*Journal of Agricultural and Applied Economics,* 45,1(February 2013):95–107


# Do Big Crops Get Bigger and Small Crops Get Smaller? Further Evidence on Smoothing in U.S. Department of Agriculture Forecasts

**Olga Isengildina, Scott H. Irwin, and Darrel L. Good**

This study sought to determine whether monthly revisions of U.S. Department of Agriculture current-year corn and soybean yield forecasts are correlated and whether this correlation is associated with crop size. An *ex-ante* measure of crop size based on percent deviation of the current estimate from out-of-sample trend is used in efficiency tests based on the Nordhaus framework for fixed-event forecasts. Results show that available information about crop size is generally efficiently incorporated in these forecasts. Thus, although this pattern may appear obvious to market analysts in hindsight, it is largely based on new information and hence difficult to anticipate.

*Key Words:* corn, crop size efficiency, fixed-event forecasts, independence, revisions, smoothing, soybeans, yield forecasts

**JEL Classifications:** Q10, Q13, E37

"There's an old saying in this business that big crops get bigger and small crops get smaller. . .I guess I'm in the camp that believes the 10.3 million bushel [USDA] corn estimate is a little bit too high."

—Tom Mueller, Taylor Ridge, IL, farmer (*Quad-City Business Journal*, 2005)

"The popular wisdom in the trade is that this corn crop may get smaller yet, tightening the fundamental structure even more. Traders are adhering to the adage, 'small crops get smaller'."

(FarmWeek, October 18, 2010)

Participants in volatile agricultural markets have relied on information from U.S. Department of Agriculture (USDA) situation and outlook reports for many decades. There is considerable evidence these reports move the markets (e.g., Baur and Orazem, 1994; Colling and Irwin, 1990; Fortenbery and Sumner, 1993; McKenzie, 2008; Sumner and Mueller, 1989) and the reports are usually widely anticipated. Because of their importance, there is a substantial body of literature devoted to analyzing the accuracy and efficiency of USDA forecasts (e.g., Bailey and Brorsen, 1998; Isengildina, Irwin, and Good, 2006; Sanders and Manfredo, 2002). A particular emphasis is placed on crop production forecasts. As described by Tom Polansek in his *Wall Street Journal* article on October 22, 2010, "Traders and analysts depend on the agency's crop reports, and while forecasting isn't an exact science, USDA estimates have historically been on target. The crop predictions, which come out at set times

---

Olga Isengildina is an associate clinical professor, Department of Economics, University of Texas at Arlington, Arlington, Texas. Scott H. Irwin is the Laurence J. Norton professor of Agricultural Marketing, University of Illinois at Urbana–Champaign, Champaign, Illinois. Darrel L. Good is a professor, Department of Agricultural and Consumer Economics, University of Illinois at Urbana–Champaign, Champaign, Illinois.

We thank the editor, Darrell Bosch, and three anonymous reviewers for their comments, which helped to improve the article substantially.

each month, are the subject of intense scrutiny, and often cause immediate price swings. They influence what a farmer will plant, how much hedging a farmer will do, and whether an investor will buy or sell."

Conventional analyses of forecast accuracy cited previously do not answer the question of how forecasts change during a forecasting cycle. Systematic under- and overadjustments are revealed through analysis of forecast revisions. The framework for analysis of efficiency in forecast revisions was developed by Nordhaus (1987). In this context, systematic underadjustments of the forecasts are termed "smoothing" and overadjustments are called "jumpiness." Detection of systematic adjustments in forecasts is of interest because it implies that: 1) if forecast revisions are correlated, then forecasts do not efficiently incorporate all available information and, therefore, may be improved; and 2) knowledge about systematic adjustments can be used by market participants to anticipate future revisions.

Two previous studies examined the revisions process for USDA crop production forecasts. Gunnelson, Dobson, and Pamperin (1972) analyzed first and second revisions for seven U.S. crops over 1929–1970 and reported, "While a relatively high percentage of the revisions was successful, the revised forecasts tended to under compensate for the errors in the previous estimate. Thus, for example, if first crop forecasts underestimated or overestimated crop size, the first revision was likely to exhibit similar characteristics" (pp. 641–42). In a more recent study, Isengildina, Irwin, and Good (2006) found that revisions to USDA corn and soybean crop production forecasts over 1970–2005 were positively correlated and consistent in their direction. For example, directional tests revealed that positive monthly revisions in corn forecasts were followed by positive revisions 79% of the time and negative revisions were followed by negative revisions 56% of the time. In soybeans, positive and negative monthly revisions were followed by revisions in the same direction 66% of the time. Positive correlation between consecutive revisions was found in all cases in corn and all but October and January revisions of soybean production forecasts. Correlation coefficients were slightly higher in corn, ranging from 0.25 to 0.68, relative to 0.14–0.26 in soybeans. Such pattern of systematic movements in forecast revisions implies that not all information available at the time the forecast is made is incorporated in the forecast and part of it is carried over into the next forecast, which results in (partial) predictability of future revisions.

The quotes at the beginning of this article illustrate that market participants believe USDA crop production forecasts evolve in a particular manner, i.e., "big" crop forecasts tend to get bigger and "small" crop forecasts tend to get smaller across the forecasting cycle. Although previous studies show evidence of smoothing in these crop production forecasts, their findings demonstrate average smoothing across all years within the study period and do not address the question of whether smoothing is concentrated in years with relatively small and large crops. Because production forecasts (especially later in the production cycle) are largely driven by yield forecasts (Good and Irwin, 2006), this study focuses on USDA yield forecasts. Therefore, the goals of this article are 1) to establish whether revisions of USDA corn and soybean yield forecasts are correlated; and 2) to determine whether this correlation is associated with crop size.

This study uses data for USDA corn and soybean yield forecasts from 1970 through 2010. Cumulative forecast revisions that summarize changes in initial forecasts throughout the forecasting cycle are used to illustrate the *ex-post* pattern of big crops getting bigger and small crops getting smaller. An *ex-ante* measure of crop size based on percent deviation of the current estimate from out-of-sample trend is used in efficiency tests based on the Nordhaus framework for fixed-event forecasts. Results show that available information about crop size is generally efficiently incorporated in these forecasts.

## U.S. Department of Agriculture Crop Forecasts

Detailed descriptions of USDA crop estimating and forecasting procedures are provided by Good and Irwin (2006, 2011), Isengildina, Irwin, and Good (2006), USDA/NASS (2006), and Vogel and Bange (1999). We present a short summary of this process as it relates to this study. The first yield and production forecast for each marketing year is released in May preceding the U.S. marketing year (September through August for corn and soybeans) within the World Agricultural Supply and Demand Estimates (WASDE) report. The yield and production forecast is updated[1] in each subsequent monthly WASDE report until the final estimate is published in the January report. May–July forecasts are methodologically very different from August–November forecasts, whereas they are based on historical trend analysis (by the World Agricultural Outlook Board) rather than National Agricultural Statistical Service (NASS) estimates. NASS yield and production forecasts are released simultaneously in WASDE and *Crop Production* reports from August through November with finalized estimates released in January.[2] In addition to yield and production forecasts, the *Prospective Plantings* report (currently released at the end of March) and the *Acreage* report (currently released at the end of June) contribute USDA information about the expected size of corn and soybean crops.

USDA forecasts of corn and soybean yields over the 1970 through 2010 crop years are examined in this study. August through November forecasts are based on data collected from the monthly Agricultural Yield Surveys (AYS) and the Objective Yield Surveys (OYS). AYS data are based on responses of a stratified sample of crop producers across the country regarding acreage they expect to harvest as well as final yield for each crop. NASS analysts are aware of the judgment-based nature of this information and try to correct for any known biases in its use for official forecasts. OYS data reflect enumerator (NASS employee) measurements from a sample of fields (with crop acreage reported within the June Acreage Survey) of factors such as plant and fruit counts, fruit size, weight, and condition as well as the final yield and harvest loss. Yield forecast is based on a regression analysis of the historical relations (15 years) between the yield factors (such as [expected] fruit count, weight, and harvest loss) and the state average yield. Forecasts are based on conditions as of the survey date and projected assuming normal weather conditions for the remainder of the growing season. Data from both survey sources described previously are compared and combined with other available data on the state and national level to provide the official yield. Data for the final yield estimate released in January are collected in the December Agricultural Survey in which respondents report actual acres harvested and the actual yield or production.

## Data and Descriptive Analysis

USDA corn and soybean yield forecasts are considered fixed-event forecasts because the series of forecasts are related to the same terminal event, $q_T^i$, where $T$ is the release month (January) for the final estimate of crop yield in the $i^{\text{th}}$ year. The forecast of the terminal event for month $t$ is denoted as $q_t^i$, where $t = 1$: August, 2: September, 3: October, 4: November, 5: January and $i = 1970, \ldots, 2010$. This layout of fixed-event forecasts and corresponding revision process is illustrated for the USDA corn and soybean yield forecasting cycle in Figure 1. To standardize for increasing crop size over time, revisions are examined in log percentage form:

$$v_t^i = 100 \times \ln\left(q_t^i / q_{t-1}^i\right) \qquad t = 2, \ldots, 5; \quad i = 1970, \ldots, 2010, \tag{1}$$

where the forecasting cycle has a length of $T = 5$, and the revision cycle has a length of $T$-1 = 4 months for both crops.

[1] NASS officials do not approach the forecasting task in terms of revisions. Instead, NASS attempts to make the best possible interpretation of production potential each month on the basis of available information to minimize forecast error. In other words, NASS starts with a clean slate and makes a new forecast each month rather than altering the previous forecast in some way.

[2] The December WASDE report does not update November yield and production estimates. The January "final" estimates are often subsequently revised. This happens most frequently in January after the end of the marketing year. As a result of the sporadic nature and long time lag of the subsequent revisions, they are not considered in this analysis.

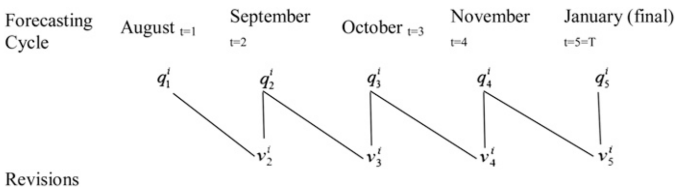


**Figure 1.** U.S. Department of Agriculture Corn and Soybean Production Forecasting Cycle and Forecast Revisions

Descriptive statistics on monthly revisions of USDA corn and soybean yield forecasts shown in Table 1 demonstrate that the absolute magnitude of forecast revisions decreased from the beginning to the end of the forecasting cycle. The first (September) revision was as large as 16% and 17.6% of final estimated U.S. yield in corn and soybeans, respectively. In general, there was a similarity in the magnitude and range of corn and soybean yield forecast revisions with soybean revisions being slightly larger than corn revisions.

Bias in revisions is examined using a Student's *t* test and a sign test. The *t* test examines if average revisions for a particular month are statistically different from zero. The nonparametric sign test examines whether revisions are equally probable to be positive or negative:

$$(2) \qquad J = \left[\frac{N^+}{N} - 0.5\right]\frac{N^{1/2}}{0.5} \sim \mathrm{N}(0,1),$$

where $N^+$ is the number of cases with positive cumulative revisions is and $N$ is a total number of cases. According to the *t* test, yield forecast revisions on average were not different from zero with one exception—November soybean revisions had an average value of 0.60%, significantly greater than zero at the 5% level. Although most of November revisions were positive for both corn and soybeans, according to the sign test, the difference was not significant at the 10% level. The general lack of bias in forecast revisions also implies no bias in forecasts themselves, because revisions are easily traced back to the forecasts (Nordhaus, 1987). However, limited evidence of bias in USDA corn and soybean yield forecast revisions

**Table 1.** Descriptive Statistics and Test of Bias for Revisions of U.S. Department of Agriculture Corn and Soybean Yield Forecasts, 1970–2010.

| Revision Month | Mean Absolute Value | Mean | Standard Deviation | Minimum | Maximum | Test of Bias *t*-statistic | Test of Bias Sign Test |
|---|---|---|---|---|---|---|---|
| Panel A: corn | | ——Percent—— | | | | | |
| September | 2.04 | –0.31 | 3.33 | –16.03 | 4.28 | –0.59 | 0.46 |
| October | 1.82 | 0.29 | 2.34 | –5.65 | 5.85 | 0.79 | 0.56 |
| November | 1.59 | 0.40 | 2.09 | –6.75 | 4.35 | 1.22 | 0.66 |
| January ($i$ + 1) | 0.83 | 0.05 | 1.08 | –2.36 | 2.76 | 0.29 | 0.51 |
| Panel B: soybeans | | | | | | | |
| September | 2.35 | –0.38 | 3.79 | –17.63 | 5.41 | –0.65 | 0.46 |
| October | 2.20 | 0.29 | 3.03 | –6.82 | 8.70 | 0.61 | 0.49 |
| November | 1.40 | 0.60 | 1.65 | –3.45 | 4.24 | 2.33** | 0.59 |
| January ($i$ + 1) | 1.20 | –0.03 | 1.54 | –5.58 | 2.76 | –0.14 | 0.46 |

Notes: Percentage revisions are calculated as the natural logarithm of the forecast in month $t$ minus the natural logarithm of the forecast in month $t - 1$, times 100. N = 41. The test of bias tests the null hypothesis that the mean percentage revision equals zero. For ease of interpretation, the sign test statistic is reported as a proportion of positive cumulative revisions within each subsample. Single, double, and triple asterisks denote statistical significance at the 10%, 5%, and 1% levels, respectively.

revealed here does not imply that revisions are independent, a necessary condition for forecast efficiency outlined by Nordhaus. Correlation in forecast revisions would imply that forecasts are not efficient, because future revisions (or forecasts) can be predicted based on knowledge of current revisions.

Table 2 shows another important characteristic of the data investigated in this study, deviations of crop yields (based on final estimate) from an in-sample trend. When the 5% deviation of the final yield from an in-sample trend is used to indicate crop size,[3] 9 of 41 years would be classified as small in both crops and 14 and 9 of 41 years would be classified as big in corn and soybeans, respectively.[4] This classification can be used to demonstrate the pattern of big crops getting bigger and small crops getting smaller investigated in this study. Figure 2 shows average cumulative revisions starting in August for big, normal, and small crops based on a 5% classification rule described previously as well as for the whole sample (dashed line). Thus, a change from zero (representing August) to September illustrates a September revision of the August forecast, and the dashed line is identical to the average revision for the whole sample shown in Table 1. October cumulative revision illustrated by a change from zero to October represents a sum of September and October revisions. A change from zero to January represents a sum of all September, October, November, and January revisions, which is identical to August forecast error. Thus, Figure 2 demonstrates that big crops

[3] Similar measures were used in previous studies to categorize crops as small, normal, or big. Wisner, Blue, and Baldwin (1998) used a 10% rule applied to deviations of the final estimates from trend yield, whereas Taylor (2003) used a 5% rule applied to deviations of the August forecasted yield from trend yield.

[4] Because "small" and "big" are defined relative to rising trends, a "small" crop yield at the end of the sample may be larger than a "big" crop yield in the beginning of the sample. For example, the 2002 corn yield is "small" because it is 7.75% below trend, whereas the 1973 corn yield is "big" because it is 7.73% above trend. Nonetheless, the 2002 corn yield estimate of 130.0 bushels per acre is larger than the 1973 corn yield estimate of 91.4 bushels per acre.

**Table 2.** Deviations of Corn and Soybean Yields from an In-Sample Trend, 1970–2010

| Corn | | | Soybeans | | |
|---|---|---|---|---|---|
| | Yield | Deviation from Trend | | Yield | Deviation from Trend |
| Year | Bu/acre | (%) | Year | Bu/acre | (%) |
| 1988 | 84.6 | –29.39 | 1988 | 26.8 | –21.31 |
| 1983 | 81.1 | –24.75 | 1983 | 25.7 | –18.65 |
| 1993 | 100.7 | –20.11 | 2003 | 33.9 | –15.93 |
| 1974 | 71.3 | –19.36 | 1974 | 23.5 | –13.94 |
| 1995 | 113.5 | –11.23 | 1984 | 28.2 | –10.77 |
| 1970 | 71.7 | –9.47 | 1980 | 26.8 | –10.11 |
| 1991 | 108.6 | –9.38 | 1993 | 32.0 | –9.99 |
| 2002 | 130.0 | –7.75 | 1976 | 25.6 | –8.58 |
| 1980 | 91.0 | –7.50 | 2008 | 39.7 | –5.52 |
| 1976 | 87.4 | –3.36 | 1999 | 36.5 | –4.02 |
| 1997 | 127.0 | –2.98 | 1995 | 34.9 | –3.77 |
| 1975 | 86.2 | –2.58 | 1989 | 32.4 | –3.65 |
| 2010 | 152.8 | –2.00 | 2002 | 38.0 | –3.40 |
| 1977 | 90.8 | –1.65 | 2000 | 38.1 | –0.88 |
| 1996 | 127.1 | –1.42 | 1991 | 34.3 | –0.53 |
| 1999 | 133.8 | –0.67 | 1990 | 34.0 | –0.13 |
| 2001 | 138.2 | –0.25 | 2007 | 41.7 | 0.45 |
| 2003 | 142.2 | –0.14 | 1981 | 30.4 | 1.02 |
| 2000 | 137.1 | 0.35 | 1978 | 29.2 | 1.48 |
| 2007 | 150.7 | 0.39 | 2010 | 43.5 | 1.55 |
| 2006 | 149.1 | 0.62 | 2001 | 39.6 | 1.84 |
| 1989 | 116.2 | 0.67 | 1996 | 37.6 | 2.48 |
| 1984 | 106.6 | 0.76 | 1987 | 33.7 | 2.93 |
| 1990 | 118.5 | 0.97 | 1975 | 28.4 | 3.39 |
| 2005 | 147.9 | 1.12 | 1998 | 38.9 | 3.51 |
| 2008 | 153.9 | 1.22 | 2009 | 44 | 3.73 |
| 1998 | 134.4 | 1.22 | 2006 | 42.7 | 3.89 |
| 2009 | 164.7 | 6.74 | 1977 | 29.6 | 4.37 |
| 1987 | 119.4 | 6.78 | 1973 | 27.8 | 4.50 |
| 1978 | 101.2 | 7.13 | 1986 | 33.8 | 4.58 |
| 1971 | 86.8 | 7.22 | 2004 | 42.2 | 4.87 |
| 1973 | 91.4 | 7.73 | 1997 | 39.0 | 4.94 |
| 1992 | 131.4 | 8.07 | 1982 | 32.2 | 5.33 |
| 1986 | 119.3 | 8.44 | 2005 | 43.0 | 5.66 |
| 1985 | 118.0 | 9.11 | 1970 | 26.8 | 5.93 |
| 1981 | 109.9 | 9.42 | 1972 | 27.9 | 6.53 |
| 1994 | 138.6 | 10.28 | 1985 | 34.1 | 6.83 |
| 1982 | 113.2 | 10.47 | 1971 | 27.6 | 7.14 |
| 2004 | 160.4 | 10.56 | 1992 | 37.6 | 7.39 |
| 1979 | 109.4 | 12.90 | 1979 | 32.2 | 9.74 |
| 1972 | 95.5 | 14.42 | 1994 | 41.9 | 15.73 |

Notes: Yields are based on January (final) estimates for each crop year and do not reflect any later USDA revisions. In-sample trend is based on 1970–2010 marketing years. Deviations are calculated in percentage log form.

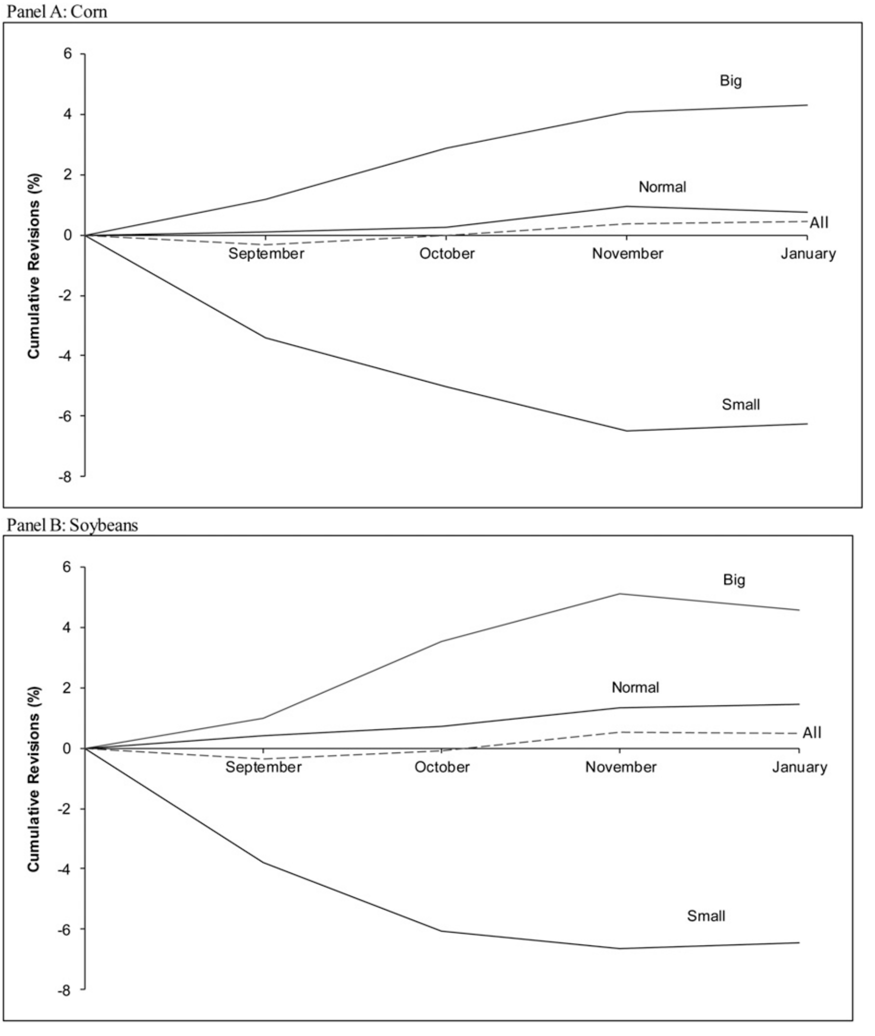


**Figure 2.** Cumulative Average Revisions of U.S. Department of Agriculture Corn and Soybean Yield Forecasts in Small, Normal, and Big Crop Years, 1970–2010

get approximately 4.3% and 4.5% bigger and small crops get 6.2% and 6.5% smaller in corn and soybeans, respectively. Further detail and the tests of bias presented in Table 3 demonstrate that average cumulative revisions in big and small years were significantly different from zero, whereas those in the normal years were not. The sign test confirmed the results of the *t* test in most cases, except small crops in corn. In addition, the sign test revealed that most revisions of corn yield in normal years as well as October and November revisions across all years were significantly more likely to be positive. This result, however, does not mean that these revisions were significantly different from zero.

Although this illustration clearly shows the pattern of big crops getting bigger and small crops getting smaller, as the quotes in the introduction suggest, it relies on the final yield estimates released at the end of the forecasting cycle for measuring crop size. Because this information is not available at the time the forecasts are made, it cannot be used in efficiency tests and to anticipate smoothing associated with crop size. Therefore, a forecast measure of crop size is explored in this article. Specifically, crop size is measured as a log percentage difference between the yield forecast made at time *t* and an out-of-sample linear trend forecast. The out-of-sample linear trend projection was constructed

**Table 3.** Descriptive Statistics and Test of Bias for Cumulative Revisions of U.S. Department of Agriculture Corn and Soybean Yield Forecasts, 1970–2010

| | Corn Revision Month | | | | Soybean Revision Month | | | |
|---|---|---|---|---|---|---|---|---|
| Category/ Statistic | September | October | November | January ($i$ + 1) | September | October | November | January ($i$ + 1) |
| Small crop | | | | | | | | |
| Mean | –3.41 | –5.03 | –6.50 | –6.23 | –3.78 | –6.04 | –6.63 | –6.46 |
| SD | 5.21 | 6.79 | 8.55 | 9.25 | 5.89 | 6.56 | 6.16 | 6.09 |
| t-statistic | –1.96* | –2.22* | –2.28** | –2.02* | –1.92* | –2.76** | –3.23*** | –3.18*** |
| Sign test | 0.22** | 0.33 | 0.33 | 0.33 | 0.11*** | 0.11*** | 0.11*** | 0.11*** |
| Normal crop | | | | | | | | |
| Mean | 0.09 | 0.25 | 0.95 | 0.74 | 0.39 | 0.70 | 1.34 | 1.46 |
| SD | 1.99 | 2.69 | 3.25 | 3.51 | 2.27 | 3.00 | 3.73 | 4.24 |
| t-statistic | 0.20 | 0.40 | 1.24 | 0.90 | 0.80 | 1.09 | 1.68 | 1.61 |
| Sign test | 0.50 | 0.72** | 0.72** | 0.72** | 0.50 | 0.59 | 0.59 | 0.64* |
| Big crop | | | | | | | | |
| Mean | 1.17 | 2.86 | 4.07 | 4.31 | 0.98 | 3.51 | 5.11 | 4.55 |
| SD | 1.74 | 2.40 | 3.28 | 3.84 | 2.41 | 4.04 | 4.29 | 5.26 |
| t-statistic | 2.53** | 4.45*** | 4.64*** | 4.20*** | 1.28 | 2.75** | 3.76*** | 2.73** |
| Sign test | 0.64 | 0.79** | 0.86*** | 0.71** | 0.70* | 0.80** | 0.90*** | 0.80** |
| Full sample | | | | | | | | |
| Mean | –0.31 | –0.02 | 0.38 | 0.43 | –0.38 | –0.09 | 0.51 | 0.47 |
| SD | 3.33 | 4.77 | 6.18 | 6.52 | 3.79 | 5.35 | 6.01 | 6.21 |
| t-statistic | –0.59 | –0.02 | 0.39 | 0.42 | –0.65 | –0.11 | 0.54 | 0.49 |
| Sign test | 0.46 | 0.61* | 0.63** | 0.59 | 0.49 | 0.56 | 0.59 | 0.56 |

Notes: Classification of crops is based on a 5% difference between final yield estimate and an in-sample linear trend. The number of observations is 41 for the full sample, nine for small crops, 18 and 22 for normal crops for corn and soybeans, respectively, and 14 and 10 for big crops for corn and soybeans, respectively. The test of bias tests the null hypothesis that the mean percentage revision equals zero. For ease of interpretation, the sign test statistic is reported as a proportion of positive cumulative revisions within each subsample. Single, double, and triple asterisks denote statistical significance at the 10%, 5%, and 1% levels, respectively.
SD, standard deviation.

using yield data starting in 1961[5] and ending in marketing year $i - 1$. Thus, a trend projection for 1980 was based on 19 previous observations (1961–1979), whereas 39 previous observations (1961–2009) were used for the 2010 trend forecast. Because this measure uses the information available at the time the forecasts are made, it is an *ex-ante* analog to the *ex-post* measure used to illustrate the pattern of smoothing in Figure 2. This *ex-ante* measure is used to test efficiency of forecast revisions with respect to crop size and to investigate whether the pattern of small crops getting smaller and big crops getting bigger can be anticipated using information available when the forecasts are made.

## Methods

Following Isengildina, Irwin, and Good (2006) and Nordhaus (1987), efficiency in yield forecast revisions is tested in this study using the following regression:

$$v_{t+1}^i = \gamma + \alpha v_t^i + \varepsilon_t^i \quad t = 1, \ldots, 4$$
$$(3) \qquad i = 1970, \ldots, 2010,$$

where all revisions made in month $t + 1$ are regressed against the previous revisions (in month $t$) during the study period. In this

[5] Previous studies of long-term yield trends (Tannura, Irwin, and Good, 2008) caution against using data before 1960 for corn and soybean yield analysis as a result of increased application of chemical fertilizers around this time.

regression, $\gamma$ measures forecast revision bias.[6] Forecasts are efficient if $\alpha = 0$. If $\alpha > 0$, forecasts are "smoothed" because part of the information known at time $t$ is being carried over into the next forecast. For example, if $\alpha = 0.6$, revisions made in month $t + 1$ are inefficient because they are based 60% on the previous revision. If $\alpha < 0$, forecasts are "jumpy" because forecasters tend to overreact to new information.

Forecast efficiency implies that not only past revisions, but any other information known at time $t$ should be fully incorporated in revision $v_t^i$ and should not be used to anticipate future revisions. Thus, we examine whether forecast revisions are associated with crop size by including a crop size indicator in the original regression:

$$v_{t+1}^i = \gamma + \alpha v_t^i + \beta c_t^i + \varepsilon_t^i, \quad t = 1, \ldots, 4 \\ i = 1970, \ldots, 2010, \tag{4}$$

where $c_t^i$ is a measure of crop size for marketing year $i$ available in forecast month $t$. Within this framework, the null hypothesis is $\beta$ = zero indicating no relationship between crop size and forecast revisions. If $\beta > 0$, larger crops are correlated with positive revisions and smaller crops are correlated with negative revisions, or big crops get bigger and small crops get smaller. If $\beta < 0$, smaller crops are correlated with positive revisions and larger crops are correlated with negative revisions, or big crops get smaller and small crops get larger.

In addition to regressions described in Equation 4 where current revisions and crop size measures are used to predict next revisions, we examine whether these factors can be used to predict all future revisions:

$$v_{t+n}^i = \gamma + \alpha v_t^i + \beta c_t^i + \varepsilon_t^i, \quad t = 1, \ldots, 4 \\ n = 1, \ldots, 3, \quad i = 1970, \ldots, 2010, \tag{5}$$

thus relaxing an assumption that correlation in revisions is present only in consecutive months. These tests allow us to examine the pattern shown in Figure 2 one step at a time by looking at correlations in revisions rather than cumulative revisions. Furthermore, predictability of forecast errors or cumulative revisions from time $t$ is examined:

$$e_t^i = \gamma + \alpha v_t^i + \beta c_t^i + \varepsilon_t^i, \quad t = 2, \ldots, 4 \\ i = 1970, \ldots, 2010. \tag{6}$$

This test looks at the whole path of forecast revisions shown in Figure 2 starting with $t$ = September and presents a combination of all monthly tests shown in Equation 5. For ease of interpretation, time $t$ used in these tests is referred to as "report month" in the remainder of the article. As explained for Equation 4, forecast efficiency requires all estimated coefficients in these regressions to be zero.

## Results and Discussion

The first set of correlation efficiency test results reported in Table 4 is based on using a log percentage difference between the yield forecast made at time t and an out-of-sample linear trend as a measure of crop size. Correlation coefficients between yield forecast revisions in two consecutive months (shown on diagonal) can be compared with the findings for production forecast revisions for a slightly shorter time period reported in Isengildina, Irwin, and Good (2006). Significant positive correlation between consecutive revisions was found in all cases except October vs. September revisions in soybeans, whereas in production forecast revisions (shown in Isengildina, Irwin, and Good, 2006), correlation between January and November soybean revisions was insignificant as well. The magnitude of the estimated coefficients was similar, ranging from 0.26–0.60 for corn yield (0.25–0.68 for corn production) and from 0.18–0.35 for soybean yield (0.14–0.26 for soybean production).

The similarity of smoothing in USDA yield and production forecast revisions is consistent with the suggested sources of smoothing in the Isengildina, Irwin, and Good (2006) study. Based

[6] Because the results of the test of bias are reported among the descriptive statistics in Table 1, they are not reported in the efficiency test results. Estimated $\gamma$ coefficients were not significantly different from zero in all cases except for November soybean revisions where $\gamma = 0.6$, as shown in Table 1.

**Table 4.** Correlation Tests Including Measure of Crop Size for U.S. Department of Agriculture Corn and Soybean Yield Forecast Revisions, 1970–2010

| Crop | Report Month | Independent Variable/Statistic | October Revision | November Revision | January ($i + 1$) Revision | Forecast Error for Report Month |
|---|---|---|---|---|---|---|
| | | | Dependent Variable | | | |
| Corn | September | Revision | 0.26** | 0.23** | 0.03 | 0.52** |
| | | Crop size | 0.01 | –0.01 | –0.04** | –0.05 |
| | | $R^2$ | 0.16 | 0.11 | 0.11 | 0.11 |
| | October | Revision | | 0.60*** | 0.24*** | 0.84*** |
| | | Crop size | | –0.01 | –0.05*** | –0.05 |
| | | $R^2$ | | 0.44 | 0.27 | 0.42 |
| | November | Revision | | | 0.35*** | 0.35*** |
| | | Crop size | | | –0.04*** | –0.04*** |
| | | $R^2$ | | | 0.39 | 0.39 |
| Soybeans | | | | | | |
| | September | Revision | 0.18 | 0.00 | –0.05 | 0.13 |
| | | Crop size | 0.00 | 0.00 | –0.04 | –0.04 |
| | | $R^2$ | 0.05 | 0.00 | 0.08 | 0.01 |
| | October | Revision | | 0.28*** | 0.15 | 0.43*** |
| | | Crop size | | –0.01 | –0.06* | –0.07 |
| | | $R^2$ | | 0.23 | 0.10 | 0.19 |
| | November | Revision | | | 0.35** | 0.35** |
| | | Crop size | | | –0.04 | –0.04 |
| | | $R^2$ | | | 0.14 | 0.14 |

Note: N = 41. Crop size is measured as percent deviation of current yield estimate from out-of-sample trend. Single, double, and triple asterisks denote statistical significance at the 10%, 5%, and 1% levels, respectively.

on the interviews of the USDA forecast providers, the authors concluded that smoothing in production forecasts likely stems from "the procedure used to translate enumerators' information about plant fruit counts early in the production season into objective yield estimates" as well as conservativeness in "farm operators' assessments of production potential" in the AYS results, both of which would affect production forecasts through yield estimates. This conclusion also concurs with the observation by Good and Irwin (2011) that "Most of the errors in production forecasts are associated with errors in yield forecasts since errors in harvested acreage forecasts are generally small." Thus, smoothing in production forecasts is likely driven by smoothing in yield forecasts.

Most estimated monthly correlations were in a narrow range between 0.23 and 0.35. Because revisions are measured in percentage form, estimated coefficients may be interpreted as point elasticities. Thus, a 0.28 coefficient in November vs. October soybean regression indicates that if the October soybean yield was revised up by 10%, one can expect the November forecast to be approximately 2.8% higher than October. In other words, 28% of the November revision is based on the previous revision and 72% on new information. Interpretation of these results should take into account the magnitude of the underlying revisions. The mean absolute values of forecast revisions reported in Table 1 indicate that the magnitude of revisions decreases over the forecasting cycle. Thus, although the correlation between November and October forecasts and January and November forecasts is very similar in soybeans (0.28 and 0.35), because October revisions are on average larger than November revisions (2.2% vs. 1.4%, as reported in Table 1), the predicted component in the November revision is larger than that in January (0.62% vs. 0.49%).

Additional tests indicate that correlations in revisions (smoothing) often carry over across several months. Thus, after a 10% September revision of corn yield, one can expect a 2.6% October revision and a 2.3% November revision, adding up to an expected September forecast error of 5.2%.[7] This anticipated component allows us to explain 11–16% of variation in future revisions. Similarly, a 10% October revision of corn yield usually precedes a 6% November revision and a 2.4% January revision, adding up to expected October forecast error of 8.4%. This anticipated component amounts to 27–44% of the variation future revisions or approximately 42% of the variation in October forecast error in corn. In soybeans, we observed smoothing at the rate of approximately 28% between October and November, which allowed us to anticipate approximately 43% of October error. This pattern represented approximately 23% of variation in November revision and 19% of variation in October error. Smoothing between November and January revisions was very similar in magnitude between corn and soybeans, 35%. November yield revisions allowed us to explain approximately 39% and 14% of the variation in January revisions or November forecast errors in corn and soybeans, respectively.

Although these tests demonstrate widespread correlation in corn and soybean yield forecast revisions, these revisions appear mostly efficient with respect to available crop size information. The only violation of rationality with respect to crop size (measured as log percentage difference between the yield forecast made at time $t$ and an out-of-sample linear trend) is observed in the January revisions of corn yield forecasts. For example, if September corn crop yield is 10% bigger (smaller) than the trend, one can expect January estimate to be revised downward (upward) by 0.4%. A similar effect of crop size information on the January revision of corn yield forecast is also shown in subsequent months. In soybeans we observe a similar pattern, but the coefficient is only significant in October. Note that the January estimate is not a forecast, but a survey-reported actual outcome, as described in the "USDA Crop Forecasts" section. Thus, the results indicate that the forecast bias is being corrected by the final estimate. This finding is consistent with the pattern illustrated in Figure 2, which shows that November forecasts tend to slightly overestimate normal corn yields and big soybean yields and underestimate small corn and soybean yields.

Because multiple additional sources of information about the crop size are available during the forecasting cycle, including private crop forecasts, weather and yield models, satellite imagery data, and crop conditions ratings, it is useful to examine the robustness of the efficiency test results using an alternative specification for the predicted crop size. The challenge with most of the alternative approaches is limited access and availability for a sufficiently long sample period to allow useful tests. One of the most comprehensive and readily available sources of crop size information is weekly USDA crop conditions ratings that are available since 1986. Crop conditions ratings may be superior to crop weather models because the ratings are more straightforward to interpret and may be a more accurate predictor of yield because they include factors other than weather that impact crop yields (such as insect and disease pressure). Previous research has shown that the sum of good and excellent ratings provides a good indication of crop conditions at various points of the growing season (Tannura, Irwin, and Good, 2008). Here, we concentrate on the crop conditions rating available at the time the USDA yield forecasts are made, around the first of the month during August, September, and October, and compute our variable by adding the percentage of crop classified as good and excellent.[8]

Table 5 presents correlation efficiency test results using crop condition ratings as a measure

[7] Note that September forecast error is a sum of October, November, and January revisions. However, only October and November revisions contribute to the expected component of the September error because they have predictable components. In contrast, January revision is not predictable according to our results; therefore, it becomes a part of the random component of the September error.

[8] For further information on crop conditions ratings, see the fact sheet available at: www.nass.usda.gov/Surveys/Crop_Progress_and_Condition/index.asp (Accessed March 12, 2010).

**Table 5.** Correlation Tests Including Measure of Crop Condition for U.S. Department of Agriculture Corn and Soybean Yield Forecast Revisions, 1986–2010

| | | | Dependent Variable | | | |
|---|---|---|---|---|---|---|
| Crop | Report Month | Independent Variable/Statistic | October Revision | November Revision | January ($i$ + 1) Revision | Forecast Error for Report Month |
| Corn | September | Revision | 0.42 | 0.43 | 0.13 | 0.98 |
| | | Crop condition | –0.02 | –0.01 | –0.02 | –0.05 |
| | | $R^2$ | 0.10 | 0.10 | 0.11 | 0.12 |
| | October | Revision | | 0.62*** | 0.24*** | 0.86*** |
| | | Crop condition | | 0.01 | –0.02 | –0.01 |
| | | $R^2$ | | 0.39 | 0.30 | 0.41 |
| | November | Revision | | | 0.28*** | 0.28*** |
| | | $R^2$ | | | 0.35 | 0.35 |
| Soybeans | | | | | | |
| | September | Revision | 0.30 | 0.10 | 0.07 | 0.47 |
| | | Crop condition | 0.00 | 0.02 | –0.01 | 0.02 |
| | | $R^2$ | 0.05 | 0.07 | 0.02 | 0.06 |
| | October | Revision | | 0.21** | 0.11 | 0.32** |
| | | Crop condition | | 0.02 | 0.00 | 0.02 |
| | | $R^2$ | | 0.29 | 0.12 | 0.27 |
| | November | Revision | | | 0.38*** | 0.38*** |
| | | $R^2$ | | | 0.30 | 0.30 |

Notes: N = 25. Crop condition is measured as a sum of good and excellent condition ratings released during the first week of the report month. Single, double, and triple asterisks denote statistical significance at the 10%, 5%, and 1% levels, respectively.

of crop size. Crop condition ratings are not available before 1986, so a shorter time period (1986–2010) is used for these tests. The results are consistent with those shown in Table 4 because significant positive correlation indicative of "smoothing" is observed in post-October revisions in both crops. Coefficient of the crop condition variable is not statistically significant in any cases indicating that this information is efficiently incorporated in corn and soybean yield forecast revisions. Thus, it appears that the pattern of "big crops getting bigger" and "small crops getting smaller" is largely based on information that is not available at the time the forecasts are made and thus would be difficult to predict and correct.

## Summary and Conclusions

There is widespread anecdotal evidence that farmers, traders, and market analysts believe that "big crops get bigger and small crops get smaller." The purpose of this article was to determine whether smoothing in USDA corn and soybean yield forecasts is concentrated in years with relatively small and large crops. The *ex-post* pattern of big crops getting bigger and small crops getting smaller over 1970–2010 is demonstrated using average cumulative revisions for crop years classified as "big," "normal," or "small" based on a 5% difference from the in-sample trend. Changes in average cumulative revisions from August to January demonstrate that big crops get approximately 4.3% and 4.5% bigger and small crops get 6.2% and 6.5% smaller in corn and soybeans, respectively. This pattern is confirmed by statistical tests, which show that average cumulative revisions in big and small years were significantly different from zero, whereas those in the normal years were not. Thus, this pattern is clearly visible when looking back and using *ex-post* data for analysis.

Data available at the time the forecasts are made are used in efficiency tests to establish whether this pattern can be anticipated and corrected. The analysis is based on incorporating information about the crop size in the framework for evaluation of efficiency in fixed-event forecast revisions developed by Nordhaus (1987).

An *ex-ante* measure of crop size is based on the percent deviation of the current estimate from an out-of-sample trend. Regression analysis is used to examine the correlation between subsequent revisions and whether it is affected by the crop size. These tests failed to reject forecast efficiency in all corn yield cases except November forecasts, which were revised down (up) in good (bad) crop years. This pattern was also observed in soybean yield forecasts, but it was only significant for information available in October. Alternative efficiency tests using crop condition ratings as a measure of crop size also failed to reject efficiency in all cases. Thus, it appears that the pattern of "big crops getting bigger" and "small crops getting smaller" is largely based on new information that is not available at the time the forecasts are made and thus cannot be predicted (or corrected) using available crop size information. In sum, this bias may appear obvious to market analysts in hindsight but it is difficult to anticipate.

An example of new information that could cause the pattern discussed here would be consistency in weather patterns in good and bad years. USDA uses a "normal weather" assumption for the remainder of the growing season to condition its crop forecasts. If the revision for a given month can be predicted based on weather conditions in the preceding month, then September revision may be predicted based on August weather, October revisions may be predicted based on September weather, and so on. Thus, a positive correlation between revisions that would result in "big crops getting bigger" and "small crops getting smaller" could be driven by unobserved weather if weather conditions are correlated during the growing season. For example, a positive (negative) September revision followed by a positive (negative) revision in October would be consistent with weather conditions in both August and September being "good" ("bad"). Although previous studies warn that long-term weather patterns are random (e.g., Hill and Mjelde, 2002), incorporating short-term (e.g. 1 month ahead) weather forecasts in lieu of the normal weather assumption could prove useful to USDA forecasters if there is consistency in weather patterns within years.

A different conclusion is reached with respect to the correlation in forecast revisions themselves; efficiency was rejected in most cases. The magnitude of positive correlation in revisions ranged from 0.26 to 0.60 for corn and from 0.18 to 0.35 for soybean yield. Furthermore, we found that correlations in corn yield revisions often carried over across several months. In the most extreme case observed here, 60% of November revision, 24% of January revision, and 84% of October forecast error were based on October revision of corn yield. The size of October and November corn yield revisions explained approximately 40% of respective forecast errors. In soybeans, smoothing at the rate of approximately 30% was limited to consecutive months between October and January. October and November soybean yield revisions allowed us to explain approximately 19% and 14% of respective forecast errors. These results suggest that USDA officials may tend to incorporate new information in their forecasts too slowly or conservatively, which results in predictable revision patterns.

Positive correlation of forecast revisions indicative of forecast smoothing was also found in corn and soybean production revisions by Isengildina, Irwin, and Good (2006). Although descriptions of USDA crop forecasting procedures (e.g., Good and Irwin, 2011) indicate that the USDA is aware of at least some of these biases (such as conservativeness of farm operators) and try to correct for them, evidence shown in this study suggests that they are not fully removed from official forecasts.[9]

These findings raise an interesting question of whether market participants understand and anticipate the smoothing in USDA corn and soybean production forecasts. If market participants are indeed aware of the smoothing process and account for it in forming expectations, economic welfare losses may be minimal. If instead market participants are unaware of or misunderstand the nature of the revisions process, welfare losses may result. The degree

[9] Isengildina, Irwin, and Good (2006) describe the implications of correction for smoothing on forecast accuracy.

to which market participants actually use this information in forming their own crop production forecasts is an interesting area for further research.

*[Received July 2011; Accepted August 2012.]*

## References

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):15–526.

Baur, R.F., and P.F. Orazem. "The Rationality and Price Effects of USDA Forecasts of Oranges." *The Journal of Finance* 49(1994):681–96.

Colling, P.L., and S.H. Irwin. "The Reaction of Live Hog Futures Prices to USDA Hogs and Pigs Reports." *American Journal of Agricultural Economics* 71(1990):84–94.

Fortenbery, T.R., and D.A. Sumner. "The Effects of USDA Reports in Futures and Options Markets." *Journal of Futures Markets* 13(1993):157–73.

Good, D., and S. Irwin. "USDA Corn and Soybean Acreage Estimates and Yield Forecasts: Dispelling Myths and Misunderstandings." Marketing and Outlook Brief 2011-02, Department of Agricultural and Consumer Economics, University of Illinois at Urbana–Champaign, March 2011. Internet site: http://www.farmdoc.illinois.edu/marketing/mobr/mobr_11-02/mobr_11-02.pdf (Accessed May 5, 2011).

Good, D.L., and S.H. Irwin. "Understanding USDA Corn and Soybean Production Forecasts: An Overview of Methods, Performance and Market Impact over 1970–2005." AgMAS Project Research Report 2006-01, Dept. Agr. Cons. Econ., University of Illinois, February 2006. Internet site: http://www.farmdoc.uiuc.edu/agmas/reports/06_01/AgMAS06_01.pdf (Accessed January 10, 2010).

Gunnelson, G., W. Dobson, and S. Pamperin. "Analysis of the Accuracy of USDA Crop Forecasts." *American Journal of Agricultural Economics* 54(1972):639–45.

Hill, H.S.J., and J.W. Mjelde. "Challenges and Opportunities Provided by Seasonal Climate Forecasts: A Literature Review." *Journal of Agricultural and Applied Economics* 34(2002):603–32.

Isengildina, O., S.H. Irwin, and D.L. Good. "Are Revisions to USDA Production Forecasts Smoothed?" *American Journal of Agricultural Economics* 88(2006):1091–104.

McKenzie, A.M. "Pre-Harvest Price Expectations for Corn: The Information Content of USDA Reports and New Crop Futures." *American Journal of Agricultural Economics* 90(2008):351–66.

Nordhaus, W.D. "Forecasting Efficiency: Concepts and Applications." *The Review of Economics and Statistics* 69(1987):667–74.

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):114–28.

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):127–34.

Tannura, M.A., S.H. Irwin, and D.L. Good. "Weather, Technology, and Corn and Soybean Yields in the U.S. Corn Belt." Marketing and Outlook Research Report 2008-01, Department of Agricultural and Consumer Economics, University of Illinois at Urbana–Champaign, February 2008. Internet site: www.farmdoc.illinois.edu/marketing/morr/morr_08-01/morr_08-01.pdf (Accessed February 3, 2010).

Taylor, S.E. "Climate Trends and Agricultural Management." Working Paper, Department of Agronomy, Iowa State University, 2003. Internet site: http://ag.arizona.edu/ANS/swnmc/papers/2003/taylor.pdf (Accessed January 23, 2010).

USDA/NASS. *The Yield Forecasting Program of NASS*. National Agricultural Statistics Service, U.S. Department of Agriculture, Staff Report SMB 06-01, May 2006. Internet site: www.nass.usda.gov/Education_and_Outreach/Understanding_Statistics/yldfrcst2006.pdf (Accessed January 29, 2010).

Vogel, F.A., and G.A. Bange. *Understanding USDA Crop Forecasts*. National Agricultural Statistics Service and World Agricultural Outlook Board, U.S. Department of Agriculture, Miscellaneous Publication, No. 1554, March 1999. Internet site: www.nass.usda.gov/Education_and_Outreach/Understanding_Statistics/pub1554.pdf (Accessed February 5, 2010).

Wisner, R.N., E.N. Blue, and E.D. Baldwin. "Pre-Harvest Marketing Strategies Increase Net Returns for Corn and Soybean Growers." *Review of Agricultural Economics* 20(1998):288–307.