



Munich Personal RePEc Archive

The Evolution of Inefficiency in USDA's Forecasts of U.S. and World Soybean Markets

MacDonald, Stephen and Ash, Mark and Cooke, Bryce

World Agricultural Outlook Board, USDA

September 2017

Online at <https://mpra.ub.uni-muenchen.de/87545/>

MPRA Paper No. 87545, posted 29 Jun 2018 18:47 UTC

The Evolution of Inefficiency in USDA's Forecasts of U.S. and World Soybean Markets

Stephen MacDonald (USDA/WAOB), Mark Ash, and Bryce Cooke (USDA/ERS)

Abstract

We derive a set of stylized facts about USDA's soybean supply and demand forecasts and draw implications from these results for efforts to improve the accuracy of these forecasts. USDA's short run soybean supply and demand forecasts are inefficient, with several key variables significantly biased throughout much of the annual forecast cycle. Bias and other characteristics have varied significantly over recent decades, and evaluation efforts intended to guide forecast improvement need to focus on very recent sample years and seasonally disaggregated data. USDA should expand the information set used in its forecasting of U.S. soybean yields, take note of its mid-season downward bias in its Brazilian production forecasts, and shift the weight of its early-season focus for China's soybean consumption estimates towards forecasts of growth rates and reduce the weight it applies to forecasting consumption and trade by China in terms of levels. During 2004-2015, downward-biased U.S. production forecasts and upward-biased foreign excess supply forecasts resulted in downward-biased U.S. export forecasts throughout much of USDA's annual forecasting cycle. Twenty years earlier, this downward bias was confined to the end of the forecasting cycle, whereas U.S. soybean ending stock forecasts have been upward biased for decades across much of the forecasting cycle. The forecasts for U.S. soybean exports are also characterized by smoothing, with strong correlation between month-to-month forecast revisions towards the later months of USDA's annual forecasting cycle.

Disclaimer—Views and conclusions expressed in this draft are the responsibility of the authors alone and do not reflect the opinions or policy of any other entity or institution.

Introduction

Soybeans are the world's leading oilseed crop, and are processed to produce protein meal for livestock feeding and vegetable oils for human consumption and industrial use. Soybeans are one of the most important field crops in the United States and typically account for about 10 percent of total cash receipts in U.S. agriculture. At \$41 billion in 2014, U.S. farmers' cash receipts for soybeans were exceeded by only two other individual commodities: cattle and corn (USDA/ERS, 2016).

Soybeans has been one of the fastest growing major field crops for several decades, most recently boosted by soaring demand from China. The United States was the world's leading producer for decades starting in the 1960s, but has been overtaken by South America. World soybean production and consumption is more geographically concentrated than it is for grains or cotton, with the United States, Argentina, and Brazil in 2015/16 accounting for 83 percent of production (89 percent of exports) and China and the European Union (EU) accounting for 75 percent of world imports (Table 1).

The U.S. Department of Agriculture is the source of a broad range of statistics concerning past, present, and future developments in U.S. and world agriculture. USDA has been providing the public with market information since the 1860s in order to improve the efficiency of agricultural markets and reduce potential asymmetries of information among market participants (Effland, 1999). Each year, USDA's forecasting cycle begins in November with the estimates for the next marketing year included in USDA's

baseline forecasts for the next 10 marketing years. This forecast is updated and published the following February. More prominently, USDA's monthly reports of U.S. production and world-wide demand and output for grains, oilseeds, and cotton play a crucial role in price discovery on futures markets in the United States and elsewhere. USDA's National Agricultural Statistics Service (NASS) publishes its *Crop Production* estimates for the current year for these commodities from August to January, and the *World Agricultural Supply and Demand Estimates (WASDE)* incorporates these estimates every month, and extends the forecasts to other balance sheet variables—such as consumption and exports—and also adds geographically dis-aggregated production and consumption forecasts for the entire world. The impact of the information provided by USDA on futures markets has been demonstrated (Fortenbery and Sumner, 1993; Adjemian, 2012; Isengildina, et al. 2008). At least a subset of market participants have a very sophisticated understanding concerning USDA's estimates, and markets have been shown to anticipate and accommodate to the inefficiencies embedded in USDA's forecasts (Good and Irwin, 2006, Isengildina, Irwin, and Good, 2013).

The U.S. demand forecasts and non-U.S. supply and demand the forecasts are the product of committees bringing together expertise from across USDA (Vogel and Bange, 1999). Each Interagency Commodity Estimates Committee (ICEC) is chaired by a member of the World Agricultural Outlook Board (WAOB) and includes representatives from a set of USDA agencies concerned with analyzing and implementing USDA programs.¹ The historical data and forecasts developed by a committee consensus are: published each month in the *WASDE*, are disseminated through PSD Online, are included in ERS's Situation and Outlook reports, are included in FAS's World Markets and Trade reports, and form the basis of policy deliberations at USDA and other government agencies.

USDA's Forecasting Cycle

Soybeans are a summer crop, and the U.S. marketing year for soybeans is September-August. Planting in the United States is most active from mid-May to mid-June and harvest during September-October. Planting in the largest producers competing with the United States—Argentina and Brazil—occurs over September-January and harvest there stretches from January to July. Exports from all of these countries are seasonal, with U.S. exports reaching a post-harvest peak during October-February and Southern Hemisphere exports peaking over April-August. Anticipating the length of these peaks is a critical component of USDA's forecasting task.

The forecasting cycle has four distinct segments, distinguished by the information set available to USDA in each and the Departmental resources devoted to the forecasts. USDA forecasting activity is organized around the U.S. marketing year, with only a few forecasts before U.S. planting and the greatest share of USDA's resources being devoted to these estimates during the period of U.S. crop development. To compare trade levels on a consistent annual basis, USDA converts the local marketing year data for Brazil and Argentina to an October-September trade year.

For each marketing year, USDA's first short-run forecast is developed 10 months before the beginning of that marketing year during preparation of USDA's long run outlook and of the President's Budget for the next 5 years (fig. 1). Each November the ICEC develops forecasts for the U.S. balance sheet in the coming marketing year, but this forecast is typically not released until the following February at USDA's

¹Other agencies represented on USDA's ICECs are: Agricultural Marketing Service (AMS), Economic Research Service (ERS), Farm Service Agency (FSA), and Foreign Agricultural Service (FAS).

Agricultural Outlook Forum. This pre-planting forecast is included in the 10 year projections released at the Forum, and an updated forecast for this marketing year is also presented at the Forum. The information set for these forecasts differs significantly from that for later segments of the forecasting cycle. Many months from finalization, estimates for many variables' activity in the preceding marketing year are subject to adjustment with observed information. No NASS survey results of U.S. farmers' spring-sown new-crop acreage intentions are yet available. Dramatic changes in crop prices and climatic conditions can quickly alter farmers' crop decisions throughout this period. Measuring the forecast horizon (h) of these earliest forecasts as the last month of the marketing year, for November $h = 22$ months (1.8 years) and for February $h = 19$ months (1.6 years). These forecasts overlap with forecasts for the previous marketing year that USDA simultaneously publishes with $h = 10$ months (November) and for $h = 7$ months (February). Following the convention used by ICEC members, we term these shorter horizon forecasts for the previous year as "back-year" forecasts. The distant horizon forecasts are termed "out-year" forecast. This distinction holds through every month of the forecasting cycle that includes forecast for overlapping years, and the forecasts for any given marketing year eventually become back-year forecasts as USDA initiates a forecasting cycle for a new marketing year.

Continuous monthly updates of USDA's forecasts commence in May, and May through July marks another short segment of USDA's forecasting cycle. During these months, the *WASDE* includes forecasts of U.S. production based on the March *Prospective Plantings* and June *Acreage* reports from NASS. Publication of forecasts for the next marketing year in the *WASDE* start in May and include production forecasts based on NASS area data and ICEC trend-based yield estimates. Forecasts for the rest of the U.S. balance sheet by the ICEC are also published monthly starting in May as well as for the rest of the world.² In addition to more accurate estimates of the previous marketing year's activities, USDA's ICEC also has in its information set at this time reports and forecasts from its overseas attaches in the Foreign Agricultural Service (FAS). USDA/FAS maintains offices in China, Argentina, and Brazil—and about 70 other countries—which provide intelligence to Washington-based members of the ICEC, as well as publicly disseminated reports and forecasts. Using information and insights from overseas-based USDA personnel, contacts in industry, public price data, remote-sensing data, meteorological analysis, contractors for macroeconomic outlook, and a centralized repository of timely international monthly trade data through contractors, USDA maintains historical data and publishes forecasts for the supply and demand for soybeans and other products in over 100 countries. Although the flow of information is constant, the frequency of each source can vary significantly.

August-January marks a more intensive segment of USDA's forecasting cycle for soybeans and similar crops, as NASS publishes its first estimate of the year's U.S. crop in August and updates it throughout the harvest period. During these months, forecasts for other aspects of the U.S. balance sheet—and the forecasts for non-U.S. balance sheets—react to changes in the U.S. production forecast and to changes in market conditions around the world. From the beginning of this segment of the forecast cycle, the planting of the U.S. crop is a past event and the issue is no longer forecasting but discovering the actual level of planting activity. Similarly, by the end of this segment yield realization and harvest of the U.S. crop is complete, and the goal of statistical agencies is to reveal this level of past activity.

²May publication of USDA's non-U.S. soybean forecasts for the coming marketing year has occurred since 2009. June publication has occurred since 2004. In the preceding decades, July marked the earliest publication in *WASDE* of non-U.S. soybean forecasts. In these earlier years, the formal reports from overseas FAS personnel were also correspondingly later.

The final segment of the USDA forecasting cycle for U.S. soybeans is February-November. During most of this period, U.S. production is known to within a relatively small margin, while Southern Hemisphere planting levels are discovered and then yield performance in the Southern Hemisphere is also determined. The proportion of the U.S. marketing year's domestic consumption and export activity that must be forecasted diminishes as monthly reports from the U.S. Bureau of the Census and NASS for past activity accumulate. For example, since 1993, U.S. soybean exports averaged 72 percent complete for the year by March 1, although the range was between 60 percent and 87 percent. Similarly, trade and crush data from industry and government sources in Argentina, Brazil, and elsewhere accumulates as the year progresses, at a rate that varies from year to year.

The finalization of USDA's estimates for U.S. activity during a given marketing year occurs several months after the marketing year's conclusion. Ending stocks, trade, and crushing data for the last month of the marketing year (August) is not available before the September *WASDE*, and NASS typically finalizes its estimate of the previous year's U.S. production in October, partly by accounting for information in the stocks and utilization statistics.³ Soybean meal is analyzed on an October-September year in the United States, so finalization of the U.S. soybean meal estimates occurs one month after the soybean finalization.

A similar process occurs in USDA's estimates of Brazilian production at the end of the forecasting cycle. While Brazil has respected and timely domestic sources of forecasts and historical estimates of its soybean production, the final estimates from these sources are typically inconsistent with the information in the utilization data. USDA's October-September estimates for Brazil and Argentina are typically not finalized until several months after the U.S. estimates due to a later ending of their local marketing years and difficulty in reconciling the data from different sources. Soybean stocks data are particularly hard to verify for both countries.

Previous Studies: Inefficiencies in USDA Soybean Forecasts

The importance of USDA's forecasts to market equilibrium and public and private policy decisions has motivated many studies of USDA forecast performance and a number of studies with a focus on U.S. soybeans. Multi-commodity studies have found inefficiencies in USDA's long run (10-year) forecasts (GAO, 1988 and 1991), but most studies focus on the more widely followed USDA monthly forecasts with a much shorter horizon (h), at most a few months greater than one year ($h = 1.3$ years). We summarize these studies first in terms of their findings and then by their methodologies.

Studies of short term USDA livestock forecasts have found inefficiencies in beef, pork (Bailey and Brorsen, 1998; Sanders and Manfredo, 2002), and poultry forecasts (Sanders and Manfredo, 2003). MacDonald (2002) and Isengildina-Massa, MacDonald, and Xie (2012) detected inefficiencies in USDA's cotton forecasts, and other studies have detected inefficiencies in orange, corn, wheat, soybean, and sugar forecasts (Baur and Orezen, 1994; Egelkraut et al., 2003; Botto et al., 2006; Isengildina, Irwin, and Good, 2006; Lewis and Manfredo, 2012). MacDonald examined USDA's trade forecasts (1992, 1999, and 2005) extending the commodities examined to horticultural products as well as livestock and grains, and also finding inefficiencies like bias and serial correlation. A recurring finding across a number of studies and

³Bureau of Census U.S. trade data for the last month of the marketing year was not consistently available for the October *WASDE* before 2013. Before 2003, release dates for this data always fell 1 or 2 weeks after publication of the October *WASDE*. Thus, in earlier samples used in this study, the U.S. soybean forecasts were finalized in November, an end date we apply to all the data through 2016.

commodities is inefficiency in the month-to-month revisions in USDA's forecasts, primarily "smoothing," as USDA publishes forecasts whose monthly revisions are often positively correlated with earlier revisions (MacDonald and Isengildina-Massa, 2012a; Isengildina-Massa, Irwin, and Good, 2013).⁴

Results specific to soybeans have included findings of bias in USDA forecasts of U.S. soybean production, exports, and ending stocks (USDA, 2004; IKI), and finding the absence of bias in U.S. ending stocks (Xao, Lence, and Hart, 2014—XLH henceforth). Other findings include: forecast smoothing in many parts of the balance sheet (MIM) and forecast smoothing in the production and yield forecasts (IIG). MacDonald (1999) found bias in USDA's oilseed value export forecasts, in contrast with a similar study using earlier data that found no such bias (MacDonald, 1992).

Volatility of the inefficiency measures over time was also found in other studies in addition to the two by MacDonald; that is the measured characteristics of USDA's soybean forecasts were to some extent determined by the samples examined. Extension of the sample period end point from 1989 in MacDonald (1992) to 1998 in MacDonald (1999) revealed bias. Irwin, Scott, Sanders, Dwight, and Darrel Good (2014) (ISG henceforth) show downward bias in USDA's October soybean production forecasts in 2001-2012 but not in 1990-2000, and IKI show that inclusion of a 2006-2009 dummy in tests of soybean export forecast accuracy indicate downward bias associated with this later period that the authors conclude is not apparent in their overall 1984-2009 sample.

Note that in the 25 past studies of all agricultural commodities reviewed here, the number of years spanned in the historical samples analyzed ranged from 4 (MacDonald, 2005) to 53 (Baur and Orazem, 1994), with a median of 20 years.⁵

Methodological Similarities Among Past Studies

There are several ways to group past studies by methodology: 1) a clear majority of these studies aim to analyze non-overlapping forecasts, i.e. forecasts with horizons not exceeding the frequency of the forecasted variable, 2) a little more than half of these studies analyzed the forecasts in their published levels form (bushels, acres, ton, etc.), and the rest first applied a transformation into the growth rate forecasts implied by the published data (typically, percent), and 3) a little more than one-third of these studies calculate evaluation measures for aggregated months of the forecasting cycle—presenting a single statistic assessing the entire annual cycle, or an aggregation of a subset of the cycle's months.

It is necessary to more fully explain the concept of overlapping forecasts before proceeding further. Overlapping forecasts are more complex to evaluate, but—as figure 1 indicates—are unavoidable in studies examining forecasts over the entire USDA forecasting cycle. For example, the cycle's first forecast for U.S. soybean exports developed in November, 10 months before the start of a given marketing year, includes in its information set a forecast for the preceding marketing year. That preceding-year forecast has a 12 month horizon before the point when USDA can finalize its estimate.

⁴In the remainder of this paper, several past study references will be abbreviated: Bailey and Brorsen, 1998—BB; Botto et al., 2006-- BIIG; Egelkraut et al., 2003--EGIG; Isengildina, Irwin, and Good, 2006--IIG; Isengildina-Massa, Karali, and Irwin, 2013—IMKI; Isengildina-Massa, MacDonald, and Xie (2012)--IMMX; MacDonald and Isengildina-Massa, 2012a—MIM; Sanders and Manfredo, 2002, 2003, and 2008—SM1, SM2, SM3.

⁵Four studies examined quarterly forecasts, over a median of 15.5 years. In most cases, each quarter's forecast was examined separately, so the sample size used in the calculation of the evaluation statistics equaled the number of years.

Thus, the first forecast of the cycle is overlapping with a contemporaneous forecast for the previous year, and the same holds true for the following 7 revisions to this forecast spread over the following 12 months, until the previous year's forecast is finalized.

Some of past studies examined were explicit in selecting to examine only non-overlapping forecasts, selecting a non-overlapping subset of a broader set of a variable's forecasts. Others studies examined forecasts that were unequivocally non-overlapping, making no reference to the issue. A third group includes studies focused on U.S. production estimates that eliminated overlapping forecasts by assumption. These studies assign the first January after the U.S. harvest as the point of finalization (BIIG; EGIG; Good and Irwin, 2006; IMIG; IMKI; ISG). Several of these studies examining U.S. production forecasts also evaluate forecasts of other balance sheet variables (BIIG; IMKI; ISG) over a span of the forecasting cycle that includes months with overlapping forecasts. None of these studies, or other studies evaluating forecasts over portions of the forecasting cycle with overlapping forecasts (BB, IMM, MIM, SM3, and XLH), thoroughly address the impact of overlapping on their evaluation statistics. The impact of overlapping forecasts on the variance-covariance matrix of the evaluation tests is accounted for where appropriate in these studies, but other impacts of overlapping on the evaluation metrics have been overlooked.

Another group of studies are those that examine forecasts as growth rates, in percent, rather than analyzing them in their published form as a level of activity, e.g. in tons or acres. An appropriate concern addressed by many past studies is accounting for the non-stationarity of the data used in the evaluations (Clements and Hendry, 1998), which is typically addressed by transforming errors from simple differences to log differences or into percent differences. But the need to correct for spurious correlation is also sometimes addressed by transforming both forecasts and errors into period-to-period differences, or percent growth rates (MacDonald, 1992, 1999, 2002; SM1, SM2; IMM; IMKI). Other studies motivate such a transformation intuitively (Clough, 1951; Gunnelson, Dobson, and Pamperin, 1972). SM3 use a hybrid approach that decomposes a forecast into a sum of a realized past level added to forecasted increases during subsequent periods.

Aggregation across the forecast cycle is generally undertaken by treating past forecasts as pooled or panel data (e.g. BB, BIIG, IIG, IKI, IMM, XLH). An advantage of this approach is increased effective sample size, and the disadvantages include homogenizing parameter estimates across potentially different forecast months and the need to account for correlations between forecast months and years. The expected seasonal similarity among forecast months' characteristics suggests the loss of month-specific parameters would be acceptable in some contexts, and all the aggregating studies examined here apply strategies to correct their standard error estimates for the interactions between forecasts and for heterogeneous variances. Correction strategies include using White's estimator (IMM, MIM), assuming a constant rate of decline in variance across the forecast cycle (BB, BIIG, XLH), and using the Karali and Thurman (2009) approach of constructing a feasible variance-covariance matrix from OLS residuals (IMKI).

Data

This study examines USDA's forecasts for every month during 1997-2016 of: U.S. soybean and supply and demand, U.S. soybean meal exports, Argentine production and trade, Brazilian production and trade, China's trade, and EU trade; for 1993-96 we also analyze monthly forecasts for these variables, using the traditional restriction that May's *WASDE* publication just ahead of the marketing year is USDA's first

forecast. Over the study period, the countries accounted for 86 percent of global production, 58 percent of global imports, and 89 percent of global exports.

One contribution of this analysis is to drop this restriction and provide the first published analysis of forecasts overlooked in other studies. USDA's first forecasts from 10 months before the start of the marketing year over 1997-2015 were drawn from USDA's publications summarizing annual updates to the Department's long run projections (USDA/OCE, 1998-2016), and the second forecasts disseminated 3 months later at USDA's Agricultural Outlook Forum were drawn from the grains and oilseeds ICECs' papers presented at the Forum (USDA/WFGOICEC, 1998-2016). The initial pre-season forecasts presented at USDA's Outlook Forum over 1993-1996 were not available, but all forecasts published in the *WASDE* or disseminated with greater international detail elsewhere were collected for these years as well.

Cornell University's Mann Library's comprehensive *WASDE* archive was the source of USDA's initial May *WASDE* U.S. forecasts over 1993-2009 and its initial June forecasts over 1993-2003; these publications were also the source of the international marketing year trade forecasts by USDA for Argentina and Brazil over 1993-2002 (USDA/WAOB, 1993-2016). During 1993-2002, USDA's Economic Research Service (ERS) disseminated monthly an updated database of USDA's global supply and demand estimates disaggregated by country, and ERS archives of these releases were drawn upon for this study (USDA/ERS, 1993-2002).⁶ Since 2002 USDA's Foreign Agricultural Service (FAS) has released these updates every month (USDA/FAS, 2002-16) and this study drew on ERS archives of these updates.

The characteristics of these data determined some aspects of our approach to analyzing the forecasts. It was noted earlier that soybean consumption, production, and trade has generally been growing rapidly in recent decades. Therefore, the expected means of these variables vary widely with time, and comparisons across multiple years of forecasts could be skewed by the behavior of forecasts for the years with the highest expected mean. Testing of the 27 variables examined in this study indicated that virtually every series had a mean that varied significantly over the sample studied.⁷ This motivated our decision to apply a log transformation to the forecasts, resulting in forecast errors in percent form.

Testing of forecast errors over 1993-2014 revealed the presence of structural breaks in these time series, with a large number of these breaks concentrated around 2002. Sensitivity testing of some characteristics of the forecasts for U.S. production and exports showed some shifts around that time as well as a degree

⁶In those years, ERS's database included local marketing year trade forecasts rather than international year forecasts for these countries, necessitating use of Cornell's *WASDE* archives for forecasts of South American trade.

⁷Data for 27 variables was tested for the presence of unit roots with and without deterministic trends using the modification to the traditional Dicky-Fuller test developed by Elliot, Rothberg, and Stock (1996). Our concern was not strictly focused on the presence or absence of stationarity, but on variation in expected mean over time for any reason, including non-stationary. The generalized least squares test developed by Elliot, Rothberg, and Stock modifies the augmented Dickey-Fuller test, de-trending the data before testing. In nearly half the variables examined, tests on the data for the 2004-2014 realizations failed to reject the null hypothesis of non-stationarity with or without the inclusion of a trend. In seven cases, the null hypothesis was rejected when the trend was included but not when testing without de-trending. We took this as evidence of the presence of a deterministic trend in the data-generating process, and therefore a time-varying mean. There were 6 variables whose tests rejected the null with and without de-trending. The forecasts of these variables could be analyzed without any transformation, but since the majority of variables need to be transformed for consistent analysis we proceed by analyzing all the variables in log difference form when testing for bias and revision efficiency.

of stabilization in years afterwards. Finally, 2004 marked the first year that USDA published its first international soybean forecasts in June rather than July, giving us a forecast cycle with a larger number of forecast months in the 2004-2014 period compared with samples beginning in earlier years. We therefore focused our analysis on USDA's 2004-2014 forecasts, although we touch on results for longer samples, particularly in the course of sensitivity analysis, and extend the sample to 2015 for the U.S. forecasts.

Finally, the national composition of the EU varied over 1993-2016. Thus USDA replaced its forecasts and historical data of EU soybean trade with those for updated aggregations of countries in 1995, 2004, 2007, and 2013. For the months following each appearance of a new aggregation, we constructed updates to earlier forecasts based on older aggregations based on the average month-to-month revisions to EU totals for the most recent past years published when USDA updated its aggregations.

Forecast Data Notation

We use the following notation for this notation for a given variable's forecasts:

F_{it} = USDA's forecast for marketing year t published in month i .

To increase comparability with previously published forecast evaluation studies, we maintain the traditional designation of the May pre-season forecast of each year as $i = 1$, even though it is not USDA's first forecast. For the earlier forecasts we then assign negative numbers as labels for the forecast months of November ($i = -5$) and February ($i = -2$) that occurred before the end of the previous marketing year ($t-1$). These values for i are equivalent to $i-12$ for overlapping forecasts of marketing year t in November ($i = 7$) and February ($i=10$).

For U.S. soybean variables,

$i = -5, -2, 1, 2, \dots n$; $n = 19$; $t = 1993-2015$;

for U.S. soybean meal exports,

$i = -5, -2, 1, 2, \dots n$; $n = 20$; $t = 1993-2015$;

and for non-U.S. variables,

$i = 3, 4, \dots n$; $n = 24$; $t = 1993-2003$,

$i = 2, 3, \dots n$; $n = 24$; $t = 2004-2008$,

$i = 1, 2, \dots n$; $n = 24$; $t = 2009-2014$.

When $i = n$, the forecast is considered finalized and becomes the variable's actual realization for that marketing year:

$$F_{nt} = A_t.$$

Following transformation into natural logs, we use the convention of lower-case notation:

$$f_{nt} = a_t,$$

and first differences (annual change) are indicated using ' :

$$\begin{aligned} a'_t &= a_t - a_{t-1}, \\ f'_{it} &= f_{it} - a_{t-1}, \\ f'_{it} &= a'_t. \end{aligned}$$

Forecast error is the difference of the forecast from actual realized value:

$$\begin{aligned} e_{it} &= a_t - f_{it} \\ e'_{it} &= a'_t - f'_{it} \end{aligned}$$

To maintain consistency with the majority of past studies, we follow this convention even though it results in a counter-intuitive outcome: when a forecast is larger than the realized value, its error is negative. This convention is consistent with several important theoretic properties of forecasts, resulting in its widespread application (Clements and Hendry, 1998).

The transformation of forecasts into growth rates has been widely applied to address the impact of a variable's trend on the calculation of important evaluation statistics, but the transformation has further implications for the analysis of overlapping forecasts. To see these implications, first note that when $i = j \geq n-12$, forecast error is the same for both the levels forecast (f_{jt}) and the change forecast (f'_{jt}) (i.e., $e_{jt} = e'_{jt}$). But, when $i = k < n-12$, the error of f_{kt} is the sum of the overlapping errors of f'_{kt} (the out-year forecast) and $f'_{k+12, t-1}$ (the back-year forecast), and $e_{kt} \neq e'_{kt}$:

$$\begin{aligned} f_{it} &= f'_{it} + a_{t-1}, & i \geq n-12 \\ f_{it} &= f'_{it} + f'_{i+12, t-1} & i < n-12 \\ f_{it} &= f'_{it} + f'_{i+12, t-1} + a_{t-2} & i < n-12 \end{aligned}$$

$$\begin{aligned} e_{it} &= a_t - f_{it} \\ e_{it} &= (a'_t + a_{t-1}) - (f'_{it} + a_{t-1}) = a'_t - f'_{it} = e'_{it} & i \geq n-12 \\ e_{it} &= (a'_t + a'_{t-1} + a_{t-2}) - (f'_{it} + f'_{i+12, t-1} + a_{t-2}) = (a'_t - f'_{it}) + (a'_{t-1} - f'_{i+12, t-1}) = e'_{it} + e'_{i+12, t-1} & i < n-12 \end{aligned}$$

It similarly follows for $i < n-12$ that,

$$f_{i+1, t} = f'_{i+1, t} + f'_{i+13, t-1} + a_{t-2}, \text{ so}$$

$$\begin{aligned} r_{i+1, t} &= f_{i+1, t} - f_{it} \\ r_{i+1, t} &= (f'_{i+1, t} + f'_{i+13, t-1} + a_{t-2}) - (f'_{it} + f'_{i+12, t-1} + a_{t-2}) \\ r_{i+1, t} &= (f'_{i+1, t} - f'_{it}) + (f'_{i+13, t-1} - f'_{i+12, t-1}) \\ r_{i+1, t} &= r'_{i+1, t} + r'_{i+13, t-1} \end{aligned}$$

This study's primary focus is on f'_{it} and its associated e'_{it} and r'_{it} . But, we include calculations utilizing e_{it} in appendix tables for some evaluation statistics. Our reasoning is that, even though it is only the change forecasts' errors (e'_{it}) that give information on USDA's utilization of information specific to the early forecast cycle ($i < n-12$), the levels forecasts are what USDA publishes and their errors are the ones more likely to be calculated by observers. The ability to make comparisons between the different error

measurements increases understanding of the forecasting process, and the summed overlapping errors ($e_{it} = e'_{it} + e'_{i+12, t-1}$, when $i < n-12$) may be more important in some contexts and for some readers.

Methodology

We examine each month's forecasts individually rather than using panel approaches to develop a single characterization across the entire forecast cycle (as in IMKI) or across subsets of months (as in IMM). USDA publishes 21 versions of a given marketing year's estimates for each U.S. balance sheet variable, and nearly as many versions of its foreign balance sheet estimates. With about 20 versions of about the same number of variables, more than 400 individual forecasts have to be analyzed if each month in the forecasting cycle is to be considered separately. Despite the resulting large number of forecasts, we retain the identity of each separate month in the forecasting cycle here for two reasons: the considerable heterogeneity in even some temporally adjacent months, and the applied nature of our analysis. Conclusions about USDA's forecasts that address aggregates of forecast months run the risk of being too abstract to contribute to USDA's efforts to improve the accuracy of its forecasts. The USDA's forecasters update their forecasts each month, and are keenly aware of the variations in the available information set and analytical resources available in different months. Furthermore, both Isengildina et al (2008) and Adjemian (2012) have highlighted how soybean markets differ from those of other major field crops in responding to WASDE releases outside of the typically crucial window during the U.S. harvest when NASS production forecasts coincide with publication of the WASDE forecasts. It is therefore particularly crucial to maintain a temporally disaggregated approach in evaluating USDA's soybean forecasts.

MAPE (mean absolute percent error) is the core summary statistic used here to characterize forecast performance.⁸ We utilize statistical tests to: 1) determine if USDA's forecasts have increasing or decreasing accuracy from month-to-month, 2) compare the forecasts in levels to a naïve forecast, 3) test the mean error for its difference from zero (bias), and 4) characterize patterns in USDA's monthly forecast revisions. Additional analysis examines USDA's record in forecasting directional change and serial correlation in errors across years. This additional analysis is used to develop deeper understanding of the behavior of selected forecasts. Finally, we also apply a rolling-window search strategy for sensitivity analysis of the forecast bias tests to distinguish between permanent and evolving characteristics of the USDA forecasts and to develop hypotheses regarding the sources of forecast bias.

To compare the change in accuracy of USDA's forecasts from month to month, we apply the Diebold-Mariano test (Diebold and Mariano, 2015) to determine in which months USDA's accuracy improves or diminishes. As a forecast's horizon (h) diminishes, the information set available to the forecaster increases. Thus, the accuracy of forecasts would be inversely proportional to the distance into the future of the events forecasted, and accuracy should improve as that distance diminishes.

The second statistical test is based on the Theil U statistic, which measures the ratio of a forecast's errors to the ratio of the errors of a random walk, or naïve forecast (Hyndman and Koehler, 2006). We apply the Diebold-Mariano test to determine the statistical significance of the forecast's U statistic difference from 1. Forecasts whose accuracy surpasses a random walk's have U statistic values less than 1 and those underperforming relative to that simple benchmark have U statistic values greater than 1.

⁸MAPE was chosen for its balanced treatment of unusually large errors. As a measure of central tendency, MAPE is less influenced by extreme values than is root-mean-squared percent error (RMSE), but is more influenced by them than median absolute percent error.

Bias in the mean of the forecast is the focus of our third test, using the variant of the traditional Mincer-Zarnowitz (1966) test. The more general inefficiency test of Mincer-Zarnowitz relied on the regression,

$$a'_t = \alpha_i + \beta_i f'_{it} + \varepsilon_{it}$$

which, by assuming $\beta = 1$, can be transformed into Holden and Peel's (1990) variant,

$$e'_{it} = \alpha_i + \varepsilon_{it},$$

where the test of the significance of the difference of $\alpha = \overline{e'_{it}}$ from zero becomes the test of forecast bias.

One revision characteristic examined is the relationship between sequential monthly revisions. Nordhaus (1987) introduced an efficiency test that does not rely on the determination of an actual realized value for a forecasted variable, examining the relationship among revisions in the sequence of forecasts addressing a given marketing year. If revisions are designated,

$$r'_{it} = f'_{it} - f'_{i-1t}$$

then revisions are inefficient if a non-zero coefficient (φ_i) describes the relationship between one month's revision and the revision from the previous month:

$$r'_{it} = \alpha_i + \varphi_i r'_{i-1t} + \varepsilon_{it}$$

If the forecast is unbiased, then we can assume $\alpha = 0$, but a biased forecast that met the low efficiency bar of reduced error with reduced distance from the forecast horizon could realize $\alpha \neq 0$. Both biased and unbiased forecasts could realize another form of inefficiency where $\varphi_i \neq 0$. The more intuitive manifestation of such inefficiency occurs when $\varphi_i > 0$, and the forecast is therefore smoothed: new information in one period is only partly accounted for in that period's revision and tends to produce a later revision in the same direction. When $\varphi_i < 0$, responses to information in one period are associated with offsetting responses in a following period: rather than smoothed, the forecasts are reverting.⁹

Bias in revisions is another form of inefficiency, indicative of failure to anticipate predictable patterns of behavior. Biased revisions could be indicative of biased forecasts, since the average direction of revisions in a biased forecast is predetermined, but they are also useful information in their own right. A test for revision bias is whether $\alpha_i \neq 0$ in,

$$r'_{it} = \alpha_i + \varepsilon_{it}.$$

Overlapping forecast analysis introduces another dimension to the characterization of revisions. Testing for the significance of the correlation (ρ) between the simultaneous month-to-month revisions of different marketing years' forecasts for a given variable does not necessarily test for an inefficiency, but can be useful information about forecasters' utilization of information. Sample correlation in this case is calculated as:

⁹Past studies have used the term jumpy, e.g. Isengildina-Massa, MacDonald, and Xie (2012).

$$\rho = \left[\sum_{t=1} (r'_{it} - \overline{r'_{it}})(r'_{i+12,t-1} - \overline{r'_{i+12,t-1}}) \right] * \left[\left(\sum_{t=1} (r'_{it} - \overline{r'_{it}}) \right)^2 * \left(\sum_{t=1} (r'_{i+12,t-1} - \overline{r'_{i+12,t-1}}) \right)^2 \right]^{-1/2}$$

To guard against the influence of outlying observations we take the precaution of only testing for the significance of the difference of ρ from zero when the null hypothesis of normality in the Shapiro-Wilk test cannot be rejected for both r'_{it} and $r'_{i+12,t-1}$.

Evaluation of directional accuracy for f_{it} based on 2x2 contingency tables provides additional insight into forecast behavior, an approach with roots going back as far as Zarnowitz (1967). Using χ -squared tests with large samples and Fisher's exact test in studies like this one, the test determines the probability that a forecaster's share of directionally accurate predictions is independent of observed events, and therefore has value (Joutz and Stekler, 2000). In some instances in this study, the test could not be applied because either every year's actual realization is in the same direction—as in China's soybean imports—or the USDA's forecast was in the same direction every year—as in forecasts for Argentine meal exports until $i = 10$. Further, a preponderance of actual changes in one direction creates a high bar for USDA to demonstrate independence, with a greater than 90 percent directional accuracy share needed to reach or surpass 5 percent significance in the forecasts for several variables. To avoid these issues, we apply Fisher's exact test on contingency tables for forecasts of the direction of change in the annual changes forecasted (f'_{it}) to gain some insight into USDA's forecasts of some variables.

Results

We first provide an overall summary of the results, and then give more detail for each evaluation measure. We used a subset of the variables' forecasts examined in this study to create summary statistics for each evaluation measure, for each month of the forecast cycle. The subset excluded U.S. area and yield, and excluded the geographically aggregated variables—such as world trade or foreign production—and the summary statistics are only calculated for forecast months when non-U.S. variables were forecast during 2004-15, i.e. no earlier than the first June of the forecast cycle, $i = 2$.

The median MAPE of our subset of forecasts ranges between 8.8 percent and 7.7 percent between June ($i = 2$) and the following March ($i = 10$), and begins falling relatively steadily after December ($i = 8$) (fig. 2). The first forecast month with median MAPE that is less than or equal to half the level of $i = 2$ is the second August ($i = 16$). Among the variables forecast, MAPE is generally greatest for variables with relatively low levels of activity. Residually forecast variables—like U.S. ending stocks and foreign excess demand also have relatively high MAPE.

Theil U statistics reveal relatively few problems, with only 7 out of all of the forecasts examined ever realizing a $U > 1$ in any month. There are only 6 examples of variables whose USDA forecasts in any month are significantly less accurate than a naïve forecast and a maximum of two such variables—or 12 percent of the forecasts—in any given month (fig.3). Months $i = 4$ and $i = 5$ are the worst for USDA with respect to evaluation using the Theil U statistic, and $i = 6$ is the last month when any USDA forecast is significantly less accurate than a naïve forecast.

In the Diebold-Mariano testing of the monthly changes in MAPE, month $i = 4$ stands out as a forecast cycle maximum of more than 40 percent of the subset of forecasts realizing a statistically significant deterioration in accuracy from the month before. The proportion of forecasts with rising MAPE falls sharply in December ($i = 8$), to 6 percent, and permanently falls to zero in month $i = 15$.

Bias is the most broadly realized inefficiency found in this set of USDA soybean forecasts. At 47 percent, its share of affected forecasts (in month $i = 10$) is tied with smoothing as the highest of any statistical test in any month. The last month any member of our subset realizes significant bias is $i = 18$, but some aggregated variables are biased to the very end of the forecast cycle. In the detailed results by evaluation measure below and in the subsequent discussions, bias receives the most attention.

Smoothing is also broadly realized, but is less prevalent in the forecasts examined here than is bias. While the $i = 13$ peak of 47 percent of our subset of forecasts realizing smoothing matches the peak month for bias, the prevalence of smoothing tends to be lower than bias in virtually any other month of the forecast cycle.

Mean Absolute Percent Error (MAPE)

Two immediate observations about USDA's forecasts (f'_{it}) are derived by comparing the more general statistical characteristics of the forecasted variables (Table 1) with their 2004-2014 MAPE (Tables 2, 3 and 4).¹⁰ One is the association of higher variability/trend-share ratios with larger errors, such as with U.S. ending stocks. Another is that trade variables' forecasts generally have higher MAPE than production variables', but that a trade variable's MAPE is generally inversely proportional to the volume of trade. Among the trade variables forecast, U.S. soybean imports and China's soybean meal import forecasts have the highest MAPE (peaking at 50 and 128 percent respectively), and the lowest annual volume, less than 1 million tons annually over 2004-2014. Similarly, Argentina soybean exports MAPE surpassed 40 percent in some months—comparable to U.S. soybean ending stocks—but these exports averaged under 10 million tons over 2004-2014. U.S. and Brazil soybean export MAPE never exceeds 13 percent, and volume averaged 30-40 million tons. Note that forecasts of some variables representing highly aggregated data—such as total world trade—have very small MAPEs (2-5 percent) thanks to offsetting errors among their components (Table 4).

Early-cycle U.S. beginning stock forecasts have MAPE levels that exceed those of most other balance sheet variables even after the completion of the previous marketing year. Beginning stocks out-year forecasts for year t in $i = 5$ (September, f'_{5t}) are equivalent to back-year ending stocks estimates for the previous year ($t-1$) in their 16th month of revision ($i = 17, f'_{17,t-1}$). This means that even though the previous marketing year's errors specific other balance sheet variables have been filtered out by converting the levels forecasts to change forecasts, back-year errors still find their way into the out-year estimates through beginning stocks.

Comparing out-year beginning stocks' annual-change forecasts (f'_{it}) MAPE with the MAPE for levels forecast of stocks (f_{it} , Appendix table 1) shows they are virtually equivalent, appropriately, since the previous year beginning stocks forecasts ($f'_{i+12,t-1}$) have virtually no error.¹¹ Most other variables have higher MAPE for forecasts of levels, but the exceptions—e.g., Argentine soybean exports, U.S. ending stocks—are instructive. MAPE is only smaller for the out-year change forecast when the overlapping forecasts have a common directional error and correlation. Negative correlation of overlapping errors is

¹⁰See Appendix tables 1, 2 and 3 for the MAPEs of the levels forecasts (f_{it}).

¹¹ The past-year change error ($e'_{i+12,t-1}$) included in the forecasts evaluated in the Appendix Tables (f_{it}) is from a sample shifted one year back from the current year sample, which introduces some differences even into variables with little or no overlapping past year forecast errors (e.g. beginning stocks, production).

an important factor in explaining how the MAPE of a sum of overlapping errors ($e_{it} = e'_{it} + e'_{i+12, t-1}$) can be smaller than the MAPE of one component of that sum (e'_{it}).

Test of Theil U Statistics

The patterns observed in testing the f_{it} Theil U statistics suggested that only a subset of forecast's results warranted detailed discussion. For example, all three production forecasts examined had U statistics significantly below 1 in every month of the forecast cycle, as did every forecast of China's imports; this is consistent with the large role that previous-year prices play in determining the level of production and the fact that China's imports rose every year in the 2004-2014 sample.¹² Table 5 is restricted to only the forecasts with U statistics that exceeded 1 in any month.¹³ As was the case with MAPE, the Theil U results for U.S. soybean ending stocks indicate this variable's forecasts have the weakest performance of any balance sheet variable. USDA's forecasts fail to prove statistically superior to a random walk until April ($i=12$). Earlier in the cycle ($i=4$), the U ratio rises to nearly 30 percent above 1, but the variability that makes forecasting this variable so difficult also prevents the superiority of the naïve forecast from reaching statistical significance.

The European Union (EU) soybean meal forecasts mirror the U.S. ending stock forecasts in that their Theil U rises from a little more than 1 at the start of the forecast cycle to about 1.3 during $i=4-6$. An important difference for the EU variable is that the naïve forecast is statistically superior to USDA's. USDA's forecasts for EU beans and meal imports don't achieve statistical superiority over the naïve forecast until February and March ($i=10$ and $i=11$). Table 1 shows that the ratio of variability to trend is relatively low for EU meal imports, and MAPE—at little more than 6 percent—is relatively low as that ratio often indicates, but compared with the naïve forecast benchmark, this is one of the most poorly performing forecasts.

Southern Hemisphere exports also show some months when USDA significantly under-performs relative to the naïve forecast. Argentine soybean and meal exports in months $i=4$ and $i=5$ are examples of this and will be examined again when reviewing the Diebold-Mariano testing of the month-to-month changes in USDA's MAPE.

Monthly Patterns of Changing MAPE

Using the Diebold-Mariano test with a mean absolute error loss factor to compare 2004-2014 forecasts in logs from one month of the forecast cycle (f'_{it}) with the same years' forecasts published in the preceding month (f'_{i-1t}) allows us to determine if the observed monthly change in MAPE is statistically significant. We observe that most latter months of the forecasting cycle are characterized by consecutive significant increases in accuracy (Tables 6 and 7), and also that accuracy declines in some months, particularly early in the forecast cycle. The longest continuous strings of consecutive monthly improvements in accuracy are 10 months (Argentina's production) and 9 months (U.S. soybean exports and China soybean imports). Early-cycle August ($i=4$) stands out as the month with the most variables' forecasts realizing negative

¹²Note that Brazil's 2004 production level was significantly affected by drought. Using the standard 2004-2014 sample USDA's forecast was not superior to a random walk until well into the marketing year. Dropping the outlying year resulted in a sample believed to be more representative of likely future behavior.

¹³Appendix table 4 has the U statistics for the entire U.S. balance sheet. The remaining calculations are available upon request.

accuracy shocks. A total of 7 variables' forecasts experience significant MAPE increases in August, compared with a total of 4 forecasts in July, and 3 such forecasts over the 3 months following August.

Earlier, we noted USDA's U.S. ending stocks forecasts had high early-cycle MAPE and failed to significantly surpass the naïve forecasts until late in the cycle. Diebold-Mariano testing shows the absence of month-to-month improvement in the ending stocks forecasts until $i = 12$ (April). The ending stocks forecasts also stand out for a one-month decline in MAPE of 14 percentage points at virtually the end of the forecast cycle, $i = 18$. This one month improvement in accuracy exceeds even the largest MAPE ever realized by several other variables at any point in the forecasting cycle. This highlights both the large initial magnitude of the errors in the ending stocks and the extent to which convergence to the actual realization tends to be postponed. A 2.5 percent mid-cycle deterioration in accuracy (January, $i = 9$) for the U.S. ending stocks forecast also stands out: note that every other U.S. balance sheet variable significantly improved in that month during 2004-15, suggesting U.S. ending stocks are forecast as a residual. Accuracy deteriorates for some variables a little later than this, such as in U.S. domestic consumption and China's imports, but these don't exceed 0.2 percentage point increases in error.

A standout for early-season poor performance is the forecast for EU soybean meal imports, which experienced declining month-to-month accuracy in 5 out of 6 months from July to December. EU meal imports and U.S. meal exports were the only variables to experience 3 consecutive months of significantly increasing MAPE; U.S. soybean exports and consumption each experienced 2 such consecutive months, concluding in $i = 5$ (September).

U.S. production forecasts in 2004-2015 experienced a 1.4 percent increase in MAPE in August ($i = 4$), a development that seems counter-intuitive since by the beginning of August a number of the factors influencing U.S. yields have run their course and are no longer subject to error-prone forecasting. The statistically significant declines in MAPE realized through July ($i = 3$) are indicative of this process. No significant improvement occurs in September, but a 2.2 percentage point decline in MAPE is realized in October, with further improvements in November and January. Notable as well is that non-trivial improvement occurs at the end of the forecasting cycle, 0.8 percent points in October and an additional small, but statistically significant, improvement of 0.1 percentage points in November. Recall that most forecast evaluations of U.S. production assume January ($i = 9$) marks the point of realization. Interactions between balance sheet components were not the focus of most of these studies, which were focused exclusively on production or yields. But exploring the sources of later-cycle errors in USDA forecasts of ending stocks and exports must consider the errors implicit in multi-million ton revisions in USDA's production estimates at the end of the forecast cycle. Marking January ($i = 9$) as the end point of the forecasting cycle is akin to discarding important information about USDA's soybean forecasts.

Comparing the Diebold-Mariano tests on the forecasts in levels form (Appendix tables 5 and 6) highlights a few characteristics of USDA's forecasts. The levels forecasts (f_{it}) of U.S. soybean meal have accuracy deteriorating less in June ($i = 2$) than the change forecasts (f'_{it}) do—and improve slightly in July ($i = 3$) while the change forecasts again deteriorate—because improvement in the previous year's forecasts ($f'_{i+12, t-1}$) offsets some of the out-year deterioration. Similarly, the EU soybean meal import forecasts are relatively unremarkable in their early-season accuracy deterioration when tested in levels form than when tested in change form for the same reason. USDA's July ($i = 3$) forecast for Brazilian soybean production was unchanged in levels from June in every year of the 2005-14 sample examined there, whereas the implied change forecast deteriorated in accuracy (slightly), reflecting the month $i = 15$ change in the preceding years' forecasts.

Bias in Forecast Means

At least one variable's forecast in the U.S. soybean balance sheet was biased in every month over 2004-15, but bias in multiple variables' forecasts was more common, often as many as 4 at one time (Table 8). U.S. trade variables are the most consistently biased, sustaining runs of consecutively biased months of 11 months (exports), 12 months (imports), and 13 months (meal exports). Bias in USDA's forecasts of EU soybean meal imports was found for 16 consecutive months, but other-wise, non-U.S. variables realized much shorter stretches of sustained significant bias through the forecasting cycle. Sensitivity testing indicates sustained upward bias in U.S. ending stock forecasts since the early 1990s, and downward bias in the U.S. soybean export forecasts that has grown in magnitude and spread to increasingly early portions of the forecast cycle in the most recent rolling-window samples. Shifts in signs and magnitude of bias in USDA's forecasts of U.S. production and foreign excess supply coincided with the changes in bias in the U.S. export forecasts.

In the first months of USDA's forecasting cycle during 2004-15, bias was virtually absent from the U.S. soybean balance sheet forecasts, confined exclusively to beginning stocks in November ($i = -5$) and then to imports in February ($i = -2$). Mean error is relatively close to zero in these initial months for ending stocks, soybean exports, and soybean meal exports, but rises in absolute terms and becomes significantly different from zero in a few months. Production mean error begins rising in month $i = 3$ but only becomes significantly different from zero in month $i = 4$, and then becomes consistently unbiased in this sample. Domestic use's bias is confined to month $i = 18$. Note that the bias in soybean export forecasts also persists until $i = 18$, the last of 11 consecutive months of downwardly biased forecasts.

Sensitivity testing through rolling-window estimation of 12-year samples over 1993-2015 shows how bias in USDA U.S. soybean ending stock forecasts has persisted over decades in some months of the forecast cycle, and diminished in others (fig. 4). In the most recent sample period, 2004-2015, only one month before month $i = 6$ had an average error different from zero with at least 10 percent significance. Earlier, in the 1997-2008 sample (centered in 2002), significant bias was observed in three early months, and in the oldest sample, bias was observed continuously from $i = 1$ to $i = 11$.¹⁴

Applying the same rolling-window analysis of mean forecast error to 12-year samples of mean errors in the U.S. production forecasts shows that the 2004-2015 bias shown in Table 8 is a recent phenomenon (fig. 5). Confining our discussion to the same 3 representative samples used for ending stocks, we see significant differences from zero for only a few months, and only in the most recent sample. A few 12-year samples outside these 3 showed small negative bias in the late months of the cycle, but none of the large positive mean errors in the early-cycle months of the oldest sample were significantly different from zero in statistical testing.

The last U.S. balance sheet variable's bias estimates we subject to sensitivity testing are those for soybean exports (fig.6). Over 1993-2015, bias was always present late in the forecast cycle, but in the oldest samples bias was absent from the forecasts of the earlier months. The sample with the largest and most extensive bias was the most recent one, 2004-2015. Comparison between figures 5 and 6 is suggestive of a link between the peak bias periods months of $i = 4$ and $i = 5$ in both the production and export forecasts,

¹⁴Samples for intermediate periods showed bias with magnitude and degree of persistence essentially consistent with steady transitions between the representative samples chosen.

but fully understanding the intra-seasonal and year-to-year variation of USDA's U.S. soybean export forecast bias requires still further examination. In the Discussion section of this paper, we address this issue further.¹⁵

Bias and USDA's non-U.S. Forecasts

Most of the non-U.S. variables analyzed had significant bias in more than one month of the forecast cycle (Table 9).¹⁶ The greatest mean error difference from zero was observed in foreign excess demand (7.6 percent, $i = 10$), but this difference was not statistically significant. The largest significant mean errors occur in $i = 9$, for EU soybean meal imports (-6.8 percent) and Argentina soybean meal exports (-6.2 percent). The 16 months of consecutive bias for 2004-15 forecasts of EU soybean meal imports was mentioned earlier. The next longest run of consecutively biased months for any variable is for foreign soybean production during the last 9 months of the forecasting cycle. Both world trade and China's imports are downwardly biased for 6 consecutive months near the end of the forecasting cycle. In addition, foreign soybean production and consumption are biased downward during 2 of those months.

Sensitivity testing finds that upward bias in the EU and Argentina meal trade is a recent phenomenon, found only in the most recent samples and reversed in the very oldest samples. In contrast, forecasts of China's soybean imports have been downwardly biased for decades, with the bias much larger in the oldest samples. But, in the 1990s when bias sometimes surpassed 30 percent, China's role in world markets (prior to its WTO accession) was negligible. More recently, while China's bias is much smaller in proportional terms, China has come to dominate both world trade and U.S. soybean exports. Since 2008/09, China has imported more U.S. soybeans than all remaining U.S. trade partners combined, so biased expectations regarding China's imports has significant implications for forecasting U.S. exports.

Sensitivity testing of the mean error in foreign excess demand shows trends that likely reflect this growing role of China's influence on USDA's global soybean forecasts. Figure 7 shows the 1993-2014 mean error in rolling samples of a simulated measure of forecasted foreign excess demand. Note that while Table 9 shows excess demand errors averaging as much as 7.6 percent, figure 7 shows that the difference between USDA's foreign consumption growth forecasts and its foreign production growth forecasts never exceeds 1.7 percent, but is statistically different from zero. Both estimates are consistent with mean errors for foreign excess demand in month $i = 10$ over 2004-2015 of 2-3 million tons, which was also the range of the mean error in USDA's U.S. soybean export forecasts over 2004-14 in the middle of the forecast cycle.

Bias and Overlapping Forecasts

¹⁵ Using the definition of domestic consumption used in the WASDE's world balance sheets, we don't observe similar patterns in bias for USDA's forecasts of U.S. domestic consumption. However, if we remove "residual" consumption from total consumption—leaving the sum of crush and seed use—we observe a number of late-cycle months with downward bias in the 2004-15 sample and a trend over time of an increasing number of months realizing bias.

¹⁶ Of the global variables analyzed, a smaller share was biased than was so for the U.S. variables. Less than half of the variables for individual countries had bias in at least one month of the forecasting cycle over 2004-2014, and only those variables are included in the tables. Variables with on bias were: Argentina soybean production and exports; Brazil soybean and soymeal exports; China soymeal imports; and EU soybean imports. World soybean meal exports had significant bias (-1.8 percent) in only one month ($i = 11$), and was not included in the tables.

A common pattern emerges when comparing a variable's mean error of forecasts in levels ($\overline{e_{it}}$) with its mean errors of forecasts of annual change ($\overline{e'_{it}}$): the absolute value of the levels forecast errors (Appendix tables 7 and 8) are almost invariably at their largest at the beginning of the forecast cycle, but the change forecasts begin the cycle with smaller absolute mean errors. These mean errors ($\overline{e'_{it}}$) then rise for several months before beginning to diminish. Even the variables whose levels forecasts (f_{it}) have the largest mean errors in absolute value terms have much smaller—and statistically insignificant—absolute values for their mean forecast errors in the first few months of the forecast cycle for their change forecasts (f'_{it}).

The relationship between the pair of overlapping change forecasts of a variable and its out-year levels forecast is more apparent with forecast bias than it is for most forecast evaluation statistics. Bias in the out-year level forecast in sample years t to T is the sum of bias in the out-year change forecast over t to T and the bias in the back-year change forecast over $t-1$ to $T-1$. Comparison of U.S. soybean import forecast mean error in Table 8 and Appendix Table 7 is illustrative: $\overline{e'_{-5}} = 31$ percent, $\overline{e'_{-7}} = 44$ percent, and $\overline{e_{-5}} = 74$ percent, approximately the sum of the first two.¹⁷ A graphical comparison for China soybean import forecast mean error illustrates the point across several months of the forecast cycle (fig. 8). For the out-year, the levels forecast is significantly (10 percent significance level) biased during the first 5 months of the cycle, whereas the out-year change forecast is not significantly biased until nearly a year after the first estimate. The back-year estimate is biased in months $i = 4$ to $i = 6$, corresponding to the significant bias for the out-year forecast in months $i = 16$ -18.

Analysis of Forecast Revisions

Forecast revision behavior by USDA was examined 3 ways: measuring the relationships between sequential monthly revisions of forecasts for the same marketing year (smoothing or reversion), testing the mean of forecast revisions in a given month of the forecast cycle over the sample period (bias), and measuring the correlations between the simultaneous revisions of forecasts for marketing years t and $t-1$ in a given month.

Only one forecast showed a pattern of sustained smoothing or reversion throughout the forecast cycle, late-cycle forecasts for U.S. soybean exports. For the overall group of forecasts, testing showed only 7 of the 20 examined had at least 2 sequential months in the forecast cycle with either significant smoothing ($\phi_i > 0$) or reversion ($\phi_i < 0$). Other than the 7 forecasts included in Table 10, only one other forecast had more than 2 months of even non-sequential smoothing or reversion.¹⁸ U.S. domestic consumption and EU soybean imports were the only variables whose forecasts ever tested for significant reversion in any month, and in each case the result was driven by an outlying observation, a problem highlighted in MacDonald and Isengildina-Massa (2012b).

The second December of the forecast cycle ($i = 8$) marks the beginning of a period when smoothing was at its most widespread in the 2004-15 sample. Two or more variables tested for smoothing every month from months $i = 8$ to $i = 14$. March ($i = 11$) and May ($i = 13$) were peak months in this respect, with 5 and 6 forecasts realizing smoothing respectively. U.S. soybean export forecasts were the most consistently

¹⁷Recall that $\overline{e'_{-7}}$ has a sample of 2004-2015, while the comparable back year average embedded in $\overline{e_{-5}}$ has a sample of 2003-2014. Using the correct samples, $\overline{e'_{-5}} + \overline{e'_{-7}} = \overline{e_{-5}}$ exactly.

¹⁸Forecasts of Brazilian soybean exports were smoothed in August ($i = 4$), December ($i = 8$), and in the following August ($i = 16$).

smoothed, with 7 consecutively smoothed months ($i = 9-15$), and a total of 9 such months over the entire forecast cycle. No other variable had more than 2 consecutive months of smoothing in this study. U.S. ending stocks was the variable with the next greatest number of smoothed months, 5 in total. Note that forecasts of production of soybeans in the United States were not smoothed in our findings, using a 5 percent threshold. USDA's November ($i = 7$) forecasts came quite close, with a smoothing coefficient of 0.028 with a significance level of 5.2 percent.

In the tests for revision bias, the U.S. soybean export forecasts again stand out, with average revisions significantly different from zero in 6 out of the 8 last months of the forecasting cycle (Table 11). U.S. soybean exports' revision bias is confined to the end of the forecast cycle, and is always upwards, whereas U.S. soybean meal export forecasts had both downwardly-biased early-cycle revisions and upwardly-biased revisions afterwards. Forecasts for China's soybean imports were similar to the U.S. soybean export forecasts in the respect that they had upwardly biased revisions in 2 out of the last 3 months of the forecast cycle. The biggest downwardly biased revision in U.S. soybean meal forecasts coincides with a downwardly biased revision in production, in month $i = 4$. Both forecasts have downwardly biased revisions in month $i = 2$, but the production bias in that month is extremely small.

Few of the variables studied here had one or more months with biased revisions, and the median number of variables with biased revisions in any given month over $i = 3$ to $i = 21$ is 1. Therefore, we do not include comprehensive tables for all variables of this statistic, including only a subset of the U.S. results in Table 11, along with their associated overlapping forecast revision correlations.¹⁹ The months with exceptional numbers of variables with biased revisions are $i = 2$ (with 6 such variables), $i = 4$ (5 variables), and $i = 19$ (4 variables). Note that an overwhelming majority of the instances of revision bias are associated with U.S. variables, including all such instances in $i = 2$.

In contrast with revision bias, instances of significant correlation between revisions to overlapping forecasts are more common in the non-U.S. forecasts (Table 12) than in those for U.S. soybean variables. For both U.S. and non-U.S. variables, such correlations are exclusively negative, even though negative annual serial correlation in the actual year-to-year changes is not a characteristic of any of the variables studied here. Since neither is positive serial correlation statistically significant in the actual year-to-year changes for any variable (other than EU soybean imports), it is not immediately obvious that this characteristic is an inefficiency. However, we can observe instances when the characteristic is associated with an increase in absolute mean error as the forecast cycle progresses. We can observe that the largest monthly deterioration in the accuracy of China's soybean import forecasts in the forecast cycle occurs in $i = 7$, and that month's revision is correlated with the overlapping forecast revision in $i = 19$, which is a biased revision. The $i = 7$ revision is not biased, but is unhelpful, the last in a string of increases in the absolute value of the mean 2004-14 error that began 3 months earlier. The absolute value of the variable's mean error peaks in month $i = 7$, and begins testing as bias 6 months later. Similarly, Argentine meal exports' revisions are negatively correlated with overlapping forecast revisions in every consecutive month from $i = 3$ to $i = 6$, a period of steadily rising absolute mean error to its peak. Mean error becomes statistically different from zero 3 months later.

¹⁹Revision bias instances excluded are U.S. beginning stocks in $i = 2$ (-5.5 percent) and $i = 5$ (-4.0 percent); U.S. domestic consumption in $i = 2$ (-0.6 percent), $i = 7$ (0.6 percent), $i = 14$ (0.5 percent), and $i = 19$ (-0.2 percent); and yield revisions coinciding with biased production revisions in $i = 2$ and $i = 4$.

Negative correlation of overlapping revisions of change forecasts is in one sense a function of rational forecasting strategy for levels forecasts. A rational appraisal of the information set available for early-cycle foreign variable forecasts leads to a small number of revisions in the months immediately following publication of USDA's first out-year forecasts for these variables. At the same time, overlapping back-year forecast revisions are inevitable, and with out-year, early-cycle levels forecasts relatively fixed, negative correlation in the change forecasts is inevitable. Recall that,

$$r_{i+1,t} = r'_{i+1,t} + r'_{i+13,t-1},$$

so, if $r_{i+1,t} = 0$, then

$$r'_{i+1,t} = -r'_{i+13,t-1},$$

and a tendency to not revise f_{it} results in the observed pattern. Since early-cycle U.S. variable forecasts in levels are typically revised about as much as late-cycle forecasts, we observe less of this negative correlation in the overlapping forecast revisions for U.S. variables.

Directional Accuracy

In general, USDA's production forecasts' directional accuracy is distinguishable from a random outcome earlier in the forecast cycle than it is for other components of the balance sheet (Table 13), consistent with the earlier physical realization of the event being forecast. An exception here is Brazilian production forecasts, which have one of the latest realizations of consistent non-randomly correct forecasts, $i = 21$. This is in part a function of the largely uni-directional nature of annual change in this variable (rising in 82 percent of the sample years), but also likely indicative of systemic issues in USDA's Brazil soybean production forecasts, as evidenced by the mid-cycle downward bias.

Other problematic forecasts see similarly delayed realization of consistent, statistically significant directional forecast accuracy for USDA: EU soybean meal imports ($i = 20$) and U.S. soybean meal exports (the longest delayed U.S. forecast, with $i = 16$).

Discussion

Evaluation can reveal forecasts with high error and/or significant inefficiencies, and a logical next step is thinking about measures to improve forecasts. But there is an implied intermediate step: understanding the sources of forecast error. Understanding the source of errors requires understanding the sources of the forecasts, which can be collected into three groups: the causes of the actual events forecasted, the information set available about these events, and the interpretation of the information set to produce the forecast.

Focusing on the interaction between bias in the U.S. ending stock and export forecasts, we will develop a description of how the underlying physical processes of world soybean markets and USDA's interpretation of available information results in inefficiency in some forecasts and its transmission to other forecasts. We want to strongly acknowledge at the beginning of this discussion that the ideas here are hypotheses, for which we provide less formal evidence than we did for our evaluation measures. This study is primarily an evaluation study and the causal relationships among the patterns discovered here should be the subject of another study.

Underlying Physical Processes

Vogel and Bange (1999) highlight the important role the balance sheet plays in USDA's forecasts of supply and demand. The balance sheet organizes commodity supply and use into categories. Its organization can be described by mathematical identities representing commodity flows between categories and countries. These identities govern the construction of forecasts and can be simplified into two statements. One identity describes the relationships that govern the annual flow of a commodity within a given country and another describes the balance of a commodity's trade flows around the world. In the first identity, for year t , in country j :

$$S_{t,j} = D_{t,j}$$

$$BS_{t,j} + M_{t,j} + Q_{t,j} = EX_{t,j} + C_{t,j} + ES_{t,j}$$

Where $S_{t,j}$ is supply and $D_{t,j}$ is demand. Supply is comprised of beginning stocks ($BS_{t,j}$), imports ($M_{t,j}$), and local production ($Q_{t,j}$). Demand is comprised of exports ($X_{t,j}$), domestic consumption ($C_{t,j}$), and ending stocks ($ES_{t,j}$). In this study's calculation of evaluation statistics, the forecasts for each variable were transformed into natural logarithms before analysis, but the above identity does not hold for the log-transformed variables. Leaving the data in its original units, the same identity holds for: the published forecasts of the variables, the implied forecasts of changes in level from the previous year, and the errors of these implied forecasts. For notational simplicity, we will continue to use the simple variable names rather than formally adding E' to the notation.

Figure 9 illustrates this identity for the U.S. soybean forecast errors in tons form, with the sum of the three sources of supply shown by the solid line. For the first two forecast months, the near-equivalence of the sum of the demand variables' errors and the supply errors is readily observable. Through the early months of the forecast cycle, the inverse correlation between the ending stocks and the production errors can be seen, as well as the positive correlation during those months between the export and production errors.

Turning to the second identity: globally, the difference must be zero between the sum of imports by every country in the world and the sum of exports by every country in the world, and this identity is easily broken down into its U.S. and its rest-of-world (row) components:

$$\sum_{j=1} M_{t,j} - \sum_{j=1} X_{t,j} = 0$$

$$(M_{t,us} + M_{t,row}) - (X_{t,us} + X_{t,row}) = 0$$

$$(M_{t,us} - X_{t,us}) = - (M_{t,row} - X_{t,row})$$

We can also rearrange the domestic identity to show the relationship between trade and domestic variables in any given country:

$$M_{t,j} - X_{t,j} = C_{t,j} - Q_{t,j} + \Delta ES_{t,j}$$

We can then substitute the rest-of-world's domestic side of the rearranged country j balance sheet, and then apply the insights that changes in rest-of-world ending stocks are only $1/10^{\text{th}}$ of its excess demand and that imports by the United States are only $1/100^{\text{th}}$ of exports (Table 1) to achieve a simplified statement:

$$M_{t,us} - X_{t,us} = - (C_{t,row} - Q_{t,row} + \Delta ES_{t,row})$$

$$X_{t,us} \approx C_{t,row} - Q_{t,row}$$

Earlier, figures 4-7 showed changes in percent forecast error across several decades that suggested similar timing over 1993-2014 in the development of bias in USDA's forecasts (f'_{it}) of U.S. soybean production, soybean exports, and foreign excess demand for soybeans. Figure 9 shows changes across the forecast cycle in forecast (F'_{it}) errors in tons during 2004-14 that suggest a relationship during that time between USDA's errors in forecasting exports and its errors in the foreign balances as well as errors in U.S. production. In Figure 9, the positive mean error in the U.S. export forecast predates the appearance of positive U.S. production error, and in July ($i = 3$), when U.S. production error's average is zero, the U.S. export error and the foreign excess demand error are roughly equivalent.

Interpreting the Available Information Set

The balance sheet identities need further elaboration before we can continue hypothesizing about error causality. The information set available for forecasting marketing year $t = m$ when USDA's first forecast is made ($i = -5$) includes all the past realizations of the soybean variables and exogenous variables, with $t = m-2$ being the last year of actual realization. The information set also includes forecasts of those variables for the incomplete year $t = m-1$. The characteristics of the physical events driving the balance sheet variables also determine that the forecasts of some of these variables are themselves part of the information set for forecasting other variables, and that the relationship is one-way. The fact that production occurs earlier in the forecast cycle than utilization does and that U.S. production occurs before production does in Argentina and Brazil point to uni-directional causal relationships between the errors in USDA's different soybean forecasts.

To describe the information set USDA uses in its soybean forecasting, we will add the following alternative restatement of the link between actual U.S. exports and foreign excess demand:

$$Q_{t,us} - C_{t,us} - \Delta ES_{t,us} \approx C_{t,row} - Q_{t,row}$$

If we simplify, and assume all non-U.S. production is in the Southern Hemisphere, then when can restate this accounting identity with each variable as a function of price and other factors:

$$Q_{t,us}(p_{t-1}, w_{t,us}, z_{t-1,1}) - C_{t,us}(p_t, y_{t,us}, z_{t,2}) - \Delta ES_{t,us}(E(\Delta p_{t+1}), z_{t,3}) \approx C_{t,row}(p_t^*, y_{t,wld}, z_{t,4}) - Q_{t,row}(p_t^*, w_{t,row}, z_{t,5})$$

Where p_t is the price of soybeans, $p_t^* = p_t * xr_t$ is the world price of soybeans, xr_t is the U.S. dollar exchange rate, and—for country j — $w_{t,j}$ is a proxy for yield-determining factors and $y_{t,j}$ is income. Other factors such as policy are represented by $z_{t,i}$. Soybeans have minimal effects on income and exchange rates generally speaking, so for all the balance sheet variables, the information set's exogenous variables include past values of $w_{t,j}$, $y_{t,j}$, xr_t , and $z_{t,i}$ (for year $t = m-2$ and earlier), and forecasts for years $t = m$ and $t = m-1$.

While we don't actually have access to the estimates and forecasts of the exogenous variables used by the ICEC, we can examine some publicly available historical forecasts for insights. The International Monetary Fund's (IMF) World Economic Outlook (WEO) is published twice annually and provides what we can reasonably be considered informed consensus forecasts concerning $y_{t,j}$. Reviewing the WEO

forecasts for world GDP published in $i = -4$ through $i = 13$ for 2004-14 finds a mean error indistinguishable from zero, and for China a mean indicating a tendency to underestimate the level of economic activity, but not to a statistically significant extent. The same is true for U.S. income, but in the opposite direction.

We apply the insight of Meese and Rogoff's (1983) seminal article on the inability of fundamental models to out-perform random-walk, naïve models of exchange rates²⁰, and assume that policy developments are so difficult to forecast that a significant tendency to make biased predictions would be difficult to achieve. In doing so, we reach the conclusion that in addition to $y_{t,j}$, the xr_t , and $z_{t,i}$ components of the exogenous members of the information set seem unlikely to be the sources of bias. Instead, we are left with the following potential sources of bias in USDA's soybean forecasts: 1) how USDA's ICEC has interpreted the relationships between these exogenous variables and the forecasted variables, 2) the transmission of bias from one forecasted variable to another, and 3) the data for and the interpretation of the data for one exogenous variable we have not yet discussed: $w_{t,j}$, the factors influencing yields, particularly in the United States.

U.S. Yield Forecasts

Different offices within USDA have responsibility for forecasting Q_{tus} and w_{tus} at different points of the forecast cycle, and they use different methodologies to develop their forecasts. For months $i = -5$ to $i = 3$, the WAOB-chaired ICEC forecasts yields with a mixture of quantitative- and judgment-based analysis of trends and weather impacts. In the rest of the forecast cycle, NASS uses survey methods that vary through the forecast cycle, with the most impact of these being the process used for $i = 4$ to $i = 7$. Then, NASS uses a combination of objective yield surveys in randomly sampled fields in the major producing states and telephone surveys of growers perceptions of the prospects for their own crops. One of the findings of this study is the recent development of a break in the U.S. soybean yield forecasts between $i = 3$ and $i = 4$, the point of transition between ICEC and NASS responsibility. The transition to the NASS forecasts has in recent years been associated with a significant increase in MAPE and the appearance of significant downward bias that persists until $i = 6$.

The U.S. soybean industry has been very dynamic in recent years, with technological change permitting extension of soybean production into new regions as demand has grown. It is possible that for the producers whose perceptions are surveyed, and for the industry in general, that past relationships between weather, crop conditions, and later-season yield developments no longer hold. It is perhaps worth noting that the mean error of the ICEC yield forecasts rise as its portion of the forecast cycle progresses, doubling in month $i = 3$.²¹ Thus, the misperception that guides the surveyed producers and NASS's interpretation of its survey results during $i = 4$ to $i = 6$ may be affecting the ICEC's earlier estimates to a lesser degree as well.²²

The U.S. production bias is very likely an important factor in the appearance of bias in the U.S. soybean export forecasts, and it is possibly a factor in the rising absolute value of the mean errors in the Brazil production and China import forecasts in the months immediately after $i = 4$. The direction of change in the mean error from month to month is appropriate in each case, and the reversal of change in months $i =$

²⁰Updated by Cheung et al., (2005).

²¹Albeit, not to a statistically significant level.

²²As one final observation on the topic of USDA's U.S. soybean yield forecasting during 2004-14, we will note that the bias in these months was confined to years when yields fell.

7 to $i = 9$ is also consistent with a response to U.S. yield forecasts, but rigorous testing of this hypothesis is needed, and other factors are clearly at work as well.

The late-cycle U.S. production mean errors are statistically different from zero in some samples (e.g. 2009-15 and 2002-13), highlighting the potential for transmission of error from the U.S. production estimates to the late-cycle U.S. export errors. This interaction between the U.S. production estimates and the export estimates may be the key to the unusual prevalence of smoothing in the export forecasts. The literature on revision inefficiency often looks at smoothing as in part a problem with forecasters not updating their information set, and Coibion and Gorodnichenko (2010) have shown how forecasts that are averages of those by a set of forecasters with different rates of updating are smoothed—a potential in a forecast produced by a committee such as the ICEC. Another way to look at it is if the information set is intrinsically flawed. In this case, the division of responsibilities between NASS and the ICEC point an institutional constraint on updating the information set. The ICEC is bound to NASS's long-standing schedule for production revisions, and reconciles the contradiction between the NASS estimate from $i = 9$ with the continued strength of export shipments late in the season gradually.²³

Forecast Error Transmission Between Variables and Marketing Years

If we take the perspective of saying our goal is to resolve the causes of U.S. ending stock and export bias, we can narrow it down to reconciling 4 biases: U.S. production, foreign excess demand, U.S. exports, and U.S. ending stocks. Clearly, U.S. exports and U.S. ending stocks interact—there's no a priori reason that each could not cause the other. But foreign excess demand has a similar relationship with U.S. exports. And, U.S. ending stock's poor forecasting performance in general and a few specific instances in particular suggest the ending stocks forecasts and their errors are largely on the receiving side of influence from exports rather than the other way around.

It is conceivable that the ICEC has maintained a hypothesized equilibrium stocks/use ratio above the actual average level since stocks/use has been trending down, reaching remarkably low levels as the market has shifted its reliance for late season supplies increasingly to the ever-expanding Southern Hemisphere crops. We did not include USDA's price forecasts in the original study, but reviewing USDA's forecasts for the prices relevant to the stocks forecast, we note that $i-12$ forecasts of the $t+1$ price changes²⁴ had mean errors averaging below zero in the months in which we can observe them (for ending stocks forecasts $i = -5$ and $i = 1$ to $i = 6$), but not significantly so. The variation of these price forecast errors over 2004-14 correlates with the ending stocks errors, suggesting an aspect of the forecasts beyond a simple residual. This suggests that at least early in the forecast cycle the ICEC applies economic theory to its ending stock forecasts. On the other hand, there are signs that in many months ending stocks forecasts have a residual character, given their poor accuracy relative to other forecasts examined here and the sharp month-to-month accuracy deteriorations in months $i = 4$ and $i = 9$.

An inefficiency in interpreting information embodied in revisions may transmit errors between marketing years. The out-year portion of a set of overlapping forecasts has changes in the back-year forecast as part

²³ Another aspect of the inconsistencies in the information set available to the ICEC is the behavior of the estimates for “residual” portion of U.S. soybean consumption. USDA's late-cycle estimates of residual consumption have fallen to a much lower average over 2004-2014 and the annual volatility of these estimates has risen. In some years, the ICEC has resorted to forecasting large negative values for residual consumption, generally restoring the estimate to more normal levels in October ($i=18$).

²⁴ Recall from above, $\Delta ES_{t,US}(E(\Delta p_{t+1}), z_{t,3})$.

of its information set. The corrections to USDA's back year growth rate forecasts should be part of the information used to revise the out-year forecasts, but for a number of forecasts this does not seem to be the case. As significant reductions in bias are implemented in the late-cycle back-year forecasts, negatively correlated revisions are made to the out-year forecasts. In other words, as new information indicates—using China imports as an example—that USDA has underestimated the back-year's growth rate—and USDA raises that back year growth rate--USDA has often simultaneously reduced its forecast of the out-year's growth rate. We believe the channel for this effect is a belief by the ICEC that the month-to-month changes observed, including the back year revisions, do not warrant a revision to their outlook for the out-year, which they signal by leaving the forecast unrevised in levels. The opposite revision in out-year growth rates that is unavoidably created when a back-year growth rate is revised with an unchanged levels forecast highlights leaving a forecast unchanged has resulted in the introduction of bias into the forecasts over 2004-14..

One of the interesting things about bias in the soybean forecasts is that it grows as the forecast cycle progresses for a number of variables. The transition to significance of the mean error's difference from zero as the forecast cycle progresses is in part a function of the declining variance, but increases in mean error are a factor as well. A remaining question is the role of revision inefficiency in transferring bias in foreign excess demand and in U.S. trade variables from back-year forecasts to out-year forecasts. This suggests a hypothesized process where an error in the estimate of an exogenous variable or a mistaken interpretation of correct variables creates a forecast bias that is endogenously carried forward into future years. Given the complete absence of change in direction for China's imports, one could posit how a reduced opportunity for introducing persistent errors of an offsetting nature could arise. Another possible example is U.S. ending stocks: note the overlapping biased revisions in U.S. ending stocks $i = 2$ and $i = 14$. The negative correlation between these revisions was not statistically significant due to the large number of other factors affecting the forecasts in those months, but coincidence of overlapping and opposite revisions could be part of a process where past bias transferred to later forecasts.

Soybean Meal Forecasts

The hypothetical nature of the potential causal mechanisms is highlighted by the exclusion of the examination of soybean meal in the above analysis. Soybean meal from Argentina competes with U.S. soybean exports in the EU and other markets, although not in China, which has virtually ceased importing meal. Soybean meal global imports are much less concentrated in the markets we examined than for unprocessed soybeans. While the principle of national and global balances extends to the relationship between soybean meal and unprocessed soybeans, we hypothesize that the greater dispersion of soybean meal markets creates more opportunities for balancing, offsetting errors to be randomly distributed. We further noted that the soybean meal trade variables examined were the most poorly performing forecasts by virtually every measure. This suggests limited opportunities for information used and created by USDA's soybean meal forecasts to influence the errors in USDA's forecasts for unprocessed soybeans.

Another factor which underlines the uncertainty of our hypotheses is that we have abstracted from the role of soybean meal markets, forecasts, and errors. In international trade, soybeans and soybean meal can substitute for each other to the extent that a country possesses crushing capacity, as soybean meal is a by-product of processed soybeans. The largest source of world soybean meal exports is Argentina, where soybean production is more volatile as well as its trade policy, macro-economic policy, and agricultural policy. Import markets for soybean meal are more globally dispersed than for soybeans as some countries

lack adequate crushing capacity. The EU and China together accounted for 38 percent of world soybean meal imports during 2004-14, compared with 83 percent for soybeans.

Conclusions

This study developed a set of stylized facts about USDA's soybean supply and demand forecasts. Compared with other studies, this study is the most comprehensive with respect to coverage of USDA's forecasting cycle and global forecasting effort.

We found that by the broadest measure of accuracy—MAPE—USDA's U.S. and global soybean forecasts are comparable to other agricultural forecasts. We also found significant and growing inefficiencies in important forecasts, such as those for U.S. soybean production, U.S. soybean exports, and China's soybean imports. The dynamic nature of these inefficiencies means that sound forecast evaluation requires relatively small, recent samples when evaluating forecasts. It also means that corrective measures need to be more carefully considered than simply applying a rote adjustment based on evaluation measures averaged over a possibly outmoded historical period. The history of bias in the estimates for a number of variables should be part of the ICEC's information set, but incorporating that information is not a trivial task. Note that the Oilseeds ICEC already covers more individual commodities than any of USDA's forecasting committees, limiting its ability to expand the scope of its inquiry.

We demonstrated the need to decompose overlapping forecasts on multiple forecast horizons in order to understand the behavior of USDA's early-cycle forecasts. Through this decomposition, we could observe the increase in the absolute level of bias in several variables as the forecast cycle progresses. Examining the forecasts in terms of growth rates assures the statistical validity of the evaluation statistics, but also reveals a potential link between a rational reluctance to revise highly uncertain early-cycle forecasts in terms of levels and the transmission of forecast bias from one marketing year's forecasts to a later year's forecasts. USDA publishes its forecasts as levels of activity, but forecasts of change are embedded in these levels forecasts, and for variables that show persistent positive change—such as China's soybean imports—increased orientation towards forecasting annual change rather than the annual level of activity might be appropriate.

Ultimately, sound corrective measures need to be grounded in a more complete understanding of the sources of forecast error than is possible based on this study. While more research is necessary to determine the extent to which errors in some forecasted variables affect the errors of others, it seems appropriate that USDA consider analyzing its methods for forecasting U.S. yields independent of the research we are proposing. In this era of "big data", corrective measures should be within reach, and would also be crucial to minimize informational asymmetries between market participants who have the big data resources to independently predict and game USDA's forecast inefficiencies and the smaller players who rely more on USDA's estimates.

References

Adjemian, M., (2012). "Quantifying the WASDE Announcement Effect," *American Journal of Agricultural Economics*, 94 (1): 238-256.

- Bailey, D.V., and B.W. Brorsen (1998). "Trends in the Accuracy of USDA Production Forecasts for Beef and Pork." *Journal of Agricultural and Resource Economics* 23: 515–26.
- Baur, R.F., and P.F. Orazem (1994). "The Rationality and Price Effects of USDA Forecasts of Oranges." *Journal of Finance* 49: 681-696.
- Botto, A, Isengildina, O, Irwin, S, and D. Good, (2006). "Accuracy Trends and Sources of Forecast Errors in WASDE Balance Sheet Categories for Corn and Soybeans," Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Long Beach, California, July 23-26.
- Cheung, Y.-W. , Chinn , M. , Garcia Pascual , A. (2005). "Empirical exchange rate models of the nineties: Are any fit to survive?", *Journal of International Money and Finance*, 24:1150 – 1175.
- Clough, Malcolm (1951). "Changes in Corn Acreage and Production After the Early Indications," *Agricultural Economics Research*, No. 4 1951.
- Coibion O., and Y. Gordonichenko (2010). "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts." Working Paper, No. 16537, National Bureau of Economic Research.
- Colling, P.L., and S.H. Irwin (1990). "The Reaction of Live Hog Futures Prices to USDA Hogs and Pigs Reports." *American Journal of Agricultural Economics* 71:84-94.
- Davies, A., and K.Lahiri (1995). "A New Framework for Testing Rationality and Measuring Aggregate Shocks Using Panel Data." *Journal of Econometrics* 68:205-227.
- Diebold, F.X., and J.A. Lopez (1998). "Forecasting Evaluation and Combination." In *Handbook of Statistics 14: Statistical Methods in Finance*, eds., G.S. Maddala and C.R. Rao, pp.241-68. Amsterdam: North-Holland.
- Effland, Anne (1999). "When the 20th Century was the Future: USDA's 'New' Outlook Program and the 'Future' of Public Forecasting," *Papers and Proceedings: The 10th Federal Forecasters Conference*, ed. Debra Gerald, U.S. Department of Education, National Center for Education Statistics.
- Egelkraut, Thorsten, Garcia, Philip, Irwin, Scott, and Darrel Good (2003). "An Evaluation of Crop Forecast Accuracy for Corn and Soybeans: USDA and Private Information Agencies," *Journal of Agricultural and Applied Economics*, 35(1): 79-95.
- Fortenbery, T. R., and D.A. Sumner (1993). "The Effects of USDA Reports in Futures and Options Markets." *Journal of Futures Markets* 13: 157-173.
- General Accounting Office (1988). *USDA's Commodity Program: The Accuracy of Budget Forecasts*. GAO.PEMD-88-8
- General Accounting Office (1991). *USDA's Commodity Forecasts: Inaccuracies Found May Lead to Underestimates of Budget Outlays*. GAO/PEMD-91-24.

Good, Darrel, and Scott Irwin (2006). Understanding USDA Corn and Soybean Production Forecasts: Methods, Performance and Market Impacts over 1970 – 2005, AgMAS Project Research Report 2006-01.

Granger, C. W. J. (1980). Testing for causality: A personal viewpoint. *Journal of Economic Dynamics and Control* 2, 329-352.

Granger C.W.J. and Paul Newbold (1986). *Forecasting Economic Time Series (Second Edition)*, New York: Harcourt Brace Jovanovich.

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

Holden K., and D.A. Peel (1990), "On Testing for Unbiasedness and Efficiency of Forecasts," *The Manchester School* 58(2): 120-7.

Irwin, Scott, Sanders, Dwight, and Darrel Good (2014) "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.

Isengildina, Olga, Irwin, Scott, and Darrel Good (2006). "Are Revisions to USDA Crop Production Forecasts Smoothed?" *American Journal of Agricultural Economics*, 88(4): 1091-1104.

Isengildina-Massa, Olga, Irwin, Scott, and Darrel Good (2013). "Do Big Crops Get Bigger and Small Crops Get Smaller? Further Evidence on Smoothing in U.S. Department of Agriculture Forecasts," *Journal of Agricultural and Applied Economics*, 45(1): 95-107.

Isengildina-Massa, O., Irwin, S., Good, D., and J. Gomez, (2008). "The Impact of Situation and Outlook Information in Corn and Soybean Futures Markets: Evidence from WASDE Reports," *Journal Agricultural and Applied Economics*, 40:89-103.

Isengildina-Massa, Olga, Karali, Berna and Scott H. Irwin (2013). "When do the USDA forecasters make mistakes?," *Applied Economics*, 45:36, 5086-5103

Isengildina-Massa, O., MacDonald, S., and R. Xie, (2012). "A Comprehensive Evaluation of USDA Cotton Forecasts," *Journal of Agricultural and Resource Economics*, 37(1): 98-113.

Lewis, K. and M. Manfredo (2012). "An Evaluation of the USDA Sugar Production and Consumption Forecasts." Paper presented at the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management, St. Louis, Missouri, April 16-17, 2012.

MacDonald, S. (1992). *The Accuracy of USDA's Export Forecasts*, Staff Report No. AGES 9224, U.S. Department of Agriculture, Economic Research Service.

MacDonald, S. (1999). "A Preliminary Evaluation of USDA's Export Forecasts," In *Papers and Proceedings: 10th Federal Forecasters Conference, 1999*, ed. Federal Forecasters Governing Board, pages 303-311

MacDonald, S. (2002). "Rational Commodity Forecasts: Improving USDA's Cotton Analysis," In *Papers and Proceedings: 12th Federal Forecasters Conference, 2002*, ed. Debra Gerald, U.S. Department of Education, National Center for Educational Statistics, pages 355-360.

MacDonald, S. (2005). "A Comparison of USDA's Agricultural Export Forecasts with ARIMA-based Forecasts." In *Papers and Proceedings: 14th Federal Forecasters Conference 2005*, ed. Federal Forecasters Governing Board, pages 151-165.

MacDonald, S. and Olga Isengildina-Massa (2012a). "Information Rigidity and Correcting Inefficiency in USDA's Commodity Forecasts," Selected Poster prepared for presentation at the Agricultural & Applied Economics Association's Annual Meeting, Seattle, Washington, August 12-14.

MacDonald, S. and Olga Isengildina-Massa (2012b). "Adjustment Strategies for Forecast Smoothing: A Soybean Production Forecasts Case Study," Presentation at Federal Forecasters Conference, Washington, DC, September 27, 2012.

Meese, Richard, and Kenneth Rogoff (1983). "Empirical Exchange Rate Models of the Seventies: Do they fit out of sample?," *Journal of International Economics*, 14:3-24.

Mincer, J, and V. Zarnowitz (1969). "The Evaluation of Economic Forecasts," In: *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, National Bureau of Economic Research.

Nordhaus, W., (1987). "Forecast Efficiency: Concepts and Applications," *Review of Economics and Statistics*, 69(4): 667-674.

Organization for Economic Cooperation and Development and Food and Agriculture Organization of the United Nations, (2016). *OECD-FAO Agricultural Outlook 2016-2025*.

Pons, J. (2000). "The Accuracy of IMF and OECD Forecasts for G7 Countries." *Journal of Forecasting* 19: 53-63.

Sanders, D. R., and M. R. Manfredo (2002). "USDA Production Forecasts for Pork, Beef, and Broilers: An Evaluation." *Journal of Agricultural and Resource Economics* 27:114-28.

Sanders, D. R., and M. R. Manfredo (2003). "USDA Livestock Price Forecasts: a Comprehensive Evaluation." *Journal of Agricultural and Resource Economics* 28:316-334.

Sanders, D. R., and M. R. Manfredo (2008). "Multiple Horizons and Information in USDA Production Forecasts," *Agribusiness*, 24(1):55-66.

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

Tversky, A. and D. Kahneman (1981). *The Framing of Decisions and the Psychology of Choice*, *Science*, 211: 453-458.

U.S. Department of Agriculture, (2004) Transcript, Release No. 0365.04.
<http://www.usda.gov/wps/portal/usda/usdamobile?contentid=2004/08/0365.xml>

U.S. Department of Agriculture, Foreign Agricultural Service, (2002-2017). Production, Supply, and Distribution, Downloadable Datasets.

U.S. Department of Agriculture, Office of Chief Economist, (1998-2016). *U.S. Agricultural Baseline Projections*. Staff Report WAOB 98-1 through WAOB 2005-1, OCE-2005-1 through OCE-2016-1.

U.S. Department of Agriculture, Economic Research Service, (2016). Farm Income and Wealth Statistics.
<https://data.ers.usda.gov/reports.aspx?ID=53581>

U.S. Department of Agriculture, Economic Research Service (1993-2002). Production, Supply and Distribution View, Stock #93002E.

U.S. Department of Agriculture, Wheat, Feed Grains, and Oilseeds Interagency Commodity Estimates Committees (1998-2016). “Grains and Oilseeds Outlook,” Presentation at USDA Agricultural Outlook Forum.

U.S. Department of Agriculture, World Agricultural Outlook Board, (1998-2016), *World Agricultural Supply and Demand Estimates*, WASDE-338 through WASDE-561.

U.S. Department of Agriculture, World Agricultural Outlook Board, (2016), World Agricultural Supply and Demand Estimates, WADE-553.

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

Williams, J.C., and B. D. Wright (1991). *Storage and Commodity Markets*, Cambridge University Press, Cambridge, U.K.

Xao, Jinzhi, Lence, Sergio, and Chad Hart (2014). “USDA And Private Analysts' Forecasts Of Ending Stocks: How Good Are They?,” Paper provided by Agricultural and Applied Economics Association in its series 2014 Annual Meeting, July 27-29, 2014, Minneapolis, Minnesota.

Figure 1—

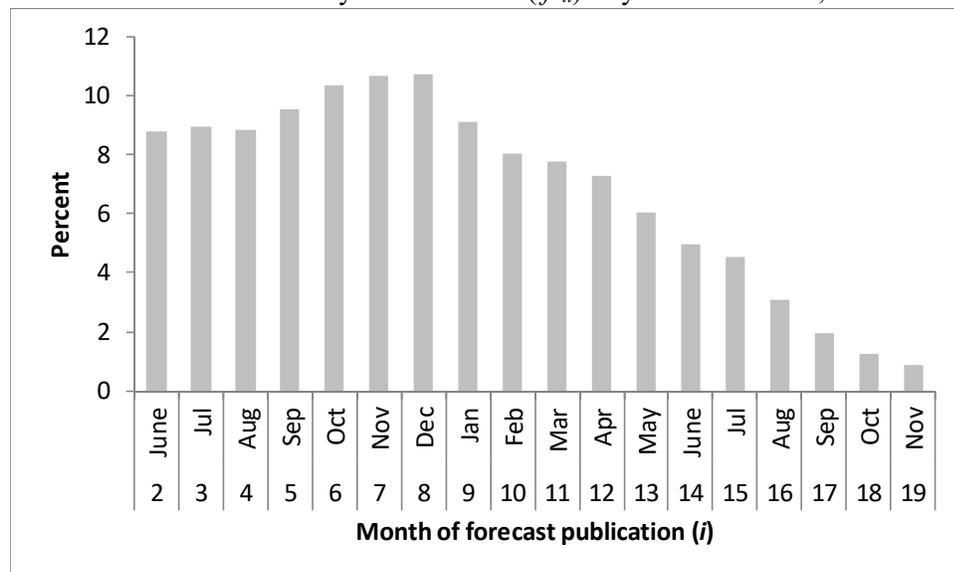
USDA's forecasting cycle for soybean marketing year $t/(t+1)$, by forecast month (fnum, i).

Year forecasted	Publication year												
		Year forecasted	Publication year										
		Year forecasted	Publication year										
$t/(t+1)$	$(t-1)/t$	--	--	-5 ¹	--	--	-2 ²	--	--	1 ³	2	3	4
$t/(t+1)$	$t/(t+1)$	5	6	7	8	9	10	11	12	13	14	15	16
$t/(t+1)$	$(t+1)/t$	17	18	19 ⁴	20	21	22	23	24 ⁴	--	--	--	--

Notes: ¹USDA's first forecast of next marketing year--not published until February. First November forecast assigned forecast number (fnum) -5 to indicate time preceding first WASDE forecast, and that 12 months will elapse until next November forecast. ²First published USDA forecast, at annual USDA Outlook Forum. ³First WASDE forecast for marketing year $t/t+1$, four months before start of marketing year. ⁴Estimate determining actual value of variables published 3 months after end of marketing year for U.S. variables, 5 months later for non-U.S. variables. ----- indicates no forecast or update published for marketing year $t/t+1$.

Figure 2—

Median MAPE of USDA soybean forecasts (f_{it})¹: by forecast month, 2004-2014.

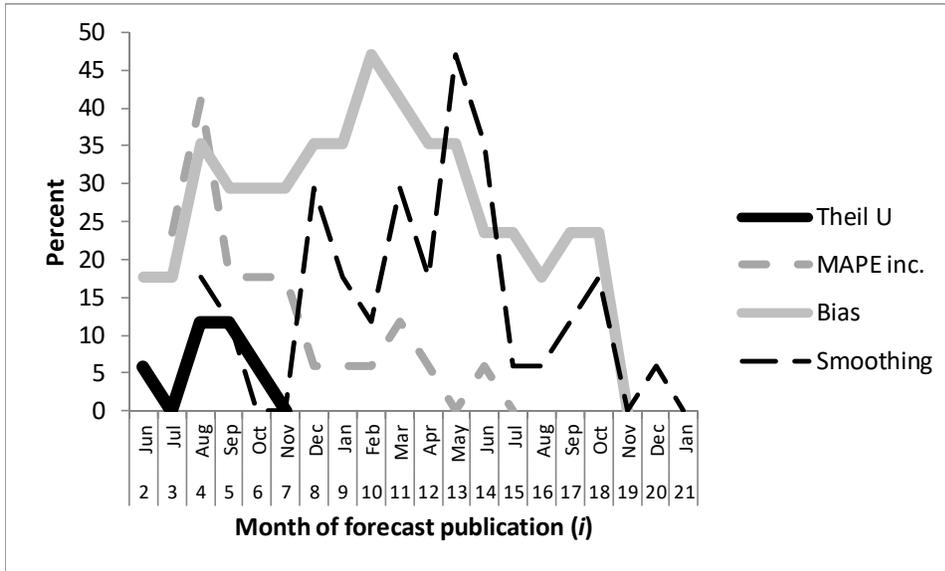


Note: ¹Median of 17 forecasts' MAPES.

Source: ERS calculations based on data from USDA/OCE (1998-2016); USDA/ WFGOICEC (1998-2016); USDA/ERS (1993-2002); USDA/FAS (2002-2016); and USDA/WAOB (1993-2016).

Figure 3—

Share of subset of USDA soybean forecasts (f'_{it}) with various inefficiencies¹: by forecast month, 2004-2014.

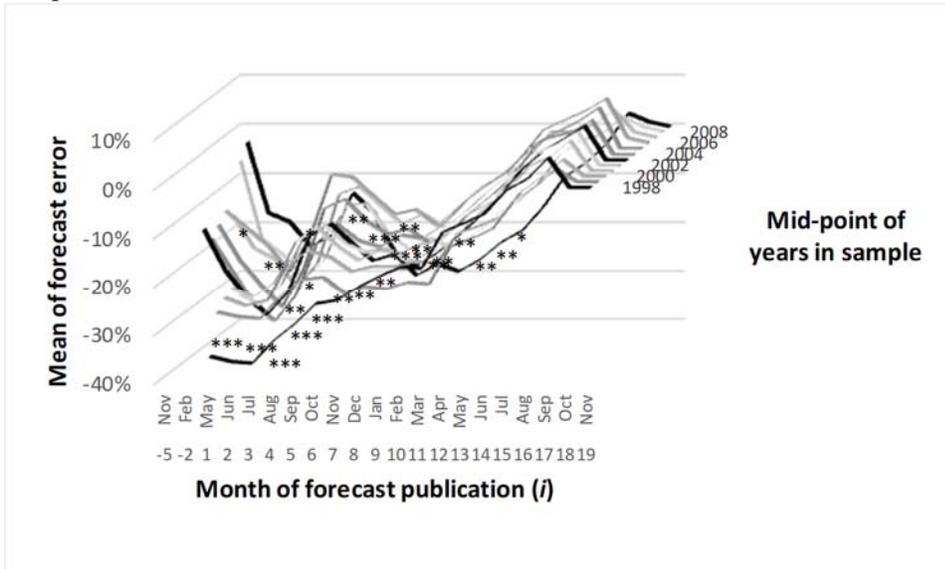


Note: ¹Share of 17-member subset whose forecasts were found inefficient in each of 4 evaluation tests.

Source: ERS calculations based on data from USDA/OCE (1998-2016); USDA/ WFGOICEC (1998-2016); USDA/ERS (1993-2002); USDA/FAS (2002-2016); and USDA/WAOB (1993-2016).

Figure 4—

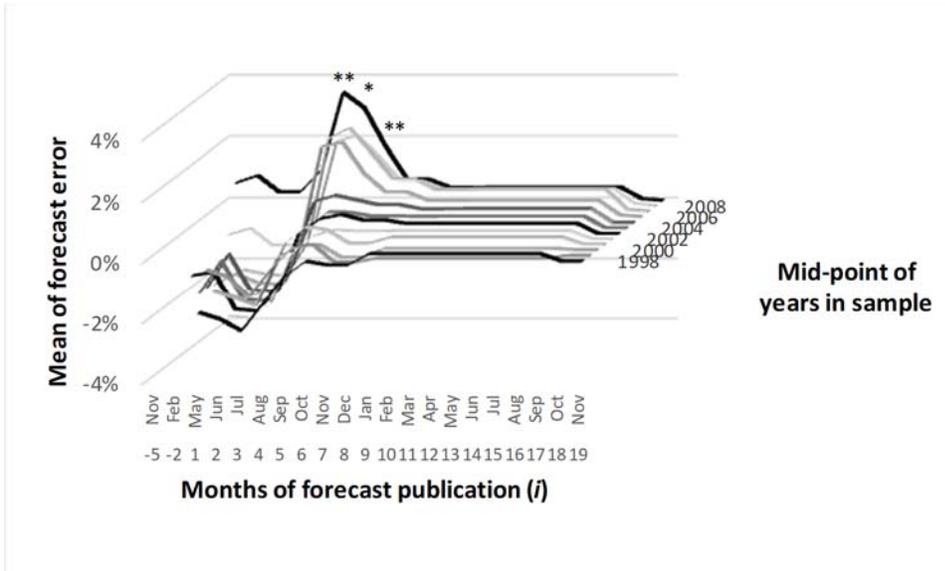
USDA U.S. soybean ending stocks forecast (f'_{it}) mean error¹: by forecast month for rolling 12-year samples, 1993-2015.



Note: ¹Bias occurs when mean difference from zero significant. (* is pval with 10% significance, ** 5%, and *** 1%).

Source: ERS calculations based on data from USDA/OCE (1998-2016); USDA/ WFGOICEC (1998-2016); USDA/ERS (1993-2002); USDA/FAS (2002-2016); and USDA/WAOB (1993-2016).

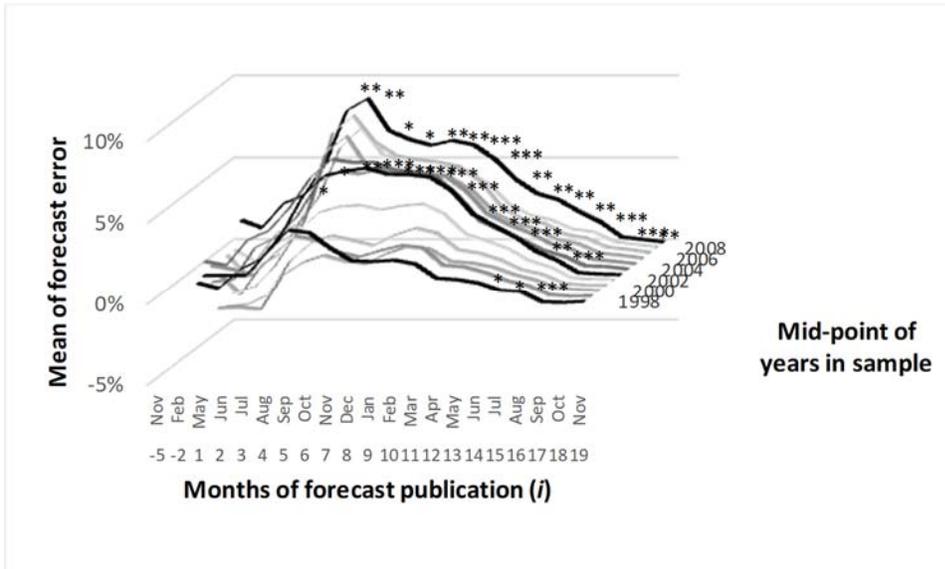
Figure 5—
 USDA U.S. soybean production forecast (f_{it}) mean error¹: by forecast month for rolling 12-year samples, 1993-2015.



Note: ¹Bias occurs when mean difference from zero significant. (* is pval with 10% significance, ** 5%, and *** 1%).

Source: ERS calculations based on data from USDA/OCE (1998-2016); USDA/ WFGOICEC (1998-2016); USDA/ERS (1993-2002); USDA/FAS (2002-2016); and USDA/WAOB (1993-2016).

Figure 6—
 USDA U.S. soybean export forecast (f_{it}) mean error¹: by forecast month for rolling 12-year samples, 1993-2015.

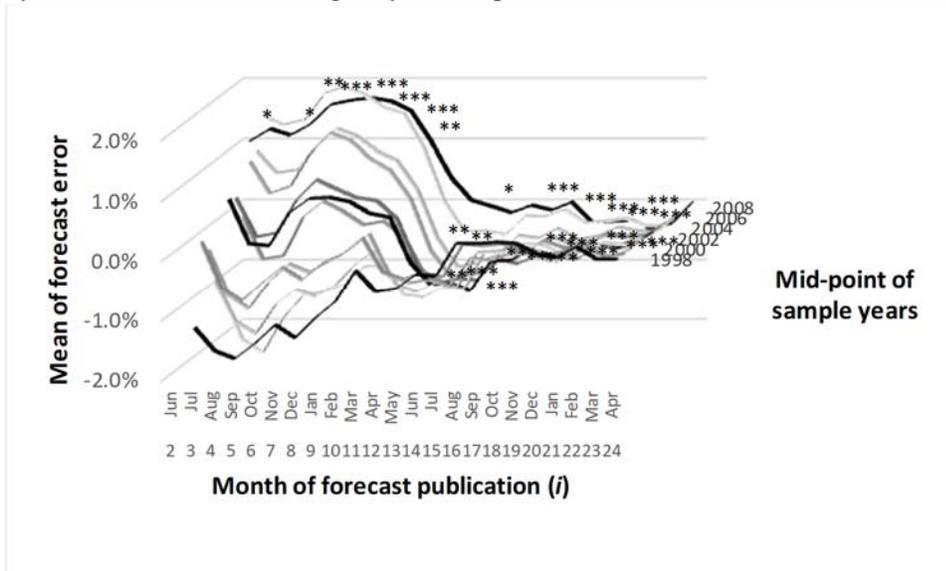


Note: ¹Bias occurs when mean difference from zero significant. (* is pval with 10% significance, ** 5%, and *** 1%).

Source: ERS calculations based on data from USDA/OCE (1998-2016); USDA/ WFGOICEC (1998-2016); USDA/ERS (1993-2002); USDA/FAS (2002-2016); and USDA/WAOB (1993-2016).

Figure 7—

USDA foreign soybean excess demand (consumption-production) forecast (f_{it}) simulated mean error^{1,2}: by forecast month for rolling 12-year samples, 1993-2015., 2004-15.

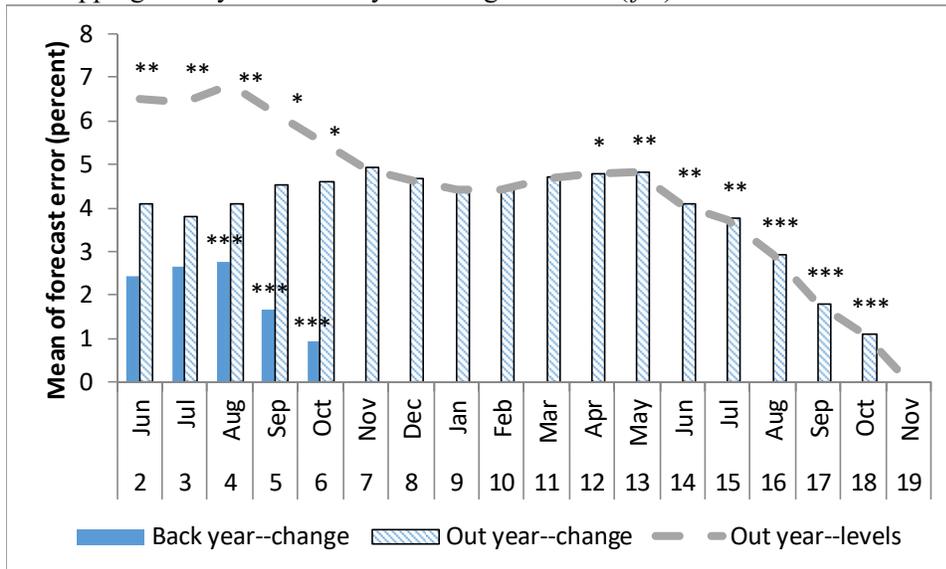


Notes: ¹Simulation is difference between errors in foreign consumption and production. ²Bias occurs when mean difference from zero significant. (* is pval with 10% significance, ** 5%, and *** 1%).

Source: ERS calculations based on data from USDA/OCE (1998-2016); USDA/ WFGOICEC (1998-2016); USDA/ERS (1993-2002); USDA/FAS (2002-2016); and USDA/WAOB (1993-2016).

Figure 8—

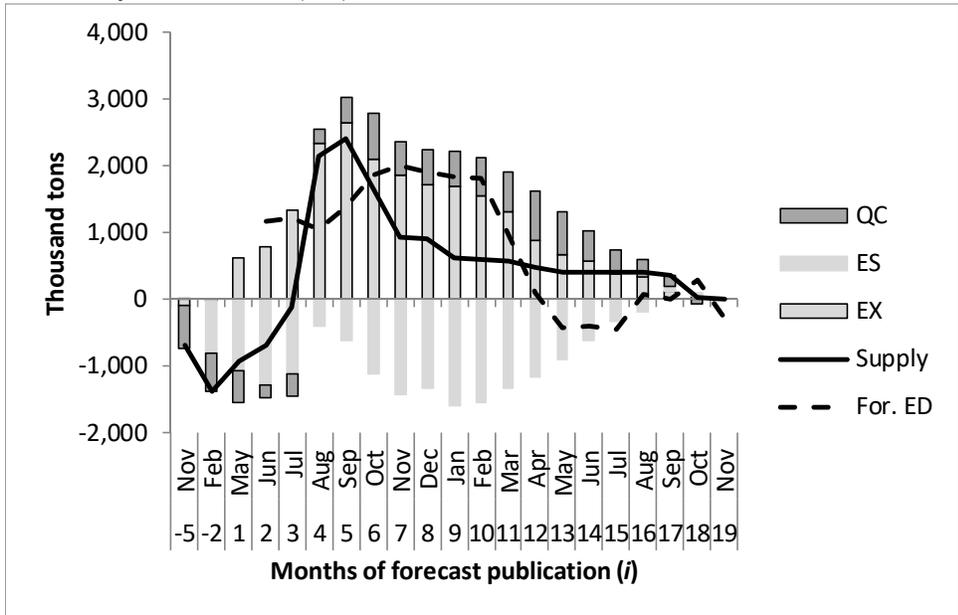
Mean errors in USDA’s 2004-14 China soybean import level forecasts (f_{it}): decomposed into overlapping back-year and out-year change forecast (f_{it}) error means.



Note: ¹Bias occurs when mean difference from zero significant. (* is pval with 10% significance, ** 5%, and *** 1%).

Source: ERS calculations based on data from USDA/OCE (1998-2016); USDA/ WFGOICEC (1998-2016); USDA/ERS (1993-2002); USDA/FAS (2002-2016); and USDA/WAOB (1993-2016).

Figure 9—
 USDA soybean forecast (F'_{it}), mean error: 2004-2014.



Source: ERS calculations based on data from

Table 1
Summary statistics by variable, 1993-2015

Country	Product	Variable	1993-2003	2004-2015		
			Mean	Mean	Coefficient of variation ¹	Trend share of variation ²
			Thousand tons			Percent
U.S.	soybean	beginning stocks	6,655	6,108	64	17
U.S.	soybean	area harvested	27,254	30,423	6	47
U.S.	soybean	yield per hectare	2.56	2.90	6	30
U.S.	soybean	production	69,236	88,552	11	47
U.S.	soybean	exports	24,273	37,925	22	84
U.S.	soybean	imports	126	567	91	49
U.S.	soybean	domestic use	45,430	51,004	5	1
U.S.	soybean	ending stocks	6,314	6,297	61	34
U.S.	soymeal	exports	6,355	8,869	18	81
World	soybean	production	157,668	249,079	15	80
World	soybean	trade	43,199	88,209	25	95
World	soymeal	trade	34,984	56,972	9	88
World	soybean	end. stock change	1,987	3,460	270	2
non-U.S.	soybean	excess demand ³	23,524	30,084	32	10
Argentina	soybean	production	20,194	46,800	18	45
Brazil	soybean	production	33,290	69,518	22	90
Argentina	soybean	exports	4,111	9,231	26	0
Brazil	soybean	exports	10,043	32,643	33	92
Argentina	soymeal	exports	11,931	25,116	10	45
Brazil	soymeal	exports	10,964	13,475	7	37
EU	soybean	imports	16,142	13,463	8	17
EU	soymeal	imports	17,671	21,275	10	59
China	soybean	imports	6,518	48,386	38	98
China	soymeal	imports	1,098	176	141	22

Notes: ¹Coefficient of variation is the ratio of the standard deviation to the mean. ²Trend share of variation is the R²-statistic from OLS regression of the the variable on a linear trend. ³non-U.S. consumption minus non-U.S. prooduction.

Source: ERS calculations based on data from USDA/FAS, 2016.

Table 2

Mean absolute percent error (MAPE), USDA forecasts of annual change (f'_{it}) in U.S. soybean supply and demand, 2004-2015.

Month	i	Area harvested	Yield	Production	Beginning stocks	Imports	Domestic consumption	Exports	Ending stocks	Soybean meal exports
Percent										
Nov	-5	3.5	4.9	5.9	31.5	49.6	4.5	7.5	38.1	15.4
Feb	-2	2.5	4.4	5.5	34.3	43.0	4.3	9.5	38.9	10.0
May	1	1.8	4.5	4.4	23.3	48.7	2.8	10.0	42.2	9.6
June	2	1.7	4.5	4.4	22.1	50.5	2.8	9.9	42.8	12.4
Jul	3	1.3	3.7	3.2	18.4	49.3	2.8	10.3	44.0	12.7
Aug	4	1.1	4.7	4.6	17.1	51.9	3.7	11.0	44.8	17.1
Sep	5	1.0	4.1	4.4	15.0	52.9	3.9	12.4	40.7	16.7
Oct	6	0.6	2.2	2.2	1.1	55.4	3.7	10.4	37.8	14.6
Nov	7	0.6	1.1	1.5	0.1	55.1	3.2	9.8	37.3	12.5
Dec	8	0.6	1.1	1.5	0.1	50.9	3.1	8.8	37.4	12.1
Jan	9	0.3	0.5	0.8	0.0	44.5	2.7	7.5	39.9	10.5
Feb	10	0.3	0.5	0.8	0.0	38.7	2.6	7.2	40.3	9.8
Mar	11	0.3	0.5	0.8	0.0	32.7	2.7	6.4	39.0	8.9
Apr	12	0.3	0.5	0.8	0.0	19.4	2.1	5.1	34.4	8.2
May	13	0.3	0.5	0.8	0.0	20.5	1.9	3.9	29.0	6.4
June	14	0.3	0.5	0.8	0.0	18.7	1.7	3.3	27.2	5.5
Jul	15	0.3	0.5	0.8	0.0	16.9	1.2	2.3	22.6	4.0
Aug	16	0.3	0.5	0.8	0.0	13.6	1.1	1.2	18.7	2.9
Sep	17	0.3	0.5	0.8	0.0	6.7	1.1	0.3	14.5	1.8
Oct	18	0.0	0.0	0.1	0.0	4.7	0.1	0.2	1.0	0.7
Nov	19	--	--	--	--	--	--	--	--	0.4

Sources: ERS calculations based on data from USDA, Office of Chief Economist (1998-2016); USDA, Wheat, Feed Grains, and Oilseeds Interagency Commodity Estimates Committees (1998-2016); USDA, Economic Research Service (1993-2002); USDA, Foreign Agricultural Service (2002-2016); and USDA, World Agricultural Outlook Board (1993-2016).

Table 3

Mean absolute percent error (MAPE), USDA forecasts of annual change (f'_{it}) in selected world soybean supply and demand variables, 2004-2014.

Month	<i>i</i>	Brazil	Brazil	Brazil	Argentina	Argentina	Argentina	EU	EU	China	China
		soybean production	soybean exports	soymeal exports	soybean production	soybean exports	soymeal exports	soybean imports	soymeal imports	soybean imports	soymeal imports
		Percent									
June	2	7.2	9.1	5.5	11.6	42.6	7.8	5.2	4.8	8.8	127.8
Jul	3	7.5	8.9	6.1	11.6	41.1	7.7	5.4	5.2	8.0	137.5
Aug	4	6.6	8.0	4.8	11.9	45.7	8.8	5.8	5.8	7.3	125.8
Sep	5	5.9	8.4	5.1	11.9	43.4	9.5	4.9	5.4	8.2	108.9
Oct	6	5.7	10.5	5.1	11.3	40.5	10.4	4.8	5.6	8.2	117.9
Nov	7	5.0	10.7	4.9	11.4	35.9	10.4	5.4	6.3	9.3	125.9
Dec	8	4.9	10.7	4.4	11.3	36.0	9.5	5.6	7.2	9.1	121.3
Jan	9	3.9	10.0	4.1	10.5	33.4	9.1	5.2	7.4	8.9	103.3
Feb	10	3.6	9.1	4.8	8.0	30.9	8.0	5.3	5.8	8.6	63.1
Mar	11	2.3	8.2	5.0	6.7	30.6	7.7	5.1	5.3	8.8	69.0
Apr	12	1.5	8.3	3.9	4.9	28.6	7.3	5.0	5.2	8.1	54.2
May	13	1.5	6.6	3.2	2.8	24.1	6.0	4.8	4.7	6.3	46.2
June	14	1.0	5.3	3.4	1.7	18.8	5.0	4.8	4.8	5.4	43.4
Jul	15	1.2	4.1	3.2	1.1	15.2	4.8	3.9	5.0	4.5	33.9
Aug	16	1.0	2.7	2.3	0.7	12.2	3.0	3.6	3.9	2.9	21.9
Sep	17	0.8	1.5	1.7	0.7	9.1	1.9	2.9	3.2	1.8	20.5
Oct	18	0.6	0.3	0.3	0.7	6.2	1.2	2.8	2.6	1.2	6.5
Nov	19	0.6	0.1	0.0	0.6	1.5	0.7	1.9	2.3	0.0	0.0
Dec	20	0.6	0.0	0.0	0.5	0.4	0.5	1.3	1.3	0.0	0.0
Jan	21	0.6	0.0	0.0	0.3	0.0	0.1	0.6	0.4	0.0	0.0
Feb	22	0.6	0.0	0.0	0.3	0.0	0.1	0.3	0.1	0.0	0.0
Mar	23	0.6	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0

Sources: ERS calculations based on data from USDA, Office of Chief Economist (1998-2016); USDA, Wheat, Feed Grains, and Oilseeds Interagency Commodity Estimates Committees (1998-2016); USDA, Economic Research Service (1993-2002); USDA, Foreign Agricultural Service (2002-2016); and USDA, World Agricultural Outlook Board (1993-2016).

Note: ¹Sample 2005-2014 used for Brazil as significant drought in 2004 rendered full sample unrepresentative.

Table 4

Mean absolute percent error (MAPE), USDA forecasts of annual change (f'_{it}) in selected world soybean forecasts, 2004-2014.

Month	<i>i</i>	Foreign soybean production	Foreign soybean consumption	Foreign soybean excess d. Percent	World meal exports	World soybean exports
Jun	2	6.9	2.3	28.2	3.0	5.5
Jul	3	6.9	2.2	26.0	2.7	5.2
Aug	4	7.2	2.5	21.5	3.3	5.8
Sep	5	7.1	2.6	22.0	3.5	5.8
Oct	6	6.9	2.4	22.4	3.6	5.3
Nov	7	6.3	2.2	22.8	3.7	5.0
Dec	8	6.3	2.2	22.7	3.9	4.8
Jan	9	5.7	2.0	22.0	3.7	4.3
Feb	10	4.5	1.6	18.8	3.1	4.2
Mar	11	3.4	1.4	16.2	2.8	4.4
Apr	12	2.0	1.3	11.0	2.4	4.7
May	13	1.2	1.0	9.3	2.1	4.4
Jun	14	0.6	1.0	8.0	2.3	4.3
Jul	15	0.8	0.9	6.1	2.5	3.1
Aug	16	0.7	0.8	4.2	1.4	2.3
Sep	17	0.6	0.6	3.7	1.3	1.3
Oct	18	0.7	0.6	3.6	0.9	0.8
Nov	19	0.6	0.3	3.6	0.7	0.5
Dec	20	0.6	0.3	3.7	0.5	0.4
Jan	21	0.4	0.3	3.0	0.3	0.4
Feb	22	0.4	0.2	3.0	0.3	0.4
Mar	23	0.4	0.1	1.8	0.1	0.3

Sources: ERS calculations based on data from USDA, Office of Chief Economist (1998-2016); USDA, Wheat, Feed Grains, and Oilseeds Interagency Commodity Estimates Committees (1998-2016); USDA, Economic Research Service (1993-2002); USDA, Foreign Agricultural Service (2002-2016); and USDA, World Agricultural Outlook Board (1993-2016).

Table 5

Theil U statistics¹, USDA forecasts of annual level (f_{it}) for selected soybean supply and demand variables, 2004-2014².

Month	i	U.S. soybean exports	U.S. ending stocks	U.S. meal exports	EU soybean imports	EU meal imports	Argentina bean exports	Argentina meal exports
Nov	-5	0.73 ***	1.06	0.74 ***	--	--	--	--
Feb	-2	0.84 *	1.22	0.76 *	--	--	--	--
May	1	0.87 *	1.25	0.71	--	--	--	--
Jun	2	0.87	1.15	0.82	0.83 ***	1.37	1.09 ^	1.31
Jul	3	0.90	1.17	0.82	0.91 *	1.33	1.20	1.28
Aug	4	1.00	1.24	1.07	0.94	1.68 ^^	1.32 ^	1.23
Sep	5	1.11	1.07	1.05	0.94 *	1.70 ^^	1.14	1.28 ^^
Oct	6	0.93 **	0.89	0.95	0.90 *	1.72 ^	1.05	1.22
Nov	7	0.87 ***	0.88	0.81	1.04	1.55	1.00	1.25
Dec	8	0.78 ***	0.88	0.78	1.00	1.62	1.00	1.18
Jan	9	0.67 ***	0.94	0.68	0.91	1.34	0.93	1.10
Feb	10	0.64 ***	0.95	0.63 *	0.89	0.93 ***	0.86 *	0.93
Mar	11	0.57 ***	0.92	0.57 *	0.86 *	0.85 ***	0.85 **	0.88
Apr	12	0.45 ***	0.81 *	0.52 **	0.77 **	0.83 ***	0.79 ***	0.83 *
May	13	0.34 ***	0.68 ***	0.41 ***	0.74 ***	0.75 ***	0.67 ***	0.71 *
Jun	14	0.29 ***	0.64 ***	0.35 ***	0.74 ***	0.76 ***	0.52 ***	0.60 ***
Jul	15	0.20 ***	0.53 ***	0.29 ***	0.65 ***	0.77 ***	0.43 ***	0.58 ***

Notes: ¹ Selected months: months omitted if they are later than the second consecutive month of highest significance in USDA superiority in worst performing forecast. ² U.S. variables' sample is 2004-15. Significance of difference from 1 based on Diebold-Mariano test: pval for superiority of USDA: *** if significance greater than 0.001, ** if greater than 0.01, and * if greater than 0.05. Pval for superiority of naive forecast: ^^ if significance greater than 0.001, ^ if greater than 0.01, and ^ if greater than 0.05.

Sources: ERS calculations based on data from USDA, Office of Chief Economist (1998-2016); USDA, Wheat, Feed Grains, and Oilseeds Interagency Commodity Estimates Committees (1998-2016); USDA, Economic Research Service (1993-2002); USDA, Foreign Agricultural Service (2002-2016); and USDA, World Agricultural Outlook Board (1993-2016).

Table 6

Statistically significant changes from previous month MAPE for USDA forecasts of annual change (f'_{it}) for selected U.S. soybean supply and demand variables (Diebold--Mariano test), 2004-15.

Month	i	Meal exports	Production	Exports	Ending stocks	Consumption	Beginning stocks
				Percent			
Feb	-2	-5.4 ***	-0.4 ***	2.0 ^^	0.8	-0.1	2.8
May	1	-0.4	-1.1 ***	0.5	3.4	-1.5 ***	-11.0 ***
Jun	2	2.8 ^^	0.0	0.0	0.5	0.0	-1.2
Jul	3	0.4 ^	-1.2 **	0.4	1.2 ^^	0.0	-3.8 ***
Aug	4	4.3 ^^	1.4 ^^	0.7 ^^	0.8	0.9 ^^	-1.3 ***
Sep	5	-0.4	-0.2	1.4 ^	-4.1	0.2 ^	-2.1 ***
Oct	6	-2.1 ***	-2.2 ***	-2.0 ***	-2.9	-0.2 *	-13.9 ***
Nov	7	-2.1 **	-0.7 **	-0.6 ***	-0.5	-0.5 ***	-1.0 **
Dec	8	-0.5	--	-1.0 ***	0.1	-0.1 ***	--
Jan	9	-1.6 ***	-0.7 ***	-1.3 ***	2.5 ^^	-0.3 ***	-0.1 ***
Feb	10	-0.7 ***	--	-0.4	0.4	-0.1	--
Mar	11	-0.9 ***	0.0 ***	-0.8 ***	-1.3	0.1 ^^	--
Apr	12	-0.7	--	-1.3 ***	-4.6 ***	-0.5 ***	--
May	13	-1.8 ***	0.0 ***	-1.2 ***	-5.4 ***	-0.2 **	0.0
Jun	14	-0.9 **	0.0 ^^	-0.6 ***	-1.8	-0.2	0.0 *
Jul	15	-1.5	--	-1.0 ***	-4.5 ***	-0.5 **	--
Aug	16	-1.1 ***	--	-1.1 ***	-3.9 ***	-0.1	0.0 **
Sep	17	-1.1 ***	--	-0.9 ***	-4.2 **	-0.1	--
Oct	18	-1.1 ***	-0.8 ***	-0.2 ***	-13.6 ***	-1.0 ***	--
Nov	19	-0.3 ***	-0.1 ***	-0.2 ***	-1.0 ***	-0.1 ***	--
Dec	20	-0.4 ***	--	--	--	--	--

Notes: Significance of difference from previous month based on Diebold-Mariano test: pval for superiority versus previous month, *** if significance greater than 0.001, ** if greater than 0.01, and * if greater than 0.05. Pval for superiority of previous month; ^^ if significance greater than 0.001, ^^ if greater than 0.01, and ^ if greater than 0.05.

Sources: ERS calculations based on data from USDA, Office of Chief Economist (1998-2016); USDA, Wheat, Feed Grains, and Oilseeds Interagency Commodity Estimates Committees (1998-2016); USDA, Economic Research Service (1993-2002); USDA, Foreign Agricultural Service (2002-2016); and USDA, World Agricultural Outlook Board (1993-2016).

Table 7

Statistically significant changes from previous month MAPE for USDA forecasts of annual change (f'_{it}) for selected non-U.S. soybean supply and demand variables (Diebold--Mariano test), 2004-2014¹.

Month	i	Brazil soybean production	Brazil soybean exports	Brazil meal exports	Argentina soybean production	Argentina soybean exports	Argentina meal exports	China soybean imports	E. Union meal imports	E. Union soybean imports
		Percent								
Jul	3	0.3	-0.2	0.6 ^{^^}	-0.1	-1.5	-0.1	-0.8 ^{***}	0.4 ^{^^}	0.1
Aug	4	-0.9 ^{***}	-1.0	-1.3 ^{***}	0.3	4.7 ^{^^^}	1.2	-0.6 ^{***}	0.6 ^{^^}	0.4 [^]
Sep	5	-0.7 ^{**}	0.5	0.3	0.0	-2.4	0.7	0.8 ^{^^}	-0.5 ^{***}	-0.9 ^{***}
Oct	6	-0.1	2.0 ^{^^^}	0.0	-0.6 ^{***}	-2.9 ^{***}	0.9 ^{^^}	0.0	0.2 ^{^^}	-0.1
Nov	7	-0.7 ^{***}	0.2	-0.2	0.1 ^{^^}	-4.6	0.0	1.1 ^{^^}	0.8 ^{^^}	0.6
Dec	8	-0.2 ^{***}	0.1	-0.5	-0.2 ^{**}	0.0	-0.8 ^{***}	-0.2 ^{***}	0.8 ^{^^}	0.2
Jan	9	-1.0 ^{***}	-0.7 ^{***}	-0.3 ^{***}	-0.8 ^{***}	-2.6 ^{***}	-0.4 ^{***}	-0.2 ^{***}	0.2	-0.3
Feb	10	-0.3 ^{***}	-0.9 ^{***}	0.7 [^]	-2.5 ^{***}	-2.5 [*]	-1.1	-0.3	-1.6 ^{***}	0.1
Mar	11	-1.3 ^{***}	-0.9 ^{***}	0.2	-1.3 ^{***}	-0.3	-0.3 ^{**}	0.2 ^{^^}	-0.5 ^{***}	-0.2
Apr	12	-0.8 ^{***}	0.1 ^{^^}	-1.1 ^{**}	-1.8 ^{**}	-2.1 ^{***}	-0.4 ^{**}	-0.8 ^{***}	-0.1	-0.2
May	13	0.0	-1.6 ^{***}	-0.6 [*]	-2.1 ^{***}	-4.5 [*]	-1.3 ^{**}	-1.7 ^{***}	-0.5 ^{***}	-0.2
Jun	14	-0.5 ^{***}	-1.3 ^{***}	0.1	-1.1 ^{***}	-5.3 ^{***}	-1.0 ^{***}	-0.9 ^{***}	0.1	0.0
Jul	15	0.1	-1.2 ^{***}	-0.2	-0.5 ^{***}	-3.5 ^{***}	-0.1	-0.9 ^{***}	0.2	-0.9 ^{***}
Aug	16	-0.1 ^{**}	-1.4 ^{**}	-0.9 ^{**}	-0.4 ^{***}	-3.0 ^{***}	-1.8 ^{***}	-1.6 ^{***}	-1.1 ^{**}	-0.3
Sep	17	-0.2 ^{***}	-1.2 ^{***}	-0.6 ^{**}	-0.02 ^{**}	-3.2	-1.1 ^{***}	-1.1 ^{***}	-0.6 ^{***}	-0.6 ^{**}
Oct	18	-0.2	-1.2 ^{***}	-1.4 ^{***}	0.0	-2.8 ^{***}	-0.7 ^{***}	-0.6 ^{**}	-0.6 [*]	-0.1
Nov	19	0.0	-0.2 ^{***}	-0.3 ^{***}	-0.1 ^{***}	-4.7 ^{**}	-0.5	-1.2 ^{***}	-0.3	-0.9 ^{***}
Dec	20	0.0	-0.1	0.0	-0.1	-1.1 ^{**}	-0.2 ^{***}	-0.001 ^{***}	-0.9 ^{***}	-0.6 ^{***}
Jan	21	0.0	0.0	0.0	-0.2 ^{***}	-0.3 ^{***}	-0.4 ^{**}	0.0	-0.9 ^{***}	-0.7 ^{***}
Feb	22	0.0	0.0	0.0	0.0	-0.04 ^{**}	-0.01 ^{***}	-0.0003 ^{**}	-0.4 ^{***}	-0.3 ^{***}
Mar	23	0.0	0.0	0.0	-0.3 ^{**}	0.0	-0.1	0.0	-0.1	-0.1
Apr	24	-0.6 ^{***}	0.0	0.0	0.0	-0.005 ^{***}	-0.01 ^{***}	0.0	-0.02 ^{***}	-0.2 ^{***}

Notes: ¹Sample covering 2005-2014 used for Brazil due to 2004 drought that rendered full sample unrepresentative. Significance of difference from previous month based on Diebold-Mariano test: pval for superiority versus previous month, ^{***} if significance greater than 0.01, ^{**} if greater than 0.05, and ^{*} if greater than 0.1. Pval for superiority of previous month; ^{^^^} if significance greater than 0.01, ^{^^} if greater than 0.05, and [^] if greater than 0.1.

Sources: ERS calculations based on data from USDA, Office of Chief Economist (1998-2016); USDA, Wheat, Feed Grains, and Oilseeds Interagency Commodity Estimates Committees (1998-2016); USDA, Economic Research Service (1993-2002); USDA, Foreign Agricultural Service (2002-2016); and USDA, World Agricultural Outlook Board (1993-2016).

Table 8

Mean error (bias¹) in USDA forecasts of annual change ($f'_{i,t}$) in U.S. soybean supply and demand, 2004-2015.

Month	i	Beginning stocks	Ending stocks	Exports	Meal exports	Imports	Domestic use	Area harvested	Yield	Production
						Percent				
Nov	-5	-24.5 *	-3.2	1.3	4.2	31.0	-1.2	0.9	-0.3	0.6
Feb	-2	-21.1	-17.5	0.9	6.4	39.6 *	0.7	0.2	0.6	0.8
May	1	-10.6	-19.3	2.3	8.3 *	42.9 **	-0.6	-0.5	0.7	0.3
Jun	2	-5.1	-24.3	2.9	11.4 **	44.7 **	0.0	-0.4	0.7	0.3
Jul	3	-2.0	-22.1	4.8	12.4 **	43.5 **	0.1	-0.5	1.5	1.0
Aug	4	-0.8	-13.5	8.0 *	16.9 **	43.8 *	1.2	-0.2	3.7 *	3.5 *
Sep	5	3.2	-17.5	8.7 *	16.7 **	44.9 *	1.2	-0.3	3.3 *	3.0
Oct	6	1.0	-25.9	6.8	14.5 **	47.6 *	1.4	-0.1	1.8 *	1.7 *
Nov	7	--	-30.3 *	6.2	12.1 **	44.0 *	0.8	0.1	0.6	0.7
Dec	8	--	-28.0 *	5.9 *	11.9 **	39.7 *	0.9	0.1	0.6	0.7
Jan	9	--	-29.5 *	6.2 *	10.3 **	35.8 *	0.7	0.2	0.3	0.4
Feb	10	--	-27.1 *	5.9 **	9.6 **	30.0 *	0.9	0.2	0.3	0.4
Mar	11	--	-23.5	5.1 **	8.6 **	23.5 *	1.0	0.2	0.3	0.4
Apr	12	--	-21.1	3.8 *	7.9 **	10.2	1.3	0.2	0.3	0.4
May	13	--	-16.3	3.0 *	6.1 *	7.5	1.1	0.2	0.3	0.4
Jun	14	--	-10.1	2.6 *	4.1	5.6	0.6	0.2	0.3	0.4
Jul	15	--	-7.3	1.9 *	2.0	5.7	0.6	0.2	0.3	0.4
Aug	16	--	-3.5	1.2 ***	0.9	3.4	0.4	0.2	0.3	0.4
Sep	17	--	2.7	0.3 ***	0.4	-5.2 *	0.2	0.2	0.3	0.4
Oct	18	--	1.0	0.1 *	0.3	-4.4 *	-0.1 *	0.0	0.0	0.1
Nov	19	--	--	--	0.3	--	--	--	--	--

Notes: ¹Mean error is the estimated value of α_i in $e'_{i,t} = \alpha_i + \varepsilon_{i,t}$ where i = forecast month and t = marketing year, and $e'_{i,t} = (a_t - a_{t-1}) - (f_{i,t} - f_{i+12,t-1})$. Logged values of the actual realization in marketing year t of the variable (A_t) and the forecast for that marketing year in month i ($F_{i,t}$) are a_t and $f_{i,t}$. Bias is present when mean error $\neq 0$. * = pvalue of 0.05, **=pvalue of 0.01, and ***=pvalue of 0.001.

Sources: ERS calculations based on data from USDA, Office of Chief Economist (1998-2016); USDA, Wheat, Feed Grains, and Oilseeds Interagency Commodity Estimates Committees (1998-2016); USDA, Economic Research Service (1993-2002); USDA, Foreign Agricultural Service (2002-2016); and USDA, World Agricultural Outlook Board (1993-2016).

Table 9

Mean error (bias¹), USDA forecasts of annual change (f'_{it}) in selected world soybean supply and demand variables, 2004-2014

Month	<i>i</i>	Argentina	Brazil	China	EU	World	U.S. share	Foreign	Foreign	Foreign
		meal exports	soybean production	soybean imports	meal imports	soybean trade	total trade	soybean production	soybean consumption	excess demand
Percent										
Jun	2	-2.4	-0.6	4.1	-3.4	1.5	0.9	-1.2	-0.2	1.4
Jul	3	-2.7	-1.1	3.8	-3.1	1.9	1.3	-1.5	-0.3	2.4
Aug	4	-4.9	-0.9	4.1	-5.0	2.3	2.3 **	-1.6	-0.5	2.6
Sep	5	-5.8	-1.8	4.5	-5.4 *	2.2	2.5 *	-2.0	-0.8	4.1
Oct	6	-6.6	-2.0	4.6	-5.3 *	1.9	2.0 *	-2.4	-0.8	6.2
Nov	7	-6.5	-1.8	4.9	-5.7 *	1.9	1.8 *	-2.4	-0.7	6.5
Dec	8	-6.0	-1.7	4.7	-6.2 *	1.8	1.7 *	-2.6	-0.9	6.7
Jan	9	-6.2 *	-1.7	4.4	-6.8 *	1.8	1.7 **	-2.5	-0.8	7.0
Feb	10	-5.3 *	-1.3	4.4	-5.5 *	2.3	1.5 **	-1.8	-0.3	7.6
Mar	11	-5.0 *	-0.6	4.7	-5.2 *	2.9	1.2 **	-1.2	-0.2	3.8
Apr	12	-4.5	0.7	4.8	-4.8 *	3.2 *	0.8 *	-0.4	0.0	1.4
May	13	-4.5 *	0.9	4.8 *	-4.6 *	3.3 *	0.5	0.2	0.3	0.6
Jun	14	-3.6 *	0.8	4.1 *	-4.7 **	3.5 *	0.2	0.4	0.3	0.4
Jul	15	-2.6	1.2 *	3.7 *	-4.8 **	2.8 **	0.1	0.5 *	0.3	0.1
Aug	16	-1.5	1.1 *	2.9 **	-3.2 **	2.0 **	0.0	0.5 *	0.4	0.8
Sep	17	-1.3	1.1	1.8 ***	-3.0 **	1.1 *	0.0	0.5 **	0.4 *	0.1
Oct	18	0.2	0.9	1.1 **	-2.4 **	0.6	-0.1	0.5 **	0.5 ***	0.7
Nov	19	-0.1	0.9	0.0	-2.1 ***	0.3	0.0	0.5 *	0.1	-1.8
Dec	20	-0.3	0.9	0.0	-1.2 *	0.1	0.0	0.4 *	0.1	-1.8
Jan	21	0.0	0.5	0.0	-0.3	0.2	0.0	0.4 **	0.1	-1.7
Feb	22	0.0	0.5	0.0	0.0	0.2	0.0	0.4 **	0.0	-2.4 **
Mar	23	0.0	0.5	0.0	0.0	0.2	0.0	0.3 **	0.0	-1.7 ***

Notes: ¹Mean error is the estimated value of α_i in $e'_{i,t} = \alpha_i + \varepsilon_{i,t}$ where i = forecast month and t = marketing year, and $e'_{i,t} = (a_t - a_{t-1}) - (f_{i,t} - f_{i+12,t-1})$. Logged values of the actual realization in marketing year t of the variable (A_t) and the forecast for that marketing year in month i ($F_{i,t}$) are a_t and $f_{i,t}$. Bias is present when mean error $\neq 0$. * = pvalue of 0.05, **=pvalue of 0.01, and *** =pvalue of 0.001.

Sources: ERS calculations based on data from USDA, Office of Chief Economist (1998-2016); USDA, Wheat, Feed Grains, and Oilseeds Interagency Commodity Estimates Committees (1998-2016); USDA, Economic Research Service (1993-2002); USDA, Foreign Agricultural Service (2002-2016); and USDA, World Agricultural Outlook Board (1993-2016).

Table 10

Revision inefficiency¹, USDA forecasts of annual change (f'_{it}) in selected soybean supply and demand variables (with constant in test equation), 2004-2015.

Month	<i>i</i>	U.S. ending stocks	U.S. domestic consumption	U.S. soybean exports	China soybean imports	Brazil soybean production	Argentina soybean production	Argentina soybean exports
May	1	0.0	-0.8 ***	0.0	--	--	--	--
Jun	2	-0.1	0.0	0.0	--	--	--	--
Jul	3	-0.2	0.7	2.1 *	--	--	--	--
Aug	4	0.3	0.6	0.9	0.6 *	-0.8	0.4	0.5
Sep	5	0.4	0.3 *	0.2	0.7	-0.1	0.3 *	-0.2
Oct	6	0.2	-0.5	-1.1	-0.1	-0.1	-0.4	0.1
Nov	7	0.3	0.1	0.3	0.0	0.2	0.0	-0.5
Dec	8	0.4 *	0.5 *	0.3	0.1	0.2 **	0.0	0.0
Jan	9	0.9 **	1.4 **	1.2 ***	0.8	0.0	0.5	0.0
Feb	10	0.2	0.4	0.7 ***	0.5	0.3	2.1 *	0.4
Mar	11	0.5 *	0.3 **	0.4 *	0.1	1.0 *	0.2	0.3 *
Apr	12	-0.1	-0.6	1.0 **	-0.1	1.0 ***	0.7	0.4
May	13	0.6	0.1	0.4 *	1.4 *	0.3 ***	1.2 ***	1.6 **
Jun	14	0.6 ***	0.2	0.5 **	0.4 **	0.6	0.4 ***	0.7 **
Jul	15	0.3	0.7	0.7 *	0.5	0.2	0.2	0.0
Aug	16	0.4	0.3	0.8	0.7	0.0	0.4	0.5
Sep	17	0.6 ***	0.3	0.5 ***	0.3	0.0	0.1	0.5
Oct	18	0.6	0.0	-0.1	0.5 *	0.0	--	0.1
Nov	19	0.0	0.0	-0.4	-0.4	--	--	-0.1
Dec	20	--	-0.1 *	--	0.0 *	--	0.2	0.0
Jan	21	--	-0.3	--	--	--	0.2	0.0
Feb	22	--	--	--	--	--	--	0.0

Notes: ¹Revision inefficiency is measured with the estimated value of φ_i in $r'_{it} = \alpha_i + \varphi_i r'_{i-1,t} + \varepsilon_{it}$,

where i = forecast month and t = marketing year, and $r'_{it} = f'_{it} - f'_{i-1,t}$. * = pvalue of 0.05, **=pvalue of 0.01, and ***=pvalue of 0.001.

Sources: ERS calculations based on data from USDA, Office of Chief Economist (1998-2016); USDA, Wheat, Feed Grains, and Oilseeds Interagency Commodity Estimates Committees (1998-2016); USDA, Economic Research Service (1993-2002); USDA, Foreign Agricultural Service (2002-2016); and USDA, World Agricultural Outlook Board (1993-2016).

Table 11

Mean revisions to USDA's forecasts of annual change (f'_{it}) in U.S. soybean supply and demand variables and the correlation of month-to-month forecast revisions with contemporaneous revisions to the past marketing year's forecast revisions¹ of the same variable ($f'_{i+12, t-1}$), 2004-2015.

Month	i	U.S. production		U.S. ending stocks		U.S. soybean exports		U.S. meal exports		China soybean imports	
		Mean	Correlation	Mean	Correlation	Mean	Correlation	Mean	Correlation	Mean	Correlation
Percent											
Feb	-2	-0.2	--	14.3	-15	0.5	-77 ***	-2.2	-22	--	--
May	1	0.5	27	1.8	-19	-1.5	-35	-2.2	-49	--	--
Jun	2	0.0 **	-12	5.0 *	-25	-0.6	-77 ***	-2.8 *	-21	--	--
Jul	3	-0.7	--	-2.2	-23	-1.9	-45	-0.8	-18	0.3	-87
Aug	4	-2.5 *	--	-8.6	-64 **	-3.2	-3	-4.3 *	-64 **	-0.3	-68 **
Sep	5	0.5	--	4.1	-46	-0.7	-8	-0.1	-51 *	-0.5	-57
Oct	6	1.3	-6	8.4	-33	2.0	-37	2.0	-35	-0.1	-48
Nov	7	1.0	31	4.4	-6	0.6	-13	2.3 *	-6	-0.3	-52 *
Dec	8	---	--	-2.4	--	0.3	--	0.2	-58 **	0.2	-15
Jan	9	0.3	32	1.5	52 *	-0.3	--	1.3	--	0.2	--
Feb	10	---	--	-2.4	--	0.3	--	0.3	--	0.0	--
Mar	11	0.0	--	-3.5	--	0.8	--	0.9 *	--	-0.2	--
Apr	12	---	--	-2.4	--	1.3 *	--	0.7	--	-0.1	--
May	13	0.0	--	-4.8	--	0.8 *	--	1.9 *	--	0.0	--
Jun	14	0.0	--	-6.2 *	--	0.4	--	2.0 ***	--	0.7	--
Jul	15	---	--	-2.8	--	0.7 *	--	1.9	--	0.3	--
Aug	16	---	--	-3.9	--	0.6	--	1.1 *	--	0.8	--
Sep	17	---	--	-6.2 *	--	0.9 **	--	0.5	--	1.1 *	--
Oct	18	0.4	--	1.8	--	0.2 *	--	0.1	--	0.7	--
Nov	19	0.1	--	1.0	--	0.1 *	--	0.1	--	1.1 **	--
Dec	20	---	--	---	--	---	--	0.3	--	0.0	--

Notes: ¹ Correlation is between revisions to forecasts for marketing year t in month i (r'_{it}) with revisions to forecasts for the same variable for marketing year $t-1$ in month $i+12$ ($r'_{i+12, t-1}$). * = pvalue of 0.05, **=pvalue of 0.01, and ***=pvalue of 0.001; in each case the null hypothesis of normality in the Shapiro-Wilk test for both of the two sets of revisions was not rejected. -- indicates one or both of the potential revisions never occurred within the sample examined.

Sources: ERS calculations based on data from USDA, Office of Chief Economist (1998-2016); USDA, Wheat, Feed Grains, and Oilseeds Interagency Commodity Estimates Committees (1998-2016); USDA, Economic Research Service (1993-2002); USDA, Foreign Agricultural Service (2002-2016); and USDA, World Agricultural Outlook Board (1993-2016).

Table 12

Correlation of USDA's month-to-month soybean change forecast (f'_{it}) revisions with the contemporaneous revisions to change forecasts for the past marketing year's forecast¹ of the same variable, selected non-U.S. forecasts, 2004-2014.

Month	<i>i</i>	Argentina soybean exports	Argentina soybean production	Argentina meal exports	Brazil soybean exports	Brazil soybean production	Brazil meal exports	China meal imports	EU meal imports	EU soybean imports
		Percent								
Jul	3	-72 **	-75 ***	-76 ***	-91 ***	-100	-89	-97 ***	-82	-96 ***
Aug	4	-48	-76	-55 *	-70	4	-88 ***	-67	-89	-53
Sep	5	-91 ***	7	-93 ***	-88 **	-64 **	-37	-9	-98	-78
Oct	6	-80	--	-92 ***	-72 **	--	-81 ***	-99	-58	-53 *
Nov	7	-93	-6	-9	12	--	-7	-89	-97 ***	-70 **
Dec	8	-78	--	-55 *	-51	--	40	--	-36	-48
Jan	9	-16	-23	-88	-49	-64	-3	--	-89	19
Feb	10	6	--	34	-12	--	--	--	2	-17
Mar	11	--	-74 ***	-44	--	--	--	--	24	-42

Notes: ¹ Correlation is between revisions to forecasts for marketing year t in month i (r'_{it}) with revisions to forecasts for the same variable for marketing year $t-1$ in month $i+12$ ($r'_{i+12,t-1}$). * = pvalue of 0.05, **=pvalue of 0.01, and ***=pvalue of 0.001; in each case the null hypothesis of normality in the Shapiro-Wilk test for both of the two sets of revisions was not rejected. -- indicates one or both of the potential revisions never occurred within the sample examined.

Sources: ERS calculations based on data from USDA, Office of Chief Economist (1998-2016); USDA, Wheat, Feed Grains, and Oilseeds Interagency Commodity Estimates Committees (1998-2016); USDA, Economic Research Service (1993-2002); USDA, Foreign Agricultural Service (2002-2016); and USDA, World Agricultural Outlook Board (1993-2016).

Table 13

Share of years with positive annual annual change ($a'_t > 0$) and USDA forecast cycle's first month of consistent directional forecast accuracy¹ (Fishers exact test) by soybean variable, 2004-2014.

Country	Product	Variable	Years with $a'_t > 0$	Month	i	Directional accuracy share
			Percent			Percent
U.S.	soybean	beginning stocks	50	May	1	100 **
U.S.	soybean	area harvested	67	Jul	3	92 *
U.S.	soybean	yield per hectare	50	Nov	7	92 *
U.S.	soybean	production	50	Jan	9	92 *
U.S.	soybean	exports	75	Mar	11	83 *
U.S.	soybean	imports	58	Apr	12	92 *
U.S.	soybean	domestic use	58	Jul	15	83 *
U.S.	soybean	ending stocks	58	May	13	100 **
U.S.	soymeal	exports	83	Aug	16	92 *
Argentina	soybean	production	64	Feb	10	100 **
Brazil	soybean	production	82	Jan	21	100 *
Argentina	soybean	exports	64	June	14	91 *
Brazil	soybean	exports	82	Sep	17	100 *
Argentina	soymeal	exports	64	Sep	17	91 *
Brazil	soymeal	exports	45	Sep	17	100 **
EU	soybean	imports	55	Aug	16	100 **
EU	soymeal	imports	55	Dec	20	91 *
China	soybean	imports	100	n.a. ²	n.a.	n.a.
China	soymeal	imports	64	Feb	10	91 *

Notes: ¹Accuracy is considered consistent starting in given month if the proportion of correct years is statistically greater than a potentially random share in every remaining month of the forecast cycle. ²Since imports never fell, 2x2 contingency tables cannot be constructed. Note that June, $i=14$, is first month directional accuracy begins sustained value of at least 91 percent. * = pvalue of 0.05, **=pvalue of 0.01, and ***=pvalue of 0.001.

Sources: ERS calculations based on data from USDA, Office of Chief Economist (1998-2016); USDA, Wheat, Feed Grains, and Oilseeds Interagency Commodity Estimates Committees (1998-2016); USDA, Economic Research Service (1993-2002); USDA, Foreign Agricultural Service (2002-2016); and USDA, World Agricultural Outlook Board (1993-2016).

Appendix table 3

MAPE of selected USDA soybean world forecasts, 2004-2014 (forecasts in levels, f_{it}).

Month	<i>i</i>	Foreign soybean production	Foreign soybean consumption	Foreign soybean excess d.	World meal exports	World soybean exports
				Percent		
Jun	2	6.9	3.2	22.1	4.8	6.0
Jul	3	6.9	2.9	22.7	4.6	5.8
Aug	4	7.0	2.6	22.0	3.6	5.8
Sep	5	7.0	2.7	24.6	3.7	5.8
Oct	6	6.7	2.5	25.1	3.6	5.6
Nov	7	6.2	2.4	24.5	3.4	5.3
Dec	8	6.2	2.4	24.7	3.7	5.0
Jan	9	5.8	2.2	24.7	3.6	4.6
Feb	10	4.6	1.8	21.7	3.1	4.4
Mar	11	3.5	1.5	17.8	3.0	4.7

Sources: ERS calculations based on data from USDA, Office of Chief Economist (1998-2016); USDA, Wheat, Feed Grains, and Oilseeds Interagency Commodity Estimates Committees (1998-2016); USDA, Economic Research Service (1993-2002); USDA, Foreign Agricultural Service (2002-2016); and USDA, World Agricultural Outlook Board (1993-2016).

Appendix table 4

Theil U statistics for USDA forecasts (f_{it}) for U.S. soybeans, 2004-2015.

Month	i	Domestic			Ending stocks	Meal exports
		Production	use	Exports		
Nov	-5	0.60 **	0.67 ***	0.73 ***	1.06	0.74 ***
Feb	-2	0.56 **	0.86 **	0.84 *	1.22	0.76 *
May	1	0.43 ***	0.63 ***	0.87 *	1.25	0.71
Jun	2	0.43 ***	0.69 **	0.87	1.15	0.82
Jul	3	0.31 ***	0.72 **	0.90	1.17	0.82
Aug	4	0.49 ***	0.87	1.00	1.24	1.07
Sep	5	0.46 **	0.86	1.11	1.07	1.05
Oct	6	0.25 ***	0.84	0.93 **	0.89	0.95
Nov	7	0.17 ***	0.73	0.87 ***	0.88	0.81
Dec	8	0.17 ***	0.70	0.78 ***	0.88	0.78
Jan	9	0.10 ***	0.62	0.67 ***	0.94	0.68
Feb	10	0.10 ***	0.59 *	0.64 ***	0.95	0.63 *
Mar	11	0.09 ***	0.61	0.57 ***	0.92	0.57 *
Apr	12	0.09 ***	0.49 **	0.45 ***	0.81 *	0.52 **
May	13	0.09 ***	0.44 **	0.34 ***	0.68 ***	0.41 ***
Jun	14	0.09 ***	0.38 ***	0.29 ***	0.64 ***	0.35 ***
Jul	15	0.09 ***	0.28 ***	0.20 ***	0.53 ***	0.29 ***

Notes: Significance of difference from 1 based on Diebold-Mariano test: pval for superiority of USDA, . *** if significance greater than 0.001, ** if greater than 0.01, and * if greater than 0.05. Pval for superiority of naive forecast; ^^^ if significance greater than 0.001, ^^ if greater than 0.01, and ^ if greater than 0.05.

Sources: ERS calculations based on data from USDA, Office of Chief Economist (1998-2016); USDA, Wheat, Feed Grains, and Oilseeds Interagency Commodity Estimates Committees (1998-2016); USDA, Economic Research Service (1993-2002); USDA, Foreign Agricultural Service (2002-2016); and USDA, World Agricultural Outlook Board (1993-2016).

Appendix table 5

Statistically significant changes from previous month in 2004-2015 MAPE, USDA U.S. soybean forecasts (Diebold--Mariano test, forecasts in levels, f_{it}).

Month	i	Meal exports	Production	Exports	Ending stocks	Consumption	Beginning stocks
				Percent			
Feb	-2	-0.5	-0.6 ***	1.3 ^^	0.6	1.3	3.2 ^
May	1	-1.7 ***	-1.2 ***	-1.1	-1.7 *	-1.9 **	-11.0 ***
Jun	2	1.3	0.0	0.0	-1.7 ***	0.1	-1.2
Jul	3	-1.2 *	-1.2 **	-0.3	0.4	-0.1	-3.8 ***
Aug	4	3.2 ^^	1.7 ^^	0.7 ^^	1.8	0.9 ^^	-1.3 ***
Sep	5	-1.0	-0.3	0.7 ^	-5.6 ***	-0.1	-2.1 ***
Oct	6	-2.2 ***	-2.1 ***	-2.1 ***	-1.0	-0.8 ***	-13.9 ***
Nov	7	-2.4 ***	-0.8 **	-0.6 ***	0.4	-0.4 ***	-1.0 **

Notes: Significance of difference from previous month based on Diebold-Mariano test: pval for superiority versus previous month, *** if significance greater than 0.01, ** if greater than 0.05, and * if greater than 0.10. Pval for superiority of previous month; ^^ if significance greater than 0.01, ^ if greater than 0.05, and ^ if greater than 0.10.

Sources: ERS calculations based on data from USDA, Office of Chief Economist (1998-2016); USDA, Wheat, Feed Grains, and Oilseeds Interagency Commodity Estimates Committees (1998-2016); USDA, Economic Research Service (1993-2002); USDA, Foreign Agricultural Service (2002-2016); and USDA, World Agricultural Outlook Board (1993-2016).

Appendix table 6

Statistically significant changes from previous month in 2004-2014¹ MAPE, selected USDA soybean forecasts (Diebold--Mariano test, forecasts in levels, f_{it}).

Month	i	Brazil soybean production	Brazil soybean exports	Brazil meal exports	Argentina soybean production	Argentina soybean exports	Argentina meal exports	China soybean imports	E. Union meal imports	E. Union soybean imports
				Percent						
Jul	3	--	-0.4 ***	-0.1	0.0	0.4	-0.02 ***	-0.1	-0.3 *	0.2 ^
Aug	4	-0.8 **	-1.4 ***	-0.4	0.3 ^^	6.2 ^^	-1.1 ***	0.1	-0.2	-0.7 **
Sep	5	-0.7 *	-0.4	0.1	0.0	2.7 ^^	0.4 ^^	0.5 ^	0.1 ^^	0.8 ^^
Oct	6	-0.1	1.1 ^^	--	-0.6 ***	-1.9 **	-0.5 **	-0.3	-0.6 ***	-0.1
Nov	7	-0.7 **	-0.1	-0.8 *	0.1 ^	-1.5	-0.2	0.6 ^^	-0.1 **	0.3 ^
Dec	8	-0.2 **	0.1	-0.5	-0.2 *	-0.6 ***	-0.4 ***	-0.2 ***	0.1	-0.6 *
Jan	9	-0.6 ***	-0.7 ***	-0.1	-1.0 ***	-2.4 ***	-0.2 **	-0.2 ***	-0.2	-0.5
Feb	10	-0.3 **	-0.8 *	0.7	-2.5 ***	-2.5 **	-1.4 ***	-0.3	-2.1 ***	-0.3
Mar	11	-1.3 ***	-0.9 ***	0.2	-0.7 ***	-0.3	-0.3 ***	0.2 ^^	-0.5 ***	-0.1
Apr	12	-1.3 ***	0.2	-1.1 *	-1.8 **	-2.1 ***	-0.1	-0.8 ***	-0.1	-0.4 ***

Notes: ¹Sample 2005-2014 used for Brazil as significant drought in 2004 rendered full sample unrepresentative. Significance of difference from previous month based on Diebold-Mariano test: pval for superiority versus previous month, *** if significance greater than 0.001, ** if greater than 0.01, and * if greater than 0.05. Pval for superiority of previous month; ^^ if significance greater than 0.001, ^^ if greater than 0.01, and ^ if greater than 0.05.

Sources: ERS calculations based on data from USDA, Office of Chief Economist (1998-2016); USDA, Wheat, Feed Grains, and Oilseeds Interagency Commodity Estimates Committees (1998-2016); USDA, Economic Research Service (1993-2002); USDA, Foreign Agricultural Service (2002-2016); and USDA, World Agricultural Outlook Board (1993-2016).

Appendix table 7

Bias¹ in USDA's U.S. soybean forecasts, 2004-2015 (forecasts in levels, f_{it})

Month	i	Beginning stocks	Ending stocks	Exports	Meal exports	Imports	Domestic use	Area harvested	Yield	Production
						Percent				
Nov	-5	-24.1 *	-27.2	6.5	16.1 **	73.8 ***	-0.1	1.1	0.3	1.4
Feb	-2	-21.1	-38.6 **	5.5	16.0 ***	68.3 **	1.9	0.4	1.0	1.4
May	1	-10.6	-29.9 *	4.3	14.3 ***	49.2 **	0.7	-0.3	1.0	0.8
Jun	2	-5.1	-29.4 *	4.5	15.4 **	49.1 **	0.9	-0.2	1.0	0.8
Jul	3	-2.0	-24.1	5.9	14.2 **	48.8 **	0.8	-0.3	1.8	1.6
Aug	4	-0.8	-14.2	8.9 *	17.6 **	46.7 **	1.7	0.0	4.1 **	4.1 **
Sep	5	3.2	-14.4	9.1 *	17.0 **	39.2 *	1.6	-0.1	3.7 *	3.6 *
Oct	6	1.0	-25.0	6.9 *	14.9 **	42.2 *	1.2	-0.1	1.9 *	1.8 *
Nov	7	--	-30.3 *	6.2	12.4 **	44.0 *	0.8	0.1	0.6	0.7

Notes: ¹Bias is estimated value of α_i in $e_{it} = \alpha_i + \varepsilon_{it}$ where i = forecast month and t = marketing year, and $e_{it} = a_t - f_{it}$.Logged values of the actual realization in marketing year t of the variable (A_t) and the forecast for that marketing year in month i (F_{it}) are a_t and f_{it} . * = pvalue of 0.05, **=pvalue of 0.01, and ***=pvalue of 0.001.

Sources: ERS calculations based on data from USDA, Office of Chief Economist (1998-2016); USDA, Wheat, Feed Grains, and Oilseeds Interagency Commodity Estimates Committees (1998-2016); USDA, Economic Research Service (1993-2002); USDA, Foreign Agricultural Service (2002-2016); and USDA, World Agricultural Outlook Board (1993-2016).

Appendix table 8

Bias¹ in selected USDA soybean forecasts, 2004-2014 (forecasts in levels, f_{it}).

Month	i	Argentina meal exports	Brazil soybean production	China soybean imports	E. Union meal imports	World soybean trade	U.S. share total trade	Foreign soybean production	Foreign soybean consumption	Foreign excess demand
						Percent				
Jun	2	-6.5	-0.1	6.5 *	-8.5 **	3.5	1.3	-1.1	-0.5	-0.3
Jul	3	-6.2	-0.1	6.4 *	-8.3 **	3.6	1.5	-1.2	-0.4	0.6
Aug	4	-7.2 *	-0.1	6.8 *	-8.7 **	3.9	2.3 **	-1.3	-0.3	2.5
Sep	5	-7.7 *	-1.0	6.2	-8.8 **	3.1	2.4 *	-1.6	-0.5	3.9
Oct	6	-7.2 *	-1.2	5.5	-8.2 **	2.4	1.9 *	-1.9	-0.4	6.5
Nov	7	-7.1 *	-1.0	4.9	-8.1 **	2.1	1.7 *	-2.0	-0.6	5.2
Dec	8	-6.8 *	-0.8	4.7	-7.6 **	1.9	1.6 *	-2.3	-0.8	5.4
Jan	9	-6.7 *	-1.2	4.4	-7.4 **	2.0	1.6 *	-2.2	-0.7	6.1
Feb	10	-5.7 *	-0.8	4.4	-5.4 *	2.5	1.4 *	-1.5	-0.3	6.0
Mar	11	-5.4 *	-0.2	4.7	-5.2 *	3.0	1.1 *	-0.9	-0.2	2.3

Notes: ¹Bias is estimated value of α_i in $e_{it} = \alpha_i + \varepsilon_{it}$ where i = forecast month and t = marketing year, and $e_{it} = a_t - f_{it}$.Logged values of the actual realization in marketing year t of the variable (A_t) and the forecast for that marketing year in month i (F_{it}) are a_t and f_{it} . Excess supply is both >0 and <0 in sample, preventing log transformation. * = pvalue of 0.05, **=pvalue of 0.01, and ***=pvalue of 0.001.

Sources: ERS calculations based on data from USDA, Office of Chief Economist (1998-2016); USDA, Wheat, Feed Grains, and Oilseeds Interagency Commodity Estimates Committees (1998-2016); USDA, Economic Research Service (1993-2002); USDA, Foreign Agricultural Service (2002-2016); and USDA, World Agricultural Outlook Board (1993-2016).