USDA Interval Forecasts of Corn and Soybean Prices:
Overconfidence or Rational Inaccuracy?

Olga Isengildina,

Scott H. Irwin,

and

Darrel L. Good [*]

Paper presented at the NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management
St. Louis, Missouri, April 21-22, 2003

[*] Olga Isengildina is a Post-Doctorate Research Associate (olga123@uiuc.edu) ; Scott H. Irwin and Darrel L. Good are Professors in the Department of Agricultural and Consumer Economics at the University of Illinois at Urbana-Champaign. The funding support of the Economic Research Service of the US Department of Agriculture under the Cooperative Agreement No. 43-3AEK-8-80106 is gratefully acknowledged. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the view of the U.S. Department of Agriculture.

USDA Interval Forecasts of Corn and Soybean Prices: Overconfidence or Rational Inaccuracy?

The USDA WASDE (World Agricultural Supply and Demand Estimates) price forecasts are published in the form of an interval, but typically analyzed as point estimates. Thus, all information about uncertainty imbedded in the forecast is ignored. The purpose of this paper is to evaluate the accuracy of WASDE price forecasts using methodology suitable for testing judgmental interval forecasts. Empirical analysis includes traditional statistical tests as well as an alternative behavioral evaluation (accuracy-informativeness tradeoff model). The results of the traditional analysis indicate overconfidence of WASDE price interval forecasts, while the results of the behavioral approach suggest rational inaccuracy.

Keywords: interval forecasts, overconfidence, rational inaccuracy, accuracy-informativeness tradeoff, WASDE.

Introduction

Agricultural prices are inherently unstable, primarily due to a combination of inelastic demand for food and production technology that is subject to the natural vagaries of weather, disease and pests. Volatility of agricultural prices causes many individuals to rely on forecasts in their decision-making. Numerous studies demonstrate that the economic value of agricultural forecasts is often substantial. For example, an investigation of the pricing performance of market advisory services in corn and soybeans over 1995-2000 revealed that revenues of producers that followed advisory service forecasts were, on average, $14/acre higher than the average revenues of producers (Irwin, Martines-Filho and Good).[1]

The need for agricultural forecasts has long been addressed by the U.S. Department of Agriculture (USDA), which has provided both quantity and price forecasts of agricultural commodities since the 1920s (Kunze). It is a commonly held belief of market participants and analysts that USDA forecasts function as the "benchmark" to which other private and public estimates are compared (e.g., Irwin, Gerlow, and Liu; Kastens, Schroeder and Plain). Because of their significance, USDA forecasts have been the subject of analytical scrutiny since the 1950s (Allen). Examples of the latest studies include investigations of the market impact of USDA forecasts (e.g,, Sumner and Mueller; McNew and Espinosa), informational content (e.g., Carter and Galopin) and accuracy (e.g. Bailey and Brorsen; Sanders and Manfredo). Previous studies analyze USDA price forecasts as point estimates, but this is not always the form in which the forecasts are published.

A prominent example of USDA forecasting efforts is the WASDE (World Agricultural Supply and Demand Estimates) program, which provides monthly forecasts for major crops, both for the US and the world. WASDE price forecasts (unlike all other WASDE estimates) are published in the form of an interval. For example, the September 2002

WASDE forecast for the 2002/03 marketing year average farm price of corn was $2.35-$2.75/bushel, and of soybeans, $5.15 -$6.05/bushel. As noted above, these interval forecasts have been reduced to a mid-point in previous empirical analyses. This conversion may cause a substantial loss of information, as interval forecasts include both the point estimate and the standard error of the forecast, thereby reflecting the uncertainty associated with the forecast. The more uncertain is the forecast, the wider the interval. The value of this information about forecast uncertainty is particularly important for decision-makers with different risk preferences. Rather than one possible outcome as in the point forecast, interval forecasts give a range of possible outcomes, thereby allowing for thorough contingency planning (Cristoffersen). Allen points out that providing "quantitative probability statements … would shift the focus from the point estimate, which will be wrong anyway, to the information content of the forecast" (p. 112). Evaluating such forecasts as point estimates ignores any information about uncertainty (or the information content in Allen's terms) associated with the forecast. Furthermore, when interval forecasts are analyzed as point estimates, it is typically assumed that the forecast corresponds to a midpoint of the range. However, O'Connor, Remus, and Griggs point out that "people generally estimate asymmetric confidence intervals where the forecast is not the midpoint of the estimated interval" (p.623). If this is the case, analysis of such forecasts based on the midpoint of the range may yield misleading results.

The need for probability and interval forecasting has been repeatedly expressed in the agricultural economics literature (e.g., Teigen and Bell; Timm; Bessler and Kling; Bessler). However, application and analysis of interval and probability forecasts has received very little attention. A few studies have been devoted to developing methods of constructing confidence intervals around point estimates (e.g., Prescott and Stengos; Bessler and Kling). Attempts to test such forecasts have been even less numerous (e.g., Bessler; Fackler and King), and have been limited to model-based or market-based interval forecasts. To the best of our knowledge, no study has investigated ex-ante judgmental interval forecasts, such as WASDE price forecasts.

Interval and probability forecasts are more common in areas outside agricultural economics. Prominent examples are weather forecasts and sports picks. These types of forecasts have been studied to a large extent in the disciplines of behavioral economics and psychology. Recently, there has been rising interest in interval forecasting among financial economists (e.g., Taylor; Meade and Islam). Application of this knowledge to agricultural forecasts, and WASDE interval price forecasts in particular, may provide a better assessment of the quality of such forecasts, aid in their interpretation by forecast recipients, and assist analysts who provide these forecasts.

The purpose of this paper is to evaluate the accuracy of WASDE interval price forecasts using methodology suitable for testing judgmental interval forecasts. All WASDE interval forecasts of corn and soybean prices for the 1985/86 through 2001/02 marketing years are examined. Corn and soybean price forecasts are of particular interest because these two crops account for about 80 percent of total U.S. grain production. Empirical analysis of the WASDE interval price forecasts includes accuracy and calibration tests.

Two confidence level benchmarks are identified: a 95 percent confidence level and a confidence level elicited from a small-scale survey of USDA analysts. Interval forecast accuracy is tested by computing hit rates, which reflect the proportion of time the forecast interval contained the final value. Calibration is tested by performing a t-test of whether published intervals are equal to either of the two benchmarks and by computing factors by which average forecast intervals should be widened in order to correspond to the two confidence level benchmarks. The last part of the paper explores alternative theories to explain any deviations from the benchmarks found in the empirical analysis.

The results of this study are expected to provide a fuller, more complete analysis of WASDE price forecasts by including analysis of uncertainty associated with interval forecasts. This insight has not been possible from the analysis of these forecasts as point estimates performed in previous studies. Interpretation of the results of this analysis based on alternative theories will provide additional insight about the possible motivations of forecast providers, which will further aid in understanding the quality and the value of WASDE price forecasts.

Data

The subjects of this investigation are corn and soybean interval price forecasts from USDA WASDE reports over the 1985/86 through 2001/02 marketing years. These forecasts are part of reports released monthly by the USDA, usually between the 9th and 12th of the month. The first price forecast for a marketing year is usually available in May preceding the US marketing year (September through August). Estimates usually are finalized by August after harvest (Figure 1).[2] Thus, 16 forecast updates of commodity prices generated in the WASDE forecasting cycle each marketing year are available for analysis. The total number of forecasts for each commodity is 272 (16 forecasts * 17 marketing years).

WASDE reports provide a commodity-by-commodity and country-by-country (selected countries) marketing year balance sheet of supply, consumption, and stocks. Supply for a particular crop year consists of carryover stocks from the previous year, production during the current year, and imports during the current year. Projections of consumption include domestic use, exports, and year-ending stocks. Prices are assumed to "tie" the two sides of balance sheets by rationing supply to competing uses. Vogel and Bange note that, "The process of forecasting price and balance sheet items is a complex one involving the interaction of expert judgement, commodity models, and in-depth research by Department analysts on key domestic and international issues" (p. 10). While this makes it clear that WASDE price forecasts are based on a variety of methods and information sources, it is nonetheless most appropriate to classify them as judgmental, or expert, forecasts.

Table 1 presents various descriptive statistics of WASDE interval price forecasts for corn and soybeans over 1985/86-2001/02. During the study period the average price of corn was $2.24/bushel and the average price of soybeans was $5.71/bushel. Prices were forecast in a form of an interval, with monthly price ranges averaging $0.38/bushel to $0/bushel for corn and $1.22/bushel to $0/bushel for soybeans. Price ranges were wider during the initial phases of the forecasting cycle and narrower after harvest, often converging to a point estimate in May after harvest. The maximum price range was 0.50 $/bu for corn and 2.25 $/bu for soybeans. The magnitude of price ranges between corn and soybean forecasts was comparable, with corn forecast price ranges averaging about 17 percent of the average forecast price prior to harvest and 8 percent of the average forecast price after harvest. Soybean forecast price ranges averaged 19 percent of the average forecast price before harvest and about 6 percent after harvest. No trends in the magnitude of ranges over the years were detected. No significant correlation between forecast errors and price ranges was detected. By definition, more uncertain forecasts have larger price ranges. Therefore, this finding means that uncertainty associated with price expectations was independent from the error structure of the forecast, suggesting that the outcome was affected by factors that could not be predicted.

Measures of Interval Accuracy

Numerous approaches to testing interval forecasts have been proposed in the forecasting and finance literature (e.g., Taylor; Baillie and Bollerslev). However, these time-series approaches analyze rolling-event forecasts that are not suitable to the analysis of fixed event forecasts, such as WASDE price forecasts. Rolling event forecasts provide a series of forecasts for events taking place one or more periods into the future (e.g., interest rates one year from the forecast date), thus each forecast describes a different event. Fixed event forecasts provide a series of forecasts of the same event in the future (e.g., 16 updates of WASDE price forecast for each marketing year). Other methods of interval forecast evaluation are aimed at testing forecasting methodology and are heavily based on forecast model specification (e.g., Chatfield). These obviously are not appropriate for non-model-based judgmental forecasts. This study concentrates on testing methods suitable for judgmental fixed-event interval forecasts. Specifically, accuracy tests of forecast coverage (percentage of times the forecast interval contained the final or "true" value) and forecast calibration (whether the forecast coverage corresponds to a stated confidence level) will be applied to the WASDE interval price forecasts for corn and soybeans.

A traditional measure of interval forecast coverage is the hit rate, which describes the proportion of times the forecast interval contained the final or "true" value (y_t). This measure indicates the "empirical confidence level" of the forecast. Another measure of forecast accuracy is forecast calibration. Interval forecasts are said to be calibrated if the proportion of times the forecast interval includes the true value corresponds to a stated confidence level. Since this study is not comparing WASDE forecasts to other forecasts of commodity prices, an absolute measure of forecast calibration is more appropriate than relative measures, such as the Brier score (Brier).

The analysis of WASDE forecast interval calibration included two steps. First, a standard t-test was used to compare the forecast confidence level with the underlying confidence level for a given announcement month:

$$t_t = \frac{\overline{X}_t - m}{\frac{\hat{s}_t}{\sqrt{N}}} \tag{2}$$

where $\overline{X}_t$ represented a mean of a binary series where 1=miss, 0=hit (equivalent to a "miss" rate) for month t, μ represented an ideal "miss" level (0.05 for 95 percent level and one minus implied value for an implied confidence level), $\hat{s}_t$ is the standard deviation of the hit-miss binary series for month t, and N is the number of observations included in the test (17).[3] Second, a factor by which the average forecast interval should be multiplied in order to achieve the underlying confidence level was computed. In order to achieve a 95 percent confidence level, 16 out of 17 intervals should include the final estimate. Therefore, the second largest distance from the interval should be added on both sides of the average interval in order to achieve this confidence level.[4]

Difficulty in assessing forecast calibration arises when no specific information on the confidence level associated with the forecast is given (as is the case with WASDE forecasts). In such cases, it is typically assumed that the forecast confidence level is fairly high, about 95-98 percent, meaning that only 2-5 percent of the time the observed values will fall outside the interval forecast. The validity of this assumption was investigated in an informal survey of USDA analysts described in the following section. Hence, two benchmarks for the confidence level were used in the empirical analysis: a 95 percent confidence level often assumed by forecast receivers, and an implied confidence level elicited from the survey of forecast providers.

Survey of Forecast Providers

In order to evaluate the common assumptions made regarding the confidence level of interval forecasts, an informal survey of USDA experts involved in compiling WASDE forecasts was conducted in August 2000. The survey was conducted via e-mail sent by an Economic Research Service (ERS) representative. The survey was sent to all ERS analysts and World Agricultural Outlook Board (WAOB) analysts involved in the WASDE corn and soybean forecasting process. The e-mail described the purpose of the survey and contained one question: "Each month, beginning in May prior to harvest, the World Agricultural Outlook Board presents a forecast of the marketing year weighted average price of corn and soybeans received by farmers. For each month, would you indicate the confidence, on average, that you have in the price forecast by indicating the percentage of time you think that the final price estimate for the marketing year will be within the forecast range presented that month." The response rate to this questionnaire was about 30 percent, which resulted in three complete responses for corn and four responses for soybeans. The respondents included the Chair of the Feed Grains committee at the WAOB, the Director of the Feed Grains and Oilseeds Division at the Farm Service Agency (FSA), the senior feed grains analyst at the Foreign Agricultural

Service (FAS), senior soybean analysts at the Economic Research Service (ERS), and the senior feed grains analyst at ERS.

According to the survey results summarized in Table 2, WASDE forecasters associate corn forecasts with an average of 77 to 97 percent levels of confidence and soybean forecasts with an average of 73 to 96 percent level of confidence. Only one respondent indicated a 95 level of confidence associated with price forecasts for each month they are published. Analyst responses differed by as much as 30 percent in the beginning of the season (65 vs 95 percent confidence level) and by as little as 5 percent close to the final forecast date (95 vs 100 percent confidence level). Average confidence levels prior to harvest were below 86 percent for both crops. Confidence levels improve after harvest to 88 to 97 percent for corn and 85 to 96 percent for soybeans. The forecast uncertainty after harvest is caused by the continued variability in the parameters involved in the forecasting process and the revisions of monthly prices published by NASS (National Agricultural Statistical Service), associated with changes in monthly marketing weights. This information indicates that the confidence levels associated with WASDE interval price forecasts were on average lower than a 95 percent benchmark. The average confidence levels generated by the survey are used in the following analysis as the confidence levels implied by the forecast providers.

Empirical Analysis of Interval Accuracy

The empirical analysis of WASDE interval forecasts of corn and soybean prices accuracy during 1985-2001 marketing years included coverage and calibration tests. The results of these tests are reported in Table 3. Coverage of WASDE forecasts was examined by computing hit rates. The hit rates of corn forecasts prior to harvest ranged from 35 to 65 percent, meaning that the published intervals contained the true value not more than 65 percent of the time. After harvest the hit rates of corn forecasts significantly improve, averaging about 83 percent. Hit rate statistics suggest that the published price interval width does not adequately reflect the uncertainty associated with the forecast. Thus, corn interval price forecasts are in general too narrow prior to harvest. The reduction of the forecast interval to the point estimate, which often happened in August after harvest, appears premature. August after harvest forecast of corn prices missed the final value 31 percent of the time.

The hit rates of soybean forecasts prior to harvest were much better, ranging from 76 to 82 percent. However, after harvest the situation is reversed, as the hit rates go down, to an average of about 73 percent. There seems to be no correlation of hit rates improving as more information becomes available, especially in soybeans. Soybean price interval price forecasts become too narrow after harvest. Soybean price interval forecasts are typically reduced to a point estimate by May after harvest, which appears to be three months too early according to hit rate statistics (37, 62, and 69 percent for May, June, and July respectively). Overall, the hit rates reveal that the accuracy of WASDE interval price forecasts for corn and soybeans generally is substantially lower than the 95 percent benchmark.

More insight on hit rate statistics can be gathered from Figures 2 and 3. Figure 2 describes the location of the forecast interval and the final price for corn in the beginning of the season (May prior to harvest), at harvest (October), and after harvest (March). Similar information is presented for soybeans in Figure 3. Figure 2 indicates that a May price interval forecast missed the actual price of corn 9 out of 17 times. It demonstrates that an October price interval forecast missed the actual price of corn 6 out of 17 times; and illustrates how forecast errors decline over the WASDE forecasting cycle. Price interval forecasts in March after harvest missed the actual price of corn 5 out of 17 times, mostly due to the fact that the forecast intervals were narrowed. Statistics for the hit rates for soybean prices are the same for May, October, and March. The price forecasts missed the actual price value 4 out of 17 times. Figure 3 illustrates variation in interval widths and error magnitudes for the forecasts of soybean prices in these months. It is interesting to point out the consistency of error for the 1987/88 marketing year for both corn and soybeans. This is not surprising, given that the severe drought of 1988 could not be forecast.

Calibration of WASDE forecasts was examined using a *t*-test of whether the published forecast intervals were equal to two benchmark confidence levels: a 95 percent confidence level and an implied confidence level elicited from the survey of forecast providers. The results of the calibration tests reported in Table 3 suggest that that prior to harvest the confidence levels of corn price interval forecasts were significantly different from both the implied and the 95 percent confidence levels. Thus, these forecasts were not calibrated. Calibration of corn forecasts improves substantially after harvest, when the confidence levels of corn price forecasts become statistically different from the implied and the 95 percent levels only in two months, March and August. These results suggest that the corn price interval forecasts were better calibrated after harvest. The August forecasts were often reduced to a point estimate, which did not adequately represent the final estimate. This resulted in the loss of calibration for August forecasts.

According to the results of soybean calibration tests, soybean price forecasts were not significantly different from the implied and the 95 percent confidence level prior to harvest. After harvest, soybean price forecasts become statistically different from both the implied and 95 percent confidence levels in May, June, and July, when these forecasts converge to a point estimate. These results suggest the soybean price forecasts were reduced to a point estimate about three month too soon. With this exception, soybean price interval forecasts appear well calibrated particularly with respect to the implied confidence level.

Calibration factors illustrate how much the published intervals should be widened in order to correspond to the two benchmark confidence levels. For example, the second largest distance from the interval for May forecasts is $0.32/bushel. This amount should be added on both sides of the $0.37/bushel average interval in order to achieve a 95 percent confidence level, which is equivalent to multiplying this average interval by the factor of 2.68. Table 4 demonstrates that calibration factors for corn price interval forecasts ranged from 1.54 to 2.36 for an implied confidence level and from 1.54 to 3.12

for the 95 percent confidence level prior to harvest. This means that the June prior to harvest corn price interval forecast should be at least 3.12 times wider in order to correspond to a 95 percent confidence level. After harvest, calibration factors became much smaller and approached two only in cases when the forecasts sometimes converged to a point estimate (May, June, and August). The calibration factors for soybean forecasts were much lower than those for corn forecasts and approached two only a few times at the 95 percent confidence level. On average, the calibration factors for soybean price interval forecasts after harvest appear to be higher than before harvest.

The results of the calibration analysis are represented graphically in Figures 4 and 5. These figures compare the average forecast intervals in respective months to the implied and 95 percent confidence levels. The average forecast price for the period of study is used as the midpoint of intervals for the graphs.[3] Average published forecast intervals are displayed symmetrically around the average forecast price for each announcement month. The implied and 95 percent confidence intervals are displayed in the same manner for comparison. The implied and 95 percent confidence intervals are computed using calibration factors reported in Table 4. Figure 4 illustrates that the first six months into the forecasting cycle (prior to harvest) corn price forecast intervals were much too narrow relative to the implied and the 95 percent confidence interval. Thus, corn price interval forecasts were $0.50/bushel too narrow for the implied confidence level in May and June and $0.78/bushel too narrow for the 95 percent confidence level in June and August prior to harvest. The situation dramatically improves from month seven (November after harvest) when the forecast intervals become close to the implied and the 95 percent confidence levels. Figure 5 demonstrates that soybean price interval forecasts match the implied confidence level fairly well throughout the forecasting cycle, but may be too narrow for the 95 percent confidence level. Soybean price interval forecasts would have to be widened by as much as $1.34/bushel in May before harvest in order to correspond to a 95 percent confidence level. The shape of the confidence intervals suggests that there was an additional increase in uncertainty in soybean price forecasts early after harvest. This jump in uncertainty was not picked up in the forecast intervals. Thus, soybean price interval forecasts would have to be widened by $1.06/bushel in November after harvest in order to correspond to a 95 percent confidence level. A premature reduction of the forecast to the point estimate in May, June and July after harvest is also illustrated. This is the main reason for the loss of calibration after harvest. Overall, both corn and soybean price interval forecasts are uncalibrated at the 95 percent confidence level, particularly prior to harvest. Soybean price interval forecasts appear better calibrated than corn price interval forecasts with respect to the implied confidence level.

The results of the empirical analysis demonstrated that WASDE price forecasts for corn and soybeans had relatively low hit rates. Calibration tests revealed that WASDE price forecasts were not calibrated at the implied and 95 percent level for corn prior to harvest and for soybeans after harvest. These results were not surprising and concurred with the findings of many previous studies. Previous empirical studies demonstrated very low hit rates associated with 90-98 percent confidence intervals in a wide variety of applications: 24-62 percent (Trip et al.), 53-81 percent (Alpert and Raiffa), 60 percent (Lichtenstein and Fischhoff), 20-58 percent (Russo and Schoemaker), 43-55 percent (Yaniv and Foster, 1997). Hit rates are usually low even among the experts: Russo and Schoemaker reported 42-61 percent hit rates for a job-relevant quiz of business managers.

A traditional approach to interpreting these results would be to suggest that forecasters produce uncalibrated forecasts due to cognitive bias, which characterizes the way they process information. Forecasts that are too narrow indicate that forecasters are overconfident. The issue of overconfidence has been studied extensively in the previous literature. Some studies demonstrate a positive relationship between forecasters' overconfidence and their experience (Yates, McDaniel, and Brown). Others suggest that overconfidence is based on ignorance of processing limitations (Pitz), or an anchor and adjust heuristic (Ferrell and McGoey). Based on this approach of interpreting poor calibration of probabilistic forecasts, USDA interval forecasts of corn and soybean prices appear overconfident. The importance of providing well-calibrated forecasts has been emphasized by Russo and Schoemaker among others. Therefore, a possible implication of the findings of this study is for USDA forecasters to improve their calibration abilities. The results of this study, which provide an estimate of these forecasters' historical performance, could be used as an aid.

However, overconfidence is not the only explanation for producing inaccurate forecasts. There is a growing body of literature that argues forecasters may intentionally provide inaccurate forecasts. A number of studies suggest that inaccurate forecasts may be a result of strategic behavior on the part of forecasters motivated by compensation structure, reputation concerns, customer base, etc. (e.g., Ehrbeck and Waldman; Ottaviani and Sorensen; Lamont; and Laster, Bennet and Geoum). These studies typically analyze rolling-event one period ahead macroeconomic forecasts against a set of forecaster characteristics. USDA price forecasts, on the other hand, are fixed-event forecasts compiled by a group of experts in the public sector; therefore these behavioral theories do not appear to be applicable in the present case.

A behavioral theory that may be particularly appealing in this case is motivated by the information provided by one of the respondents to the survey of forecast providers. This respondent declined a request to provide specific confidence levels and argued that because of all the uncertainty involved in the forecasting process an interval would have to be huge to correspond to even a 66 percent confidence level. This statement suggests that providing well calibrated forecasts may not be the primary goal of forecasters. This argument is fundamental to the behavioral theory developed by Yaniv and Foster (1995) who argued that poor calibration may be a result of rational behavior. This theory is reviewed and applied in the following section.

Accuracy-Informativeness Trade-Off Model

Yaniv and Foster's (1995) model results from their observations in the area of experimental psychology. Numerous studies have found that judgmental intervals are often poorly calibrated (overconfident). Rather than taking a normative approach to calibration accuracy and exploring corrective procedures for it, Yaniv and Foster attempted to interpret judgment under uncertainty as a part of the communication process between forecast "senders" and "receivers." They point out that "conversational norms suggest that judges should be appropriately informative as well as accurate" (Grice). This implies that excessively wide intervals (that are more likely to be accurate) are not preferred in communication terms. Therefore, Yaniv and Foster suggested that the evaluation of uncertain judgments involves a trade-off between two competing objectives: accuracy and informativeness. Experiments presented in their work demonstrate that sometimes people accept errors in the interest of securing more informative, or specific, judgments. Assuming that forecast producers respond to recipients' expectations, their objectives may be viewed as a continuous function of two dimensions: accuracy, expressed as a continuous measure of distance of an interval from the truth, and informativeness, measured as the specificity of the estimate. Following Yaniv and Foster, the tradeoff between accuracy and informativeness can be captured by a formal model of the form:

$$L = f\left[\left|\frac{t-m}{g}\right|, \ln(g)\right] \tag{3}$$

where *t* is the true outcome, *m* is the best point estimate, typically measured as the midpoint of an interval forecast, and *g* is the width of the interval. Yaniv and Foster assume that *f* is a monotonically increasing function of its arguments and propose an additive form:

$$L = f_1\left|\frac{t-m}{g}\right| + f_2\left[\ln(g)\right] \tag{4}$$

The tradeoff occurs because as interval width (*g*) increases, the first element $\left(\left|\frac{t-m}{g}\right|\right)$ describing accuracy decreases (improves), while the second element describing informativeness (ln(g)) increases (diminishes). Note that improvement in either direction would be indicated by a lower value of the element; thus a lower L score would indicate a forecast interval that is likely to be preferred by receivers. For simplicity, the individual functions f_1 and f_2 were substituted with the identity function and the coefficient a, respectively, resulting in a specific model form:

$$L = \frac{|t-m|}{g} + a\ \ln(g) \tag{5}$$

where a is a tradeoff parameter that reflects the weights placed on accuracy and informativeness of the estimates (a = 0). Experimental data suggested (Yaniv and Foster, 1995) that the value of a was close to one (from 0.6 to 1.2). The authors also demonstrated that the fit of their tradeoff model to respondents' rankings of interval forecasts was superior to that of alternative models.[5] The correlation of the model rankings to respondents' rankings was 84 percent.

The accuracy-informativeness tradeoff model (5) was applied to three different scenarios for WASDE forecasts: (1) actual WASDE price interval forecasts for corn and soybeans published from 1985/86-2001/02, (2) WASDE interval forecasts with ranges increased to correspond to 95 percent confidence levels, and (3) WASDE interval forecasts with ranges increased to correspond to implied confidence levels. In order to correspond to the 95 percent and the implied confidence levels, forecast intervals were multiplied by the respective calibration factors reported in Table 4. This analysis assumed that a=1, following Yaniv and Foster's (1995) experimental observations.[6] L scores were computed for each observation and then the average L score for each month was used for the final comparison.

The results of this analysis are reported in Table 5. The L scores based on actual forecast ranges (first scenario) were consistently lower for both crops before and after harvest. These scores are equivalent to the scores based on the implied confidence levels (third scenario) for the months when forecast ranges did not need to be increased to correspond to the implied confidence level (soybeans before harvest and corn early after harvest). The scores for the second scenario based on the 95 percent confidence level forecast ranges were consistently higher (worse) than the scores for other scenarios. These findings suggest that if the accuracy-informativeness tradeoff model accurately reflects the preferences of WASDE forecast receivers, the actual forecasts published by the USDA appear superior to well-calibrated forecasts. This implies that the public would prefer the combination of accuracy and informativeness associated with these forecasts to any gains in accuracy that would inevitably cause losses in informativeness. Hence, WASDE forecasts may be rationally inaccurate in response to the precision requirements of forecast receivers.

Summary and Conclusions

This study sought to examine the characteristics of WASDE price forecasts that are given in an interval form. WASDE interval forecasts of corn and soybean prices for the 1985/86 through 2001/02 marketing years are analyzed. Principal aspects of interval forecast quality include accuracy measured by hit rates and calibration measured by *t*-tests of whether published intervals are equal to the benchmark confidence levels, and average calibration factors. Two benchmarks for confidence level were identified: a 95 percent confidence level and an implied confidence level elicited from a small-scale survey of forecast providers. Hit rates for corn ranged from 35 to 65 percent prior to harvest and averaged 83 percent after harvest. Hit rates for soybeans ranged 76 to 82 percent prior to harvest and averaged 73 percent after harvest. The results of the t-test indicate that both corn and soybean price forecast intervals were not calibrated at the 95 percent confidence level. Corn price interval forecasts were statistically different from the implied confidence levels prior to harvest. Soybean price interval forecasts were significantly different from the implied confidence levels in May, June and July after harvest, when they were converged to a point estimate. Calibration factors illustrate that corn price interval forecasts would have had to be widened by as much as 2.36 times and 3.12 times to correspond to the implied and the 95 percent confidence levels, respectively. The highest calibration factors for soybean price interval forecasts were 1.50 and 2.33 on the implied and 95 percent confidence levels, respectively. Overall, this analysis revealed that WASDE price forecasts for corn and soybeans were not calibrated at the 95 percent level and at the implied confidence level for corn prior to harvest and for soybeans after harvest.

A traditional approach to interpreting these results would be to suggest that forecasters produce uncalibrated forecasts due to cognitive bias, which characterizes the way they process information. Forecasts that are too narrow indicate that forecasters are overconfident. Based on this approach, a possible implication of the findings of this study is for USDA forecasters to improve their calibration abilities. The results of this study, which provide an estimate of these forecasters' historical performance, could be used as an aid.

It was recognized, however, that widening of forecast intervals would cause a loss of specificity, or informativeness, in WASDE forecasts. There is a growing body of literature that argues forecasters may intentionally provide inaccurate forecasts. One such theory had been proposed by Ianiv and Foster (1995) based on their experiments in psychology. They suggest that the evaluation of uncertain judgments involves a trade-off between two competing objectives: accuracy and informativeness. Relative to the interval width, as accuracy improves, informativeness diminishes. Ianiv and Foster's tradeoff model was reviewed and applied to three alternative scenarios: first including the actual forecast intervals from WASDE price forecasts for corn and soybeans (1985-2001), second involving forecast intervals for these forecasts that would correspond to 95 percent confidence levels, and third including forecast intervals equivalent to the implied confidence intervals elicited from the survey of forecast producers.

An alternative interpretation of our empirical results based on Ianiv and Foster's model is that when the combination of accuracy and informativeness is considered, published forecast intervals from WASDE price forecasts appear superior to those of alternative intervals. These results imply that if the accuracy-informativeness tradeoff accurately describes the preferences of WASDE forecast receivers, the public would prefer the lack of accuracy associated with the current WASDE forecasts due to the informativeness that they provide. Or, in other words, the public would prefer the combination of accuracy and informativeness associated with these forecasts to any gains in accuracy (better calibration) that would inevitably make these forecast intervals less specific. Thus, we concluded that taking into account the combination of accuracy and informativeness, the USDA produces good forecasts that are superior to calibrated forecasts.

Regardless of either interpretation, overconfidence or rational inaccuracy, it is important to point out that confidence levels associated with these forecasts are not very high. It may be misleading for public to assume that published WASDE forecasts include the true outcome some high percentage of time (95 percent). Therefore, it may be recommended that approximate confidence levels should accompany the forecast intervals published by WASDE board. The information on historical confidence levels associated with WASDE forecasts from 1985/86 to 2001/02 contained in this paper may help identify these confidence levels.

Interestingly, similar recommendations were repeatedly proposed (e.g. Timm; Crowder) in 1960s and 1970s during the discussion of the adequacy of agricultural outlook programs in the U.S. The difficulty of providing such specific forecasts was also recognized by these authors. It was probably due to this difficulty that the recommendations were never taken into account. However, the level of confidence associated with any given forecast interval becomes very important when certain decisions are being made. While for some activities coarse approximate forecasts may suffice, certain investment decisions require precise knowledge of future prices. In these cases, failure to provide realistic confidence levels associated with price forecasts may incur significant social costs (Russo and Schoemaker). Therefore, it may be beneficial if USDA analysts provided forecasts accompanied by approximate confidence levels.

Finally, it is important to recognize the limitations of the behavioral approach investigated in this study. Yaniv and Forster's model was calibrated to the responses of the university students. Their preferences may not accurately represent the preferences of WASDE forecast receivers. Therefore, further research on preferences of WASDE forecast receivers and calibration of the applied model is needed.

References

Allen, G. P. "Economic Forecasting in Agriculture." *International Journal of Forecasting*, 10(994):81-135.

Alpert, M. and H. Raiffa. "A Progress Report on the Training of Probability Assessors", in D. Kahneman, P. Slovic, and A. Tversky (eds), *Judgment under uncertainty: Heuristics and Biases*, Cambridge University Press, New York, 1982, pp. 294-305.

Bailey, D.V. and B.W. Brorsen. "Trends in the Accuracy of USDA Production Forecasts for Beef and Pork." *Journal of Agricultural and Resource Economics*, 23(1998):515-526.

Baillie, R.T. and T. Bollerslev. "Prediction in Dynamic Models with Time-Dependent Conditional Variances." *Journal of Econometrics*, 51(1992):91-113.

Bessler, D.A. "Subjective Probability." in P.J. Barry (ed.) *Risk Management in Agriculture*. Iowa State University Press: Ames, Iowa, 1984; pp. 43-52.

Bessler, D. A. and J.L. Kling. "The Forecast and Policy Analysis." *American Journal of Agricultural Economics*, 71(1989):503-506.

Brier, G. W. "Verification of Forecasts Expressed in Terms of Probability." *Monthly Weather Review*, 75(1950):1-3.

Carter, C.A. and C.A. Galopin. "Informational Content of Government Hogs and Pigs Reports." *American Journal of Agricultural Economics*, 75(1993):711-718.

Chatfield, C. "Calculating Interval Forecasts." *Journal of Business and Economics Statistics*, 11(1993):121-135.

Christoffersen, P.F. "Evaluating Interval Forecasts." *International Economic Review*, 39(1998):841-62.

Crowder, R.T. "Statistical vs. Judgment and Audience Considerations in the Formulation and Use of Econometric Models." *American Journal of Agricultural Economics*, 54(1972):779-783.

Ehrbeck, T. and R. Waldman. "Why are Professional Forecasters Biased? Agency Versus Behavioral Explanations." *Quarterly Journal of Economics*, 111(1996):21-40.

Fackler, P.L. and R.P. King. "Calibration of Option-Based Probability Assessments in Agricultural Commodity Markets." *American Journal of Agricultural Economics*, 72(1990):73-83.

Ferrel, W.R. and McGoey, P.J. "A Model of Calibration for Subjective Probabilities." *Organizational Behavior and Human Performance*, 26(1980):32-53.

Grice, H.P. “Logic and Conversation”, in P. Cole and J.L. Morgan (eds), *Syntax and Semantics 3: Speech Acts*, Academic Press, New York, 1975, pp. 41-58.

Irwin, S.H., M.E. Gerlow, and T.R. Liu. “The Forecasting Performance of Livestock Futures Prices: A Comparison to USDA Expert Predictions.” *Journal of Futures Markets*, 14(1994):861-875.

Irwin, S.H., J.Martines-Filho and D.L. Good. “The Pricing Performance of Market Advisory Services in Corn and Soybeans Over 1995-2000.” AgMAS Project Research Report 2002-01, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, April 2002.

Kastens, T.L., T.C.Schroeder and R. Plain. “Evaluation of Extension and USDA Price and Production Forecasts.” *Journal of Agricultural and Resource Economics*, 23(1998):244-261.

Kunze, J. “The Bureau of Agricultural Economics’ Outlook Program in the 1920s as Pedagogical Device.” *Agricultural History*, 64(1990):252-261.

Lamont, O.A. “Macroeconomic Forecasts and Microeconomic Forecasters.” *Journal of Economic Behavior and Organization*, 48(2002):265-280.

Laster, D., P. Bennett, and I.S. Geoum. “Rational Bias in Macroeconomic Forecasts.” *Quaterly Journal of Economics*, 114(1999):293-318.

Lichtenstein, S. and B. Fischhoff. “Training for Calibration.” *Organizational Behavior and Human Performance*, 26(1980):149-171.

McNew, K. and J. Espinosa. “The Informational Content of USDA Crop Reports: Impacts on Uncertainty and Expectations in Grain Futures Markets. *The Journal of Futures Markets*, 14(1994):475-492.

Meade, N. and T. Islam. “Prediction Intervals for Growth Curve Forecasts.” *Journal of Forecasting*, 14(1995):413-430.

O’Connor, M., W. Remus, and K. Griggs. “The Asymmetry of Judgmental Confidence Intervals in Time Series Forecasting.” *International Journal of Forecasting*, 17(2001):623-633.

Ottaviani, M. and P.N. Sorensen. “The Strategy of Professional Forecasting.” Unpublished working paper, London Business School, London, England, February 2002.

Pitz, G. F. “Subjective Probability Distributions for Imperfectly Known Quantities.” in Gregg, L.W. (ed.), *Knowledge and Cognition*, Hillsdale, Erlbaum, New Jersey, 1974.

Prescott, D.M. and T. Stengos. "Bootstrapping Confidence Intervals: an Application to Forecasting the Supply of Pork." *American Journal of Agricultural Economics*, 9(1987):266-273.

Russo, J.E. and P.J. Schoemaker. "Managing Overconfidence." *Sloan Management Review*, 33(1992):7-17.

Sanders, D.R. and M.R. Manfredo. "USDA Production Forecasts for Pork, Beef, and Broilers: An Evaluation," *Journal of Agricultural and Resource Economics*, 27(2002):114-128.

Sumner, D.A. and R.A.E. Mueller. "Are Harvest Forecasts News? USDA Announcements and Futures Market Reactions." *American Journal of Agricultural Economics*, 71(1989):127-134.

Taylor, J.W. "Evaluating Volatility and Interval Forecasts." *Journal of Forecasting*, 18(1999):111-128.

Teigen, L.D. and T.M. Bell. "Confidence Intervals for Corn Price and Utilization Forecasts." *Agricultural Economic Research*, 30(1978):23-29.

Timm, T.R. "Proposals for Improvement of the Agricultural Outlook Program in the United States." *Journal of Farm Economics*, 48(1966):1179-1184.

Trip, G., R.B.M. Huirne and J.A. Renkema. "Price-Predicting Ability of Farm Managers: Empirical Findings with Flower Producers in the Netherlands." *Review of Agricultural Economics*, 22(2000):464-476.

Vogel, F.A. and G.A. Bange. "Understanding USDA Crop Forecasts." Miscellaneous Publication No. 1554, US Department of Agriculture, National Agricultural Statistics Service and Office of the Chief Economist, World Agricultural Outlook Board, 1999.

Yaniv, I. and D.P. Foster, "Graininess of Judgment Under Uncertainty: An Accuracy-Informativeness Trade-Off." *Journal of Experimental Psychology: General*. 124(1995):424-432.

Yaniv, I. and D.P. Foster, "Precision and Accuracy of Judgmental Estimation." *Journal of Behavioral Decision Making*, 10(1997):21-32.

Endnotes

[1] Calculated based on average prices received by farmers published by NASS and respective monthly marketing weights.

[2] The final WASDE price estimate for the forecasting period is published in November, but this estimate rarely changes after August. Because changes in the estimate are very small and infrequent August was chosen as the final forecast month for the purpose of this study.

[3] This test was not performed on a pooled data set because 1) monthly forecasts are associated with implied confidence levels that vary significantly from the beginning to the end of the forecasting period, 2) accuracy patterns across the forecasting cycle are of interest, 3) there is a potential heteroscedasticity problem in the hit-miss series across the forecasting cycle, 4) hit-miss observations are not independent from month to month because of the substantial overlap in the forecasting horizon.

[4] This analysis assumes that forecast intervals are symmetric. This assumption is consistent with a survey response which indicated that "Each month a midpoint is forecast using the U.S. and global supply and use and then a range is put on each side of the midpoint." Additionally, forecast intervals were tested for bias. A standard *t*-test revealed that average forecast errors were not statistically different from zero, thus interval forecasts were unbiased during the period of study.

[5] Results were compared to absolute error plus half width model, nearest boundary model, lexicographic semiorder model, absolute error model, normalized error model, interval width model, and inclusion model.

[6] The results of this analysis were not sensitive to a=0.6 and a=1.2, the extremes found in Yaniv and Foster's experiments.

Table 1. Descriptive Statistics for WASDE Interval Forecasts of Corn and Soybean Prices, 1985/86-2001/02 Marketing Years.

Month	Corn			Soybeans		
	Average Interval	Minimum Interval	Maximum Interval	Average Interval	Minimum Interval	Maximum Interval
	---$ per bushel---			---$ per bushel---		
Prior to harvest						
May	0.37	0.20	0.40	1.22	0.40	2.25
June	0.37	0.20	0.40	1.15	0.40	2.00
July	0.38	0.20	0.50	1.13	0.30	2.50
August	0.38	0.20	0.45	1.11	0.30	2.50
September	0.37	0.20	0.40	1.00	0.30	2.50
October	0.37	0.20	0.40	0.85	0.30	2.00
After harvest						
November	0.37	0.20	0.40	0.80	0.30	2.00
December	0.34	0.20	0.40	0.66	0.30	1.00
January	0.31	0.20	0.40	0.60	0.20	1.25
February	0.26	0.20	0.40	0.51	0.15	1.25
March	0.20	0.10	0.30	0.40	0.15	1.00
April	0.15	0.10	0.30	0.24	0.10	0.40
May	0.12	0.00	0.25	0.00	0.00	0.00
June	0.09	0.00	0.20	0.00	0.00	0.00
July	0.00	0.00	0.10	0.00	0.00	0.00
August	0.01	0.00	0.10	0.00	0.00	0.00

Table 2. Confidence Levels for WASDE Corn and Soybean Price Interval Forecasts Based on the Survey of USDA Analysts.

Month	Corn			Soybeans		
	Average	Minimum	Maximum	Average	Minimum	Maximum
		---%---			---%---	
Prior to harvest						
May	77	65	95	73	65	95
June	77	65	95	73	65	95
July	80	70	95	76	70	95
August	85	80	95	81	75	95
September	85	80	95	83	75	95
October	85	80	95	84	80	95
After harvest						
November	88	80	95	85	80	95
December	88	80	95	85	80	95
January	90	85	95	88	85	95
February	92	85	95	89	85	95
March	92	85	95	89	85	95
April	94	90	97	93	90	95
May	94	90	98	93	90	95
June	94	90	98	95	90	100
July	97	95	100	96	95	100
August	97	95	100	96	95	100

Table 3. Accuracy Tests of WASDE Interval Forecasts of Corn and Soybean Prices, 1985/86-2001/02 Marketing Years.

Month	Corn			Soybeans		
	Hit Rate	***t*-test for Implied CL**	***t*-test for 95% CL**	**Hit Rate**	***t*-test for Implied CL**	***t*-test for 95% CL**
	---%---			---%---		
Prior to harvest						
May	47	2.47*	3.95**	76	-0.29	1.53
June	35	3.44**	4.92**	76	-0.29	1.53
July	53	2.23*	3.47**	82	-0.52	1.04
August	59	2.16*	2.98**	82	-0.11	1.04
September	65	1.67	2.50*	82	0.05	1.04
October	65	1.67	2.50*	76	0.62	1.53
After harvest						
November	88	-0.02	0.56	76	0.70	1.53
December	94	-0.50	0.07	82	0.22	1.04
January	88	0.15	0.56	71	1.44	2.01*
February	88	0.31	0.56	76	1.03	1.53
March	71	1.77*	2.01*	76	1.03	1.53
April	76	1.45	1.53	82	0.88	1.04
May	82	0.96	1.04	37	4.58**	4.74**
June	82	0.96	1.04	62	2.68**	2.68**
July	94	0.27	0.10	69	2.25**	2.16*
August	69	2.33*	2.16*	94	0.19	0.10

Notes: The *t*-test for Implied CL is a test of whether the forecast confidence level equals the confidence level derived from the survey of USDA analysts. The *t*-test for 95% CL is a test of whether the forecast confidence interval equals the 95% confidence level. Two stars indicates statistically significant difference at the one percent level and one star indicates significant difference at the five percent level.

Table 4. Calibration Factors for WASDE Interval Forecasts of Corn and Soybean Prices, 1985/86-2001/02 Marketing Years.

Month	Corn		Soybeans	
	Calibration Factor 1	Calibration Factor 2	Calibration Factor 1	Calibration Factor 2
Prior to harvest				
May	2.36	2.68	1.00	2.10
June	2.36	3.12	1.00	1.90
July	2.06	2.32	1.00	1.48
August	1.79	3.05	1.00	1.49
September	1.59	2.56	1.05	1.54
October	1.54	1.54	1.12	1.35
After harvest				
November	1.00	1.05	1.50	2.33
December	1.00	1.00	1.46	2.16
January	1.00	1.33	1.33	1.93
February	1.39	1.39	1.20	1.51
March	1.50	1.50	1.25	1.40
April	1.64	1.64	1.42	1.42
May	1.80	1.80	n/a	n/a
June	2.14	2.14	n/a	n/a
July	n/a	n/a	n/a	n/a
August	n/a	n/a	n/a	n/a

Notes: Calibration Factor 1 is the factor by which the average interval should be multiplied in order to achieve an implied confidence level. Calibration Factor 2 is the factor by which the average interval should be multiplied in order to achieve a 95% confidence level. Calibration Factors are not avalable (n/a) for the months in which forecasts converged to point estimates.

Table 5. Accuracy-Informativeness Tradeoff Scores for WASDE Interval Forecasts of Corn and Soybean Prices, 1985/86-2001/02 Marketing Years.

	Corn			**Soybeans**		
Month	**Published Intervals**	**Implied Confidence Intervals**	**95% Confidence Intervals**	**Published Intervals**	**Implied Confidence Intervals**	**95% Confidence Intervals**
Prior to harvest						
May	**-0.23**	0.17	0.27	**0.52**	**0.52**	1.08
June	**-0.23**	0.18	0.38	**0.46**	**0.46**	0.94
July	**-0.39**	0.03	0.12	**0.53**	**0.53**	0.73
August	**-0.39**	-0.07	0.34	**0.51**	**0.51**	0.72
September	**-0.51**	-0.22	0.14	**0.37**	0.41	0.64
October	**-0.60**	-0.31	-0.31	**0.19**	0.26	0.39
After harvest						
November	**-0.77**	**-0.77**	-0.73	**0.11**	0.40	0.76
December	**-0.88**	**-0.88**	**-0.88**	**-0.12**	0.16	0.49
January	**-0.95**	**-0.95**	-0.73	**-0.20**	0.01	0.31
February	**-1.12**	-0.87	-0.87	**-0.38**	-0.25	-0.07
March	**-1.30**	-1.01	-1.01	**-0.67**	-0.50	-0.40
April	**-1.59**	-1.20	-1.20	**-1.17**	-0.89	-0.89
May	**-1.74**	-1.35	-1.35	n/a	n/a	n/a
June	**-1.92**	-1.52	-1.52	n/a	n/a	n/a
July	n/a	n/a	n/a	n/a	n/a	n/a
August	n/a	n/a	n/a	n/a	n/a	n/a

Note: Scores are based on accuracy-informativeness tradeoff from Yaniv and Foster's model (equation 5). Scores were computed for each observation. Averages for all years are reported in the table. Lowest (best) scores are highlighted in Bold. Implied confidence intervals are calculated by multiplying published intervals by calibration factors reported in Table 4 in order to correspond to the confidence level elicited from the survey of forecast providers. 95% confidence intervals are computed by multiplying published intervals by calibration factors reported in Table 4 in order to correspond to the 95 percent confidence level. Scores are not available (n/a) for the months in which forecasts converged to point estimates.

Figure 1. The 2001/2002 WASDE Forecasting Cycle for Corn and Soybeans Relative to the US Marketing Year.

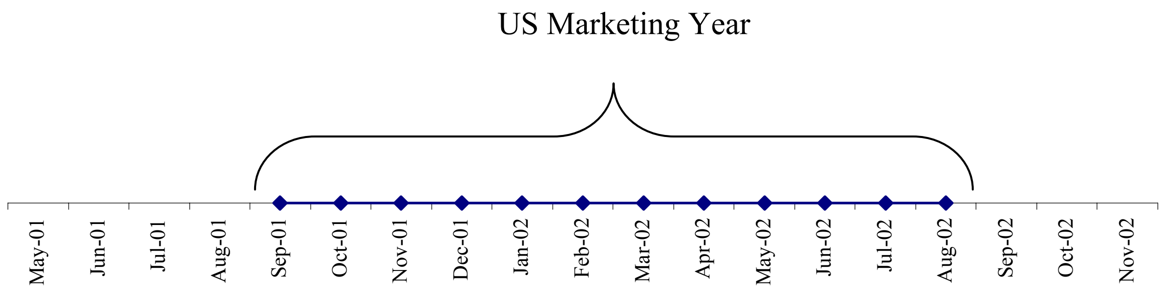

Figure 2. USDA Forecast Price Intervals in Selected Months and Actual Marketing Year Average Price for Corn (1985/86-2001/02).

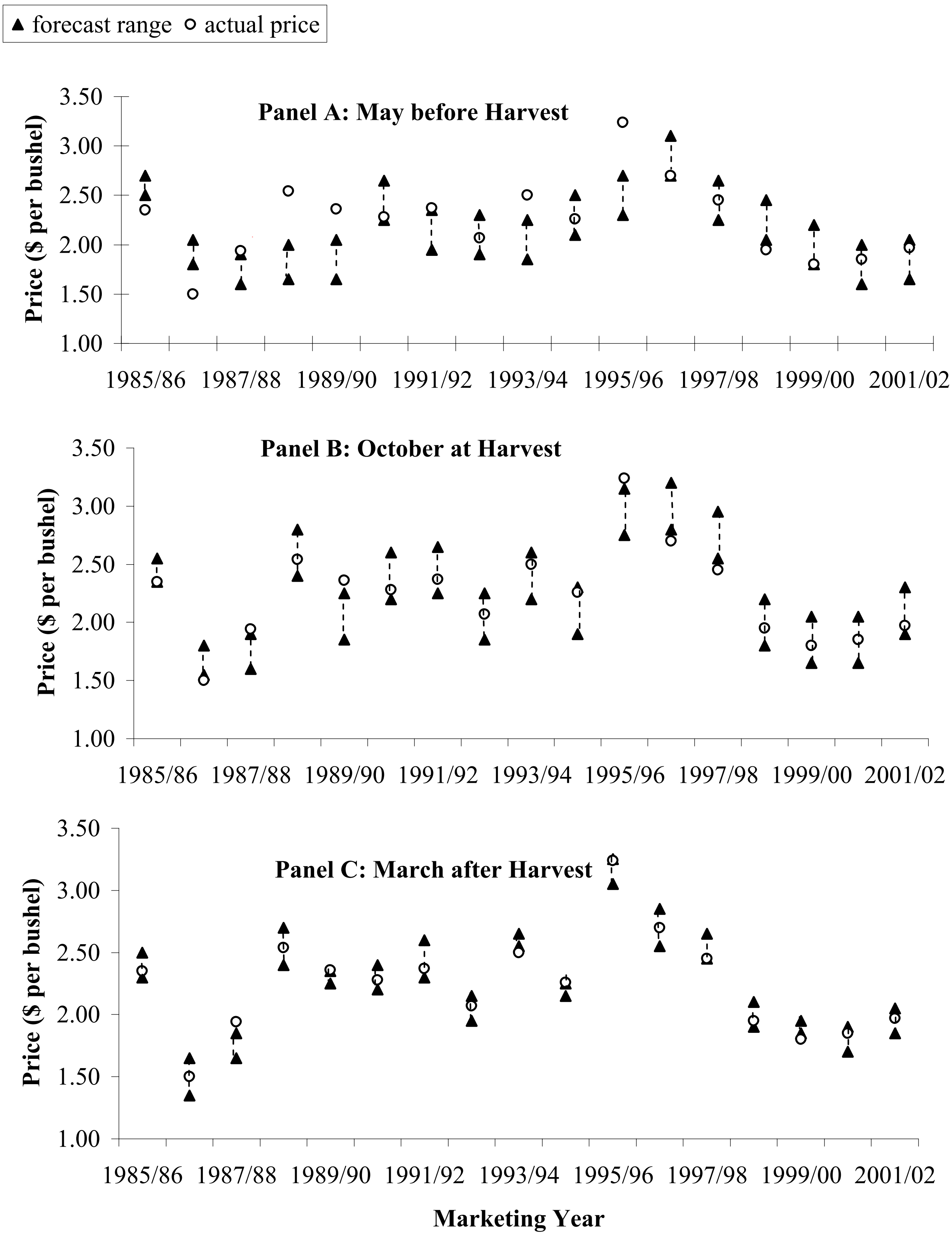

Figure 3. USDA Forecast Price Intervals in Selected Months and Actual Marketing Year average Price for Soybeans (1985/86-2001/02).

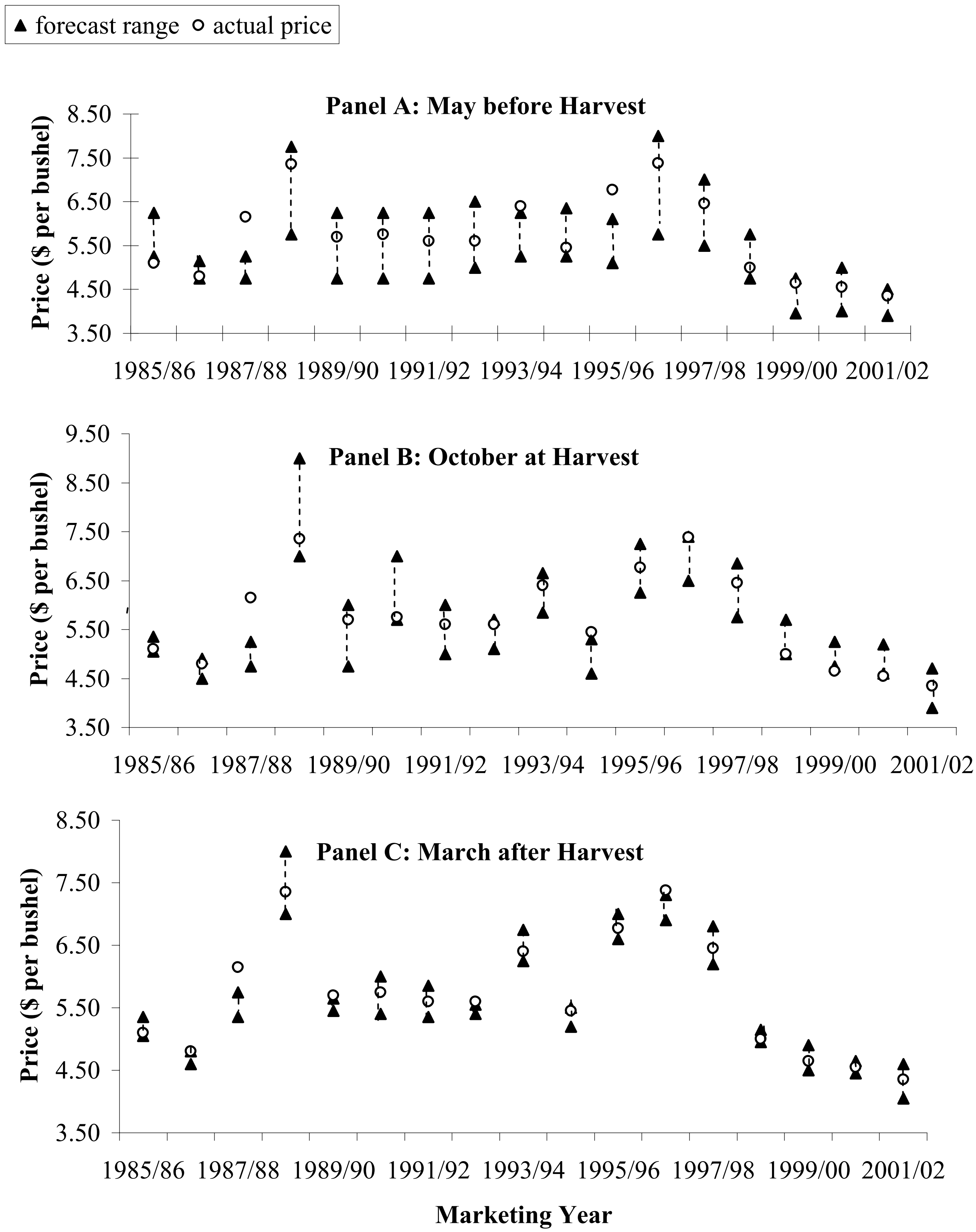

Figure 4. Confidence Intervals for WASDE Forcasts of Corn Prices, 1985/86-2001/02 Marketing Years.

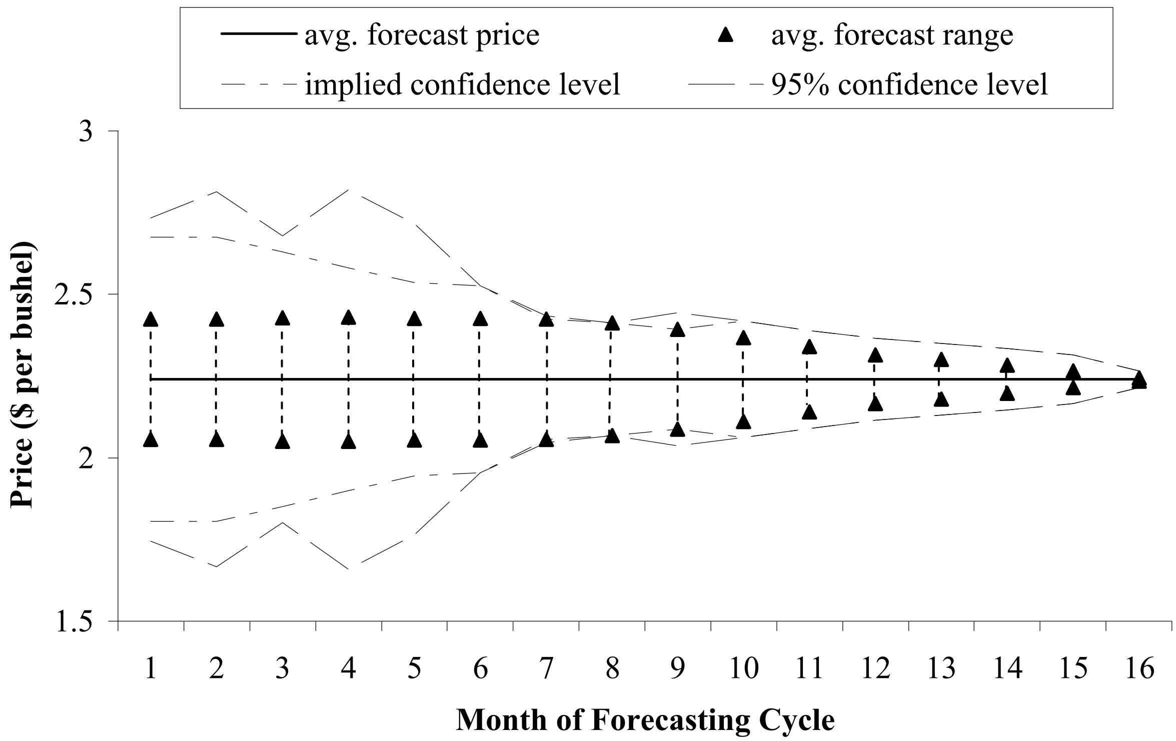

Figure 5. Confidence Intervals for WASDE Forecasts of Soybean Prices, 1985/86-2001/02 Marketing Years.