MODIFYING TRADITIONAL OPTION PRICING FORMULAE FOR OPTIONS ON SOYBEAN FUTURES

Robert J. Hauser
Assistant Professor
University of Illinois

Dane K. Andersen
Graduate Student
University of Illinois

Presented at the 1984 AAEA Annual Meeting, 5-8 August 1984, Cornell University.

MODIFYING TRADITIONAL OPTION PRICING FORMULAE FOR OPTIONS ON SOYBEAN FUTURES

Robert J. Hauser and Dane K. Andersen

Abstract. The option pricing assumptions that (a) the logarithmic price return on soybean futures is distributed normally and (b) the variance of the instantaneous return is constant throughout the option contract's life are investigated. Systematic variance changes are then incorporated into an option pricing formula and the resultant premia are compared to constant-variance premia.

Modifying Traditional Option Pricing Formulae for Options on Soybean Futures

Based on discussions with exchange executives, trading of options on select agricultural futures will probably begin by late 1984 at many of the futures exchanges. The option contract will offer agricultural producers and merchandisers an additional marketing tool; however, the extent of its use will depend largely on the option premium. Thus the appropriateness of traditional option pricing formulae for pricing options on agricultural futures becomes a very important issue to agricultural economists studying the price and use of agricultural options.

Contemporary theory of option pricing is based on works done in 1973 by Black and Scholes and by Merton. In short, equilibrium pricing models were developed under the assumption that an option trader can form a risk-free portfolio of options, underlying commodity, and bonds. Given this "risk-free hedge" and the assumption that the underlying commodity price is distributed log-normally[1] at the end of any finite period, then a closed form solution for the option price can be found which is valid for all risk preference structures.

The basic Black-Scholes model is for options on non-payout physicals. Black develops a pricing formula for options on futures by adjusting the Black-Scholes model to account for the reduced cost of holding the underlying commodity; i.e., implicit in the Black-Scholes model is that the underlying commodity is held at a cost rate equal to the risk-free interest rate, as opposed to Black's model for futures in which the cost of holding the underlying commodity (futures) is zero. For an excellent review of the theory of pricing options on physicals, see Smith; for a review of pricing options on futures, see Asay (1982, 1983).

This paper investigates two assumptions of Black's option pricing model: (a) the logarithmic return in the futures price is distributed normally and (b) the variance of the instantaneous return is constant over the life of the option contract. Given the empirical results of this investigation with respect to soybean futures prices, Black's option pricing formula is modified to provide for systematic changes in the instantaneous variance and the resulting premium estimates for case examples are compared to premium estimates based on constant variances.

Log-Normality and Constant Variance

As discussed below, the option pricing assumptions regarding log-normality and constant variance have already received a fair amount of attention in past studies. Examination of log-normality or normality has been in "indirect" and "direct" forms. The indirect tests have usually been associated with price efficiency studies which investigate serial independence of first differences. If serial independence is found, then normality is implied by invoking the central limit theorem under the assumption that the differences have a finite variance. Testing linear dependencies through the use of trading rules, Houthakker, Leuthold, Peterson and Leuthold, Smidt, and Stevenson and Bear find non-random behavior. Cargill and Rausser (1972, 1975), Labys and Granger, Leuthold, Mann and Heifner, and Rocca find mixed results when using time-domain and frequency-domain statistical tests.

Direct tests examine the shape of the distribution and usually investigate the finite variance hypothesis. The null hypothesis (implicitly or explicitly) is usually that the log-price differences are drawn from a Gaussian (normal) distribution of the stable Paretian family versus the alternative hypothesis that the changes are drawn from a non-Gaussian (non-normal) stable Paretian distribution. For agricultural commodities, Houthakker, Mandelbrot, Mann and Heifner, and Stevenson and Bear find, in general, leptokurtotic distributions.

Many authors have suggested that leptokurtocity is often found because series are not defined "correctly". Clark contends that the price distribution is subordinate to a normal distribution where the observations are defined on a trade-to-trade basis. A common reason offered for finding leptokurtocity is that a non-constant variance exists (Houthakker, Cootner, Mann and Heifner, and Stevenson and Bear). An interesting theoretical model developed recently by Neftci suggests that seasonality within an efficient market will be exhibited in price variance and not in expected price.

Kamara identifies two general hypotheses attempting to explain variance changes over a futures contract's life: (a) the well known time-to-maturity effect (e.g., Samuelson) which causes price volatility to increase as maturity date approaches and (b) the state variable effect (e.g., Anderson and Danthine) which reflects changes in underlying state variables such as supply and demand uncertainty.

The empirical analysis below investigates the distribution and variance characteristics of soybean futures prices.

The Case of Soybean Futures

Soybean futures are chosen for analysis because of the high probability that options on soybean futures will be traded (based on discussions with exchange executives). Twenty-four March contracts (1960-1983), 23 July contracts (1960-1982), and 23 November contracts (1960-1982) are examined.[2] The first differences of the natural logarithms of daily closing prices are analyzed. Normality tests are conducted on this "raw" series and on a "standardized" series.[3] Observations are standardized in attempt to offset the effect of changing variances on tests of distribution and is done by dividing each logarithmic difference by the estimated standard deviation calculated from the observations of the respective month.

Three direct tests of normality are performed using Statistical Analysis System's calculations of kurtosis, a skewness estimate, and the Kolmogorov D statistic. The results are summarized in Table 1. For the raw series, the kurtosis and D tests, in general, do not support the hypothesis that the differences are distributed normally, particularly for the 1960-1972 period. The skewness tests, however, tend to support the hypothesis although it is possible for a distribution to not be skewed yet not be normal. Thus, based on the kurtosis and D tests we conclude that the raw series is not distributed normally and, because of the significantly positive kurtosis estimates, leptokurtocity is indicated. In contrast, the tests on the standardized series are much more supportive of the normality hypothesis. By simply standardizing for a non-constant variance, stronger support for normality than has been previously found for most agricultural price series is established, suggesting that normal distributions may indeed exist but, through time, exhibit different variances.

In an attempt to explain the changes in the variance, linear ordinary least squares models were developed under four categories distinguishing independent variable specifications as (a) single factors reflecting either the maturity effect, a seasonal effect, volume, or open interest, (b) two factors representing the maturity and seasonality effects, (c) three factors representing the maturity, seasonality, and year effects and (d) in addition

Table 1. Signs and Significance of Normality Tests

Contract Year	March Contracts: Raw Series K[a]	S[a]	D[a]	March Contracts: Standardized Series K	S	D	July Contracts: Raw Series K	S	D	July Contracts: Standardized Series K	S	D	November Contracts: Raw Series K	S	D	November Contracts: Standardized Series K	S	D
1960	+**[b]	–	*	**	–		+**	–*	*	+	–		+**	+*	**	+	+	*
1961	+**	+**	**	+	+		+*	–	**	–	+		+**	–**	**	+*	–	*
1962	+**	–	**	+	–	*	+*	+	**	+	+		+**	+**	**	+*	+	**
1963	+**	+*	**	+*	+	*	+**	+	*	+*	+		+**	+	**	+**	+	
1964	+**	–	**	+*	+		+**	+	**	+*	–		+**	+**	**	+*	–	
1965	+**	–	**	+*	–	*	+**	–	**	+	–		+**	–	**	+	–	
1966	+**	+*	**	+	–		+**	+**	**	+	+		+**	+	**	+	+	**
1967	+**	+*	**	+	+		+**	–	**	+*	+		+**	+*	**	+**	–	*
1968	+**	+**	**	+	+	*	+**	–	**	+*	–		+**	–*	**	+*	–	**
1969	+**	+	**	+	–	*	+**	+	**	+	+	*	+**	–	**	+	–	
1970	+**	–**	**	+	–**	**	+**	+	**	+*	–*	*	+**	+**	**	–	+	
1971	+**	+	*	+	+		+**	–	**	+**	+	*	+**	–	**	+	–	
1972	+**	–		+	+		+**	–	**	+	+	*	+**	–		+*	+	
1973	–	+		–	–		+**	–**	**	–**	–	**	–	+	*	–**	–	**
1974	–	+		–**	–	**	–	+		–**	–	**	–**	–	**	–**	+	**
1975	–**	+	**	–**	+	**	–**	+	**	–**	–	**	–**	+		–**	–	
1976	–	–		–	–		+	–	**	–	–	*	–	–	*	–	–	*
1977	–	–	*	–	+		–*	–	**	–**	+	**	+*	–	**	–	–	*
1978	+*	–	**	+	–		+	–	*	+	–		+**	–**	*	+	–	
1979	+	–**	*	+	–		+*	–		–	–*		+**	–**	**	–	–	
1980	+**	–**	**	+	–		+**	–	*	+	–		+**	+	**	–	–	
1981	+	–		–	–		–	–	*	–	–		+*	–	*	+	–	
1982	+**	–	*	+	–		+**	–	**	+	–		+	–		–	–	
1983	+	–		–	–		c	c	c	c	c	c	c	c	c	c	c	c

[a] Tests using kurtosis, skewness, and Kolmogorov D statistics are labeled K, S, and D, respectively. The D statistic must be positive.

[b] Significance levels for the hypothesis of normal characteristics are: ** = 2%; * = 10%.

[c] The July and November 1983 contracts were not tested.

to those effects under (c), volume and/or open interest. Model results under specifications (a) and (b) had low R^2's, low Durbin-Watson statistics, and coefficient estimates rarely significant at the .05 level. A tremendous improvement in the performance statistics is realized when year effects are combined with maturity and seasonality effects whereas open interest and volume variables add little to model performance. Thus, the model chosen as "best" is very similar to Anderson's (p. 14) and can be expressed as:

(1) $$V_{it} = a + b_1M + b_iS_i + b_tY_t + e_{it},$$

where V_{it} is the variance of log-price first differences of observations in month i and year t; M is the number of months until maturity; S_i is a binary variable for month i (seasonality effect, i=2,3,...,12); Y_t is a binary variable for year t (year effect, t=61,62,...,83); and e_{it} is the error term under classical assumptions. A major difference between this model and Anderson's is that Anderson used observations across all contracts within one model whereas we run separate regressions for individual contracts because of the intertemporal price dependence of a storable commodity. Another difference is that Anderson's dependent variable is the natural logarithm of the variance of prices.

Regression results are presented in Table 2. Three general conclusions are drawn. First, the time to maturity effect does not seem to be strong; although its coefficient is of expected sign for all three contracts, it is significant at the .05 level for only the November contract. Second, the seasonal effect is largest during June, July, and August, corresponding with the time of year in which supply and demand uncertainty is relatively high. Third, the annual level or year effect can be divided into three general periods: 1960-1972, 1973-1978, and 1979-1983. These period classifications

Table 2. Regression Results for Explanatory Models[a]

Variable[b]	March Coefficient	March t-value	July Coefficient	July t-value	November Coefficient	November t-value
Intercept	2.634	0.40	3.230	0.27	4.369	0.59
TTM	-0.785	-1.87	-0.429	-0.57	-0.866	-2.04*
S2	-4.452	-1.05	-3.278	-0.37	-3.411	-0.68
S3	-0.814	-0.13	0.454	0.05	-1.690	-0.33
S4	2.242	0.40	0.608	0.07	-2.051	-0.40
S5	1.494	0.28	-2.196	-0.26	-3.629	-0.69
S6	12.347	2.43*[c]	15.962	1.88	7.089	1.31
S7	19.506	3.98**	43.216	3.56**	14.453	2.58*
S8	11.393	2.39*	11.482	1.18	6.366	1.10
S9	4.321	0.93	3.326	0.36	-1.114	-0.21
S10	2.759	0.60	1.627	0.18	-0.800	-0.16
S11	2.244	0.49	1.092	0.12	0.433	0.07
S12	-0.040	-0.23	-1.217	0.14	-0.212	-0.04
Y61	3.398	0.43	14.555	1.09	4.784	0.65
Y62	0.174	0.02	-2.334	-0.18	1.265	0.17
Y63	0.575	0.08	0.596	0.05	7.405	1.00
Y64	8.297	1.08	5.351	0.41	5.821	0.79
Y65	5.725	0.74	8.284	0.64	3.130	0.43
Y66	1.857	0.24	2.149	0.17	8.260	1.12
Y67	5.415	0.70	0.822	0.06	0.848	0.12
Y68	-1.284	-0.17	-2.290	-0.18	0.195	0.03
Y69	-1.698	-0.22	-2.457	-0.19	0.559	0.08
Y70	-0.457	-0.06	-1.080	-0.01	6.147	0.85
Y71	6.382	0.84	0.933	0.07	6.217	0.87
Y72	9.349	0.69	1.269	0.10	5.624	0.78
Y73	9.349	1.23	54.906	4.22**	67.331	9.55**
Y74	78.718	10.52**	60.334	4.72**	41.415	5.88**
Y75	39.225	5.09**	34.497	2.65**	36.382	5.16**
Y76	29.311	3.92**	21.897	1.78	28.544	4.00**
Y77	24.659	3.30**	31.408	2.50*	35.089	3.92**
Y78	29.571	4.03**	30.381	2.47*	17.824	2.50*
Y79	15.604	2.06*	12.048	0.97	15.517	2.14*
Y80	15.729	2.13*	8.505	0.69	20.386	2.89**
Y81	23.268	3.14*	18.804	1.46	17.982	2.55*
Y82	13.047	1.76	3.488	0.28	9.926	1.41
Y83	9.958	1.36	---	---	---	---

[a] For the March, July, and November models, R^2=.59, .34, and .56; Durbin-Watson=1.00, 1.15, and 1.24; and number of observations=287, 275, and 281, respectively.

[b] TTM=time to maturity in months; S2-S12 are binary variables for February-December, respectively; and Y61-Y83 are binary variables for 1961-1983, respectively.

[c] Significance levels for hypothesis of zero parameter are: **=1%; *=5%.

are most appropriate for the July contract and perhaps reflect uncertainty emanating from the export market.

Pricing Options on Soybean Futures

Results reported above indicate that, if log-price differences are generated by a normal distribution, there is a systematic change in the variance of the distribution--particularly during the growing season. This suggests that traditional option pricing models in which the underlying commodity price is from a log-normal distribution should be adjusted to account for a changing variance.

The formula for calculating premia under Black's option pricing model contains the overall variance term, $\sigma^2 t$, where t is the time to option expiration and σ^2 is the instantaneous variance assumed known and constant. If, theoretically, the instantaneous variance is a known function of time, $\sigma^2(y)$, then the overall variance term should be replaced with $U = \int_{y'}^{y^*} \sigma^2(y)dy$ (Ingersoll, p. 112), where y' is the time at which the option is priced and y* is the time of expiration ($y^*-y'=t$). Whether constant or a function of time, the instantaneous variance must of course be estimated. Thus one of the model assumptions (known variance) is always broken when the pricing formula is used in practice. Estimating σ^2 as a function of time is an attempt to find a better estimate for the overall variance and has received virtually no attention in past work on options because there is little theoretical justification for this type of systematic change in relation to stocks and other underlying commodities on which options have recently been traded. To exemplify the errors in option pricing caused by a non-constant variance in soybean futures, we define the case in which put options on November soybeans are being priced at the end of April. Thus σ^2 must be forecasted over the May-October period.

For this example, three types of forecasting models are used. The first model (MOD I) is an adaptation of (1) in which individual year-effect dummy variables are dropped so that a year-effect coefficient does not have to be estimated for the forecasting period; one binary variable for the period 1973-1975 is added on the basis of model performance; and the sample variance is lagged one month to represent current year effects. The second model (MOD II) is an autoregressive integrated moving average (ARIMA) model based on past variances specified as $(0,1,3) \times (0,0,1)_7$. The third forecasting model (MOD III) is a naive model which assumes that the current sample variance (April's variance) is the best variance forecast for any future month between April and November.

Variance forecasts for each month within the May-October period, inclusively, are made for 1976-1982, inclusively. The coefficients for MOD I and MOD II are estimated from observations during the period of 1960 to April of the year being considered. Thus, since seven years are considered, seven sets of coefficients for both MOD I and MOD II are estimated. For the naive model, a variance estimate from April's observations is calculated for each year. When using MOD I for forecasting, April's variance is used as the lagged variable to forecast May's variance; the forecasted May variance is then used as the lagged variable to forecast June's variance; and so on.

For the non-constant variance scenarios, it is assumed that variance changes from month to month are continuous and linear and that a forecast is for mid-month. Linear functions connecting the mid-month forecasts are developed. The overall variance is calculated by integrating each of these functions over the mid-month to mid-month period and then summing the five

(May-June, June-July,..., September-October) integrals.[4]

The overall variances resulting from each forecasting method are compared to the actual[5] overall variances in Table 3. The root mean squared error (RMSE) values indicate that the MOD II model performs much better than the other two forecasting models on average; however, the MOD II forecasts are closer to the observed variance in only three of the seven years.

Another perspective is given when examining the premium estimates implied by the different variance estimates. Premia are calculated for at-the-money put options using closing futures prices on May 20 (or the first subsequent trading day) and interest rates of low-risk securities[6] for each year. Since the options are at the money and on futures, these premium estimates are also the premium estimates for call options. The results are in Table 4.

If the best premium estimate is defined as that being closest to the premium calculated using the "actual" variance, then MOD I performs best in one of the seven years, MOD II performs best in three years, and MOD III performs best in two years (both MOD I and III yield the best premium forecast in the remaining year). We conclude that, in general, MOD II performs better than the other models because, except for 1978, the "errors" of its resulting premium estimates are not extreme when compared to the errors of the other models.

These results suggest that the use of models incorporating systematic variance changes will improve premium estimates. However, there is certainly a large need for more work in the area of forecasting these systematic changes. This work may involve the use of composite models and, in practice, will undoubtedly involve the development of criteria for more subjective judgement.

Table 3. Actual and Forecasted Variances (times 10^5) for May-October.

	Actual	Forecasts MOD I	MOD II	MOD III
1976	4538.958	1128.438	4477.708	670.833
1977	5453.542	2385.625	4243.229	4354.792
1978	1832.188	1962.708	4067.500	2642.083
1979	2887.500	1611.979	2816.875	507.813
1980	3135.000	1749.063	2282.396	1175.104
1981	1607.708	2166.250	2651.354	2206.250
1982	1237.396	1781.875	1963.333	1094.479
RMSE		1898.006	1110.571	1953.325

Table 4. Calculated Option Premia Using Actual and Forecasted Variances, in Cents Per Bushel.

Year	Futures Price ($)	Exercise Price ($)	Annualized Interest Rate (%)	Method of Variance Estimation: Actual	MOD I	MOD II	MOD III
1976	5.53	5.53	6.40	44.0	22.0	43.7	16.9
1977	7.22	7.22	6.53	62.7	41.6	55.4	56.2
1978	6.30	6.30	8.15	31.3	32.4	46.6	37.6
1979	7.28	7.28	10.37	44.4	33.2	43.8	18.6
1980	6.55	6.55	11.10	41.3	30.9	35.3	25.3
1981	7.85	7.85	14.08	34.4	40.0	43.4	40.0
1982	6.81	6.81	10.69	27.1	32.5	34.1	25.5

Footnotes

[1] The log-normal distribution is traditionally used for continuous-price formulae because it implies that the underlying commodity price cannot be less than zero and because a closed form solution can be found under the risk-free hedge conditions.

[2] Price data were obtained from tapes provided by the Chicago Board of Trade Foundation, the Commodity Futures Trading Commission, and MJK Associates, Inc.

[3] Observations during a contract's expiration month were not used because options on futures will often expire close to or during the preceding month.

[4] The use of these functions illustrates the point that, theoretically, the variance changes must be continuous for the model to hold. However, since a linear relationship is assumed, the overall variance found when calculated by using the simple average of the five variance estiamtes as a constant instantaneous variance is only slightly different than the overall variance found when summing the integrals.

[5] "Actual" in the sense that each month's variance during the forecasting period is computed and, under the same assumptions of continuity and linearity as stated above, the overall variance is computed.

[6] For each year during 1976-1980, intermediate credit bank loan rates are used (U.S. Department of Agriculture); rates for U.S. Governmental three month bills (Bureau of Economic Analysis) are used for 1981 and 1982. These rates were chosen because they were on low-risk instruments and readily available. It can be shown that relatively large changes in interest rates, given the range of other parameters used in this example, cause very small changes in option premia.

References

Anderson, R. W. "The Determinants of the Volatility of Futures Prices." Working Paper Series #CSFM-33, Center for the Study of Futures Markets Columbia University, New York, December 1981, revised June 1982.

Anderson, R. W., and J. P. Danthine. "The Time Pattern of Hedging and the Volatility of Futures Prices." Working Paper Series #CSFM-7, Center for the Study of Futures Markets, Columbia University, New York, 1980.

Asay, M. R. "A Note on the Design of Commodity Option Contracts." The Journal of Futures Markets 2(1982):1-7.

Asay, M. R. "A Note on the Design of Commodity Option Contracts: A Reply." The Journal of Futures Markets 3(1983):335-338.

Black, F. "The Pricing of Commodity Contracts." Journal of Financial Economics 3(1976):167-79.

Black, F., and M. Scholes. "The Pricing of Options and Corporate Liabilities." Journal of Political Economy 81(1973):637-59.

Bureau of Economic Analysis. Survey of Current Business. Washington, D.C.: United States Department of Commerce, 63(October, 1983):14.

Cargill, T. F., and G. C. Rausser. "Temporal Price Behavior in Commodity Futures Markets." Journal of Finance 30(1975):1043-53.

Cargill, T. F., and C. G. Rausser. "Time and Frequency Domain Representation of Futures Prices as a Stochastic Process." Journal of American Statistical Association 67(1972):23-30.

Clark, P. K. "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices." Econometrica 41(1973):135-55.

Cootner, P. H. The Random Character of Stock Market Prices, Cambridge, MA: MIT Press, 1964.

Houthakker, H. S. "Systematic and Random Elements in Short-Term Price Movements." American Economic Review 51(1961):164-72.

Ingersoll, J. A. "A Theoretical and Empirical Investigation of the Dual Purpose Funds." Journal of Financial Economics 3(1976):83-123.

Kamara, A. "Issues in Futures Markets: A Survey." The Journal of Futures Markets 2(1982):261-94.

Labys, W. C., and C. W. J. Granger. Speculation, Hedging and Commodity Price Forecasts, Lexington, MA: D. C. Heath, 1970.

Leuthold, R. M. "Random Walk and Price Trends: The Live Cattle Futures Market." Journal of Finance 27(1972):879-89.

Mandelbrot, B. B. "The Variation of Certain Speculative Prices." Journal of Business 36(1963):394-419.

Mann, J. S., and R. G. Heifner. "The Distribution of Short Run Commodity Price Movements." USDA-ERS Technical Bulletin 1536, March 1976.

Merton, R. C. "Theory of Rational Option Pricing." The Bell Journal of Economics and Management Science 4(1973):141-83.

Neftci, S. N. "Some Econometric Problem in Using Daily Futures Price Data." Working Paper Series #CSFM-55, Center for the Study of Futures Markets, Columbia University, New York, March 1983.

Peterson, P. E., and R. M. Leuthold. "Using Mechanical Trading Systems to Evaluate the Weak Form Efficiency of Futures Markets." Southern Journal of Agricultural Economics (1982):147-51.

Rocca, L. H. "Time Series Analyis of Commodity Futures Prices." Unpublished Ph.D. thesis, University of California, Berkeley, 1969.

Samuelson, P. A. "Proof That Properly Anticipated Prices Fluctuate Randomly." Industrial Management Review 6(1965):41-9.

Smith, C. W., Jr. "Option Pricing: A Review." Journal of Financial Economics 3(1976):3-51.

Smidt, S. "A New Look at the Random Walk Hypothesis." Journal of Financial and Quantitative Analysis 3(1968):223-35.

Stevenson, R. A., and R. M. Bear. "Commodity Futures: Trends or Random Walks?" Journal of Finance 25(1970):65-81.

U. S. Department of Agriculture. Agricultural Statistics, 1981. Washington, D.C.: U. S. Government Printing Office, 481, 1982.