Accuracy and Asymmetry of Corn, Soybean and Wheat Interval Forecasts

Olga Isengildina-Massa, Julia L. Sharp, Jing Zhang and Wallace Campbell[1]

Poster prepared for presentation at the Agricultural & Applied Economics Association 2010 AAEA, CAES, & WAEA Joint Annual Meeting, Denver, Colorado, July 25-27, 2010

[*] Olga Isengildina-Massa and Julia L. Sharp are Assistant Professors in the Department of Applied Economics and Statistics at Clemson University, Jing Zhang is a graduate research assistant in the Department of Mathematics and Statistics at Bowling Green State University and Wallace Campbell is a graduate research assistant in the Department of Applied Economics and Statistics at Clemson University.

Christoffersen, P. F. (1998) "Evaluating Interval Forecasts", International Economics Review 39, 841-862.

Accuracy and Asymmetry of Corn, Soybean and Wheat Interval Forecasts

Olga Isengildina-Massa[1], Julia L. Sharp[1], Jing Zhang[2], and Wallace Campbell[1]
[1]Clemson University [2]Bowling Green State University

Abstract

• This study evaluates the accuracy of USDA interval forecasts for corn, soybean, and wheat prices using Christoffersen's (1998) tests for unconditional coverage, independence and conditional coverage adjusted for asymmetries in tail probabilities. The findings of this study demonstrate that due to uneven distribution of forecast misses around the interval, calibration of soybean price forecasts in several cases was rejected by basic coverage tests (suitable for symmetric intervals) but not rejected by the tests adjusted for asymmetry. Thus these forecasts were asymmetric but accurate. Symmetry was not a limiting assumption for corn and wheat interval forecast accuracy.

Motivation

• To the best of our knowledge, only two previous studies evaluated USDA price forecasts as intervals rather than reducing them to a point estimate.

• The analysis of accuracy of USDA interval forecasts in these previous studies is limited to calculation of hit rates and basic conditional and unconditional coverage tests, i.e., the proportion of times the intervals contained the final value (hit rate), whether this proportion was equal to stated (or implied) probability level with (conditional coverage) and without (unconditional coverage) taking into account independence (absence of clustering) in forecast sequence.

• Previous analyses do not examine weather forecast misses were unevenly distributed around the forecast interval. The asymmetry in forecast misses is only given in descriptive manner in Isengildina, Irwin and Good (2004) and not included in accuracy analysis.

•Information about asymmetry in misses will help interpret the information contained in USDA interval forecasts.

Objective

•The goal of this study is to expand the evaluation of USDA interval forecast accuracy by including examination of asymmetries in the tail probabilities.

Methods

•Asymmetry in tail probabilities is introduced in the analysis of forecast intervals through indicator sequences I_t (suitable for symmetric intervals) and S_t (suitable for asymmetric intervals):

$$I_t = \begin{cases} 1, if\ y_t \in (L_{t/t-1}(p), U_{t/t-1}(p)) \\ 0, if\ y_t \notin (L_{t/t-1}(p), U_{t/t-1}(p)) \end{cases} \qquad S_t = \begin{cases} 1, y_t < L_{t/t-1}(\alpha_L) \\ 2, L_{t/t-1}(\alpha_L) \le y_t \le U_{t/t-1}(\alpha_U) \\ 3, y_t > U_{t/t-1}(\alpha_U) \end{cases}$$

where $L_{t|t-1}(p)$ and $U_{t|t-1}(p)$ are the lower and upper limits of the interval forecast for time t made at time $t-1$ for coverage probability p (the confidence levels), respectively. The accuracy condition is changed from $\alpha_L=\alpha_U=(1-p)/2$ (used in basic tests suitable for symmetric forecasts) to $p=1-\alpha_L-\alpha_U$ (for asymmetric forecasts); this test is thus less restrictive in cases when forecast misses are distributed unevenly in the tails. This adjustment is applied to the tests of unconditional coverage, independence and conditional coverage developed by Christoffersen (1998).

•Based on results of a survey of forecast providers (Isengildina, Irwin and Good, 2004), the stated prior to harvest and after harvest confidence levels are 80% and 90%, respectively.

•***The LR Tests for Asymmetric Intervals:***

Assuming independence of the hit sequence (S_t), unconditional coverage is examined with the hypotheses:

$$H_0 : E[S_t] = \alpha_M = p$$
$$H_1 : E[S_t] \neq \alpha_M = p$$

Where α_M is the stated probability that the final price falls within the forecast interval. If the null hypothesis is not rejected, the forecasts are said to be calibrated – the hit rate is equal to the specified confidence level.

• The likelihood ratio test statistic for unconditional coverage is given below.

$$LR_{UC} = -2\log\left\{\frac{\alpha_L^{n_L}(1-\alpha_L-\alpha_U)^{n_M}\alpha_U^{n_U}}{\pi_L^{n_L}(1-\pi_L-\pi_U)^{n_M}\pi_U^{n_U}}\right\} \xrightarrow{asy} \chi^2(2)$$

where π_L is the observed probability that the final price falls below the forecast interval.

•Independence of the hit sequence is tested using the likelihood ratio test given below.

$$LR_{ind} = -2\log\left(\frac{\pi_L^{n_L}(1-\pi_L-\pi_U)^{n_M}\pi_U^{n_U}}{\pi_{LL}^{n_{LL}}(1-\pi_{LL}-\pi_{UU})^{n_{LM}}\pi_{LU}^{n_{LU}}\pi_{ML}^{n_{ML}}(1-\pi_{ML}-\pi_{MU})^{n_{MM}}\pi_{MU}^{n_{MU}}\pi_{UL}^{n_{UL}}(1-\pi_{UL}-\pi_{UU})^{n_{UM}}\pi_{UU}^{n_{UU}}}\right) \xrightarrow{asy} \chi^2(4)$$

where π_{ML} is the observed probability that the final price falls within the forecast interval in the previous year and below the forecast interval in the current year. Conditional coverage can then be tested by combining the likelihood ratio test statistics for unconditional coverage and independence.

Data

• **Table 1: Summary Statistics for Corn, Soybean and Wheat Forecast Intervals for 1980/81-2007/08 Marketing Years.**

Month	Average Forecast Price ($/bu.)			Average Range of Interval ($/bu.)			Percentage (%)			Hit Rate (%)			Misses Below (%)			Misses Above (%)		
May	2.33	5.77	3.36	0.4	1.26	0.47	0.17	0.22	0.14	36	50	39	18	18	32	46	32	29
June	2.35	5.78	3.37	0.4	1.22	0.46	0.17	0.21	0.14	29	54	36	25	14	36	46	32	29
July	2.38	5.84	3.34	0.39	1.19	0.44	0.17	0.2	0.13	43	64	57	21	7	21	36	29	21
August	2.42	5.95	3.37	0.39	1.18	0.43	0.16	0.2	0.13	54	79	64	25	4	14	21	18	21
September	2.42	6.03	3.39	0.38	1.07	0.37	0.16	0.18	0.11	54	75	71	25	7	7	21	18	21
October	2.41	6.02	3.43	0.38	0.97	0.32	0.16	0.16	0.09	54	68	75	21	11	7	25	21	18
November	2.42	6.04	3.44	0.37	0.9	0.26	0.15	0.15	0.07	71	68	64	11	11	14	18	21	21
December	2.42	6.08	3.45	0.35	0.8	0.21	0.14	0.13	0.06	79	82	71	7	7	14	14	11	14
January	2.44	6.08	3.46	0.31	0.69	0.18	0.13	0.11	0.05	86	79	71	7	4	14	7	18	14
February	2.44	6.07	3.46	0.26	0.6	0.13	0.11	0.1	0.04	82	82	71	7	0	14	11	18	14
March	2.44	6.05	3.45	0.21	0.45	0.11	0.09	0.07	0.03	75	82	75	11	0	11	14	18	14
April	2.46	6.06	3.45	0.15	0.27	0.08	0.06	0.04	0.02	75	79	64	11	4	14	14	18	21
Average Prior to Harvest							0.16	0.19	0.14	45	65	44	23	10	30	33	25	26
Average after Harvest							0.11	0.1	0.07	78	79	70	9	4	12	13	17	18

Notes: Black represents data for corn, red represents soybean data, and blue represents wheat data. Pre-harvest months are May-October for corn and soybeans and May –July for wheat, post-harvest months are November – April for corn and soybeans and August – April for wheat.

• For soybean price forecast intervals, the proportion of misses above the interval was 2.5 times greater than the proportion of misses below the interval prior to harvest and 4.25 times greater after harvest, reflecting a systematic tendency to underestimate soybean prices.

• Hit rates for both crops are low, but improve after harvest (from 45%, 65%, and 44% to 78% , 79%, and 70%), reflecting greater certainty of the final price.

Results

• **Table 2: Unconditional Coverage Test Statistics for Corn, Soybean and Wheat Forecast Intervals, 1980/81-2007/08 Marketing Years.**

Month	α_L =0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1	LR_{ind}
May	35.92***	30.40***	27.83***	26.53***	25.98***	25.95***	26.33***	27.08***	28.16***	29.59***	1.95
	25.13***	19.17***	16.14***	14.36***	13.29***	12.71***	12.50***	12.60***	12.99***	13.65***	10.95**
	53.35***	41.73***	35.35***	31.14***	28.16***	25.98***	24.39***	23.27***	22.54***	22.17***	12.49**
June	51.82***	43.52***	39.33***	36.88***	35.43***	34.68***	34.44***	34.66***	35.27***	36.27***	2.84
	18.71***	14.14***	11.92***	10.71***	10.09***	9.87***	9.97***	10.34***	10.97***	11.84***	10.67**
	61.91***	48.91***	41.72***	36.93***	33.50***	30.96***	29.06***	27.67***	26.71***	26.13***	7.2
July	34.42***	27.18***	23.46***	21.22***	19.83***	19.03***	18.66***	18.66***	18.98***	19.63***	7.86*
	6.52**	4.61*	3.9	3.72	3.86	4.24	4.81*	5.55*	6.47**	7.58***	4.81
	27.45***	19.78***	15.60***	12.88***	10.98***	9.62*	8.66**	8.01**	7.64***	7.52***	7.24
August	34.48***	25.42***	20.43***	17.13***	14.78***	13.06***	11.79***	10.88***	10.27***	9.94***	4.18
	1.13	0.29	0.05	0.08	0.28	0.6	1.04	1.57	2.2	2.95	1.42
	19.57***	15.43***	13.79***	13.35***	13.35***						7.96*
September	34.48***	25.42***	20.43***	17.13***	14.78***	13.06***	11.79***	10.88***	10.27***	9.94***	2.4
	4.53	2.3	1.25	0.71	0.46	0.42	0.54	0.81	1.21	1.74	2.6
	9.03**	7.67**	7.65**	8.35**	8.35**						6.58
October	28.59***	21.03***	16.96***	14.36***	12.58***	11.36***	10.55***	10.07***	9.87***	9.94***	7.2
	9.42***	5.91*	4.16	3.16	2.6	2.33	2.3	2.46	2.79	3.3	4.74
	7.06**	5.46*	5.18*	5.57*	5.57*						4.76
November	11.84***	8.86***	7.76***	7.57***	7.57***						12.98**
	13.91***	11.16***	10.33***	10.46***	10.45***						4.41
	19.57***	15.44***	13.80***	13.35***	13.35***						8.14*
December	5.59*	3.76	3.2	3.28	3.28						12.05**
	4.71*	2.64	1.82	1.6	1.59						2.69
	15.73***	11.12***	8.95**	7.88**	7.88**						5.53
January	4.6	2.3	1.21	0.67	0.67						6.88
	3.42	3.21	3.74	4.70*	4.70*						3.97
	15.73***	11.12***	8.95**	7.88**	7.88**						5.53
February	4.71	2.64	1.82	1.6	1.6						7.65
	NA	NA	NA	NA	NA						2.69
	15.73***	11.12***	8.95**	7.88**	7.88**						5.53
March	10.27*	7.05*	5.69*	5.19*	5.19*						8.01*
	NA	NA	NA	NA	NA						2.69
	10.27***	7.05**	5.69*	5.19*	5.19*						4.83
April	10.27*	7.05*	5.69*	5.19*	5.19*						8.98*
	3.42	3.21	3.74	4.70*	4.70*						6.45
	19.57***	15.44***	13.80***	13.35***	13.35***						4.91

Notes: Black represents data for corn, red represents soybean data, and blue represents wheat data. *,**,and *** indicate significance at the 10%, 5% and 1% level, respectively. α_M is the coverage probability (80% prior to harvest and 90% after harvest). For asymmetric intervals, $\alpha_L \neq \alpha_U$. For symmetric intervals, $\alpha_L=\alpha_U= 0.1$ pre-harvest and $\alpha_L=\alpha_U= 0.05$ post-harvest. Pre-harvest months are May-October for corn and soybeans and May –July for wheat, post-harvest months are November – April for corn and soybeans and August – April for wheat. The unconditional coverage test statistics follow a Chi-square distribution with two degrees of freedom.

Results, cont.

• The unconditional coverage test results for α_L=0.1 (prior to harvest) and α_L=0.05 (after harvest) in Table 2 corresponds to the basic test suitable for symmetric intervals. These results suggest that USDA forecasts are not calibrated at the implied confidence level in most cases except December-February after harvest in corn, and August-October prior to harvest and December after harvest in soybeans.

• The unconditional coverage test results for α_L< 0.1 (prior to harvest) and α_L<0.05 (after harvest) in Table 2 represent cases when the condition $\alpha_L = \alpha_U$= (1-p)/2 (used in basic tests suitable for symmetric forecasts) is relaxed to $p = 1-\alpha_L-\alpha_U$ (for asymmetric forecasts) for various levels of α_L. These results are generally unchanged for corn and wheat indicating that symmetry is not a limiting assumption for USDA corn interval forecast accuracy.

• For soybeans, differently from the first set of results, calibration cannot be rejected in July prior to harvest for $0.03 \le \alpha_L \le 0.06$, in October prior to harvest for $0.03 \le \alpha_L$, December after harvest $0.02 \le \alpha_L$, January and April after harvest for $\alpha_L \le 0.03$. These results illustrate that for July prior to harvest forecasts, the probability of misses above the interval was much higher than the probability of misses below the interval (0.03-0.06), but the total coverage probability was 0.8 consistent with the stated confidence level. Other results indicate that soybean price forecasts in October prior to harvest, December after harvest, and January and April after harvest have been even more skewed to the right with the probability of misses below the interval less than 0.03 for most of these forecasts. Since these forecasts' coverage is not significantly different from the target confidence level, the probability of misses above the interval can be calculated as *1- p –* $\alpha_L = \alpha_U$, implying $\alpha_U \ge 0.17$ for October prior to harvest forecasts, $\alpha_U \ge 0.08$ for December after harvest forecasts and $\alpha_U \ge 0.07$ for January and April after harvest forecasts. These findings suggest that soybean price forecasts in these months should not be interpreted as symmetric.

• The tests of independence are not affected by asymmetry in the tails and reveal clustering in corn forecasts in July prior to harvest and November, December, March and April post harvest, soybean forecasts in May and June prior to harvest, and wheat forecasts in May pre-harvest and August and November post harvest.

• Conditional coverage test statistics are the combination of the unconditional coverage and independence test statistics. The results of the conditional coverage tests (not presented here) are consistent with the unconditional coverage test results shown in Table 2.

Conclusions

•Although forecast intervals published by the USDA for corn, soybean and wheat prices are constructed symmetrically, we have shown that these intervals should not always be interpreted as symmetric. The findings of this study demonstrate that due to uneven distribution of forecast misses around the interval, calibration of soybean price forecasts in several cases was rejected by basic coverage tests (suitable for symmetric intervals) but not rejected by the tests adjusted for asymmetry. Fore these forecasts, the proportion of misses above the interval was much higher than the percentage of misses below the interval, but the total coverage probability was consistent with the stated confidence level. In other words, these forecasts were asymmetric but accurate.

• Accuracy tests adjusted for asymmetry also illustrate the degree of asymmetry in the tails as demonstrated in the previous example.

• If the goal of USDA is to provide symmetric forecasts, several soybean price forecasts should be revised based on the evidence provided in this study.

References

•Christoffersen, P. F. (1998) "Evaluating Interval Forecasts", *International Economics Review 39*, 841-862.

•Isengildina, O., Irwin, S. H., and Good, D. L. (2004), "Evaluation of USDA Interval Forecasts of Corn and Soybean Prices", *Amer. J. Agr. Econ. 86(4)*, 990-1004