NCCC-134
APPLIED COMMODITY PRICE ANALYSIS, FORECASTING AND MARKET RISK MANAGEMENT

Simulating the Value of Directional Information in Futures Markets

by

Jeffrey H. Dorfman and Christopher S. McIntosh

Simulating the Value of Directional Information in Futures Markets

Jeffrey H. Dorfman and Christopher S. McIntosh*

Forecasts are often evaluated by either quantitative precision or qualitative reliability. However, consumers purchase forecasts for the potential utility gains from utilizing the forecasts, not for their accuracy. This is analogous to household production theory where goods are purchased for their derivative consumption services. Using Monte Carlo techniques to incorporate the temporal heteroscedasticity inherent in asset returns, the expected utility of qualitative forecasts are simulated. The associated monetary values for directional forecasts of various reliability levels are then derived. The method goes beyond normal forecast evaluation and allows forecast consumers to price the information value of a set of forecasts given their own utility function and trading system.

Introduction

In a recent article, Leitch and Tanner provide evidence suggesting that economists would be better off if they evaluated forecasts of economic time series using economic rather than statistical criteria. For example, they suggest that for financial asset prices, profits resulting from trading in the asset based on a set of forecasts would be a better guide to forecast performance than such statistical measures as mean squared error (MSE). Leitch and Tanner go on to show a lack of significant correlation between profits and root mean squared error (RMSE) in forecasts of T-bill rates. In fact, the correlation often has a perverse sign.

Several studies have done what Leitch and Tanner suggest and evaluated forecasts of financial asset prices by the profit earned in trading with the forecasts. Brandt and Bessler computed the net returns to hedging using seven different hog price forecasts. Lukac, Brorsen, and Irwin (1988a, 1988b), in a pair of related papers, compare the performance of twelve technical trading systems by monthly returns in simulated trading. None of these papers computes the correlation between profits and MSE, although Brandt and Bessler do discuss the differences in rankings of the seven models when ranked by profit versus MSE. Figlewski and Urich and Hein and Spudeck both compute profit-related measures for forecasts which they find to be unrelated to point forecast error measures.

Leitch and Tanner find that directional accuracy, the ability of a forecasting model to predict the upward or downward movement of the series, is significantly correlated with profits. This leads them to suggest this might be a better indicator of forecast quality than MSE. In this article, we set out to achieve three tasks. First, Leitch and Tanner's suggestion are taken one step further, leading to two metrics for evaluating forecasts based directly on

---

* The authors are assistant professors in the Department of Agricultural & Applied Economics, The University of Georgia. The authors thank Dave Bessler, Wade Brorsen, Jack Houston, William Tomek, Hal White, seminar participants at UC Davis, and attendees of NCR-134 for helpful comments.

economic criteria: profit and utility of profit. Second, the metrics are demonstrated with an application to futures markets trading. Third, the *ex post* correlation between profit, MSE, directional accruacy, and the two metrics proposed here are examined for comparative usefulness to forecast consumers.

In order to properly assess the value of forecasts of financial price series, the problem of the temporal heteroscedasticity of the price changes must be dealt with. That the variance of price changes is not constant for many financial assets is well-known and has spawned a large literature (a short list includes: Bollerslev; Bollerslev, Engle, and Woolridge; Engle, Lilien, and Robins; Fama; and Nelson. This temporal (or conditional) heteroscedasticity implies that <u>when a forecast system is correct</u> can have a larger influence on actual trading profits than <u>how often the forecasts are correct</u>. It also implies that the statistical expectation of profits from a trading system will not be of a standard form which can be easily handled analytically. To deal with this difficulty, numerical simulations will be employed to, in effect, integrate over the heteroscedasticity, arriving at a relation between profits and directional accuracy which accounts for the time-varying variance of the price changes.

The paper procedes by first describing a simulation method for computing the expected profit from trading in a financial asset market with a given forecasting system and set of trading rules. Next, the simulations are employed to develop the two forecast evaluation metrics. Five years of data on corn and soybean futures are then used to demonstrate the metrics. Finally, for several simple forecasting models, several common measures of forecast accuracy are computed along with the two new metrics and correlations between actual profit and the various criteria are computed. These results allow comparison with earlier findings by Brandt and Bessler, Figlewski and Urich, Hein and Spudeck, and Leitch and Tanner.

A Method for Profit Simulation

The calculation of the expected returns from trading are complicated by the temporal heteroscedasticity of asset price changes. Because the variance of price changes is not constant over time, analytical methods are not easily applied to the computation of expected profits. Instead, a simulation method was developed which integrates out the heteroscedasticity by randomly varying the distribution of correct and incorrect forecasts across the trading period.

The profit simulations performed in this paper are based on the corn and soybean harvest contract futures markets for 1984 through 1988. The harvest contract is the contract which matures closest to the crop's harvest date. In order to simulate the profit characteristics of a forecasting system with a given degree of directional accuracy, a set of trading rules must be specified. The rules used in this paper are quite simple and are similar to profit rule A in Leitch and Tanner. Given a forecast that the price will rise (fall), a long (short) position is taken by buying (selling) one contract. If already in the desired position, the current contract is held, thus minimizing transaction costs. Given a forecast of no-change, one remains in the current position. On the final day of the trading period, the position is closed out. Thus, except on the first and last day of the period, either no transaction takes place (hold pat) or the position is switched from long to short or vice versa.

The costs of trading were calculated assuming a reasonably large trader. Transaction costs of buying or selling were taken to be $50 per roundturn (one roundturn consists of both buying and selling one contract). The margin requirement was assumed to be 10 percent of the contract price, and the lost interest income from the margin requirement was calculated using an interest rate of 8 percent. Profits are accrued daily, since commodity futures contracts are repriced to the market each day. Profits for one year of daily trading (250 days) were calculated as discounted back to the first day of the year using the same 8 percent interest rate. Each contract represents 5000 bushels.

Given these trading rules and transaction costs, the simulation of expected profit can proceed. Using the actual daily closing prices, a set of forecasts is generated using a random number generator to distribute the correct and incorrect forecasts across the 250 trading days. Let f_t be the forecast direction of revision for the futures price on day t, i.e., $f_t = E[\text{sign}(p_t - p_{t-1})]$. Let u_t be a random variable generated from the uniform [0,1] distribution. Then, for a forecasting system with a directional accuracy of d, $0 \leq d \leq 1$, the simulated forecasts are generated by the rule

$$f_t = \text{sign}[(p_t - p_{t-1}) \times (d - u_t)].$$

This rule will generate forecasts which in the long run forecast the direction of revision (up or down) correctly d percent of the time. Such forecasts are artificial, but they serve the necessary purpose of allowing an expected profit to be calculated for forecasts with time-varying payoffs.

The simulations performed in this paper assume daily trading. A set of 250 f_t's ($t = 1, 2, \ldots, 250$) are generated to constitute a single year of simulated trading. This set of f_t's represents the up and down forecasts of a forecast system with a d percent reliability in predicting price change direction. It is used to calculate the discounted profit earned from one year of trading with the forecasts according to the trading rules outlined above.

The profit earned over one year is recorded, and the process is repeated for a given accuracy level d until reliable estimates of the average profit and the variance of profit are obtained. The expected profit from trading with a forecasting system with an accuracy of d percent is the sample mean of profits obtained in the simulation. That is, $E\pi(d) = (1/n)\Sigma\pi_i(d)$ where $i = 1, 2, \ldots, n$ represents the number of year-long simulations conducted for a given d. In application presented below, n was set to 5000. The variance of profit was estimated by the sample variance of the 5000 simulated profit values; thus, it represents the variability of actual trading profit around the mean due to the temporal distribution of returns, not a day-to-day measure of risk. After the average profit and variance of profit were calculated for a given d, d was changed and the whole process repeated. We simulated forecasts with directional accuracies ranging from d = 50 percent to d = 80 percent.

A Money Metric From Expected Utility

After completing the simulations, the relation between directional accuracy and expected profit (and variance of profit) has been established for a range of accuracy levels. While it is tempting to think that the profit obtainable using a set of forecasts is the same as the value of the forecasts, this is not the case. First, what if money could be made with a forecasting method which was correct 50 percent of the time? Clearly, one would only pay forecasters for the additional profit earnable above what could be made by simply flipping a coin. Further, the risk inherent in the trading process must be considered in the valuation process. Finally, we argue that it is the utility of profit which matters, not the profit itself, where the utility of profit accounts for the risks involved and the opportunity costs of partaking in the trading venture.

An Expected Utility Metric

For these reasons, an expected utility approach is taken to calculating the information value in a set of forecasts. The utility function used is the negative exponential, $U(\pi) = 1 - \exp(-\phi\pi)$, where ϕ is the Arrow-Pratt absolute risk aversion coefficient. This is a common utilty function for evaluating the utility of risky returns (dating back to Freund).[1] One advantage of this utility function is that the expected utility of a risky return is equal to a function of its mean and variance as long as the return is distributed normally. Maintaining that assumption results in the following relation for expected utility of profit,

$$EU(\pi) = 1 - \exp(-\phi[E(\pi) - \frac{\phi}{2}var(\pi)]). \tag{1}$$

If the average profit and variance of profit calculated from the simulations are inserted into (1), the result is the expected utility from using a set of forecasts with a given degree of accuracy. Denote this by EU(d) where d is the percent reliability of the forecasting system being evaluated.

Calculating values of EU(d) for a range of plausible directional accuracy levels provides a utility metric for judging a set of forecasts. Such a criterion is in line with the suggestion of Leitch and Tanner that forecasts be judged on a basis related to their inherent profit potential. In this sense, a utility metric is superior to a statistical metric such as MSE for evaluating economic forecasts. The superiority derives from choosing a metric which measures what the forecast user actually cares about (e.g., gains utility from). However, while an expected utility metric is an appealing way to measure forecast quality, it does not provide direct information on the value of information contained in the forecast set.

A Money Metric

A money metric is necessary to price the information contained in a set of forecasts. Yet such a metric should not depend solely on the profit potential of a forecasting system, since utility may not be a function of profit alone. What is needed is a money metric which is a transformation of the expected utility metric derived above. The transformation suggested here is based on a feature of the class of utility function employed. It would be appropriate for any utility function which depends on the moments of the profit distribution.

If one could purchase a set of forecasts (or forecasting system) which had an accuracy level of d = 100 percent--i.e., a perfect forecast--there would be no risk. The profit potential of this perfect forecasting system is certain (given the trading rules). The profit distribution is a single point at the

---

[1] Hal White has correctly suggested that many professional traders and account managers probably have utility functions which vary considerably from the one used here. The simple function chosen here is still a useful choice for the introduction of this procedure. We agree that more complex functions would be beneficial in applied situations.

maximum profit obtainable for the given trading rules. Because of the lack of risk in using such a forecast, the value of a perfect forecast should be equal to the profit obtainable from the forecasts, $\pi(d)$. A forecast user would pay up to the full profit potential to obtain such a forecasting system because no risk premium enters into the calculation. This fact fixes one point of the transformation between the expected utility metric and the money metric.[2]

To obtain the remaining values for the money metric for forecast accuracy from the expected utility metric, one employs the standard conditions for consumer utility maximization when purchasing two or more goods. Denote the money value of a set of forecasts which are correct d percent of the time by V(d). Then since the forecasts must be consumed in a discrete quantity (either you buy them or you do not), the expected marginal utility from "consuming" a set of forecasts is simply the expected utility from using the forecasts, MU(d) = EU(d). Therefore, for any accuracy level d, the value of the information contained in the set of forecasts is given implicitly by

$$\frac{V(d)}{MU(d)} = \frac{V(100)}{MU(100)}. \tag{2}$$

Solving for V(d), inserting EU(d) for MU(d), and noting that V(100) = $\pi(100)$ gives a simple rule for calculating the value of information for a set of qualitative forecasts which are correct d percent of the time,

$$V(d) = \frac{\pi(100)EU(d)}{EU(100)}. \tag{3}$$

Calculating V(d) across a range of values for d provides the money metric. The V(d)'s associated with several competing forecasting systems can be computed based on their historic accuracy levels. Instead of the forecasting system with the lowest MSE being declared superior, the V(d)'s would serve as the criterion by which the systems are judged. This provides a more logical basis to evaluate economic forecasts. Since the forecasts are intended to be used in an economic arena (investment and speculation), it makes sense to judge forecasts by an economic (and utility) based measure, rather than a statistical one.

The V(d) are not invariant to the set of trading rules employed in calculating the expected profits, a feature we believe is an advantage. In particular, the V(d) calculated in this paper depend on the trading rules outlined in Section I and the assumption of trading only a single contract. Multiple contract positions will produce different values of V(d), as will different methods for deciding when to go long or short. This accounts for the fact that the forecasts' value to the consumer changes based on how the information is utilized, following the Beckerian view of valuing commodities for the products derived or produced from the initially purchased good or service (Becker). In this sense, the forecasts are purchased in order to "produce" buy and sell decisions which then hopefully produce profit. While the dependence of the V(d) on an underlying set of trading rules makes comparison across studies somewhat more difficult than for MSE, a fairly standard set of single contract trading rules does not seem hard to achieve for academic comparison purposes.

---

[2] Since all expected utility functions are arbitrary up to a linear transformation, this is equivalent to choosing the scale of the utility function.

User specific trading rules would still remain important to correctly assessing individual applications.

A Demonstration of the Metrics

To demonstrate how a forecast consumer could use the utility-based money metric of section II in valuing forecasting systems with varying degrees of directional accuracy, an example using the futures markets for corn and soybeans is developed. While, the calculations are performed for the 1984 through 1988 harvest contracts, only 1988 results are discussed in this section.

The expected utility of forecasts with varying degrees of accuracy d were calculated from the average profit and variance of profit simulated in section I and the negative exponential utility function of section II. The risk aversion coefficient ϕ was varied across a range of plausible values to provide some sensitivity analysis to the results. The values of ϕ chosen are 0.0001, 0.00002, 0.00001, 0.000002, and 0.000001. Because the risk aversion coefficient can be interpreted as the inverse of the largest amount one is willing to lose (Pratt), the values of ϕ are all fairly small. They represent willingnesses to lose between \$10,000 and \$1,000,000. It is unlikely that anyone trading in the commodity futures markets is more risk averse than the low end of this range.

The expected utility of accuracy rates ranging from 50 to 80 percent and for 100 percent are shown in Tables 1 and 2 (for corn and soybeans, respectively) for the five different risk aversion coefficients. As can be seen, the utility of a forecasting system does not rise linearly with increases in qualitative accuracy, but shows a diminishing return to accuracy (as one would expect of any consumption good). For a constant d, expected utility decreases as ϕ decreases (the consumer becomes less risk averse); however, this is simply a result of the scaling of the utility function, not of any real effect.

The values of EU(d) from Tables 1 and 2 can be used in equation (3) to solve for the value of information in a set of forecasts, the V(d). These values are shown in Tables 3 and 4 for the same range of d and ϕ. Figures 1 and 2 show the V(d) for corn and soybeans for three values of ϕ: 0.0001, 0.00001, and 0.000001. The most important conclusion which can be drawn from these tables is that more risk averse consumers have a lower willingness-to-pay for forecasts at the low end of the accuracy scale (d close to 50 percent), but place a higher value on accurate forecasts (d close to 80 percent) than do less risk averse consumers. This makes intuitive sense since the more risk averse consumer is willing to pay a higher premium to avoid uncertainty. Purchasing a forecast lowers the consumer's risk, and a risk averse consumer will trade off the lower resulting net profit against the reduction in uncertainty. A consumer more tolerant of risk (with a small ϕ) has a certainty equivalent closer to the expected value outcome and will not pay as much to move towards the certainty equivalent. As expected, the more risk averse the forecast consumer, the more nonlinear is the relation between directional accuracy and forecast value.

Although the less risk averse consumer places a much lower value on the information in a forecast with an accuracy rate in the 60 to 80 percent range, the elasticity of value with respect to accuracy is much higher at d = 80 percent for these less risk averse traders. As the accuracy of the forecasting system approaches 100 percent, the less risk averse consumers' valuation is rising much more rapidly than that of the more risk averse. This is because the valuations for all levels of risk aversion must converge to the same point at d = 100.

Table 1. Expected Utility of 1988 Corn Forecasts

d[a]	ϕ = 0.0001	0.00002	0.00001	0.000002	0.000001
50	-1.0086	-0.1352	-0.0646	-0.0125	-0.0062
51	-0.8347	-0.1146	-0.0549	-0.0106	-0.0053
52	-0.6591	-0.0926	-0.0445	-0.0086	-0.0043
53	-0.5294	-0.0755	-0.0363	-0.0070	-0.0035
54	-0.4081	-0.0577	-0.0277	-0.0053	-0.0027
55	-0.2858	-0.0383	-0.0181	-0.0035	-0.0017
56	-0.1942	-0.0235	-0.0109	-0.0020	-0.0010
57	-0.0884	-0.0046	-0.0015	-0.0002	-0.0001
58	-0.0035	0.0116	0.0066	0.0014	0.0007
59	0.0764	0.0283	0.0150	0.0032	0.0016
60	0.1563	0.0452	0.0236	0.0049	0.0025
61	0.2233	0.0607	0.0316	0.0065	0.0033
62	0.2869	0.0771	0.0401	0.0083	0.0042
63	0.3444	0.0924	0.0480	0.0099	0.0050
64	0.4032	0.1089	0.0567	0.0117	0.0059
65	0.4478	0.1225	0.0639	0.0132	0.0067
66	0.4964	0.1382	0.0723	0.0150	0.0075
67	0.5414	0.1544	0.0811	0.0169	0.0085
68	0.5770	0.1675	0.0882	0.0184	0.0093
69	0.6082	0.1802	0.0952	0.0199	0.0100
70	0.6453	0.1960	0.1040	0.0218	0.0110
71	0.6773	0.2109	0.1123	0.0236	0.0119
72	0.6984	0.2215	0.1183	0.0250	0.0126
73	0.7266	0.2364	0.1267	0.0268	0.0135
74	0.7517	0.2507	0.1349	0.0287	0.0144
75	0.7680	0.2608	0.1408	0.0300	0.0151
76	0.7897	0.2749	0.1490	0.0318	0.0161
77	0.8077	0.2874	0.1563	0.0335	0.0169
78	0.8229	0.2989	0.1632	0.0351	0.0177
79	0.8390	0.3120	0.1710	0.0369	0.0186
80	0.8530	0.3242	0.1784	0.0386	0.0195
100	0.9752	0.5226	0.3090	0.0713	0.0363

[a] d is the directional accuracy, ϕ is the risk aversion coefficient.

Table 2. Expected Utility of 1988 Soybean Forecasts

d[a]	ϕ = 0.0001	0.00002	0.00001	0.000002	0.000001
50	-2.5287	-0.1616	-0.0709	-0.0128	-0.0063
51	-1.7891	-0.1084	-0.0461	-0.0080	-0.0039
52	-1.1133	-0.0474	-0.0168	-0.0023	-0.0011
53	-0.7209	-0.0024	0.0054	0.0021	0.0011
54	-0.3217	0.0476	0.0304	0.0072	0.0037
55	-0.0165	0.0980	0.0565	0.0126	0.0064
56	0.2235	0.1415	0.0794	0.0174	0.0088
57	0.3975	0.1869	0.1042	0.0228	0.0115
58	0.5328	0.2230	0.1240	0.0271	0.0137
59	0.6434	0.2648	0.1480	0.0325	0.0164
60	0.7157	0.2967	0.1666	0.0368	0.0186
61	0.7884	0.3357	0.1900	0.0422	0.0214
62	0.8402	0.3708	0.2116	0.0473	0.0240
63	0.8712	0.3965	0.2277	0.0513	0.0260
64	0.9006	0.4261	0.2468	0.0560	0.0285
65	0.9242	0.4580	0.2682	0.0614	0.0313
66	0.9411	0.4844	0.2862	0.0661	0.0337
67	0.9551	0.5096	0.3037	0.0707	0.0360
68	0.9652	0.5335	0.3209	0.0753	0.0384
69	0.9733	0.5571	0.3382	0.0801	0.0409
70	0.9799	0.5802	0.3556	0.0849	0.0435
71	0.9841	0.5998	0.3708	0.0893	0.0458
72	0.9879	0.6197	0.3865	0.0939	0.0481
73	0.9906	0.6390	0.4023	0.0986	0.0506
74	0.9929	0.6573	0.4176	0.1032	0.0531
75	0.9945	0.6727	0.4307	0.1072	0.0552
76	0.9959	0.6912	0.4470	0.1124	0.0579
77	0.9969	0.7067	0.4609	0.1169	0.0603
78	0.9976	0.7206	0.4738	0.1211	0.0626
79	0.9981	0.7344	0.4869	0.1255	0.0649
80	0.9985	0.7459	0.4980	0.1294	0.0670
100	1.0000	0.9099	0.6998	0.2139	0.1134

[a] d is the directional accuracy, ϕ is the risk aversion coefficient.

Table 3. Value of Information in 1988 Corn Forecasts

d[a]	ϕ = 0.0001	0.00002	0.00001	0.000002	0.000001
50	-38236.14	-9563.67	-7728.57	-6469.25	-6324.04
51	-31643.21	-8106.49	-6566.30	-5506.06	-5383.62
52	-24984.29	-6554.10	-5319.81	-4467.56	-4369.00
53	-20070.32	-5343.56	-4341.36	-3648.07	-3567.83
54	-15471.27	-4082.18	-3308.07	-2773.34	-2711.50
55	-10835.43	-2706.98	-2170.27	-1802.42	-1760.06
56	-7363.73	-1660.48	-1302.74	-1060.98	-1033.36
57	-3351.20	-324.27	-181.20	-92.88	-83.32
58	-131.52	821.51	788.53	749.77	744.30
59	2896.99	2000.06	1797.28	1634.26	1613.99
60	5926.34	3195.04	2821.76	2533.70	2498.53
61	8463.17	4295.34	3776.21	3379.54	3331.32
62	10876.28	5454.84	4795.02	4291.77	4230.64
63	13056.69	6533.46	5746.35	5146.18	5073.28
64	15283.77	7703.54	6785.98	6085.38	6000.22
65	16976.59	8664.85	7648.33	6870.40	6775.72
66	18817.56	9774.44	8650.96	7788.41	7683.27
67	20524.29	10925.07	9705.34	8764.67	8649.75
68	21872.96	11850.59	10555.19	9552.86	9430.21
69	23057.02	12747.59	11389.04	10333.83	10204.46
70	24462.87	13866.62	12435.62	11318.79	11181.54
71	25675.67	14920.83	13432.33	12264.98	12121.17
72	26475.94	15671.14	14148.42	12949.88	12801.96
73	27543.67	16725.27	15160.89	13923.22	13770.08
74	28497.03	17733.07	16136.95	14867.76	14710.33
75	29115.29	18450.00	16839.31	15553.70	15393.93
76	29936.76	19443.41	17817.53	16512.94	16350.39
77	30617.26	20333.27	18702.18	17387.06	17222.80
78	31196.07	21145.63	19516.82	18197.58	18032.43
79	31805.91	22071.96	20454.87	19138.19	18972.92
80	32337.63	22934.01	21334.80	20026.13	19861.46
100	36968.60	36968.60	36968.60	36968.60	36968.60

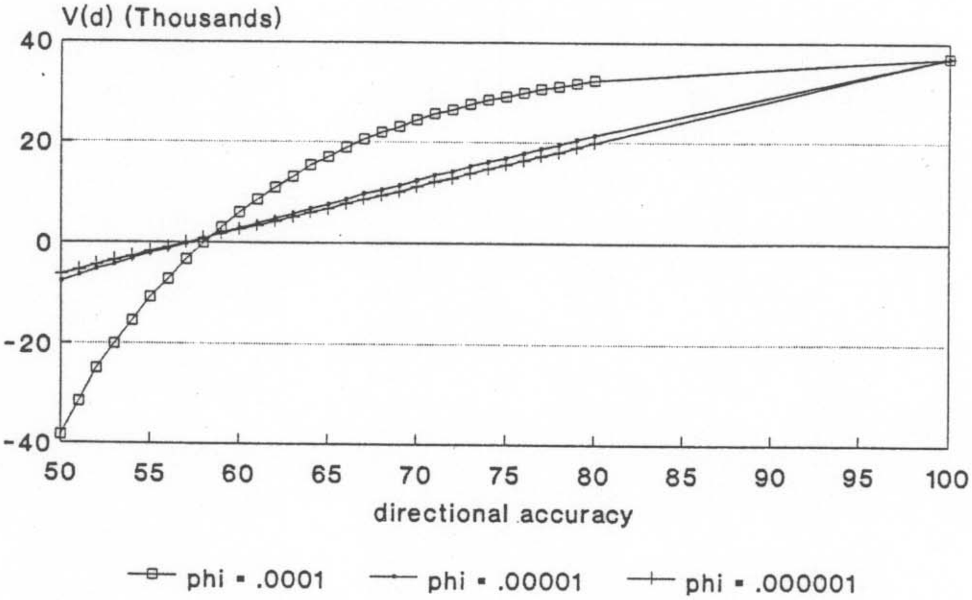
[a] d is the directional accuracy, ϕ is the risk aversion coefficient.

Table 4. Value of Information in 1988 Soybean Forecasts

d[a]	ϕ =	0.0001	0.00002	0.00001	0.000002	0.000001
50		-304277.46	-21376.87	-12195.04	-7179.15	-6682.40
51		-215290.31	-14338.90	-7929.12	-4515.36	-4182.81
52		-133969.01	-6264.00	-2891.81	-1296.44	-1153.59
53		-86745.89	-315.43	933.81	1208.78	1211.23
54		-38705.16	6291.11	5232.28	4050.23	3896.62
55		-1989.92	12958.03	9723.54	7102.92	6791.70
56		26888.58	18716.71	13644.30	9790.02	9342.67
57		47830.99	24720.05	17921.00	12829.24	12241.02
58		64115.15	29486.97	21321.79	15248.24	14548.15
59		77423.52	35018.13	25443.52	18284.00	17456.14
60		86116.01	39234.78	28647.11	20680.36	19756.09
61		94870.05	44398.95	32664.77	23742.87	22702.31
62		101107.25	49041.53	36378.97	26637.54	25494.75
63		104837.08	52435.58	39160.35	28846.91	27631.16
64		108374.66	56349.71	42444.81	31505.47	30207.95
65		111206.90	60566.54	46118.11	34569.26	33188.67
66		113244.80	64060.51	49218.64	37194.29	35747.34
67		114931.68	67388.25	52218.42	39765.36	38257.20
68		116142.81	70556.96	55174.58	42369.20	40807.72
69		117117.84	73670.55	58150.98	45042.26	43432.39
70		117913.72	76733.34	61143.73	47776.92	46123.33
71		118413.36	79318.85	63764.22	50242.38	48558.23
72		118877.90	81951.87	66460.22	52799.13	51085.76
73		119206.23	84509.38	69183.54	55463.84	53730.34
74		119473.20	86929.49	71808.94	58071.07	56322.64
75		119668.69	88969.28	74054.13	60325.64	58567.35
76		119836.02	91409.93	76853.77	63231.15	61472.09
77		119953.85	93454.19	79247.47	65756.16	64001.54
78		120040.90	95304.45	81465.38	68139.27	66394.29
79		120104.97	97117.40	83716.52	70627.05	68900.94
80		120153.29	98644.49	85636.75	72770.20	71062.97
100		120331.08	120331.08	120331.08	120331.08	120331.08

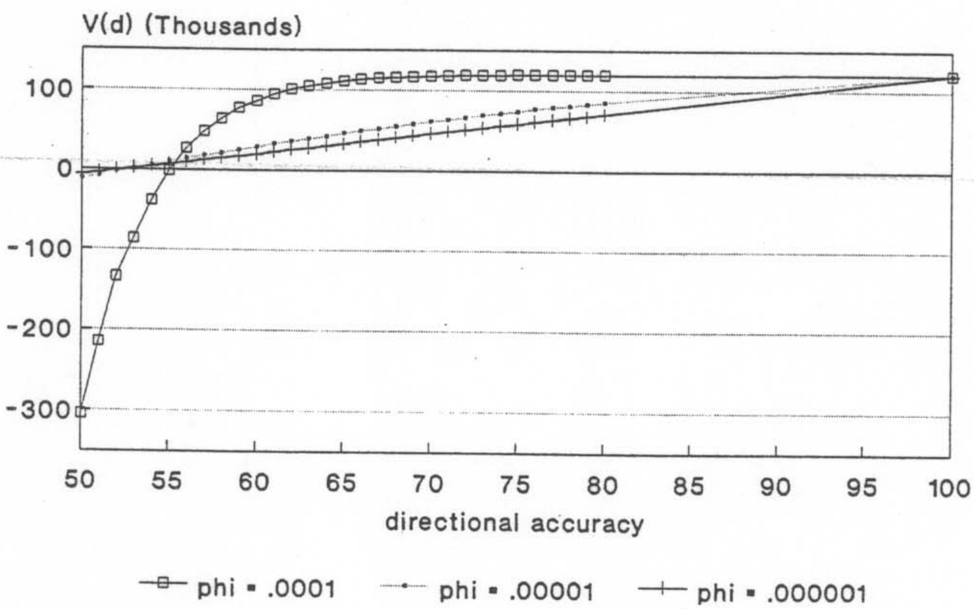
[a] d is the directional accuracy, ϕ is the risk aversion coefficient.

Figure 1. Value of Information V(d) for 1988 Corn Forecasts



phi is the Arrow-Pratt risk aversion coefficient.

Figure 2. Value of Information V(d) for 1988 Soybean Forecasts



phi is the Arrow-Pratt risk aversion coefficient.

Do the Metrics Work in Practice?

Following the guide of Leitch and Tanner, several simple forecasting models were selected and used to generate forecasts of soybean and corn futures prices for trading simulations. The models employed were a random walk with drift, an AR(1) model, an AR(3) model, and the "best" ARIMA model where "best" is used only to denote that autocorrelation profiles and statistical tests were employed to choose the model which seemed most appropriate.

For both corn and soybeans, all four models were first estimated using the data from 1984. The model specification of the two ARIMA models was fixed at the end of 1984. Forecast generation commenced at the start of 1985 and continued through the end of 1988 using the same data as in the above simulations. The coefficients of each model were updated as each new observation became available. In this manner, 1000 forecasts were produced by each of the eight models.

The forrecasts' mean squared errors, the percentage of correct directional forecasts, and the profit from trading with the forecasts under the rules outlined above were calculated on an annual basis. If directional accuracy is indeed important as Leitch and Tanner suggest, then the directional accuracy of the forecasting models should be correlated with the profit resulting from trading with them. Similarly, if MSE is not relevant, then there is no *a priori* reason for MSE to be correlated with profit. In fact, Leitch and Tanner find the correlation between RMSE (the square root of the MSE) and profit to have a perverse (positive) sign.[3] A positive correlation is perverse because a large MSE or RMSE is bad, indicating a lack of point predictive ability.

To analyze the two new metrics proposed here, the expected utility metric, EU(d), and the money metric, V(d), these values were also calculated for each crop and year using the respective directional accuracy and the simulated trading experiment results already discussed. That is, actual profit is not used in the formula, but rather the expected profit and variance of profit from the simulation for the appropriate accuracy level are employed. For the results in this section, only a risk aversion coefficient of $\phi = 0.00001$ was employed. This yields a total of four evaluation measures for each year-crop combination -- MSE, d, EU(d), and V(d). The correlations between these measures and the calculated profits are displayed in Table 5. These allow direct comparison to the results of Brandt and Bessler, Figlewski and Urich, Hein and Spudeck, and Leitch and Tanner. The individual values for profit, directional accuracy, MSE, EU(d), and V(d) used to compute the correlations are shown in Table 6.

With the data used here, somewhat different results are obtained from those of earlier researchers. First, directional accuracy is positively correlated with profit for corn and soybeans with the correlation being significant for soybeans and for corn and soybean toegether (see Table 5). This agrees with Leitch and Tanner's results and suggests that it would be useful to present directional accuracy measures along with goodness-of-fit statistics such as MSE.

---

[3] Fligewski and Urich and Hein and Spudeck also find no consistent relationship between profits and point accuracy. Brandt and Bessler, compare the net returns per period from hedging using seven different forecasting models of hog prices to the MSEs of the models. Their data shows a correlation between returns and MSE of -.479, which has the expected sign, but is not significant.

Table 5. Correlation of Evaluation Measures with Profit

Data	MSE	d	EU(d)	V(d)	n
soybeans	-.6239*	.7132*	.8542*	.8471*	16
corn	-.4058	.0403	.0365	.0755	16
both	-.4596*	.3626*	.6225*	.6362*	32

* Correlation is significantly different from zero at 5% significance level for a one-tailed test. Forecasts included are an AR(1), AR(3), the "best" ARIMA model, and a random walk with drift. Calculations are based on trading for the years 1985-88.

However, we also find MSE to be negatively correlated with profit (the correct sign). Again, the correlation coefficients are significant for soybeans and for corn and soybeans together, but not for corn alone (Table 5). This presents evidence opposite to the findings of Figlewski and Urich, Hein and Spudeck, and Leitch and Tanner. In our data, MSE is still apparently a useful guide for evaluating models which will be used in trading situations.[4]

What about the new metrics? Table 5 shows that the new measures are significantly correlated with profit for soybeans and corn and soybeans together, the same as MSE and directional accuracy. In fact, the correlation coefficients are larger than for either of the other two measures. This suggests that these metrics could be valuable tools in forecast evaluation because they relate directly to the relevant economic criterion -- utility of profit.

Conclusions

Two new metrics, based on the expected utility of profit, have been proposed for evaluating forecasts of economic time series. These metrics extend the suggestion of Leitch and Tanner to evaluate economic forecasts by economic, rather than statistical, criteria.

The metrics were demonstrated using futures markets data for corn and soybean over a five year period. The method of computer simulation necessary to compute the new metrics is somewhat computer intensive, but not unreasonably so, taking approximately three hours on a 486/33MHz computer. This seems a small time investment for a forecast consumer who will be risking thousands of dollars on the advice of a forecasting system.

[4] Although a random walk model without drift is not included in the results presented here, the results are not sensitive to this exclusion. MSE, EU(d), and V(d) are all still significantly correlated with profits for soybeans and for corn and soybeans together. Directional accuracy is not significantly correlated when a random walk model is included which is not unexpected since profit with a random walk model will be much higher than with a model which is right less than 50% of the time.

Table 6. Values used in the Calculation of Profit Correlations

Year	Profit	MSE	d	EU(d)	V(d)
			soybeans		
1985	-7266	0.00350	0.42970	-0.13969	-17084
	-5227	0.00349	0.48190	-0.08671	-10604
	6723	0.00344	0.55820	-0.00516	-630
	3710	0.00344	0.54620	-0.01487	-1818
1986	-2035	0.00272	0.46180	-0.10383	-12499
	-11218	0.00271	0.49800	-0.06481	-7802
	1546	0.00273	0.50600	-0.05530	-6657
	1940	0.00273	0.49000	-0.07484	-9008
1987	1695	0.00502	0.51810	-0.04295	-5403
	-182	0.00503	0.48590	-0.07739	-9736
	-1689	0.00505	0.46590	-0.10053	-12648
	-5143	0.00509	0.49400	-0.07739	-9736
1988	-19721	0.02346	0.42970	-0.27814	-47825
	-4020	0.02332	0.48190	-0.12844	-22084
	-6950	0.02326	0.48590	-0.09985	-17168
	-13891	0.02314	0.47390	-0.15804	-27174
			corn		
1985	-3179	0.00043	0.44180	-0.08456	-8925
	-7862	0.00041	0.54220	-0.04883	-5153
	1594	0.00042	0.51000	-0.05954	-6284
	-7088	0.00040	0.52210	-0.05588	-5897
1986	2797	0.00057	0.57030	-0.03483	-3713
	-4433	0.00056	0.51810	-0.05463	-5824
	2822	0.00056	0.57030	-0.03483	-3713
	-4184	0.00056	0.50600	-0.05847	-6234
1987	-237	0.00081	0.45780	-0.08300	-9037
	-7397	0.00083	0.55820	-0.03329	-3624
	-237	0.00081	0.45780	-0.08300	-9037
	-6832	0.00083	0.52210	-0.05291	-5761
1988	-103	0.00280	0.44980	-0.11090	-13266
	-7644	0.00267	0.53010	-0.03629	-4341
	-11579	0.00279	0.43370	-0.12983	-15529
	-8869	0.00266	0.54220	-0.02765	-3308

Models are in the same order for each year: AR(1), AR(3), random walk with drift, and ARIMA model in rows 1 through 4 respectively.

For a set of four simple forecasting models, the correlation between actual trading profits and the two new metrics was postive and significant for four years of daily data on corn and soybeans harvest contract futures prices. The correlation between profit and MSE was negative and significant, while profit and directional accuracy were significantly positively correlated. Thus, for these data, all four measures proved to be a reasonable measure of forecast quality.

However, only the money metric proposed here can tell a forecast consumer not only whether a forecasting system is worth buying, but also how much it is worth. The method proposed here allows a price to be placed on a set of forecast which is specific to the risk attitudes and trading characteristics of the individual forecast user (such as the number of contracts traded). This is a significant advantage of the money metric which allows it to be used in a variety of situations while still correctly valuing the forecasts based on how they will be used by the forecast purchaser.

References

Becker, G. S. "A Theory of the Allocation of Time." *Economic Journal* 75(1965):493-517.

Bollerslev, T. "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics* 31(1986):307-327.

Bollerslev, T., R. F. Engle, and J. M. Woolridge. "A Capital Asset Pricing Model with Time Varying Covariances." *Journal of Political Economy* 96(1988):116-131.

Brandt, J. A. and D. A. Bessler. "Price Forecasting and Evaluation: An Application to Agriculture." *Journal of Forecasting* 2(1983):237-248.

Engle, R. F., D. M. Lilien, and R. P. Robins. "Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model." *Econometrica* 55(1987):391-407.

Fama, E. F. "Forward and Spot Exchange Rates." *Journal of Monetary Economics* 14(1984):319-338.

Figlewski, S. and T. Urich. "Optimal Aggregation of Money Supply Forecasts: Accuracy, Profitability, and Market Efficiency." *Journal of Finance* 38(1983):695-710.

Freund, R. J. "The Introduction of Risk into a Programming Model." *Econometrica* 24(1956):253-263.

Hein, S. E., and R. E. Spudeck. "Forecasting the Daily Federal Funds Rate." *International Journal of Forecasting* 4(1988):581-591.

Leitch, G., and J. E. Tanner. "Economic Forecast Evaluation: Profits Versus the Conventional Error Measures." *American Economic Review* 81(1991):580-590.

Lukac, L. P., B. W. Brorsen, and S. H. Irwin. "A Test of Futures Market Disequilibrium Using Twelve Different Technical Trading Systems." *Applied Economics* 20(1988a):623-639.

Lukac, L. P., B. W. Brorsen, and S. H. Irwin. "Similarity of Computer Guided Technical Trading Systems." *Journal of Futures Markets* 8(1988b):1-13.

Nelson, D. B. "Conditional Heteroskedasticity in Asset Returns: A New Approach." *Econometrica* 59(1991):347-370.

Pratt, J. W. "Risk Aversion in the Small and in the Large." *Econometrica* 32(1964):122-136.