



The Role of Sampling Errors in Explaining Recent Surprises in USDA Corn Stocks Estimates: Part II

Scott Irwin and Darrel Good

Department of Agricultural and Consumer Economics
University of Illinois

Dwight Sanders

Department of Agribusiness Economics
Southern Illinois University

February 14, 2014

farmdoc daily (4):28

Recommended citation format: Irwin, S., D. Sanders, and D. Good. "[The Role of Sampling Errors in Explaining Recent Surprises in USDA Corn Stocks Estimates: Part II](#)." *farmdoc daily* (4):28, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, February 14, 2014.

Permalink:

<http://farmdocdaily.illinois.edu/2014/02/the-role-of-sampling-errors-in-usda-corn-stocks-estimates-part2.html>

In our [farmdoc daily article yesterday](#), we investigated whether sampling errors in USDA quarterly stocks estimates could explain the large surprises in some stocks estimates since the 2006-07 marketing year. We showed that sampling errors associated with USDA estimates of on-farm corn stocks can explain some of the surprises in the quarterly stocks estimates, but the sampling errors cannot explain the large surprises. In addition, the presence of sampling errors would not satisfy the criteria needed to explain the recent surprises in the corn stocks estimates. In today's article we take up the issue of sampling errors associated with the final production estimates released in January following the corn harvest. This is the fifth in a series of *farmdoc daily* posts discussing the findings of our recent report, which can be found [here](#). The research was funded by the [Office of the Chief Economist of the USDA](#).

Market analysts have noted for some time that large market surprises for stocks may be related to large sampling errors for January production estimates. For example, assume that market analysts under-estimate March 1 corn stocks by 200 million bushels. The under-estimate, or surprise, could be due to analysts over-estimating usage for the December-February quarter by 200 million bushels, or alternatively, the USDA could have under-estimated corn production by 200 million bushels in January. In the latter scenario, the estimate of usage made by market analysts would be correct but there is still a market surprise associated with the March 1 stocks estimate due to the 200 million bushel sampling error in the January "final" corn production estimate.

The USDA is always careful to include a detailed discussion of potential sampling errors in production reports. For example, the *Crop Production: 2012 Summary* released in January 2013 included this discussion of the reliability of production estimates:

We request all readers, electronic media and others follow our citation guidelines when re-posting articles from farmdoc daily. Guidelines are available [here](#). The farmdoc daily website falls under University of Illinois copyright and intellectual property rights. For a detailed statement, please see the University of Illinois Copyright Information and Policies [here](#).

“The surveys used to make the acreage, yield, and production estimates contained in this report are subject to sampling and non-sampling type errors that are common to all surveys. Reliability of the objective yield and farmer survey must be treated separately because the survey designs for the two surveys are different. The objective yield indications (corn, cotton, and soybeans) are subject to sampling variability because all acres of a given commodity are not included in the sample.”

“The farm operator survey indications are also subject to sampling variability because not all operations with commodities of interest are included in the sample. This variability, as measured by the relative standard error at the National level, is approximately 1.1 for corn, 2.3 for Upland cotton and 1.1 for soybeans. This means that chances are approximately 95 out of 100 that survey estimates for production will be within plus or minus 2.2 percent for corn, 4.6 percent for Upland cotton, and 2.2 percent for soybeans.”

“Survey indications are also subject to non-sampling errors such as omission, duplication, imputation for missing data, and mistakes in reporting, recording, and processing the data. These errors cannot be measured directly, but they are minimized through rigid quality controls in the data collection process and a careful review of all reported data for consistency and reasonableness.” (p.94)

The previous information implies that the range of sampling errors for corn production estimates can be surprisingly wide. With a point estimate for 2012 corn production of 10.780 billion bushels, the 95 percent confidence interval based only on the farm operator survey reliability is 10.543 to 11.017 billion bushels (10.780 +/-2.2%), or a range of 474 million bushels. Likewise, with a point estimate for 2013 corn production of 13.925 billion bushels, the 95 percent confidence interval is 13.535 to 14.315 billion bushels (13.925 +/-2.8%), or a range of 780 million bushels.

This discussion should make it clear that the sampling error for any particular January corn production estimate could easily be several hundred million bushels, and therefore, unresolved errors in production estimates for corn could logically account for the magnitude of surprises in the quarterly stocks estimates observed over the 2006-07 through 2012-13 marketing years. Under such circumstances, surveys for the USDA stocks estimates would presumably reveal actual crop size, while analyst expectations would be based on inaccurate production estimates or forecasts. While unresolved corn production errors can in general explain the stocks and usage surprises, a credible explanation needs to satisfy four criteria related to the specific pattern of surprises since 2006-07. We consider each of these criteria below.

- 1) **Why corn and not soybeans?** This criteria is the most straightforward to meet because of differences in usage accounting across corn and soybeans. Since all but a small fraction of soybean use is measured, the more definitive accounting of soybean usage allows for a relatively precise final estimate of soybean production once the stock estimate for the end of the marketing year becomes available, while the lack of definitive accounting of feed and residual use for corn does not provide for such a precise estimate. So, it is quite reasonable to expect stocks/usage surprises to be larger in corn than soybeans.

Since corn and soybeans in the U.S. are grown in similar areas and USDA uses the same estimation procedures for the two crops, the history of soybean production revisions should provide a reasonable indication of corn production revisions if USDA were able to make such revisions. Table 1 shows the history of September soybean production revisions over the 1990-91 through 2012-13 marketing years. In these years, the January soybean production estimate was revised in the following September stocks report in all but five of the years and the revisions ranged from -41 to 91 million bushels. In proportional terms, the range was -1.6 to 3.5 percent. Revisions to corn production estimates are simulated by applying the percentage soybean revisions to the average January corn production estimate over 2006-2012, 12.06 billion bushels. We use this procedure in order to scale the simulated corn revisions to production levels over 2006-07 through 2012-13, which is the period of greatest concern about the accuracy of USDA corn stocks estimates. As an example of the calculations, the September revision of the soybean production estimate for the 2003-04 marketing year is 1.5 percent. The simulated corn production revision of 180 million bushels is found by multiplying 12.06 times 1.5 percent. Figure 1 presents the simulated corn production revisions ordered from smallest to largest irrespective of marketing year. The range of the hypothetical revisions for corn is relatively large at -193 to 425 million bushels. The simulated

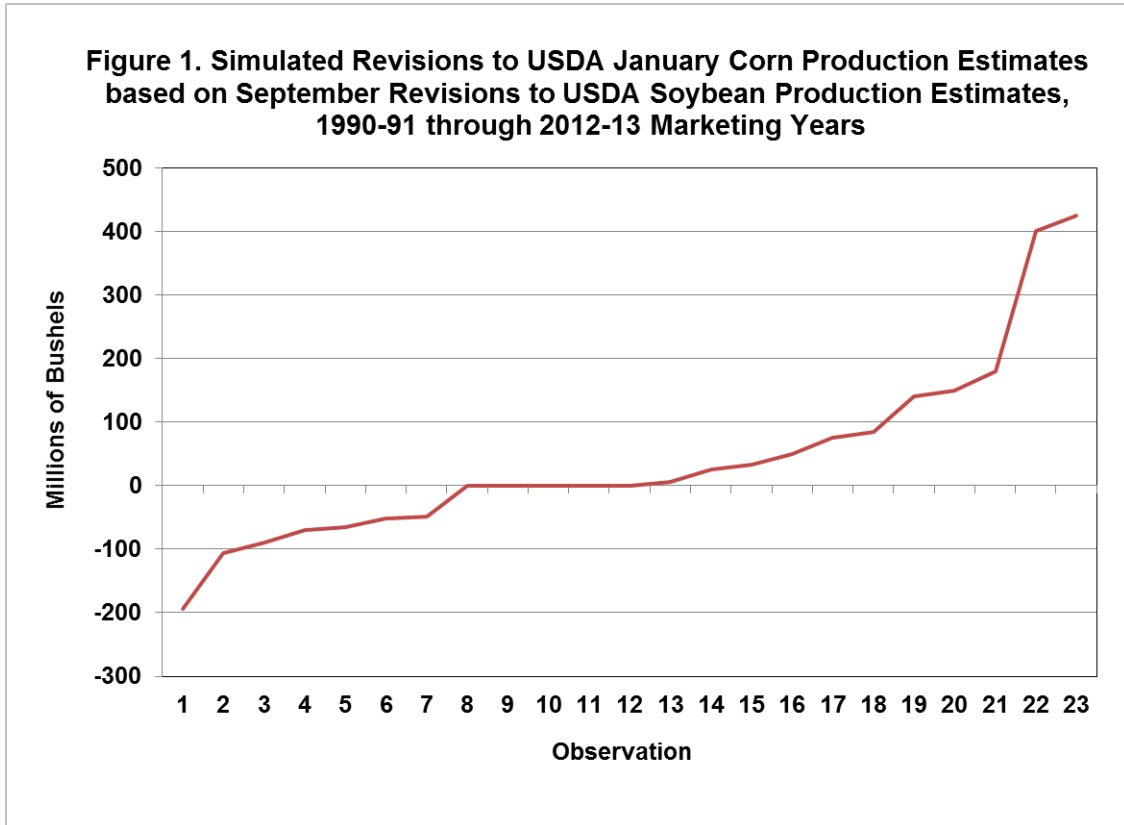
corn production revisions clearly could be large enough to potentially explain implied usage surprises of 300-400 million bushels for corn in recent years.

Table 1. USDA January and Following September Estimates of U.S. Soybean Production, 1990-91 through 2012-13 Marketing Years

Marketing Year	January Estimate (mil. bu.)	Revised		
		September Estimate (mil. bu.)	September Revision (mil. bu.)	September Revision (%)
90	1,922	1,926	4	0.2
91	1,986	1,987	1	0.1
92	2,197	2,188	-9	-0.4
93	1,809	1,869	60	3.3
94	2,558	2,517	-41	-1.6
95	2,152	2,177	25	1.2
96	2,382	2,382	0	0.0
97	2,727	2,703	-24	-0.9
98	2,757	2,741	-16	-0.6
99	2,643	2,654	11	0.4
00	2,770	2,758	-12	-0.4
01	2,891	2,891	0	0.0
02	2,730	2,749	19	0.7
03	2,418	2,454	36	1.5
04	3,141	3,124	-17	-0.5
05	3,086	3,063	-23	-0.7
06	3,188	3,188	0	0.0
07	2,585	2,676	91	3.5
08	2,959	2,967	8	0.3
09	3,361	3,361	0	0.0
10	3,329	3,329	0	0.0
11	3,056	3,094	38	1.2
12	3,015	3,034	19	0.6
		Min.	-41	-1.6
		Max.	91	3.5
		St. Dev.	29.0	1.2

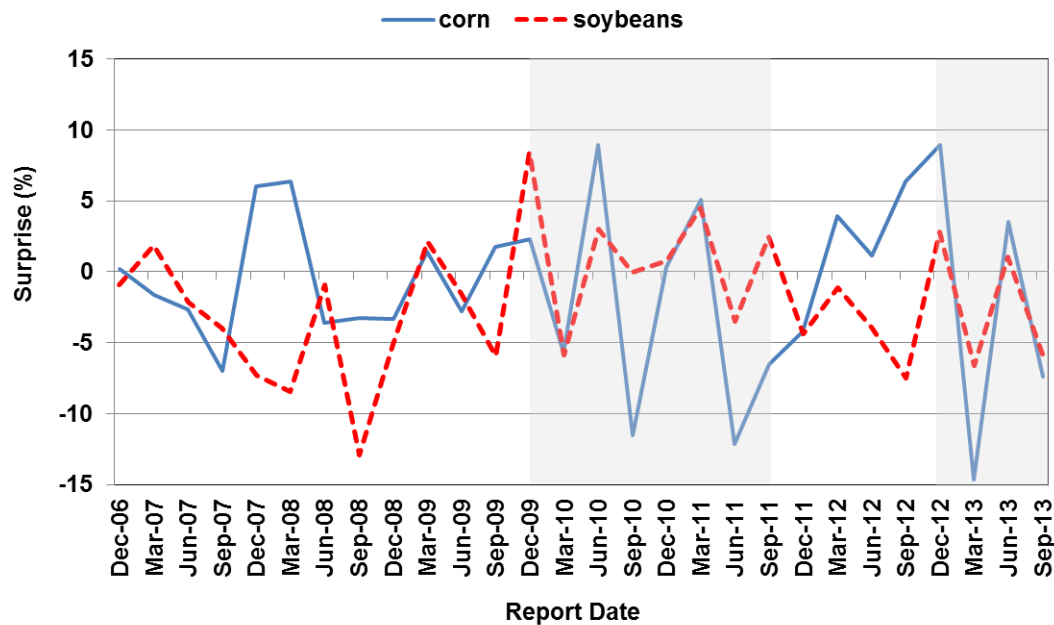
- 2) **Why 2006-07 through 2012-13 and not earlier?** There was a notable increase in the volatility of market surprises in corn stocks estimates/implied usage starting with the 2006-07 marketing year and a credible explanation of the surprises needs to account for this fact. One possibility is that unresolved corn production estimation errors increased over 2006-07 through 2012-13 due to a series of unusual growing season weather conditions that either increased the sampling variability of production estimation or simply caused large sampling errors for particular years. Summer weather conditions in the U.S. Corn Belt for all but a few years over 1990-2005 [were rather "benign"](#)

and there were relatively few large deviations from trend yields. It is also undoubtedly true that growing season weather during 2006-2012 tended to be more extreme. For example, 2009 saw record corn yields but an exceptionally late harvest due to excessive rainfall during the normal harvest period. As another example, [the 2012 drought](#) was among the most severe of the last century and drove corn yields far below trend. It does not seem unreasonable to conclude that sampling errors for corn production estimates would be larger under these circumstances. But, the fact remains that there is uncertainty why sampling errors increased over the 2006-07 through 2012-13 marketing years if these errors are in fact the correct explanation for the large surprises in corn stocks/implied usage.



- 3) **Why only in particular marketing years?** The largest surprises in corn stocks estimates/implied usage were concentrated in the 2009-10, 2010-11, and 2012-13 marketing years. The tendency for surprises to occur in these particular years must be accounted for by a valid explanation. The most logical possibility is again larger sampling errors due to unusual growing season weather conditions during these three years. Beyond the weather observations, there is additional evidence that some kind of unique conditions were present in these three years that led to the problematic stocks/implied usage surprises. Figure 2 plots implied usage surprises in both corn and soybeans for each USDA *Grain Stocks* report during the 2006-07 through 2012-13 marketing years. The two highlighted areas correspond to the 2009-10, 2010-11, and 2012-13 marketing years. What is remarkable is the complete lack of correlation between the implied usage surprises outside of these three marketing years and then nearly perfect synchronization of the sign of the implied usage surprises across corn and soybeans within these years. Specifically, 11 out of the 12 quarters in 2009-10, 2010-11, and 2012-13 have the same sign for corn and soybean implied usage surprises. Something truly unique occurred during these years and the conditions impacted both corn and soybean stocks estimates. One possibility is that there were unresolved production errors in both corn and soybeans during these years, but the errors were simply larger in corn than soybeans.

Figure 2. Surprise in USDA Stocks Estimates for U.S. Corn and Soybeans, December 2006-September 2013 *Grain Stocks Reports*



Note: A positive (negative) surprise occurs when analyst estimates are lower (higher) than the actual.

- 4) **Why a pattern of reversals during marketing years?** There was a distinct tendency towards reversals in the surprises in corn stocks estimates/implied usage within the 2009-10, 2010-11, and 2012-13 marketing years. This is undoubtedly a difficult criterion to meet given the specificity of the pattern. However, the reversal pattern of implied usage surprises could be generated by market analysts mistakenly viewing a stock surprise as a usage surprise when in fact the stocks surprise is the result of unresolved sampling errors in production estimates. This leads to a reversal in the sign of the subsequent surprise as the true rate of usage is revealed. (A detailed discussion of how this could occur is found on pp.49-50 of the [full report](#)) While this is a logical possibility, one also has to acknowledge that the varying magnitude of surprises in implied usage through the marketing year could be associated with other factors, particularly sampling errors in the stocks estimates themselves.

Implications

We conclude that unresolved errors in production estimates for corn is the most likely explanation for the large surprises in quarterly corn stocks and implied usage estimates that were observed over the 2006-07 through 2012-13 marketing years. USDA stocks estimates undoubtedly encompass sampling errors for both production and stocks estimates and it is highly likely that unresolved sampling errors for corn production estimates are large enough to explain even the largest surprises. It is more difficult to pin down exactly why unresolved sampling errors for corn production were concentrated in the 2009-10, 2010-11, and 2012-13 marketing years and caused the quarter-to-quarter reversal pattern in surprises, but reasonable arguments can be put forward. Our analysis highlights the potential value of adding a survey of corn feed use that would allow a fuller accounting of corn usage and revision of January corn production estimates similar to what has been historically possible for soybeans.

References

Irwin, S. "[Was 2012 Really the 'Big One' for U.S. Corn Yields?](#)" *farmdoc daily* (3):72, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, April 17, 2013.

Irwin, S., D. Sanders, and D. Good. "[The Role of Sampling Errors in Explaining Recent Surprises in USDA Corn Stocks Estimates: Part I.](#)" *farmdoc daily* (4):27, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, February 13, 2014.

Irwin, S. H., D. R. Sanders, and D. L. Good. "[Evaluation of Selected USDA WAOB and NASS Forecasts and Estimates in Corn and Soybeans.](#)" Marketing and Outlook Research Report 2014-01, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, January 2014.

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.