market that impact model formulation, (2) data considerations and their effect on the model, and (3) findings and conclusions.

Hedonic Price Model for Cotton

Most domestic cotton producers sell cotton on the basis of a class card, or officially, the Smith-Doxey classification system. Each producer's cotton is graded by USDA employees who examine a sample from each bale and assign values for three quality attributes — grade, staple, and micronaire — which are then recorded on the class card.[1] The producer sells the cotton on the basis of the official values. The first attribute, grade, is a two-digit index that depicts the color of the cotton lint and its trash content. The first digit of the grade index is a scale of 1 to 8 which indicates the trash content and a characteristic called "condition" [USDA, 1980]. The higher the digit, the more trash contained in the sample and the worse the condition. The second digit of the grade index may assume a value from 0 to 7 which refers to the color characteristics of the lint.[2] Pure white cotton would be assigned a low number, whereas yellow, gray or discolored cotton would receive higher values, with the values increasing for the less desirable colors, indicating lower quality cotton. The second quality variable, staple, signifies the length of cotton fiber in 32nds of an inch, thus staple 32 denotes a 1-inch fiber length. Micronaire is an index of fiber fineness and maturity. Micronaire values, such as 3.1, 3.8, 5.0, etc. are determined by an instrument with fineness decreasing (coarseness increasing) as the numbers rise. Micronaire values range from 2.5 to 5.4. Grade and staple are determined visually by the federal grader.

Producers generally sell cotton in mixed lots. That is, individual bales are combined into lots of varying numbers of bales of differing quality characteristics. These lots are sold to merchants on the basis of a "recap", which is a one-page summary of the number of bales in each quality category [Ethridge, Shaw, and Ross]. The merchant pays one price per pound for all bales in the lot. This practice complicates model formulation because sale prices for individual bales of cotton are observable only in the case of one-bale lots. Thus, the model was adjusted to accommodate data on mixed-lot sale observations by using lot averages for the meaningful quality variables. Lot size and variation in quality within lots were hypothesized to affect the

[1]The Cotton Division, Agricultural Marketing Service, USDA is the federal agency responsible for classing cotton. The explanation of cotton classing presented here is greatly simplified. For more information, see USDA, "The Classification of Cotton".

[2]This interpretation of the grade code is a simplification of a highly complex set of descriptive standards and is used with the assumption that the simplifications are realistic for purposes of quantification. A problem with utilizing the second digit of the grade code is that the numbers 6 and 7 on the scale refer to grayness and are not a true continuum from the previous numbers. However, for the cotton sampled in this study area (Lubbock, Texas, Classing Territory), an inconsequential amount falls in these groups (about .01 percent). Therefore, the model formulation below may not be usable as a generalized model, *i.e.*, when considering *all* ranges of color.

average cotton price paid per lot. Lot size was expected to directly affect price since the paper work and time required for merchants to purchase a large lot of cotton is about the same as for a small lot. Quality variation within lots was expected to affect lot price for two reasons. First, merchants need less time and effort to organize standardized mill orders with more homogeneous lots, and second, homogenous lots should tend to reduce merchants' uncertainty about market outlets.

From the considerations discussed above and excluding the market forces which affect general price levels and vary with time, the hedonic price model for cotton was specified as:

$$(1) \qquad P = h(G_1, G_2, L, M, LS, VG_1, VG_2, VL, VM)$$

where

P = producer price for a lot of cotton in cents per pound

G_1 = average first digit of the grade code for the lot of cotton

G_2 = average second digit of the grade code for the lot of cotton

L = average staple length code for the lot in 32nds of an inch

M = average micronaire reading for the lot

LS = lot size in number of bales

VG_1, VG_2, VL, VM = variation about G_1, G_2, L, and M within the lot of cotton

Data Considerations

Data for all variables in equation (1) are available on or can be computed from recap sheets. The data set consisted of 992 recap sales observations from eight gin pionts in a localized area of the Texas High Plains for the 1976/77 and 1977/78 seasons.[3] Because average cotton quality appeared to be different for each year, the model was estimated for each year as well as for both years combined. For each sales observation, the average price for the lot of cotton and the lot size were observed directly; average lot values for G_1, G_2, L, and M were calculated. Since micronaire values were reported as the number of bales in micronaire groupings, the midpoint of each group was used to calculate a mean value for M. Standard deviations of quality variables within lots were used as indices of VG_1, VG_2, VL, and VM.

Examination of the data suggested multicollinearity problems among the indices of quality variability within a lot of cotton.[4] Two alternatives for adjusting the model were considered: (1) constructing a composite variability index and (2) using only one of the variability indexes in the estimating equation. The first approach required a method of assigning weights to the various indices (VG_1, VG_2, VL, and VM) and was deemed to be arbitrary. The second approach required some means of selecting a single index. Discussions with individuals knowledgeable in cotton quality, merchandizing, and textile mill use suggested that of the four indices, VM would be the most appropriate single

[3]The data were gathered for a study of instrument testing of cotton [Ethridge, Shaw, and Ross; Robinson, et al.], although not all of the data gathered were used for that purpose.

[4]The first suggestion that a problem might exist was that VG_1, VG_2, L, and VM were all intercorrelated with simple correlation coefficients between .31 and .55; individual simple correlation coefficients did not necessarily signify a problem, but the fact that they were all intercorrelated did indicate a potential problem. Earlier analysis [Ethridge, Shaw, and Ross] which found high standard errors and low t values of coefficients coupled with a high F statistic for the equation and that coefficients were sensitive to the inclusion of other variables further supported the existence of multicollinearity [Intriligator, pp. 153-56].

index because (1) the spinning process is more sensitive to micronaire than to grade and fiber length characteristics and (2) variations in micronaire pose a major problem in the finishing (dyeing) process. Thus, if price is sensitive to variation of quality within a lot of cotton, one might expect more sensitivity to variation in micronaire than to other quality attributes. The confirming rationale for using VM was that initial regression analysis showed that of the variability indexes, only VM had a coefficient with a significant t value at the .05 level.

The model was specified as linear for all variables except micronaire within the range of values contained in the sample of cotton used. There is a range of micronaire values considered to be most desirable for textile manufacturing and the value of cotton tends to decrease as micronaire deviates both ways from that range. This pattern is also reflected in the government loan rates. Thus, the relationship between price and micronaire was expected to be curvilinear and was formulated as quadratic. Again excluding market impact, the resulting hedonic price regression model was:

$$(2) \quad P = B_0 + B_1G_1 + B_2G_2 + B_3L + B_4M + B_5M^2 + B_6LS + B_7VM + \varepsilon$$

where

B_i = parameters

VM = standard deviation of micronaire readings within a lot of cotton

ε = stochastic error term

Expected signs for the regression coefficients are as follows:

1. The coefficients of G_1 and G_2, the two components of grade should both be negative since higher values for these variables represent lower grades and price should fall as grade deteriorates.
2. The coefficient of L should be positive since longer staple cotton will bring a higher price, other things equal.
3. The coefficients of M and M^2 should be positive and negative, respectively, since a particular quadratic relationship among price and these variables is expected. As mentioned earlier, price will tend to be low for very low and very high micronaire values and high for intermediate values. Thus a mound-shaped quadratic concave to the origin is expected.
4. Price is expected to vary directly with LS since transactions costs are diminished as lot size increases and the merchant can afford to pay a slightly higher price for large lots, other things equal.
5. The coefficient of VM is expected to be negative. The greater the variability of the micronaire the greater the processing problems with dyeing and finishing and the lower the value or price of the cotton.

Results

The model was estimated using both ordinary and generalized least squares.[5] Generalized least squares were used because the data comprised a time series for each of nine gin locations and autocorrelation problems were expected both because this is common with time series data in general and because the cotton data tend to exhibit seasonal price patterns over time which give rise to autocorrelation. The ordinary least squares procedure was used to verify whether the expected autocorrelation was present. Equations were estimated using data for each crop year separately and for both years combined with year as a dummy variable.

The OLS equations, numbers (1), (2) and (3), Table 1, included date of sale within the season as an independent variable coded in sequential order beginning with November 1

[5]The use of weighted regression was considered on the basis that price variation might increase as lot size increases. However, to use lot size to weight observations assumes that observed lot prices are weighted averages of individual bale prices, which is not the industry practice. The mixed lot is the smallest unit of observation and a per pound price for the lot is established between the trading parties with the buyer having only the summary of quality information available on the recap.

TABLE 1. Hedonic Price Regression Coefficients for Producer Sales of West Texas Cotton; 1976/77, 1977/78 and Combined Crop Years.[a]

	Dependent Variable: Cotton Price					
	Ordinary Least Squares			Generalized Least Squares		
Independent Variable	1976/77 (1)	1977/78 (2)	Both Years (3)	1976/77 (4)	1977/78 (5)	Both Years (6)
Constant	11.370	−7.947	−16.440	0.137	0.306	0.513
G_1	−1.965	−1.506	−1.453	−1.060	−1.259	−0.9599
	(−6.43)	(−14.09)	(−8.13)	(−7.04)	(−19.31)	(−6.83)
G_2	−2.970	−1.011	−1.991	−3.121	−0.840	−2.270
	(−11.10)	(−8.10)	(−10.77)	(−20.13)	(−10.94)	(−15.59)
L	0.867	1.070	0.917	0.351	1.059	0.442
	(6.10)	(16.11)	(9.55)	(7.54)	(23.21)	(8.89)
M	15.860	10.420	26.670	31.580	6.928	25.971
	(5.94)	(6.35)	(18.20)	(29.34)	(8.95)	(28.93)
M^2	−1.424	−1.182	−2.887	−3.986	−0.755	−2.975
	(−3.68)	(−5.96)	(−14.50)	(−23.75)	(−7.79)	(−23.10)
VM	−3.791	−0.246	−3.319	−3.335	−0.740	−2.688
	(−4.60)	(−0.72)	(−6.42)	(−7.88)	(−3.56)	(−6.32)
LS	0.0075	0.0041	0.0049	0.0040	0.0045	0.0033
	(1.96)	(2.97)	(2.12)	(1.41)	(5.13)	(1.69)
Date	−0.0140	0.0433	0.0245	-	-	0.0228
	(−2.43)	(30.39)	(9.08)			(5.54)
Year			−21.44			−20.296
			(−66.12)			(−67.85)
R^2	0.763	0.843	0.860	0.900	0.897	0.905
F	224.52	279.73	670.48	705.11	514.04	1035.73
n	566	426	992	566	426	992
d	0.62	0.81	0.75	2.02	2.39	2.11

[a]Numbers in parentheses below coefficients are parameter t-values.

as 1. This variable was used to adjust for any trend of prices within years. All OLS coefficients had expected signs and high t-values; however, the Durbin-Watson d-statistic indicated significant autocorrelation with all three equations (Table 1).[6] The signs on the coefficients for date of sale were opposite in equations (1) and (2) indicating different seasonal patterns for price for the two crop years. The coefficient for the date variable in equation (3) was positive but half the size of that for 1977/78, while the dummy variable for year was negative, indicating lower prices in 1977/78. The dummy variable for year was used to adjust for the effect of other factors which influence the general level of price between the two years.

The generalized least squares equations, numbers (4), (5) and (6), Table 1, also had statistically significant coefficients with expected signs, and as expected, the d-values showed no autocorrelation. Only the coefficient for lot size (LS) in equation (4) appeared to have a somewhat low t-value, but it too, is significant at the .10 level of probability. A Chow test [Intriligator, p. 194] was con-

[6]The data set did not constitute a true time-series because multiple sales occurred on some dates and no sales occurred on other dates. However, since the observations were arranged in time sequence in the data set, the d-statistic was used.

ducted to see if the regression coefficients in equations (4) and (5) were equal. The resulting computed F of 879.5 greatly exceeded the table value of 2.5 at the .01 level of significance. Thus the hypothesis of equality of regression coefficients for the two years was rejected, signifying that the quality attributes affected price differently in each year and the equation from pooled data was not a realiable indicator of their effects.

From an examination of the regression coefficients in equations (4) and (5), micronaire, variability in micronaire and the color of the cotton fiber tended to be much more important in affecting price in the 1976/77 crop year than in the 1977/78 crop year, while fiber length was less important. Lot size and the trash component of grade seemed to have about the same influence in each year. Elasticity estimates, computed at the mean values of the variables, indicate similiar relationships (Table 2). If the sizes of the elasticities are meaningful, then micronaire, fiber length and color seemed to have the most impact on price, but the absolute amount of the impact varied by crop year. Weather and other factors affect the maturity and color of the crop as measured by these three important quality variables. For the two years under consideration, in 1976/77 cotton was of relatively longer staple, but lower in micronaire and more off-color than in the following year. The model indicates those factors contributed to the lower prices in 1976/77. Given the variability in growing conditions, this study shows that it may not be practical to obtain one general equation that will show the effect of quality on price for a period of years, rather a more complex system of equations may be needed. Quality appears to affect price in any given year, with the effect dependent upon the particular quality attributes present in that year's crop, and upon the demand for the attributes.

If, as the hedonic price hypothesis suggests, there is an implied market for the individual quality attributes then relative scarcity of a quality attribute will raise its price, or in the context of this analysis, its relative importance in the determination of the market price of the product. For example, consider the coefficients for fiber length in columns (4) and (5), Table 1. There was a relative abundance of cotton with longer fiber length in 1976/77 (compare mean fiber lengths, Table 2), so the value of length was lower in 1976/77. Conversely, there was a relative scarcity of high micronaire cotton in 1976/77 (mean value of 3.35 versus 4.37 in 1977/78), so the relative impact of micronaire values was greater in 1976/77.

Implications

Many factors other than quality attributes affect producer cotton prices. However, the range of prices implied by the variation in explanatory variables in this study has a substantial affect on producer prices over and above those from formal market fluctuations. When all variables in equation (4) are varied one standard deviation in both directions from their means, cotton price varies from 43 to 67 cents per pound. For the variables in equation (5) the same type of variation produces a range of cotton prices from 46 to 75

TABLE 2. Estimated Elasticity Coefficients at Mean Values of Variables.[a]

Elasticity of Expected Price	With Respect to:					
	G_1	G_2	L	M	VM	LS
1976/77	−.078	−.134	.206	1.920	−.016	.002
(55.11)	(4.04)	(2.37)	(32.35)	(3.35)	(0.27)	(29.0)
1977/78	−.073	−.019	.556	.499	−.002	.003
(60.66)	(3.53)	(1.35)	(31.85)	(4.37)	(0.20)	(45.3)

[a]Mean values of variables are shown in parentheses.

cents per pound.

Once the relative values of the quality attributes are known, producers, ginners, policy makers and others can influence the variables to some extent. Producers have substantial influence on the values associated with LS and VM through their approach to formulating mixed lots for sale. However, these two variables have the least relative impact on price. Variety selection may be the best way that producers have of affecting G_2, M and L, while G_1 may be influenced through ginning practices and processes. Political decisions to support research in plant genetics and/or educational programs and market information on the value of quality attributes are positive ways that policy makers can help producers in the long term.

In interpreting these results, the reader is reminded that measured price impacts occurred at the point of first sale of the commodity and applied only to producer prices. To the extent that the final users of cotton as a fiber look for other quality attributes not currently observed at the first pricing point, the model applied here would need to be adjusted to take those other factors into account.

References

Bailey, M. J., R. F. Muth, and H. O. Nourse. "A Regression Method for Real Estate Price Index Construction", *Journal of the American Statistical Association*. 58(1963):933-942.

Cagan, Phillip. "Measuring Quality Changes and the Purchasing Power of Money: An Exploratory Study of Automobiles." *National Banking Review*. 3:217-236.

Dean, C. R., and H. J. DePodwin. "Product Variation and Price Indexes: A Case Study of Electrical Apparatus." *Proceedings of the Business and Economic Statistics Section*, American Statistical Association, 1961, 271-279.

Dhrymes, Phoebus J. "On the Measurement of Price and Quality Changes in Some Consumer Capital Goods." *American Economic Review*. 57(1967):501-518.

Ethridge, Don E. "A Computerized Remote-Access Commodity Market: Telcot." *Southern Journal of Agricultural Economics*. 10(1978):177-182.

Ethridge, Don E., and Dale L. Shaw. "Evaluating the Impact of Instrument Test Line Values on Cotton Marketing and Use." *1977 Proceedings of the Beltwide Cotton Production Research Conferences*. January 10-12, 1977, National Cotton Council, 209-212.

Ethridge, Don E., Dale L. Shaw, and John 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, U.S.D.A., Aug. 1977.

Fettig, Lyle P. "Adjusting Farm Tractor Prices for Quality Changes, 1950-1962." *Journal of Farm Economics*. 45(1967):599-611.

Fisher, Franklin M., Zvi Griliches, and Carl Kaysen. "The Costs of Automobile Model Changes since 1949." *Journal of Political Economy*. 79(1971):433-51.

Garett, Thomas W. "Quality and a Pure Price Index." *Monthly Labor Review*. 90(1967):16-20.

Griliches, Zvi. "Introduction: Hedonic Price Indexes Revisited." Zvi Griliches, ed., *Price Indexes and Quality Change*. Cambridge, Mass.: Harvard Univ. Press 1971.

Griliches, Zvi. "Hedonic Price Indexes for Automobiles: An Econometric Analysis of Change." *The Price Statistics of the Federal Government*. New York: National Bureau of Economic Research. General Series, No. 73 (1961):137-96.

Intriligator, Michael D. *Econometric Models, Techniques, and Applications*. Englewood Cliffs, N.J.: Prentice-Hall, Inc. 1978.

Kravis, Irving G., and Robert E. Lipsey. "International Price Comparisons." *International Economic Review*. 10(1969):233-46.

Murphy, James L. *Introductory Econometrics*. Homewood, IL: Richard D. Irwin, Inc. 1973.

Murphy, Michael M. "Quality and Consumer Demand." *The American Economist*. 24(1980):42-5.

Musgrave, J. C. "The Measurement of Price Changes in

Construction." *Journal of the American Statistical Association*. 64(1969):771-86.

Officer, Lawrence H. "Demand Conditions Under Multidimensional Pricing." *Econometrics and Economic Theory: Essays in Honour of Jan Tinbergen*. Willy Sollekarerts, ed. White Plains, N.Y.: International Arts and Sciences Press. 1974.

Perrin, Richard K. "The Impact of Component Pricing of Soybeans and Milk." *American Journal of Agricultural Economics*. 62(1980):445-55.

Robinson, Jess A., Don E. Ethridge, Dale L. Shaw, and Clarence D. Rogers, "Market Impact Evaluation of the Instrument Testing of Cotton: 1977 Crop Progress Report.: CED Working Paper, Economics, Statistics, and Cooperatives Service, U.S.D.A., May 1979.

Rosen, Sherwin. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy*. 82(1974):34-55.

U.S. Department of Agriculture. "Cotton Quality, Crop of ______." Agricultural Marketing Service, Cotton Division, various annual issues.

U.S. Depatment of Agriculture. "Summary of Cotton Fiber and Processing Test Results, Crop of ______." Agricultural Marketing Service, Cotton Division, various annual issues.

U.S. Department of Agriculture. "The Classification of Cotton." Agricultural Marketing Service, Agriculture Handbook No. 566, April, 1980.

Waugh, Frederick V. *Quality as a Determinant of Vegetable Prices*. New York: Columbia University Press. 1929.