Accuracy Trends and Sources of Forecast Errors in

WASDE Balance Sheet Categories for Corn and Soybeans

Augusto C. Botto, Olga Isengildina, Scott H. Irwin, and Darrel L. Good[1]

Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Long Beach, California, July 23-26, 2006

[1] Augusto C. Botto is a former Graduate Research Assistant for the AgMAS Project; Olga Isengildina is a Post-Doctorate Research Associate; Scott H. Irwin is the Laurence J. Norton Professor of Agricultural Marketing and Darrel L. Good is a Professor in the Department of Agricultural and Consumer Economics at the University of Illinois at Urbana-Champaign.

Accuracy Trends and Sources of Forecast Errors in

WASDE Balance Sheet Categories for Corn and Soybeans

Introduction

Commodity markets are known to be highly volatile mainly due to the stochastic nature of agricultural production and the inelastic properties of the underlying supply and demand. These markets are also characterized by the existence of important time lags which are responsible for exposing participants to different kinds of risks. The dynamic and volatile nature of agricultural markets causes individuals to rely on forecasts in their decision-making. Consequently, the US Department of Agriculture (USDA) devotes substantial resources to agricultural situation and outlook programs (Offutt, 2002). The WASDE (World Agricultural Supply and Demand Estimates) program is an especially prominent example of this effort. It is a commonly held belief of agricultural market participants and analysts that WASDE forecasts function as the "benchmark" to which other private and public forecasts are compared. The dominant role of WASDE forecasts is not surprising given the classic public goods problem of private underinvestment in information (Wolf, Just and Zilberman, 2001), and the critical role that public information plays in coordinating the beliefs of market participants (Morris and Shin, 2002).

WASDE forecasts released to the market between the 9th and 12th of each month provide a country-by-country and commodity-by-commodity balance sheet for each marketing year (Spilka, 1983; Vogel and Bange, 1999). Separate balance sheets are simultaneously prepared and published for 90 countries. The supply-side of these balance sheets consists of estimates for carryover stocks from the previous year, prospective production (planted acreage times estimated yield) and imports from current year. Components included in the consumption side are projections for exports and for domestic use (which is further subdivided depending on each crop). Ending stocks are estimated as the difference between total supply and total consumption. For U.S crops, the balance sheets also contain a projection of the average price received by farmers.

Previous studies of WASDE forecasts have mostly concentrated on forecast accuracy of two major components of WASDE balance sheets, production (e.g., Gunnelson, Dobson and Pamperin,1972; Thomson, 1974; Isengildina, Irwin, and Good, 2006) and price (e.g., Marquardt and McGann, 1977; Just and Rausser, 1981; Irwin, Gerlow and Liu, 1994; Sanders and Manfredo, 2002; Egelkraut et al., 2003; Isengildina, Irwin, and Good, 2004). The importance of price forecasts is obvious, given the role price expectations play in decisions on resource allocation. Production forecasts are important because they are a major determinant of future supply. Interestingly, the accuracy of most other categories describing supply and demand forces in WASDE forecasts have been overlooked in the previous literature. The importance of these other categories can be illustrated with corn food, seed and industrial use. This category has grown rapidly during the last 20 years mainly as a consequence of the increase in the usage of corn for ethanol production, representing at present about 20% of total use, thus having a substantial impact on price discovery.

The importance of WASDE forecasts is well illustrated by a recent controversy about the reliability of such estimates. In August 2004, Senator Tom Harkin of Iowa requested that the General Accounting Office (GAO) of the U.S. Congress review the accuracy of these forecasts. Harkin stated that he had numerous complaints from constituents that bad USDA forecasts have cost them money:

> "It is absolutely essential that farmers in Iowa and across the country have confidence in USDA's public reports on commodity production and supply and demand. Farmers rely heavily on this information in making decision about marketing their crops. That is why I have asked a review of USDA's practices to ensure their numbers are accurate and reliable. With ever tightening bottom lines for many farmers, there is no room for error. We have to make sure any information USDA makes public is as accurate and unbiased as possible." (Doane's Agricultural Report, 2004).

USDA's forecasts for soybean ending stocks were highlighted as a particular problem area. In recent years, early projections for ending stocks have been above actual levels, tending to depress prices early in the season.

This controversy also illustrates the interrelationship between WASDE categories. Because of the balance sheet structure of WASDE forecasts, the errors in the aggregate categories, such as total supply, total use, ending stocks and price are likely to be caused by the errors in the individual categories. In the above case it is argued that the errors in the ending stocks estimates, which are calculated as the difference between total supply and total use, cause biases in the price forecasts. Separate examination of production and price forecast accuracy provided in the previous literature gives little guidance on sources of errors in price forecasts. To the best of our knowledge, no evidence exists about potential sources of errors in other aggregate WASDE categories, such as ending stocks. Furthermore, the exact impact of ending stocks errors on price forecast errors is not known.

The purpose of this study is twofold: 1) to analyze trends in forecast accuracy of all WASDE balance sheet categories for U.S. corn and soybeans and 2) to identify possible sources of errors in ending stocks and average price forecasts. This study uses data from monthly WASDE balance sheets for U.S. corn and soybeans over 1980/1981 through 2003/2004 to calculate percent forecast errors for each category. WASDE forecast accuracy is examined in a dynamic framework using Bailey and Brorsen's (1998) approach to analyze trends in the mean and the variance of percentage forecast errors for each category. This analysis is based on a two-equation model, where the first equation regresses percentage errors for a given category and crop on forecast horizon, marketing year, and an interaction term. The estimated squared residuals obtained from the first equation are used as a proxy for the error variance. In the second equation, the natural log of the estimated squared residuals is regressed against the same explanatory variables. Consequently, in this framework the first equation evaluates forecast bias while the second equation evaluates trends in the variability of errors over the forecast horizon and over the sample period. Parameters of the equations are estimated using OLS, but the standard errors of the estimates are corrected using a panel White estimator that allows for both period heteroskedasticity and autocorrelation.

In the second part of the study the source of corn and soybean price and ending stocks forecast errors is analyzed by regressing these categories against errors in all individual supply and consumption categories, which include beginning stocks, production, crush, exports and feed, seed & residual for soybeans; and beginning stocks, production, exports, feed & residual and food, seed & industrial use for corn. These regressions are estimated for each report month in order to evaluate if sources of errors varied within the forecasting cycle.

WASDE Forecast Generating Process[1]

Several USDA agencies are responsible for preparing crop statistics for WASDE reports. The World Agricultural Outlook Board (WAOB) coordinates an interagency process that prepares monthly forecasts of supply and demand for major crops, both for the U.S. and the World. Analysts from the Economic Research Service (ERS), Foreign Agricultural Service (FAS), Agricultural Marketing Service (AMS) and Farm Service Agency (FSA) meet to evaluate current forecasts and new information. Several information sources are used when making the forecast. USDA's own resources include weather analysis, country reports, evidence from satellite imagery and also private and public information sources. Throughout the growing season, as new information on production and consumption becomes available, revisions are made and new reports are released. Available information is reviewed by analysts from several agencies in order to arrive at a consensus forecast.

This study focuses on WASDE balance sheets for U.S. corn and soybeans. The first marketing year forecast for U.S. corn and soybeans is usually available in May preceding the marketing year (September through August). WASDE estimates for U.S. corn and soybeans are typically finalized by November of the following marketing year. Thus, 19 forecast updates for U.S. corn and soybean markets are generated in the WASDE forecasting cycle each marketing year. WASDE forecasts are considered fixed-event forecasts because the series of forecasts are related to the same terminal events, namely supply and consumption categories for a specific marketing year.

WASDE balance sheets for U.S. corn and soybeans consist of several supply and consumption categories. The supply-side of these balance sheets consists of estimates for beginning stocks, imports, and prospective production (planted acreage times estimated yield). Components included in the consumption side are projections for exports and for domestic use (crushings, seed and residual for soybeans; and feed and residual, food, seed and industrial for corn). Ending stocks are estimated as the difference between total supply and total use. This category is widely followed by market participants since it is a key indicator of the fundamental conditions of a given market, illustrating the relative strength of the consumption components with respect to the supply components. The balance sheets also contain a projection of marketing year average price received by farmers, which is based on commodity models reflecting the supply and demand conditions via stock-to-use ratios, lagged prices and other variables (Labys, 1973; Wescott and Hoffman, 1999). The average marketing price differently from all other WASDE forecasts is published in the form of an interval to reflect the uncertainty associated with this forecast. Because analysis of interval forecast accuracy is different from point estimate accuracy (e.g., Isengildina, Irwin, and Good, 2004), corn and soybean average price forecast errors were computed using the midpoint of the published interval to be consistent with the rest of the analysis.

Data

The subjects of this investigation are monthly WASDE balance sheets for U.S. corn and soybeans for the 1980/1981 through 2003/2004 marketing years.[2] For each category, monthly announcement and marketing year percentage forecast errors were calculated according to the following equation:

(1) $$PE_{kt} = \frac{(y_{19t} - y_{kt})}{y_{19t}}, \quad k=1,\ldots,19; \quad t=1,\ldots,24$$

where PE_{kt} corresponds to the percentage error for a given report month *k*, and marketing year *t* and y_{19t} corresponds to the final estimate for a given marketing year published in the November report 19 months after the first forecast was released for that marketing year (for the May report, *k*=1). Finally, y_{kt} corresponds to the forecast for a given report month and marketing year. Thus, for each balance sheet category a total of 432 forecast errors were computed (18 report months times 24 marketing years). Because final estimates for corn and soybean production were typically released in January or February, errors for production forecasts were computed only until the December WASDE report.

Tables 1 and 2 show mean percentage errors (MPEs) for all WASDE balance sheet components for U.S. corn and soybeans, respectively. MPEs were computed across years for each report month in order to measure if the forecasts were biased. A standard t-test was used to examine if MPEs were significantly different from zero. Negative forecast errors suggest a tendency for overestimation, while positive errors imply underestimation in underlying forecasts. Table 1 demonstrates that overall the WASDE forecasts for U.S. corn were unbiased. A single significantly different from zero MPE (1.40) was associated with October forecasts for Food, Seed and Industrial use. Underestimation in this consumption category may be due to the fact that it reflects the use of corn for ethanol production, which experienced rapid growth during the last 20 years.

The picture is very different for WASDE forecasts for U.S. soybeans. As shown in table 2, although the absolute value of MPEs of soybean crush was never larger than 2% for all report months, p-values indicate that a significant bias toward underestimation was present in March through August forecasts. A significant tendency for overestimation is observed in July and August forecasts of soybean use for feed, seed and residual. However, since the proportion of soybeans used for seed is relatively stable and predictable across marketing years, the bias in this category may be due to the residual component. Consistent with the arguments presented in the introduction, a significant bias toward overestimation was found in soybean ending stocks forecasts from May through March. Interestingly, this bias in ending stocks forecasts did not necessarily result in biased price forecasts. Even though soybean average price MPEs from May through July were 2.81, 2.67, and 2.25, suggesting underestimation, they were not significantly different from zero.

Finally, some common patterns in percentage errors were observed both for corn and soybean forecasts across all categories. First, the size and variability of the errors diminished approaching the final report, resulting in a clear heteroskedastic pattern in the variance of the errors across report months. Second, a high autocorrelation in the errors was observed across percentage errors for each marketing year, where early positive errors tended to be followed by positive errors and vice-versa. This pattern is likely due to the fact that USDA forecasts for all the components included in the balance sheets are fixed-event forecasts. While the descriptive accuracy statistics discussed here are intended to provide a general idea about WASDE forecast accuracy during the study period, the next section presents the formal accuracy tests in a dynamic framework.

Trends in USDA Forecast Accuracy

Bailey and Brorsen (1998) developed a model that allows testing whether forecast accuracy has changed over time. Following Bailey and Brorsen (1998), a two-equation model is used for estimating trends in forecast accuracy. The first equation of the model estimates trends in the mean of percentage errors for each category while the second equation estimates trends in the variance of the errors for that category as follows:

(2) $$\frac{(y_{19t} - y_{kt})}{y_{19t}} = \beta_0 + \beta_1 k + \beta_2 t + \beta_3 kt + e_{kt} \,,$$

$$e_{kt} \sim N(0, \sigma_{kt}^2) \,,$$

(3) $$\sigma_{kt}^2 = \alpha_0 + \alpha_1 k + \alpha_2 t + \alpha_3 kt + u_{kt} \,,$$

$$k=1,\ldots,18; \qquad t=1,\ldots,24$$

where, the independent variables in both equations are report month (*k*), marketing year (*t*) and an interaction term (*kt*). As shown in equation 2, percentage forecast errors for a given marketing year *t* are computed as the difference between the final estimate (y_{19t}) and the forecasted value in a given report month *k* (y_{kt}). The error terms in each equation are represented by e_{kt} and u_{kt} , respectively. Consequently, trends in the mean and in the variance of percentage forecast errors for each category are estimated across forecast horizons, marketing years and an interaction term.

All parameters in the mean equation (the β vector) should be zero if forecasts are unbiased since optimal forecast errors must cancel out across horizons and also across marketing years. In the variance equation (3) α_1 is expected to be negative since this would indicate that the variance of the errors becomes systematically smaller within the forecast period. Also, if $\alpha_2 <0$, then USDA's forecasts have become more accurate (less variable) since 1980. Finally, $\alpha_3 >0$ indicates that the variances of the percentage errors during the early months of a forecast are smaller in the final years than during the initial years of the period under study. The intuition behind the interaction term can be gained by rearranging one segment of equation 3 as follows:

(4) $$(\alpha_1 k + \alpha_3 kt) = k(\alpha_1 + \alpha_3 t)\ .$$

If α_1 is negative and α_3 is positive, then for a given report month *k* the entire term becomes less negative over time and the slope of the variance function flattens as a result. In other words, the term presented in equation (3) changes over time and α_3 determines how much it changes. Consequently, a significant negative slope estimate for the marketing year variable and a significant positive interaction term would indicate that the variability of the forecast errors has decreased over the sample period and that the improvement in accuracy is more important early in the forecasting cycle of a marketing year.

The model is estimated in two steps: First, the mean equation (2) is estimated by regressing percentage forecast errors for a given category and crop on the forecast horizon, marketing year, and an interaction term. Second, the variance equation (3) is estimated by regressing the log of the estimated squared residuals from the first step regression for a given category and crop on the forecast horizon, marketing year, and an interaction term. However, the OLS estimator assumes that residuals are homoscedastic and i.i.d. Both of these assumption are violated in the data used for this study according to statistically significant values of Goldfield-Quandt (GQ) statistics, which indicate the presence of heteroscedasticity in the residuals of the mean equation across forecast horizons for all categories in both corn and soybean balance sheets and Durbin Watson tests, which indicated that the null of zero first order autocorrelation across report months is rejected at the 99% level for all categories in both the mean and variance equations.[3]

In the presence of heteroscedasticity and autocorrelation, the least squares estimator for the slopes remains unbiased, but the least squares estimator for the sampling variance is biased and consequently hypothesis testing can be misleading. Therefore, standard errors of the estimates were corrected using a cross-section panel variant of the White's estimator (White, 1980) in which variances and covariances are replaced by residual moment estimators. White's panel estimator uses a covariance structure that allows for arbitrary period serial correlation and period heteroskedasticity between the residuals for a given cross section, but restricts residuals in different cross-sections to be uncorrelated (Wooldridge, 2002, pp. 148-153; Arellano, 1987).

The results of empirical estimation of equation 2 for soybeans and corn shown in table 3 include the slope estimates obtained in OLS regressions and error estimates obtained using White's panel estimator. No significant bias was detected for corn WASDE categories. The results indicate that only soybean ending stocks and average price forecasts are significantly biased at the 1% and 5% confidence levels, respectively. The slope estimate for the marketing year variable for soybean ending stocks forecasts was -3.66%, indicating that each marketing year the soybean ending stocks percentage forecast error is estimated to decrease by 3.66%. Negative errors indicate an overestimation for ending stocks category. Also, a positive and significant interaction term was observed, indicating that the bias is more important for early report months. These results are consistent with the descriptive statistics of forecast accuracy presented in the previous section. In addition, a positive estimate for the horizon and marketing year variables and a negative estimate for the interaction term suggest that during the last marketing years included in the sample period soybean average price forecasts were biased toward underestimation at the 1% confidence level. These findings provide formal evidence in support of the arguments regarding biases in soybean ending stocks estimates, which may be associated with depressed prices early in the season.

Differently from the results for the mean equation, the results for the variance equation (3) shown in table 4, indicate that almost all slope estimates for the horizon variable were significant at the 5% level, suggesting an increase in forecast accuracy across forecast horizons. However, improvement in forecast accuracy over the sample period was significant only for soybean total use and for corn average price forecasts at the 5% and 1% levels, respectively. Nonetheless, the fact that the slope coefficients for the marketing year variable were negative for almost all categories suggests that in general forecast accuracy has improved for other categories as well.

Results presented in table 4 indicate that both corn and soybean production forecast accuracy improved over the sample period since the estimated slopes for marketing year are around -10%. This means that for a given report month, the variane of production forecast errors is estimated to decrease by approximately 10% from one marketing year to the next. Similarly, results indicate that the variability of USDA forecast errors for soybean total use is estimated to decrease by 17% from one marketing year to the next. Likewise, the variance of corn average price forecast errors is estimated to decrease by 11%. The fact that for these two categories a significant and positive interaction term (α_3) was observed indicates that the reduction in the variability of forecast errors across marketing years occurs primarily early in the forecast period. Similar results were observed for soybean exports and for corn feed & residual, ending stocks and average price forecasts. On the other hand, the model does not indicate a significant improvement in soybean average price forecast accuracy over the sample period, since although the estimated α_2 coefficient is negative, it is not statistically significant.

Predicted standard deviations for corn average price forecast errors for 1980/1981 and 2001/2002 marketing years are shown in Figure 1 to clarify the above results. These two years were arbitrarily selected to compare results across marketing years near the beginning and the end of the sample period. The estimated slope coefficients were used for modeling the variance of the forecast errors for these marketing years. Figure 1 illustrates that the model predicts an improvement in corn average price forecast accuracy over the sample period since the estimated standard deviation of the forecast errors is significantly lower for the 2001/2002 marketing year. However, the reduction in the variability of forecast errors occurs primarily early in the forecast period when forecast errors are the largest (interaction term). On the other hand, the lack of improvement in soybean average price forecast accuracy over the sample period, is illustrated in Figure 2, which shows the predicted standard deviations for soybean average price forecast errors for the same marketing years 1980/1981 and 2001/2002. Figure 2 demonstrates that no significant difference in the predicted standard deviation of soybean average price forecast errors is observed across the two marketing years.

Overall, the analysis of trends in USDA forecast accuracy revealed that over the period of study only soybean ending stocks and soybean average price forecasts were biased toward overestimation and underestimation respectively, while the rest of the categories for both corn and soybeans were unbiased. Results of the analysis of the trends in variance of forecast errors suggest that, in general, forecast accuracy for most categories has improved over the sample period. However, this improvement in accuracy is only significant for soybean total use and corn average price forecasts and tends to be more pronounced early in the forecasting cycle.

Sources of USDA Forecast Errors

Analysis of forecasts accuracy presented in the previous section revealed the presence of biases in categories, such as ending stocks and average price forecasts. Interestingly, these are aggregate categories, ending stocks reflecting the difference between supply and demand components and average price resulting from the supply-demand relationship. Therefore, the presence of biases in these categories may be caused by accumulation of errors in underlying supply and demand components. The purpose of this section is to attempt to track down the errors in ending stocks and average price forecasts to the errors in individual supply and demand categories. This section also examines the extent to which errors in ending stocks forecasts affect average price forecast errors.

Multiple regression equations are estimated for each crop and report month to evaluate which categories are significantly related to forecast errors in ending stocks and average price. First, percentage forecast errors in ending stocks are regressed over percentage errors in other balance sheet categories:

$$(5) \quad PEes_{kt}^{soy} = \lambda_0 + \lambda_1 PEbs_{kt} + \lambda_2 PEprod_{kt} + \lambda_3 PEcrush_{kt} + \lambda_4 PEex_{kt} + \lambda_5 PEfsr_{kt} + e_{kt}$$

$$PEes_{kt}^{corn} = \lambda_0 + \lambda_1 PEbs_{kt} + \lambda_2 PEprod_{kt} + \lambda_3 PEfr_{kt} + \lambda_4 PEex_{kt} + \lambda_5 PEfsi_{kt} + e_{kt}$$

where, *PEbs, PEes, PEprod, PEcrush, PEex, PEfsr,* $PEfr_k$ and $PEfsi_k$ are percentage forecast errors for beginning stocks, ending stocks, production, crush, exports, feed, seed & residual, feed & residual, and food, seed & industrial use, respectively and e_{kt} is the error term for report month *k* and marketing year *t*. Percentage forecast errors for total supply and total use are not included as independent variables since they are perfectly correlated to other explanatory variables, which would create perfect multicolinearity and estimation would be impossible. Correlations between the other explanatory variables used in the multiple regression analysis are small and not likely to cause multicolinearity problems. Because ending stocks are calculated as the difference between supply and demand components, the signs of the estimated parameters should be positive for beginning stocks and production forecast errors, and negative for the consumption categories.

Similarly, percentage forecast errors for average price are regressed over the other balance sheet categories:

$$(6)\ PEp_{kt}^{soy} = \lambda_0 + \lambda_1 PEbs_{kt} + \lambda_2 PEprod_{kt} + \lambda_3 PEcrush_{kt} + \lambda_4 PEex_{kt} + \lambda_5 PEfsr_{kt} + e_{kt}$$

$$PEp_{kt}^{corn} = \lambda_0 + \lambda_1 PEbs_{kt} + \lambda_2 PEprod_{kt} + \lambda_3 PEfr_{kt} + \lambda_4 PEex_{kt} + \lambda_5 PEfsi_{kt} + e_{kt}$$

where, *PEp* are percentage forecast errors for average price, and other variables are as defined above. In this case, a negative relationship between average price forecast errors and beginning stocks and production forecast errors and a positive relationship with the consumption independent variables is expected. All regressions were estimated for each report month to evaluate if the source of errors changed across forecast horizons. For both crops beginning stocks percentage forecast errors were only computed until the September report, since no revisions were made from October on, and consequently forecast errors are zero after the September report. Similarly, corn production forecast errors were only computed until the January report, since no revisions were made thereafter. Soybean production forecast errors were computed until the final report since for some marketing years, small revisions were made close to the final report in November.

Figures 3 and 4, show the estimated elasticities obtained from the multiple regressions of soybean and corn ending stocks percentage forecast errors over errors in other balance sheet categories (equation 5). These graphs indicate that forecast errors in almost all balance sheet categories contributed to forecast errors in soybean and corn ending stocks. The major determinants of forecast errors in ending stocks for both crops and for almost all report months, were production forecast errors. The graphs show that a 1% overestimation in production in May reports results in approximately a 9% overestimation in soybean ending stocks and a 4% overestimation in corn ending stocks. Another major contributor to ending stocks forecast errors throughout the forecasting cycle were the major use categories, crush for soybeans and feed and residual for corn. Errors in ending stocks forecasts for both commodities were also significantly affected by errors in export forecasts throughout the forecasting cycle. Errors in beginning stocks, although small in magnitude had a significant impact on ending stocks forecast errors in soybeans, but not in corn. Overall, the absolute value of the estimated elasticities was much higher for soybeans than for corn in explaining forecast errors in ending stocks. For example, a 1% overestimation in exports in May is estimated to result in a 3.3% underestimation in soybean ending stocks, while only in a 0.2% underestimation for corn ending stocks. This finding implies that ending stocks errors in soybean forecasts are more sensitive to errors in individual balance sheet categories than those in corn forecasts.

A much different picture is presented in figures 5 and 6, which show the estimated elasticities obtained from the multiple regressions of soybean and corn average price percentage forecast errors over errors in other balance sheet categories (equation 6). The major difference from the previous results is that all estimated elasticities are relatively small. While all estimated coefficients had expected signs, with the exception of feed seed & residual use, only production and exports forecast errors were significant in explaining errors in soybean average price forecasts in the first half of the forecasting cycle. Similarly, in corn the signs of estimated coefficients were correct, but most variables, except exports and feed & residual use (in 3 and 5 out of 17 forecast months, respectively), were insignificant at the 5% confidence level in explaining errors in average price forecasts. The estimated elasticities indicate that during early reports, when soybean production is overestimated by 1%, average price tends to be underestimated by 1.3%. Also, when soybean exports are overestimated b 1%, average price is also overestimated by approximately 0.4%. For corn, all estimated elasticities are less than 1, illustrating that an impact of errors in balance sheet categories on average price forecast errors is relatively small.

Furthermore, the explanatory power of the average price equations was much smaller than that of ending stocks equations. As illustrated in figures 7 and 8, joint variation in the supply and demand independent variables explained around 85% of the variation in ending stocks forecast errors. In contrast, estimated R^2 for average price equations averaged only about 37%. These results suggest while errors in USDA ending stocks forecasts are directly traceable to errors in individual supply and demand categories, errors in average price forecasts for the most part appear to be a consequence of the factors other than forecast errors in balance sheet categories.

Summary and Conclusions

The purpose of this study was to provide a comprehensive evaluation of WASDE forecast accuracy over the 1980/81 through 2003/04 marketing years. The comprehensive nature of this study was based on including all supply and consumption components of U.S. corn and soybean balance sheets. Specifically, the analysis focused on two issues:

1. To test whether USDA's forecast accuracy (in terms of bias and variability) of each balance sheet category has changed during the study period.
2. To identify whether errors in individual balance sheet categories caused errors in ending stocks and average price forecasts.

This study used the data from monthly WASDE balance sheets for U.S. corn and soybeans over 1980/1981 through 2003/2004 to calculate forecast errors for each balance sheet category. Trends in the mean and the variance of percentage forecast errors for each category were analyzed using Bailey and Brorsen's (1998) approach. A two-equation model of forecast errors was estimated, one equation for the mean of percentage forecast errors and one equation for the variance of percentage forecast errors. Results obtained from the analysis were consistent with those obtained in the descriptive analysis of forecast accuracy and suggest that soybean ending stocks were biased toward overestimation during the last years of the study period and particularly early in the forecast cycle. Additionally, this analysis revealed that soybean average price forecasts were biased toward underestimation during the last years of the study period.

A significant downward trend in the variance of the forecast errors was observed for all categories with respect to the forecast horizon. The absolute size and consequently, the variability of the forecast errors significantly diminish as approaching the final report and consequently, the variance of percentage errors was a decreasing function of the forecast horizon. Early forecasts of soybean total use and for corn average price were found to be more accurate, or less variable, during the last years of the sample period than during the initial years. Although the improvement in forecast accuracy was only significant for these two categories, the fact that negative slope estimates for the marketing year variable were observed in almost all the categories included in the balance sheets suggests that forecast accuracy has improved.

The second objective of this study was to identify whether errors in individual balance sheet categories caused errors in ending stocks and average price forecasts. This analysis was carried out by regressing percentage forecast errors in individual balance sheet categories against ending stocks and average price forecast errors. The findings reveal that only during early reports were soybean production and exports percentage errors significantly related to forecast errors in average price. On the other hand, almost all the individual categories were significant in explaining errors in soybean ending stocks. Similarly, only corn exports and feed & residual use were sporadically significant in explaining errors in corn average price forecasts, while almost all categories were significant in explaining errors in ending stocks. Interestingly, for a given category, the absolute value of the estimated elasticities was higher for soybeans than for corn both for average price and for ending stocks errors. In addition, errors in corn and soybean ending stocks were found to be significantly and negatively related to errors in average price early in the forecasting cycle. Furthermore, the joint variation in the individual balance sheet categories explained around 85% of the variation in ending stocks forecast errors, and only 35% to 50% of the variation in average price forecasts.

Overall, the results of this study suggest that USDA performed reasonably well in generating supply and demand estimates for U.S. corn and soybeans. However, soybean ending stocks forecasts errors have significantly increased in absolute size during recent years. A tendency to overestimate soybean ending stocks was observed during this period. Furthermore, it is likely that the observed bias in early soybean average price forecasts is a consequence of the bias in ending stocks forecasts. Limited impact of the individual balance sheet categories and the low explanatory power of the average price regressions, suggest that forces other than errors in balance sheet estimates were affecting USDA price forecast performance. The unexplained 50% to 65% of the variation in price forecast errors can be a consequence of judgmental biases on the part of USDA analysts and/or of a mis-specification in the model used for forecasting prices. No matter which is the main reason, these results illustrate the difficulties entailed in forecasting U.S. corn and soybean prices.

References

Arellano, M. “Computing Robust Standard Errors for Within-groups Estimators.” *Oxford Bulletin of Economics and Statistics*, 49(1987):431-434.

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):515-525.

Doane’s Agricultural Report, Vol. 67, No. 35-1, August 27, 2004.

Egelkraut, T.M., P. Garcia, S.H. Irwin, and D. Good. “An Evaluation of Crop Forecast Accuracy for Corn and Soybeans: USDA and Private Information Agencies.” *Journal of Agricultural and Applied Economics*, 35(2003):79-95.

Gunnelson, G., W. Dobson, and S. Pamperin. “Analysis of the Accuracy of USDA Crop Forecasts.” *American Journal of Agricultural Economics*, 54(1972):639-645.

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.

Isengildina, O., S.H. Irwin, and D.L. Good. “Evaluation of USDA Interval Forecasts of Corn and Soybean Prices.” *American Journal of Agricultural Economics*, 86(2004):990-1004.

Isengildina, O., S.H. Irwin, and D.L. Good. “Are Revisions to USDA Crop Production Forecasts Smoothed?” *American Journal of Agricultural Economics,* in press.

Just, R.E., and G.C. Rausser. “Commodity Price Forecasting with Large-Scale Econometric Models and the Futures Market.” *American Journal of Agricultural Economics,* 63(1981):197-208.

Marquardt. R., and A.F. McGann. “Forecasting Commodity Prices.” *Commodity Journal,* 10(1975):29-33.

Morris, S. and H. Y. Shin. "Social Value of Public Information." *American Economic Review* 92(2002):1521-1534.

Offutt, S. "Finding the Keys to Federal Statistical Programs," *The Exchange: The Newsletter of the American Agricultural Economics Association*, March/April 2002, p.1.

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):114-127.

Spilka, W. "An Overview of the USDA Crop and Livestock Information System." *Journal of Futures Markets*, 3(1983):167-176.

Thompson, J.M. "Analysis of the Accuracy of USDA Hog Farrowings Statistics." *American Journal of Agricultural Economics,* 34(1974):1213-1217.

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

Westcott, P.C. and L.A. Hoffman. "Price Determination for Corn and Wheat: The Role of Market Factors and Government Programs." US Department of Agriculture, Economic Research Service, 1999.

White, H. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica,* 48(1980):817-838.

Wooldridge, J.M. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Massachusetts: The MIT Press, 2002.

Wolf, S., D. Just, and D. Zilberman. "Between Data and Decisions: the Organization of Agricultural Economic Information Systems." *Research Policy* 30 (2001):121-141.

Endnotes.

[1] This section draws heavily from Spilka (1984) and Vogel and Bange (1999).

[2] WASDE reports became available in the marketing year 1973/1974. However, the first WASDE forecasts did not include all the categories currently available. For example, the price forecasts were first published in 1976/1977. The sample period included in this analysis goes from 1980/1981 through the 2003/2004 marketing year so as to avoid years with missing observations.

[3] Results not presented here but available from authors upon request.

Table 1. Mean Percentage Forecast Errors for U.S. Corn WASDE Balance Sheet Components 1980/81-2003/04 Marketing Years

	Production	Total Supply	Feed & Residual	Food, Seed & Industrial	Exports	Total Use	Ending Stocks	Average Price
May	-5.40	-2.80	-1.40	0.83	-6.00	-1.70	-24.00	2.41
	-(1.63)	-(1.20)	-(0.77)	(0.90)	-(1.52)	-(1.28)	-(1.92)	(0.79)
June	-4.80	-2.30	-1.00	0.94	-6.00	-1.50	-20.00	1.39
	-(1.47)	-(1.03)	-(0.60)	(1.06)	-(1.52)	-(1.14)	-(1.70)	(0.46)
July	-2.60	-1.10	-0.09	0.95	-5.70	-0.87	-15.00	-0.28
	-(0.99)	-(0.54)	-(0.06)	(1.16)	-(1.47)	-(0.72)	-(1.43)	-(0.10)
August	-0.21	0.26	0.34	1.21	-5.30	-0.49	-3.50	-2.20
	-(0.13)	(0.20)	(0.27)	(1.67)	-(1.44)	-(0.45)	-(0.44)	-(0.98)
September	0.93	0.86	0.56	1.25	-4.50	-0.22	1.39	-2.20
	(0.91)	(0.98)	(0.49)	(1.85)	-(1.28)	-(0.22)	(0.25)	-(1.15)
October	0.62	0.50	0.16	1.40 *	-4.90	-0.54	0.94	-1.60
	(0.91)	(0.84)	(0.14)	(2.23)	-(1.44)	-(0.53)	(0.17)	-(0.95)
November	0.33	0.33	0.52	1.36	-4.30	-0.24	0.32	-0.47
	(1.03)	(1.19)	(0.46)	(2.59)	-(1.43)	-(0.25)	(0.06)	-(0.31)
December	0.33	0.33	0.46	1.24	-3.70	-0.17	0.58	-0.28
	(1.03)	(1.17)	(0.41)	(2.33)	-(1.28)	-(0.18)	(0.11)	-(0.22)
January			-0.40	0.67	-3.20	-0.72	5.24	-0.47
			-(0.43)	(1.34)	-(1.30)	-(0.89)	(1.54)	-(0.44)
February			-0.66	0.27	-2.50	-0.84	6.25	-0.46
			-(0.76)	(0.60)	-(1.13)	-(1.13)	(2.11)	-(0.50)
March			-0.66	0.27	-2.00	-0.73	5.88	-0.10
			-(0.76)	(0.60)	-(1.20)	-(1.12)	(2.18)	-(0.13)
April			-0.69	-0.06	-1.20	-0.71	5.55	-0.48
			-(1.03)	-(0.21)	-(0.94)	-(1.42)	(2.37)	-(0.71)
May			-0.64	-0.05	-1.00	-0.67	4.86	-0.46
			-(1.09)	-(0.17)	-(0.91)	-(1.62)	(2.45)	-(0.77)
June			-0.66	0.03	-0.81	-0.64	4.69	-0.41
			-(1.13)	(0.10)	-(0.84)	-(1.57)	(2.50)	-(0.81)
July			-0.46	0.06	-0.45	-0.44	3.79	-0.21
			-(0.96)	(0.24)	-(0.58)	-(1.14)	(1.86)	-(0.79)
August			-0.22	-0.06	0.05	-0.14	1.66	-0.13
			-(0.52)	-(0.29)	(0.11)	-(0.46)	(1.15)	-(0.66)
September			-0.06	0.00	-0.01	-0.01	0.58	-0.17
			-(0.16)	(0.02)	-(0.03)	-(0.06)	(0.64)	-(1.07)
October			0.12	0.01	-0.23	0.03	-0.09	0.00
			(0.34)	(0.06)	-(1.92)	(0.13)	-(0.12)	(0.00)

Notes: A single and double asterisks (*) denote significantly different from zero at the 5% and 1% levels respectively.
Numbers in parenthesis are standard errors

Table 2. Mean Percentage Forecast Errors for U.S. Soybean WASDE Balance Sheet Components 1980/81-2003/04 Marketing Years

	Production	Total Supply	Crush	Exports	Feed, Seed & Residual	Total Use	Ending Stocks	Average Price
May	-2.70	-2.90	0.14	-2.20	0.86	-0.31	-36.00 **	2.81
	-(1.26)	-(1.49)	(0.11)	-(0.59)	(0.22)	-(0.16)	-(3.46)	(0.94)
June	-2.80	-2.90	-0.03	-2.10	0.71	-0.43	-35.00 **	2.67
	-(1.32)	-(1.50)	-(0.02)	-(0.60)	(0.18)	-(0.23)	-(3.51)	(0.90)
July	-2.40	-2.40	0.22	-1.30	0.58	-0.02	-32.00 **	2.25
	-(1.39)	-(1.44)	(0.18)	-(0.44)	(0.15)	-(0.01)	-(3.43)	(0.87)
August	-1.40	-1.40	0.79	0.47	0.05	0.91	-27.00 **	0.53
	-(1.05)	-(1.11)	(0.78)	(0.18)	(0.01)	(0.73)	-(3.28)	(0.24)
September	-0.20	-0.15	1.15	1.16	0.82	1.40	-16.00 *	-1.00
	-(0.20)	-(0.16)	(1.48)	(0.47)	(0.25)	(1.39)	-(2.33)	-(0.51)
October	0.11	0.00	1.15	0.97	1.91	1.37	-15.00 *	-0.88
	(0.17)	(0.01)	(1.75)	(0.42)	(0.68)	(1.71)	-(2.24)	-(0.52)
November	-0.25	-0.30	1.15	0.54	1.69	1.20	-16.00 *	-0.51
	-(0.56)	-(0.89)	(1.87)	(0.25)	(0.61)	(1.66)	-(2.45)	-(0.33)
December	-0.25	-0.30	1.06	0.05	1.69	0.96	-13.00 *	-0.75
	-(0.56)	-(0.89)	(1.72)	(0.02)	(0.61)	(1.41)	-(2.25)	-(0.63)
January			1.03	-0.08	1.46	0.86	-11.00 *	-0.25
			(1.63)	-(0.04)	(0.54)	(1.29)	-(2.34)	-(0.30)
February			1.01	0.08	-0.44	0.78	-8.90	0.12
			(1.89)	(0.05)	-(0.13)	(1.21)	-(1.80)	(0.18)
March			1.09 *	0.59	-0.33	1.00	-10.00 *	0.73
			(2.22)	(0.44)	-(0.10)	(1.62)	-(2.13)	(1.43)
April			1.11 *	-0.06	-4.80	0.46	-5.00	0.55
			(2.73)	-(0.05)	-(1.79)	(0.94)	-(1.28)	(1.28)
May			0.86 *	-0.01	-6.30	0.24	-2.90	0.40
			(2.38)	-(0.02)	-(1.75)	(0.54)	-(0.70)	(1.03)
June			0.76 *	0.08	-6.30	0.20	-2.30	0.29
			(2.52)	(0.10)	-(1.75)	(0.50)	-(0.61)	(0.92)
July			0.51 *	0.09	-4.80 *	0.10	0.75	0.15
			(2.39)	(0.15)	-(2.19)	(0.35)	(0.24)	(0.83)
August			0.32 **	0.28	-5.20 *	0.02	1.94	0.08
			(2.91)	(1.22)	-(2.36)	(0.12)	(0.71)	(0.99)
September			0.02	-0.06	-4.60	-0.24	4.18	0.08
			(0.41)	-(0.40)	-(1.85)	-(1.87)	(1.72)	(1.75)
October			0.02	-0.15	0.37	0.00	0.00	0.02
			(0.94)	-(1.60)	(0.63)	(0.06)	(0.00)	(1.00)

Notes: A single and double asterisks (*) denote significantly different from zero at the 5% and 1% levels respectively.
Numbers in parenthesis are standard errors

Table 3. Estimated Trends in the Mean of Percentage Forecast Errors for U.S. Corn and Soybean WASDE Balance Sheets Components, 1980/81-2003/04 Marketing Years.

	Soybeans				Corn			
	Intercept	**Horizon**	**Year**	**Interaction**	**Intercept**	**Horizon**	**Year**	**Interaction**
Production	-2.17	0.12	0.02	0.00	-3.41	0.28	0.09	-0.01
	(2.91)	(0.24)	(0.20)	(0.02)	(4.32)	(0.32)	(0.30)	(0.02)
Crush	-0.72	0.04	0.12	0.00	-	-	-	-
	(2.35)	(0.15)	(0.16)	(0.01)				
Feed Seed & Residual	7.40	-0.82	-0.40	0.03	-	-	-	-
	(9.22)	(0.58)	(0.65)	(0.04)				
Feed & Residual	-	-	-	-	-2.02	0.10	0.14	-0.01
					(3.38)	(0.20)	(0.24)	(0.01)
Food, Seed & Industrial	-	-	-	-	1.98	-0.10	-0.04	0.00
					(1.90)	(0.13)	(0.13)	(0.01)
Exports	-9.25	0.55	0.68	-0.04	-13.31	0.75	0.52	-0.03
	(6.81)	(0.42)	(0.48)	(0.03)	(9.23)	(0.58)	(0.65)	(0.04)
Total Use	-3.08	0.17	0.30	-0.02	-4.32	0.23	0.26	-0.01
	(3.14)	(0.21)	(0.22)	(0.01)	(2.67)	(0.16)	(0.19)	(0.01)
Ending Stocks	11.72	-0.65	-3.66 **	0.23 **	-2.62	0.87	-0.77	0.02
	(14.84)	(1.11)	(1.04)	(0.08)	(19.45)	(1.48)	(1.36)	(0.10)
Average Price	-7.98	0.59	0.74 *	-0.05	-3.49	0.21	0.26	-0.02
	(4.28)	(0.30)	(0.30)	(0.02)	(4.93)	(0.34)	(0.34)	(0.02)

Notes: A single and double asterisks (*) denote significantly different from zero at the 5% and 1% levels respectively.
Numbers in parentheses are standard errors.

Table 4. Estimated Trends in the Variance of Percentage Forecast Errors for U.S. Corn and Soybean WASDE Balance Sheets Components, 1980/81-2003/04 Marketing Years.

	Soybeans				Corn			
	Intercept	**Horizon**	**Year**	**Interaction**	**Intercept**	**Horizon**	**Year**	**Interaction**
Production	4.26 **	-0.34 **	-0.11	0.01	4.09 **	-0.32 **	-0.10	0.00
	(0.92)	(0.08)	(0.06)	(0.01)	(0.96)	(0.07)	(0.07)	(0.00)
Crush	3.58 **	-0.34 **	-0.06	0.01	-	-	-	-
	(0.91)	(0.05)	(0.06)	(0.00)				
Feed Seed & Residual	2.79 *	0.08	0.11	-0.01	-	-	-	-
	(1.17)	(0.08)	(0.08)	(0.01)				
Feed & Residual	-	-	-	-	3.81 **	-0.13	-0.08	0.00
					(1.07)	(0.07)	(0.07)	(0.00)
Food, Seed & Industrial	-	-	-	-	0.84	-0.12	0.05	-0.01
					(1.08)	(0.08)	(0.08)	(0.01)
Exports	6.80 **	-0.39 **	-0.11	0.00	5.35 **	-0.26 **	0.03	0.00
	(0.82)	(0.05)	(0.06)	(0.00)	(1.12)	(0.09)	(0.08)	(0.01)
Total Use	5.70 **	-0.42 **	-0.17 **	0.01 *	3.59 **	-0.19 *	-0.05	-0.01
	(0.91)	(0.07)	(0.06)	(0.00)	(1.15)	(0.09)	(0.08)	(0.01)
Ending Stocks	6.86 **	-0.17 *	-0.02	0.00	7.55 **	-0.23 **	-0.08	0.00
	(0.76)	(0.07)	(0.05)	(0.01)	(0.84)	(0.06)	(0.06)	(0.00)
Average Price	5.17 **	-0.47 **	-0.05	0.01	6.57 **	-0.54 **	-0.11 **	0.01 **
	(0.80)	(0.08)	(0.06)	(0.01)	(0.53)	(0.05)	(0.04)	(0.00)

Notes: Simple and double asterisks (*) denote significantly different from zero at the 5% and 1% levels respectively.
Numbers in parentheses are standard errors.

Figure 1. Predicted Standard Deviation for Corn Average Price Percentage Forecast Errors, 1981/82 and 2001/02 Marketing Years

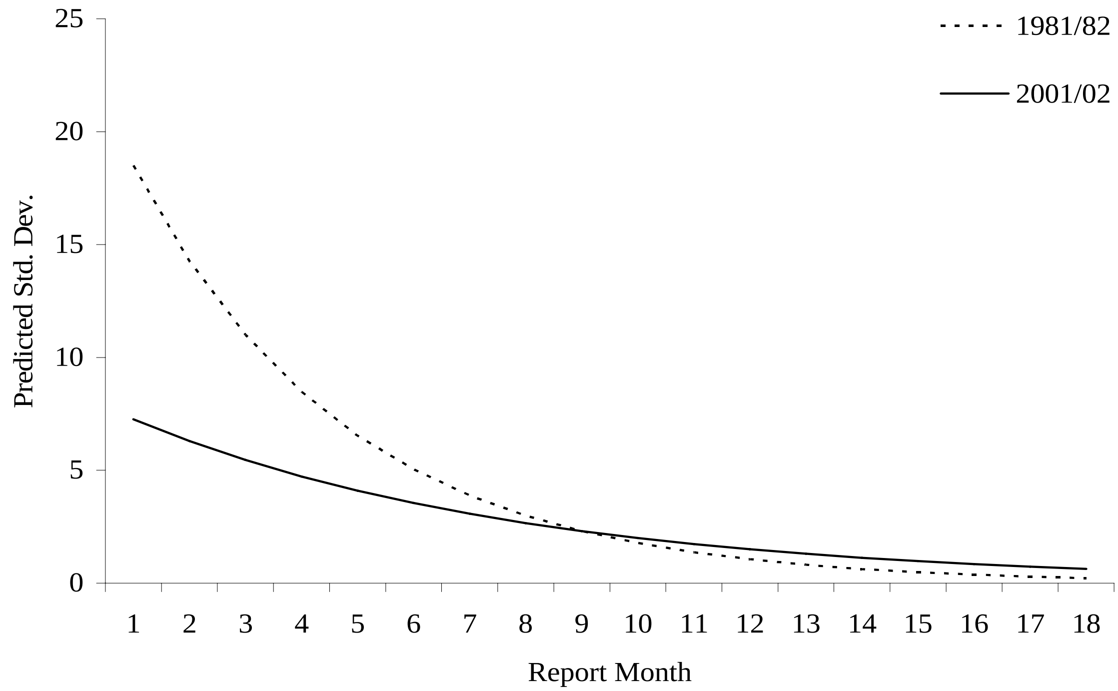

Figure 2. Predicted Standard Deviation for Soybean Average Price Percentage Forecast Errors, 1981/82 and 2001/02 Marketing Years

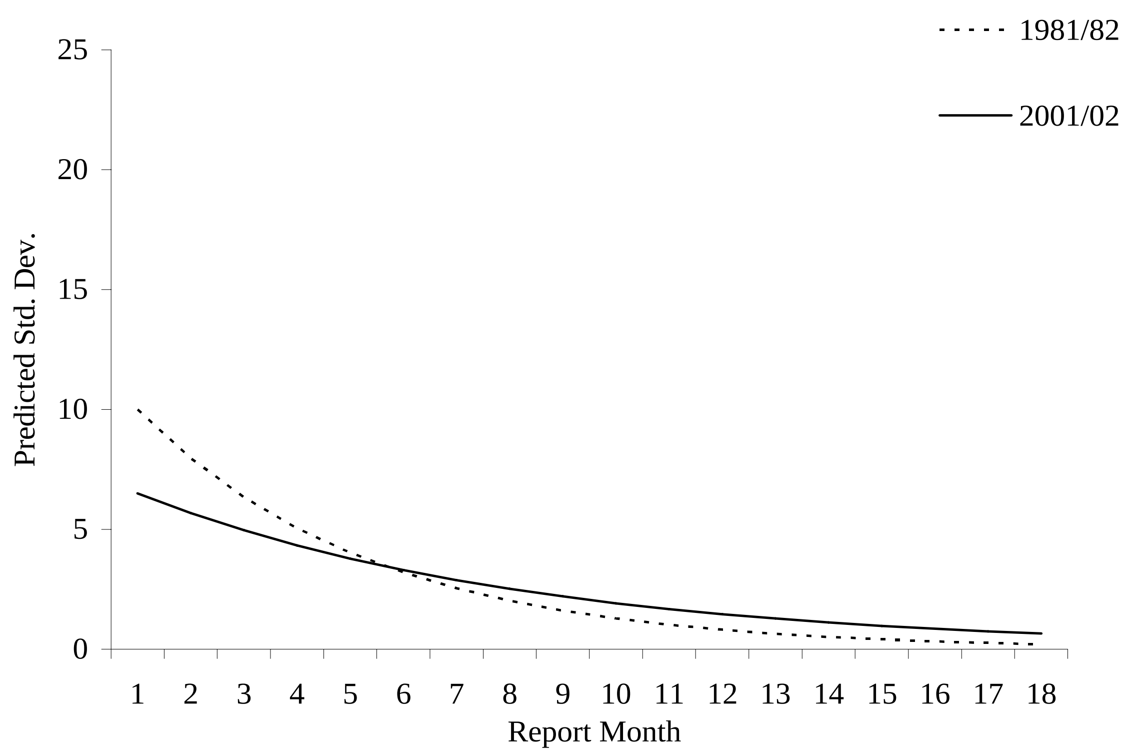

Figure 3. WASDE Supply and Consumption Forecast Error Elasticities for Soybean Ending Stocks Forecast Errors, 1980/81-2003/04

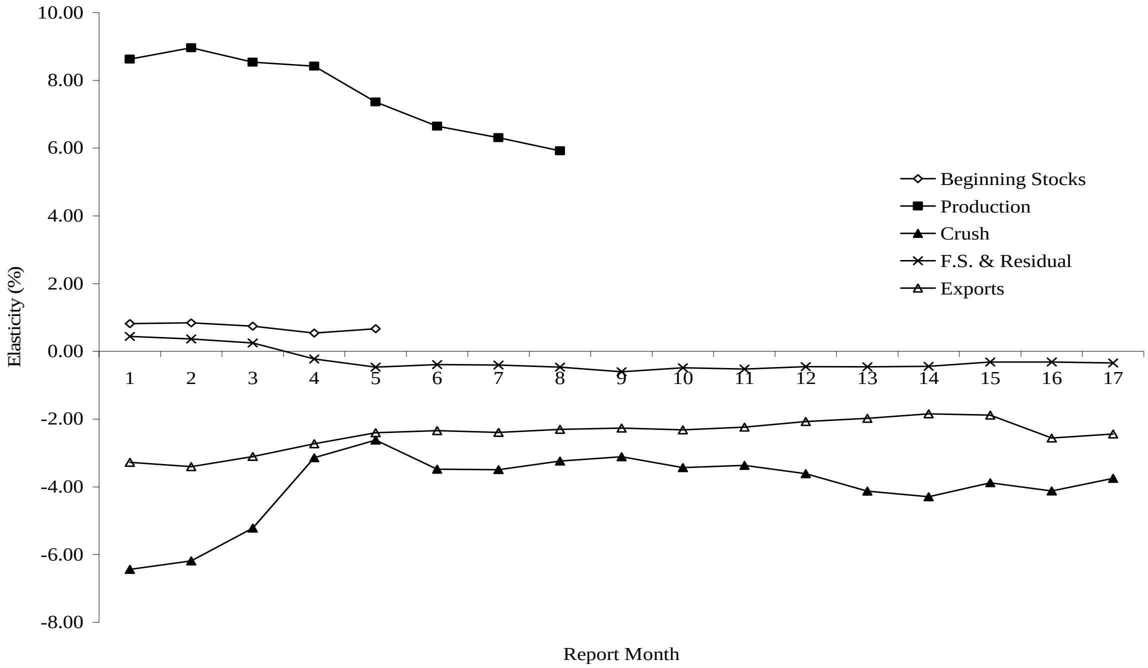

Figure 4. WASDE Supply and Consumption Forecast Error Elasticities for Corn Ending Stocks Forecast Errors, 1980/81-2003/04

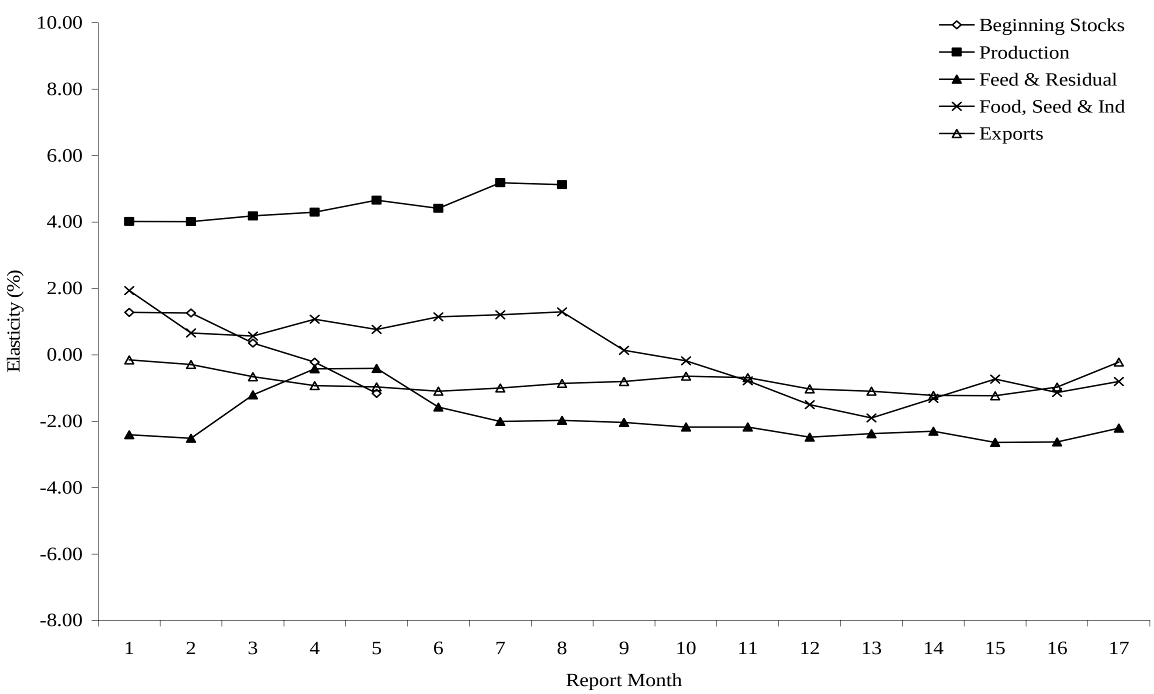

Figure 5. WASDE Supply and Consumption Forecast Error Elasticities for Soybean Price Forecast Errors, 1980/81-2003/04

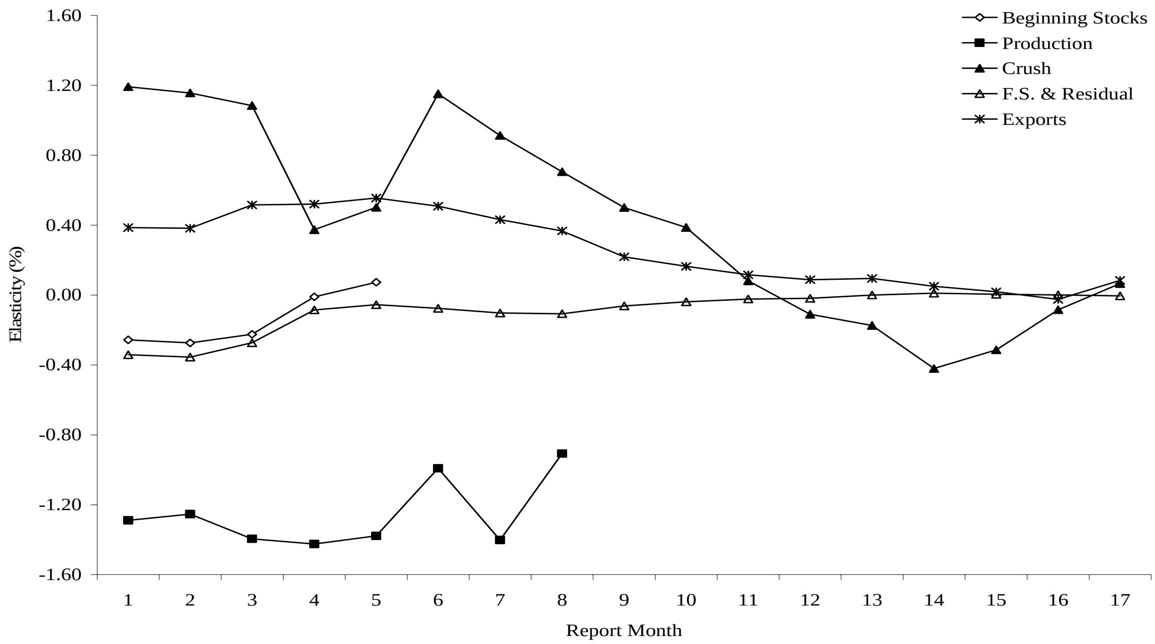

Figure 6. WASDE Supply and Consumption Forecast Error Elasticities for Corn Price Forecast Errors, 1980/81-2003/04

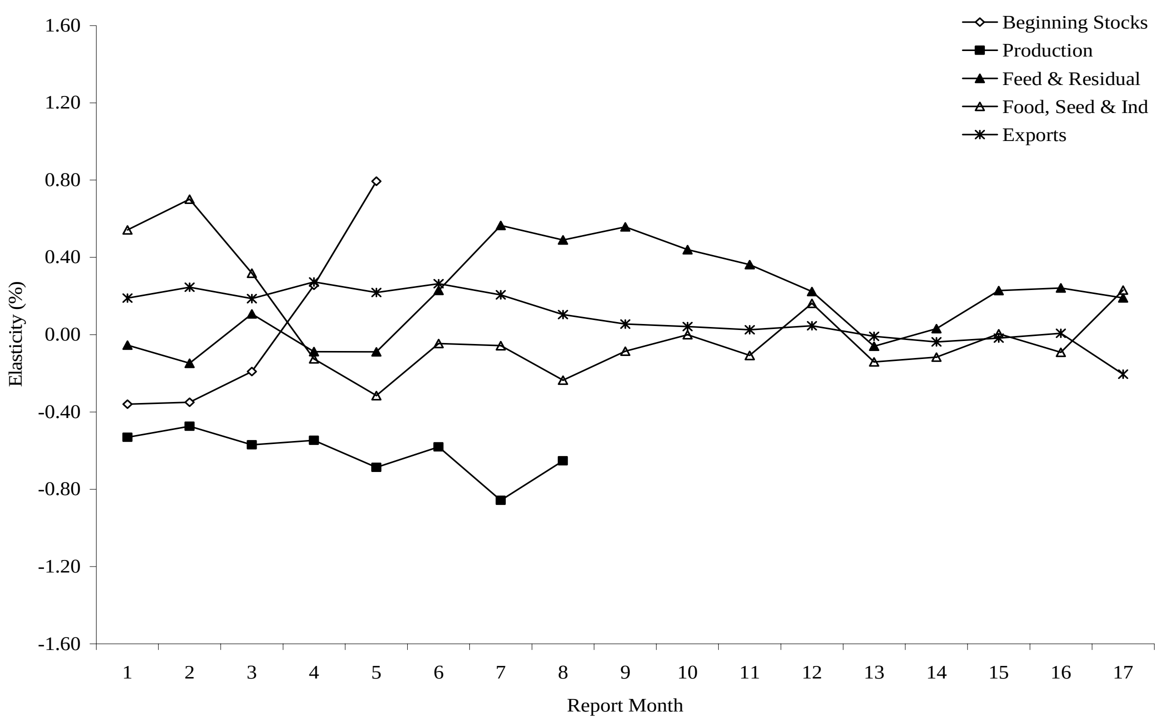

Figure 7. Coefficients of Determination for OLS Regressions of Sources of Errors in Soybean Ending Stocks and Price Forecasts, 1980/81-2003/04

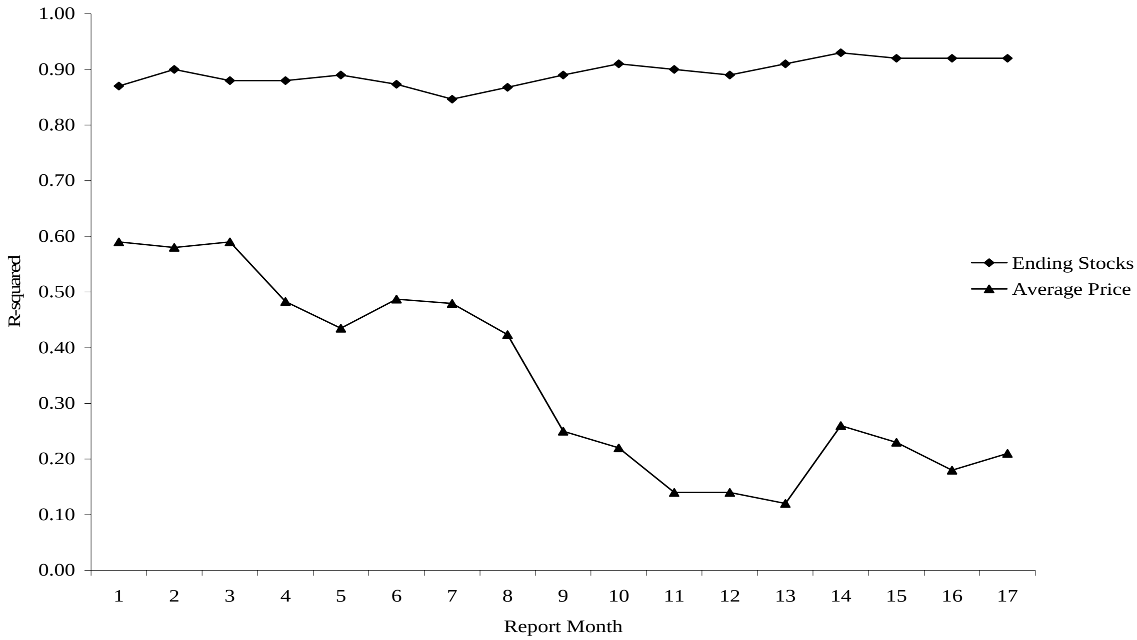

Figure 8. Coefficients of Determination for OLS Regressions of Sources of Errors in Corn Ending Stocks and Price Forecasts, 1980/81-2003/04