Staff Paper P85-28

August 1985

Staff Papers Series

PROBABILISTIC PRICE FORECASTS BASED ON COMMODITY OPTION VALUES

by

Robert P. King
Associate Professor

and

Paul L. Fackler
Research Assistant

Department of Agricultural and Applied Economics

University of Minnesota
Institute of Agriculture, Forestry and Home Economics
St. Paul, Minnesota 55108

Staff Paper P85-28 August 1985

PROBABILISTIC PRICE FORECASTS BASED ON COMMODITY OPTION VALUES*

by

Robert P. King
Associate Professor

and

Paul L. Fackler
Research Assistant

Department of Agricultural and Applied Economics
University of Minnesota

* Selected paper presented at the annual meeting of the American Agricultural Economics Association, Ames, Iowa, August 1985. This research was funded by the Minnesota Agricultural Experiment Station and by the Graduate School of the University of Minnesota.

The recent institution of agricultural commodity options trading has given participants in agricultural commodity markets a new and potentially valuable tool for risk management. Options allow market participants to insure against adverse price movements without giving up the opportunity to benefit from favorable price changes, and it is anticipated that they will be widely used by both buyers and sellers of agricultural commodities.

By expanding the range of risk management (and speculative) alternatives, commodity options markets benefit only those who choose to trade in them. As Gardner notes, however, the establishment of commodity options markets also creates a valuable source of public information about anticipated price movements:

> Just as futures prices generate information about expectations of commodity prices, an option's price generates information about expectations of the variability of commodity prices.
>
> (Gardner, p. 989)

Gardner goes on to show how a price volatility estimate can be derived from an option value. His derivation is based on Black's commodity options pricing model and assumes that commodity prices are lognormally distributed.

In this paper we present a new method for deriving information about commodity price distributions from option values. This approach is relatively easy to implement, and it imposes no restrictions on the form of the price distribution. It also yields a straightforward statistical test that can be used to evaluate the reliability of the price forecasts it generates.

In the sections which follow, we first briefly review Gardner's derivation of implied commodity price variance levels. We then derive our method for constructing non-parametric representations of commodity price

distributions based on options values. We then illustrate our approach, using recent commodity options trading data. Finally, we outline an approach for evaluating the reliability of options-based probability representations.

Volatility Assessments Based on Options Pricing Models

Options pricing theory has been an area of considerable interest in the finance literature, especially since the publication of equilibrium pricing models by Black and Scholes and by Merton in 1973. Assuming frictionless markets, stock price dynamics that can be described by a continuous time Gauss-Weiner process, and uniform investor assessments of the instantaneous price variance, Black and Scholes demonstrated that the value of a European call option depends only on its exercise price, the time to maturity, a risk-free interest rate, the current price of the security, and its instantaneous variance. All but the latter of these can be observed directly. Merton generalized these results to allow for a stochastic interest rate and the payment of dividends. Using dominance arguments, he also demonstrated that the value of an American call option is identical to that of a European call option, and he derived bounds on the difference between the values of American and European put options. Black later extended the theory of stock option pricing to the problem of pricing commodity options and showed that rational pricing formulas are quite similar in these two cases.

It was soon recognized that estimates of the variability of returns on a security could be derived from these option valuation models. Studies by Latané and Rindleman, by Trippi, and by Chiros and Manaster--all published in the late 1970s--explored procedures for deriving implied stock price standard deviations from stock option values and compared the reliability of

these estimates to volatility estimates based on historical data. At about the same time, Gardner showed how commodity option values can be used to derive estimates of commodity price variability--estimates that can be used to construct probabilistic commodity price forecasts.

Gardner's derivation of implied commodity price standard deviations begins with an expression that equates the value of a European call option to the discounted expected return associated with holding it to maturity (his equation 1). Assuming the commodity price on the date the option matures is lognormally distributed, he derives an expression (his equation 4) that equates the value of the option to a function of its exercise price, the time to maturity, the riskless interest rate, and the mean and variance of the commodity price on the date the option matures. If one assumes the expected value of the commodity price on the option maturity date is equal to the current futures price for the commodity, this equation can be shown to be identical to the commodity option valuation formula derived by Black (equation 16, p. 177). Given this expression and current values of the time to maturity, interest rate, option value, and current futures price, an estimate of the standard deviation of the commodity price can be derived. This information, combined with the assumption of lognormality, can, in turn, be used to construct any desired confidence interval for the future commodity price.

This approach, which parallels that used in the stock market studies, has two important limitations. First, it imposes the assumption that commodity prices are lognormally distributed. This may be an unrealistic assumption, especially in commodity markets influenced by government programs designed to reduce both up-side and down-side variability. When prices are near price support levels, the assumption of lognormality (which

implies positive skewness) may be reasonable. When prices are near price ceilings or perceived maximum levels, however, the assumption of log-normality may result in a serious misrepresentation of the price distribution. The second shortcoming of this approach is that it can yield a different standard deviation estimate for each exercise price. Chiros and Manaster found substantial differences in implied standard deviations across a range of stock option exercise prices and noted that the choice of an appropriate weighting scheme for determining an "average" standard deviation is an important and difficult problem in an applied setting.

Non-Parametric Options-Based Price Forecasts

In deriving non-parametric price forecasts based on options values, we assume, as Gardner does implicitly, that transactions costs and tax impacts associated with options trading can be ignored and that money can be borrowed or invested at a risk-free rate of return. Our derivation begins with the following expression for $V_c(P^*,t)$, the value of a call option with exercise price P^* that expires in t time periods:

$$(1) \qquad V_c(P^*,t) = e^{-rt}\int_{p^*}^{\infty} f(P_t)(P_t-P^*)dP_t.$$

where r is the risk free rate of return, P_t is the commodity price on the option expiration date, and $f(P_t)$ is its probability density function. This equates the value of the option to the present value of its expected return and is identical to Gardner's equation (1). It reflects the fact that the return on a call option is zero when $P_t < P^*$ and is P_t-P^* when $P_t > P^*$. We will show that directly observed option values at several exercise prices, together with t and r, can be used to construct a representation of $f(P_t)$. If commodity options are an actuarially fair form of price insurance, which we will assume as a working hypothesis, this will be an accurate

representation of the distribution of P_t conditioned on currently available information.

Given the general properties of a probability density function, (1) can be integrated without substituting a specific expression for $f(P_t)$. The result is:

$$(2) \qquad V_c(P^*,t) = e^{-rt}\int_{p*}^{\infty}(1-F(P_t))dP_t,$$

where $F(P_t)$ is the cumulative distribution function (CDF) of P_t and all other variables are defined as above. A more complete derivation is presented in the Mathematical Appendix. From (2), it follows that

$$(3) \qquad V_c(P_1^*,t) - V_c(P_2^*,t) = e^{-rt}\int_{P_1^*}^{P_2^*}(1-F(P_t))dP_t,$$

where $P*_2$ and $P*_1$ are exercise prices such that $P*_2 > P*_1$. Applying the mean value theorem for integrals to the right-hand side of (3) and recognizing that $F(P_t)$ is monotonic, there exists a unique exercise price, $P**$, such that:

$$(4) \qquad e^{-rt}\int_{P_1^*}^{P_2^*}(1-F(P_t))dP_t = e^{-rt}(P_2^*-P_1^*)(1-F(P**)).$$

For sufficiently small differences between $P*_1$ and $P*_2$, $P**$ can be approximated by $(P_1{}^*+P_2{}^*)/2$. Combining (3) and (4) under these conditions and rearranging terms yields the following expression:

$$(5) \qquad F((P_1^*+P_2^*)/2) = 1 - [(V(P_1^*,t) - V(P_2^*,t))/e^{-rt}(P_2^*-P_1^*)]$$

Given the values of two call options that expire in t periods, their respective exercise prices, and an appropriate interest rate, then, equation (5) can be used to approximate one point on the CDF for P_t.

If this information is available for a number of exercise prices, a rough approximation of the entire CDF can be constructed by linearly interpolating between known points. Because no assumptions have been made about the form of the distribution of P_t, this CDF is a non-parametric representation of that distribution. It can be skewed in either direction and its expected value need not equal the current futures contract price. This overcomes one of the limitations of earlier price distribution representations based on option pricing models. Furthermore, because information about option values at all active exercise price levels is incorporated into a single representation, this method does not yield conflicting implied standard deviation estimates. Finally, from a practical standpoint, the calculations needed to construct such a non-parametric representation of the price distribution are simple, and the representation they yield is in a form that can be easily used by decision makers.

A similar set of relationships can be derived for European put options. Again, equating an option value to the present value of its expected returns, the current value of a European put option with exercise price P* and a maturity date t periods in the future, $V_p(P^*,t)$, is given by the following expression:

$$(6) \qquad V_p(P^*,t) = e^{-rt}\int_0^{P^*} f(P_t)(P^*-P_t)\,dP_t,$$

where P_t is, again, the price of the commodity when the option expires and $f(P_t)$ is the probability density function of P_t. This reflects the fact that the value of a put option is P^*-P_t when $P_t < P^*$ and zero when $P_t > P^*$. Integrating by parts yields the following expression:

$$(7) \qquad V_p(P^*,t) = e^{-rt}\int_0^{P^*} F(P_t)\,dP_t.$$

A more complete derivation of this expression is given in the Mathematical Appendix. Following the steps used to derive equations (3), (4), and (5) yields:

$$(8) \quad F((P_1^*+P_2^*)/2) = e^{-rt}(V_p(P_2^*,t) - V_p(P_1^*,t))/(P_2^*-P_1^*).$$

In a strict sense, this expression holds only for European put options, since, as demonstrated by Merton, the values of European and American put options are not equivalent. If the difference between European and American put option values is relatively constant for adjacent exercise price levels, however, equation (8) should yield an adequate approximation of a point on the CDF for P_t even when it is based on American put option values.

Since active exercise price levels for put and call options typically coincide under current trading rules for commodity options, it is possible to construct a second approximation of the commodity price CDF using put option values. Values on this CDF should correspond closely to those for comparable price levels on the CDF based on call option values. In addition, since the range of active exercise price levels may differ for put and call options, it will often be possible to combine CDF values based on the two sets of option values to construct a more detailed approximation of the price CDF.

Empirical Examples

Option values for May soybeans, June live cattle, and June treasury bonds and the CDF values derived from them are presented in Table 1. The option values are from the close of trading on Friday, February 22, 1985. The risk-free interest rates in Table 1 are based on certificate of deposit interest rates reported for February 22, 1985, in the February 25, 1985, edition of the Wall Street Journal. The CDF values clearly conform to

Table 1. Non-parametric CDFs for Three Futures Price Distributions

Contract	Contract Settle	Interest Rate[a]	Strike Price	Call Value	Put Value	Price $(P^*_2+P^*_1)/2$	CDF (Call)	CDF (Put)	CDF (Combined)[b]
May Soybeans	586 1/4	8.85%	550	39	3				
		t=.142	575	19	9	562.5	.190	.237	.214
		(.988)	600	8 1/4	22	587.5	.565	.514	.540
			625	2 7/8	41 1/2	612.5	.782	.771	.776
			650	7/8	64	637.5	.919	.889	.904
			675	1/4	88 1/2	662.5	.975	.968	.972
June Cattle	67.37	8.97%	62	--	0.20				
		t=.260	64	--	0.40	63	--	.098	.098
		(.977)	66	2.35	0.97	65	--	.278	.278
			68	1.35	1.90	67	.488	.454	.471
			70	0.65	3.22	69	.642	.645	.644
			72	0.30	4.80	71	.821	.772	.796
June T-Bonds	68-27	8.97%	64	4-58	0-15				
		t=.260	66	3-20	0-35	65	.184	.153	.169
		(.977)	68	2-01	1-11	67	.336	.305	.321
			70	1-00	2-10	69	.480	.481	.480
			72	0-31	3-38	71	.736	.702	.719
			74	0-13	5-24	73	.856	.870	.863

a/ The value of e^{-rt} is in parentheses, with t being the difference between the first business day following the end of option trading for the contract and February 22 divided by 365.

b/ This is the simple average of the put and call CDF values if both are available.

theoretical expectations--i.e. they are strictly increasing with price level, and they are always between zero and one. Of particular interest here is the close conformity between CDF values based on call and put options.

Graphical representations of the combined CDF's are presented in Figure 1. The tails of each distribution were constructed by linear extrapolation from known points on the CDF. Graphical representation of the probability density functions (PDF) implied by these CDF's are also presented in Figure 1. PDF values over a price range are simply the slope of the CDF.

Probabilistic statements about a futures price on its option maturity date are easily derived from the CDF graphs. For example, the following statements are based on the May soybean CDF. They could have been made on February 22, 1985, and they refer to P_t, the May contract futures price on April 15, 1985, the first business day following the end of trading on the May contract option.

(1) The probability that P_t will be below 586-1/4, the current contract price, is .54.

(2) The probability that P_t will be above 600 is .33.

(3) There is a .75 probability that P_t will be below 610-3/4.

Since new CDF's can be constructed each trading day, such probability statements can be regularly updated to reflect changes in market conditions.

The PDFs in Figure 1 are of interest because they make it easy to see the direction of skewness, if any, implied by the options-based CDF's. The implied soybean price distribution is clearly positively skewed. That for cattle is relatively symmetrical. In fact, the implied cattle price distribution is nearly uniform. Finally, the treasury bond price distribution is negatively skewed. It is interesting that this apparent

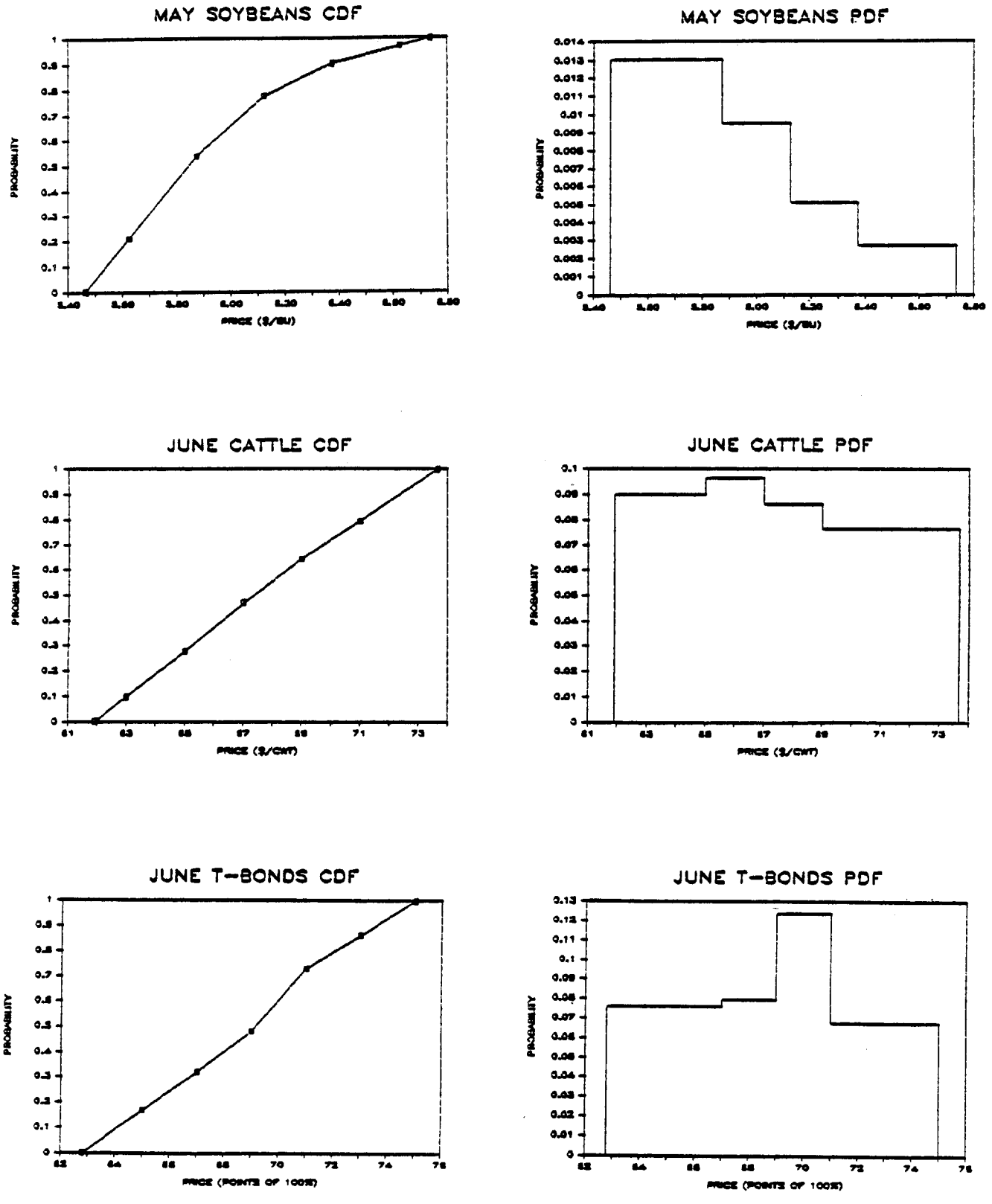

Figure 1. Graphical Representations of Price Distributions

contradiction to the assumption of lognormality occurs in a very active market where many traders would be expected to rely on standard rational option pricing models.

Evaluating Forecast Performance

If options-based probability distribution representations are to be of practical use, some assessment of their performance is needed. The assumptions underlying our derivation of non-parametric price forecasts are relatively unrestrictive. If they do not hold, however, forecasts based on this approach may be biased or may systematically overstate or understate price variability. In this section, we outline a procedure for evaluating the performance of options-based probabilistic forecasts, which will be implemented in a future study.

A CDF, $F(P_t)$, is a complete description of the probability distribution of P_t. Since P_t is a random variable, $F(P_t)$--the value of the CDF associated with P_t^a, the futures price actually realized on the date the option matures--is, in an ex ante sense, also a random variable. Given the properties of a CDF, if $F(P_t)$ is an accurate description of the distribution of P_t, $F(P_t^a)$ should be a uniformly distributed random variable defined on the interval [0,1].

Of course, there is only one realized price, P_t, for each option contract. While a new CDF can be constructed for each trading day, the values of $F(P_t^a)$ will not necessarily be independent. A random sample may be obtained by ensuring that the time periods between the trading and expiration dates are non-overlapping for each sample observation. Given such a sample, standard goodness of fit tests (Law and Kelton, pp. 192-204) can be used to test the null hypothesis that they were drawn from a [0,1] uniform distribution.

If the null hypothesis is rejected in such a test, the distribution of sample $F(P_t^a)$ values can be used to diagnose the shortcomings of options-based probabilistic forecasts. For example, if sample values of $F(P_t)$ are too tightly clustered around .5, the options-based CDF's overstate the actual level of price volatility. Conversely, if sample values of $F(P_t)$ are concentrated near zero and one, options-based CDF's understate price variability. Similarly, a concentration of sample values below or above .5 may indicate a positive or negative bias is options-based forecasts. If such patterns are observed, it may be possible to attribute them to income tax effects, transactions costs, or market inefficiencies.

MATHEMATICAL APPENDIX

To derive equation (2) from equation (1) note that, for any continuous, non-negative random variable, z,

$$E[z] = \int_0^\infty zf(z)dz = \int_0^\infty\int_0^z dyf(z)dz = \int_0^\infty\int_z^\infty f(z)dzdy = \int_0^\infty(1-F(y))dy.$$

Also used in the derivation is the following result, which makes use of integration by parts.

$$\int_0^{P*} f(P_t)(P*-P_t)dP_t = (P*-P_t)F(P_t)\Big|_0^{P*} + \int_0^{P*} F(P_t)dP_t = \int_0^{P*} F(P_t)dP_t$$

Note that this result demonstrates the equivalence of equation (6) and (7) as well.

Using these results, it takes only a little manipulation to show that

$$\int_{P*}^\infty (P_t-P*)f(P_t)dP_t = \int_0^\infty (P_t-P*)f(P_t)dP_t - \int_0^{P*}(P_t-P*)f(P_t)dP_t$$

$$= \int_0^\infty P_t f(P_t)dP_t - \int_0^{P*} dP_t + \int_0^{P*}(P*-P_t)f(P_t)dP_t$$

$$= \int_0^\infty (1-F(P_t))P_t - \int_0^{P*}(1-F(P_t))dP_t$$

$$= \int_{P*}^\infty (1-F(P_t))dP_t,$$

which, ignoring the e^{-rt} term, demonstrates the equivalence of equations (1) and (2).

REFERENCES

Black, F. "The Pricing of Commodity Contracts." Journal of Financial Economics. 3(1976): 167-179.

_____ and M. Scholes. "The Pricing of Options and Corporate Liabilities." Journal of Political Economy. 8(1973): 637-654.

Chiras, D.P. and S. Manaster. "The Information Content of Option Prices and a Test of Market Efficiency." Journal of Financial Economics. 6(1970):213-234.

Gardner, B.L. "Commodity Options for Agriculture." American Journal of Agricultural Economics. 59(1977):986-992.

Latané, H. and R. Rendleman, Jr. "Standard Deviation of Stock Price Ratios Implied in Option Prices." Journal of Finance. 31(1976): 369-381.

Law, A.M. and W.D. Kelton. Simulation Modeling and Analysis. New York: McGraw Hill Book Company, 1982.

Merton, R.C. "The Theory of Rational Option Pricing." Bell Journal of Economics and Management Science. 4(1973):141-183.

Trippi, R. "A Test of Market Efficiency Using a Random-Walk Valuation Model." Journal of Economics and Business 29:(1977): 93-98.