Journal of Agricultural and Resource Economics 37(1):98–113

A Comprehensive Evaluation of USDA Cotton Forecasts

Olga Isengildina-Massa, Stephen MacDonald, and Ran Xie

This study evaluates all USDA cotton supply and demand estimates for the United States and China (including unpublished price forecasts) from 1985/86 through 2009/10 for accuracy and efficiency. Results reveal that at every stage of the forecasting cycle forecast smoothing was the most widespread and persistent type of inefficiency observed in most U.S. variables. Correlation with past errors indicated the tendency to repeat past errors in most cases. Tendency to overestimate growth was also found. Bias was uncommon and limited to several cases of overestimation of China's exports and U.S. price and underestimation of China's domestic use. While forecasts of China's imports and endings stocks improved, U.S. price and ending stock forecast errors became larger toward the end of the study period.

Key words: cotton, forecast accuracy, forecast efficiency, forecast evaluation, forecast smoothing, USDA forecasts

Introduction

Many agricultural market participants and analysts share a common belief that USDA forecasts function as a benchmark for other private and public forecasts, which is not surprising given the classic public goods problem of private underinvestment in information and the critical role that public information plays in coordinating the beliefs of market participants. As a result, there is a vast body of literature devoted to analyzing the accuracy and efficiency of USDA forecasts, which focuses mainly on production (e.g., Gunnelson, Dobson, and Pamperin, 1972; Sanders and Manfredo, 2002; Isengildina, Irwin, and Good, 2006a) and price (e.g., Irwin, Gerlow, and Liu, 1994; Sanders and Manfredo, 2003; Isengildina, Irwin, and Good, 2004) forecasts. The accuracy of most other USDA forecasts describing supply and demand forces has been largely overlooked. To the best of our knowledge, only one previous study investigated the accuracy of all supply, demand and price forecasts for U.S. corn and soybeans published within WASDE (*World Agricultural Supply and Demand Estimates*) reports (Botto et al., 2006).

Price forecast accuracy depends heavily on the accuracy of supply and demand forecasts. The USDA's commodity forecasting follows a balance sheet approach, accounting for each component of supply and utilization in the countries for which the USDA creates a commodity forecast (see Vogel and Bange (1999) for a detailed description of the USDA crop forecast generation process). For each country's commodity forecast the USDA forecasts all variables included in a balance sheet, including beginning stocks, production, and imports on the supply side and domestic use (or consumption), exports, and ending stocks on the demand side. The balance sheet approach

Olga Isengildina-Massa is Assistant Professor in the John E. Walker Department of Economics at Clemson University, Stephen MacDonald is Senior Economist in the USDA ERS, and Ran Xie is a graduate student in the John E. Walker Department of Economics at Clemson University. The funding support of the Economic Research Service of U.S. Department of Agriculture under the Cooperative Agreement No. 58-3000-9-0077 is gratefully acknowledged. The authors thank David Tysinger and Patrick Gerard for their input and contribution to the earlier drafts of the paper and two anonymous reviewers and the Journal editor for their helpful comments. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the view of the U.S. Department of Agriculture.
Review coordinated by Gary Brester.

requires internal consistency among the variables as a group. In other words, "total supply must equal domestic use plus exports and ending stocks. Prices tie both sides of the balance sheet together by rationing available supplies between competing uses" (Vogel and Bange, 1999, p. 10). WASDE price estimates describe the marketing year average prices received by farmers, which are averages of monthly prices weighted by the amounts marketed at these prices. While the USDA's WASDE reports affect markets (e.g. Fortenbery and Summer, 1993; Isengildina, Irwin, and Good, 2006b; Isengildina-Massa et al., 2008), little is known about the accuracy of forecasts beyond production and price. Even less is known about the accuracy and efficiency of WASDE forecasts of the foreign supply and demand categories that may affect U.S. markets through trade.

The objective of this study is to provide a comprehensive examination of the accuracy and efficiency of all supply and demand categories of the USDA's WASDE cotton forecasts for the United States and China. The United States is one of the largest cotton producers and exporters in the world, with average production of 18% and exports of 31% of world totals during the study period. China is the world's largest producer, consumer, and importer of cotton. China accounts for about 40% of the world consumption and 30-40% of world trade. Sound forecasts of supply and demand for both countries are therefore crucial for policy-makers, farmers, and other decision-makers in the United States and around the world. However, only a few studies have investigated a subset of USDA forecasts for cotton (MacDonald, 2002) or included cotton in studies of USDA export forecasts for a number of commodities (MacDonald, 2005). This study will use the data from monthly WASDE balance sheets for cotton for the U.S. and China between 1985/86 through 2009/10, including unpublished price forecasts.[1]

The analysis is comprehensive as it does not focus on a single aspect of forecast evaluation, but incorporates multiple tests of forecast performance including: 1) accuracy, 2) bias, 3) efficiency with respect to forecast levels, 4) efficiency with respect to past errors, 5) efficiency in forecast revisions, and 6) forecast improvement over the study period. Understanding various aspects of WASDE cotton forecast performance will help USDA analysts identify areas that need improvement and will assist forecast users in efficient interpretation and application of the information contained in these forecasts.

Data

This study focuses on monthly WASDE cotton forecasts for the U.S and China from 1985/86 through 2009/10 (U.S. Department of Agriculture, Office of the Chief Economist, 1985-2010). Means of forecast levels of these variables shown in table 1 demonstrate that both China and the United States are major cotton producers, jointly producing over 43% of cotton in the world. All quantities are measured in million 480-pound bales of cotton, including both upland and extra-long staple varieties. Price is measured in cents per pound of upland cotton (which accounts for 96% of U.S. cotton). China is also a major consumer of cotton, with the growing textile sector supported by domestic production and increasingly supplemented by imports since the early 2000s. The demands of China's textile sector are also facilitated by relatively high levels of stocks. In contrast, the U.S. textile industry has been shrinking since the mid-1990s, as reflected in declining domestic use and growing exports. The nominal U.S. cotton price averaged about 56.19 cents/lb. during the study period. Similar to other major U.S. farm commodities, cotton's price was supported by U.S. farm programs prior to 1985, but U.S. programs and price determination have become more market oriented since. Due to the increased export orientation of the U.S. cotton industry, the price of U.S. cotton has been increasingly affected by international market forces (Isengildina and MacDonald, 2009). While price forecasts have been published in interval form since 2008, earlier, unpublished

[1] Although cotton price forecasts were not published, the USDA's Interagency Commodity Estimates Committee (ICEC) for cotton calculated unpublished price forecasts each month. The accuracy of these unpublished forecasts should be evaluated as the USDA moves forward with its cotton price forecasting mission.

Table 1. Descriptive Statistics for WASDE Cotton Forecasts, 1985/86-2009/10 Marketing Years

Category	Production	Imports	Domestic Use	Exports	Ending Stocks	Price
U.S.						
Mean level (y_t^I), mil. bales and c/lb[a]	16.77		7.93	8.93	5.09	56.19
Std Deviation (y_t^I)[a]	3.49		2.42	3.77	2.17	10.38
Coefficient of variation (y_t^I)[a]	0.21		0.30	0.42	0.43	0.18
Mean absolute percent forecast (f_t^i)[b]	13.85		5.60	22.65	31.84	16.20
Mean absolute percent error (e_t^i)[b]	3.69		5.49	10.98	20.63	6.63
Max percent error (f_t^i)[b]	27.84		24.69	29.48	72.31	20.39
Min percent error (f_t^i)[b]	−19.76		−20.35	−93.65	−80.35	−45.56
RMSPE[b]	5.37		6.44	14.27	23.35	8.53
China						
Mean level (y_t^I), mil. bales and c/lb[a]	23.30	4.01	27.72	0.73	12.66	
Std Deviation (y_t^I)[a]	6.16	4.87	11.40	0.90	4.78	
Coefficient of variation (y_t^I)[a]	0.26	1.21	0.41	1.23	0.38	
Mean absolute percent forecast (f_t^i)[b]	11.29	67.16	5.00	63.84	16.59	
Mean absolute percent error (e_t^i)[b]	6.33	52.45	4.71	62.89	22.59	
Max percent error (f_t^i)[b]	25.53	519.30	18.23	128.09	81.59	
Min percent error (f_t^i)[b]	−23.36	−230.26	−18.23	−299.57	−85.91	
RMSPE[b]	7.70	56.90	5.34	82.26	25.30	

Notes: [a] Level statistics are calculated using the final November (I=19) estimate for each category.
[b] Forecasts and forecast errors are measured in log percent form. Statistics calculated using data for stages 1-4.

price forecasts were point estimates. To overcome this inconsistency and keep the analysis consistent across all forecasts, midpoints of the price forecast intervals published since 2008 are used.

Coefficients of variation of the forecast levels shown in table 1 demonstrate that ending stocks and exports are the most volatile categories on the U.S. balance sheet and exports and imports are the most volatile on the Chinese balance sheet. High variability of China's forecasts relative to all other forecasts illustrates challenges associated with obtaining reliable data from China. As described by Skelly, Colby, and Johnson (2010, p. 415):

> Until 2007, USDA and most other cotton forecasting agencies relied mainly on statistics released by the NBS (National Bureau of Statistics) to estimate China's cotton production . . . However, by mid-2007, sources in China were examining information on rail shipments of cotton from Xinjiang to eastern China and concluded that the NBS production estimates for Xinjiang were too low. In late September 2007, the high-level National Development and Reform Commission (NDRC), an agency under China's State Council, confirmed higher production estimates for the 2006 and 2007 crops.

These challenges also resulted in an important data reporting change that affected production and indirectly influenced other categories in the Chinese balance sheet. Since NDRC estimates were deemed more realistic, the USDA switched to forecasting NDRC rather than NBS production estimates in 2007. This switch resulted in "adjustments in monthly releases for July and October of 2007 which raised estimates for the 2004/05 through 2007/08 crops by a cumulative total of 14.4 million bales" (Skelly, Colby, and Johnson, 2010, p. 415).

Typically, WASDE reports are released between the 9th and the 12th of each month and contain forecasts of supply and demand for most major crops. Supply and demand estimates are forecasted on a marketing year basis (August through July for U.S. cotton). The first forecast for a marketing year is released in the May preceding the U.S. marketing year. USDA forecasts for China were

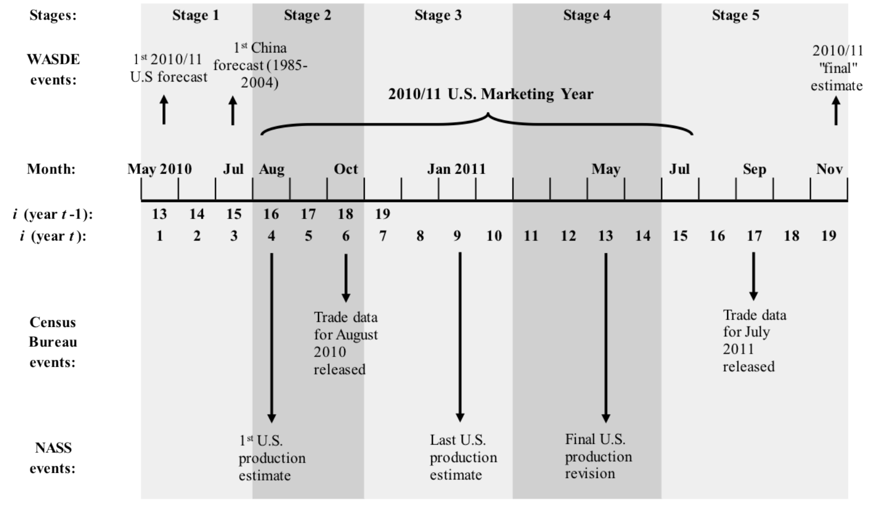

Figure 1. WASDE Forecasting Cycle for Cotton Relative to the 2010/11 U.S. Marketing Year

historically not released until the July preceding the U.S. marketing year.[2] Estimates for the United States are largely finalized 18 months later, by November after the marketing year (figure 1). U.S. production forecasts are an exception as they are finalized by May (month 13 of the forecasting cycle). The 19 month schedule is also adopted for the analysis of China variables. Revisions of non-U.S. variables' estimates are common through the 25th month, but are small enough to justify the simplification of choosing a common forecast horizon for both U.S. and non-U.S. variables. The USDA's WASDE forecasts are usually viewed as fixed-event forecasts, since during each year (t) there are a series of forecasts related to the same terminal event occurring in year t (y_t^I), where I is the release month of the final estimate. The forecasted value of the year t terminal event published in month i of year t is denoted as y_t^i, where $i = 1, \ldots, I$, and $I = 19$. Thus, each subsequent forecast is an update of the previous forecast describing the same terminal event. Based on our definition of a 19 month forecasting cycle, WASDE generates 18 updates for each U.S. variable except production (12 updates) within each marketing year. The years covered in this study are t = 1(1985/86), ..., 25 (2009/10). Historical series of WASDE forecasts also have rolling-event characteristics as 18 ($I - 1$) different horizon forecasts are available for all 25 target dates (marketing years) for all U.S. variables and 16 ($I - 3$) different horizons for China variables.

Forecast evaluation is derived from accuracy tests based on the size and direction of forecast errors and optimality tests based on the idea of forecast error and revision unpredictability (Diebold and Lopez, 1998; Nordhaus, 1987). To standardize for changing forecast size over time, errors, revisions, and forecasts are examined in log percentage form.[3] The forecast error e_t^i is calculated as:

$$(1) \qquad e_t^i = 100 \times \ln\left(\frac{y_t^I}{y_t^i}\right) \qquad i = 1, \ldots, I-1; t = 1, \ldots, 25.$$

and forecast revision is measured as:

$$(2) \qquad r_t^i = 100 \times \ln\left(\frac{y_t^i}{y_t^{i-1}}\right) \qquad i = 1, \ldots, I-1; t = 1, \ldots, 25.$$

Percent forecasts, f_t^i, were computed as percent changes in forecasted values from the previous year's values. Since the forecasting cycle spans 19 months, the final estimate for the previous

[2] The USDA began publishing June forecasts for China in 2004 and May forecasts in 2005.

[3] Unit errors ($y_t^I - y_t^i$) and revisions ($y_t^i - y_t^{i-1}$) were also included in the original analysis. Results are available from the authors upon request.

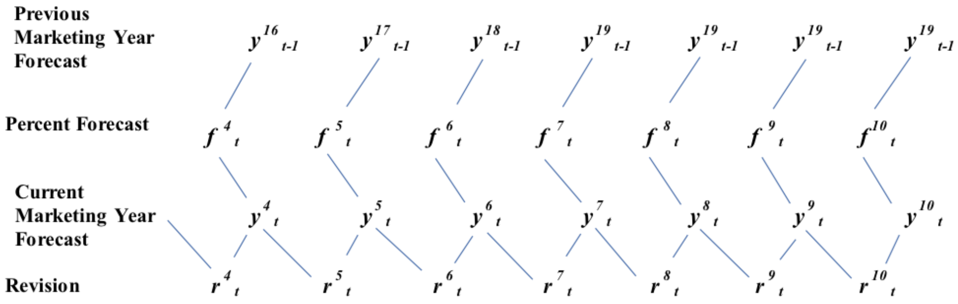

Figure 2. WASDE Cotton Forecasts, Percent Forecasts and Revisions for Stages 2 ($i = 4-6$) and 3 ($i = 7-10$) of the Forecasting Cycle

Notes: Percent Forecast (f_t^i) measures a change between previous year's final (y_{t-1}^{19}) or most recently available (y_{t-1}^{y+12}) estimate and the current year forecast (y_t^i): $f_t^i = [\ln(y_{t-1}^{i+12}) - \ln(y_t^i)] \times 100$.

marketing year is not known during the first six months of the forecasting cycle and the most recently available estimate is used:

$$(3)\quad \begin{aligned} f_t^i &= 100 \times \ln\left(\frac{y_t^i}{y_{t-1}^{i+12}}\right) && \text{where} \quad i = 1, \dots, 6; t = 1, \dots, 25, \\ f_t^i &= 100 \times \ln\left(\frac{y_t^i}{y_{t-1}^{I}}\right) && \text{where} \quad i = 7, \dots, I; t = 1, \dots, 25, \end{aligned}$$

The layout of fixed event forecasts and corresponding revisions and percent forecasts are illustrated for stages 2 and 3 in figure 2.

Figure 3 shows changes in forecast errors over the forecasting cycle using box-and-whisker plots. For each month, the distance from the top of the upper whisker to the bottom of the lower whisker captures the entire range of forecasting errors over the past 25 years. The box reflects the middle 50% of the errors and the upper and lower whiskers show the upper and lower 25% quantiles. These graphs demonstrate that the variability of the forecast error generally declines during the forecasting cycle as more information becomes available, thus satisfying this forecast efficiency criteria outlined by Diebold and Lopez (1998). Difference between means (marked with an x) and medians (marked with a line) of forecast error distributions suggest that distributions are asymmetric. For example, negative outliers in the errors for China's export forecasts pull the mean below the median, and positive outliers have the opposite effect on mid-cycle U.S. ending stock forecast errors.

This illustration of how forecast errors change within the forecasting cycle reflects the differences in information sets available at the time the forecasts are made (as illustrated in figure 1). May-July forecasts are largely based on historical information as limited information about the development of new crop is available and the previous marketing year estimates are not finalized. August marks better information from NASS about U.S. crop progress (NASS is not responsible for the May-July U.S. production forecasts) and the onset of the marketing year. November marks the end of the northern hemisphere harvest and finalized information about the previous marketing year. Forecasts between February and July are a combination of the observed marketing year activities to date (e.g., actual U.S. consumption and export data for the first several months of the marketing year) and forecasts for the later months. China's forecasts follow a similar pattern, with generally longer lags representing greater difficulty acquiring timely, sound data. Thus, the forecasts may be largely differentiated into highly uncertain stage 1 forecasts during May-July (i =1, 2, 3), less uncertain stage 2 forecasts during August-October (i =4, 5, 6), mid-cycle stage 3 forecasts during November-February (i =7, 8, 9, 10), stage 4 estimates during March-June (i =11, 12, 13, 14), and

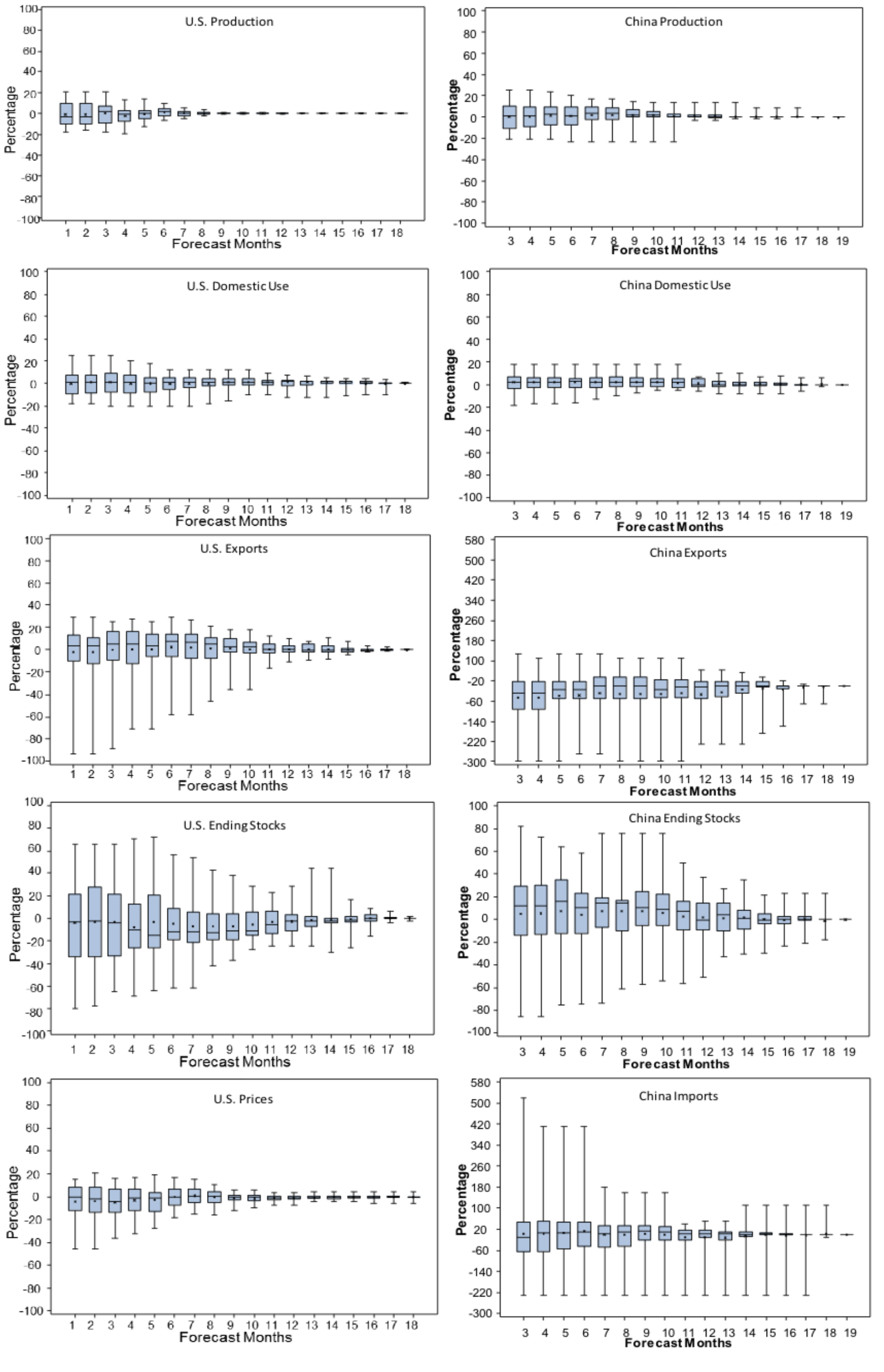

Figure 3. USDA Errors in Forecasting U.S. and China Cotton Variables, 1985/86-2009/10 Marketing Years.

Table 2. Forecast Accuracy Tests for WASDE Cotton Forecasts, 1985/86-2009/10 Marketing Years

Country/ Stage of forecasting cycle	Test	Production	Imports	Domestic Use	Exports	Ending Stocks	Price
U. S.							
Stage 1	Theil's U	0.68		1.05	0.78	0.84	0.70
	Directional χ^2	3.55*		3.38*	9.15***	0.50	9.24***
Stage 2	Theil's U	0.39		0.26	0.65	0.78	0.61
	Directional χ^2	9.15***		9.15***	11.53***	3.38*	7.17***
Stage 3	Theil's U	0.08		0.22	0.47	0.53	0.27
	Directional χ^2	21.42***		15.37***	12.09***	4.97**	17.78***
Stage 4	Theil's U			0.19	0.21	0.25	0.10
	Directional χ^2			15.37***	15.64***	14.64***	25.15***
China							
Stage 1-2	Theil's U	0.71	0.77	0.77	0.84	1.03	
	Directional χ^2	4.97**	5.42**	1.62	3.38*	0.21	
Stage 3	Theil's U	0.47	0.60	0.56	0.64	0.80	
	Directional χ^2	11.92***	11.70***	4.97**	2.10	6.94***	
Stage 4	Theil's U	0.26	0.19	0.42	0.54	0.49	
	Directional χ^2	21.29***	18.19***	10.62***	9.15***	9.92***	

Notes: All tests are based on the average forecasts for each stage across years, thus N=25. Directional χ^2 test of independence of actual and forecasted direction uses Yates-corrected χ^2 statistic.

stage 5 estimates during July-November (i =15, 16, 17, 18) (figure 1). This study focuses on stages 1-4 of the forecasting cycle.

Accuracy Analysis

Descriptive statistics shown in table 1 demonstrate that, in absolute terms, errors were the largest for the categories that are the most variable: China's exports, imports, and ending stocks, and U.S. exports and ending stocks. When forecasts are measured as a rate of change from the previous year, the largest average magnitude of change is observed in China's imports and exports and the U.S. ending stocks. Large maximum and minimum errors in China's imports and exports should be interpreted with care as they represent changes from near-zero levels. RMSPE and MAPE provide consistent signals about the relative accuracy of forecasts. Ending stocks forecasts are the least accurate among U.S. categories and exports, imports, and ending stocks are the least accurate forecasts on the Chinese balance sheet. Comparison of similar categories across countries highlights inaccuracy in forecasting China's trade variables.

Forecast accuracy evaluation relative to a naïve alternative was conducted using a Theil's U statistic, calculated as the ratio of the mean squared error (MSE) of a variable's forecast relative to the MSE of a random walk for the variable $U = \sum_{t=1}^{T}(y_t^I - y_t^i)^2 / \sum_{t=1}^{T}(y_t^I - y_{t-1}^I)^2$. Interpretation of Theil's U is based on a value of 0 for perfect forecasts and a value of 1 for forecasts with accuracy equivalent to that of naïve, "no-change" forecasts (Leuthold, 1975). Statistics shown in table 2 indicate that USDA was most successful with U.S production and price and China production forecasts. On the other hand, early U.S. domestic use and China's ending stocks forecasts seem most problematic. These problems were likely due to challenges with forecasting structural change in these categories as described in Isengildina and MacDonald (2009).

The evaluation of directional accuracy shown in table 2 is based on the timing test developed by Henriksson and Merton (1981) and applied by McIntosh and Dorfman (1992) and Schnader and

Stekler (1990), among others. The test is based on 2 x 2 contingency tables, reflecting the direction of year-to-year change in each variable forecast for each stage's average forecast. The frequency with which forecasts and actual realizations of the variable decrease or increase together is compared with the expected frequency of independent directional changes using a χ^2-statistic. Our findings highlight the difficulty the USDA faces in forecasting ending stocks of cotton: both China's and the United States' stage 1 ending stock forecasts fare poorly, failing to provide information about the direction of change in the coming year. The only other variable faced with similar challenges was China's domestic use. Directional accuracy improves for virtually every variable as the forecasting cycle progresses, but at different rates. By stage 2, all the U.S. forecasts provide information about direction, but problems persist in the China's ending stock forecasts into stage 2 and in the export forecasts into stage 3. A surprising result was found for China's export forecasts directional accuracy, with the null hypothesis of independence not rejected for stage 3 while rejected (although marginally) for stage 1-2 forecasts, thus demonstrating the only violation of the efficiency condition that accuracy should improve across the forecast horizon found in this study. This finding likely illustrates challenges with obtaining and analyzing data for this highly variable category.

Forecast Optimality Evaluation Framework

Theil (1958) pioneered the framework for rolling-event forecast efficiency testing, which was extended by Mincer and Zarnowitz (1969) and Pons (2000). Nordhaus (1987) introduced the utilization of these tests into a fixed-event framework, and Clements (1997) and Isengildina, Irwin, and Good (2006a) extended this to the pooling of rolling sets of fixed-event forecasts. Following Elliott and Timmermann (2008), this study assumes the forecaster's loss function is of the mean squared error (MSE) type, so that the forecasts minimize a symmetric, quadratic loss function.

The fundamental measures of optimal forecasts are bias and efficiency (Diebold and Lopez, 1998). Traditionally the test of bias is conducted with a regression of forecast errors on an intercept term. Changes in the direction of bias during a study period are captured in this study by a trend term. The test of bias is performed here using the following regression:

$$(4) \qquad e_t^i = \alpha_0 + \alpha_1 T + \varepsilon_t^i \qquad \text{where} \quad i = 1,\ldots,I-1; t = 1,\ldots,25,$$

where α_0 is a constant and T is a centered linear time trend. The null hypothesis for an unbiased forecast is $\alpha_0 = 0, \alpha_1 = 0$. If $\alpha_0 > 0$, then forecasts will consistently underestimate the final estimate. If $\alpha_0 < 0$, forecasts consistently overestimate the final estimate. If $\alpha_1 \neq 0$, the direction of bias has changed over time.

Using absolute errors as the dependent variable in the above test of bias allows testing for forecast improvement over time as suggested by Bailey and Brorsen (1998) and Sanders and Manfredo (2003):

$$(5) \qquad |e_t^i| = \beta_0 + \beta_1 T + \varepsilon_t^i \qquad \text{where} \quad i = 1,\ldots,I-1; t = 1,\ldots,25,$$

In this equation, β_0 measures average absolute forecast error. The null hypothesis is $\beta_1 = 0$, which indicates that there is no systematic change in the size of forecast error. If $\beta_1 > 0$, the forecasts become less accurate over time as evidenced by larger errors. If $\beta_1 < 0$, the forecasts improve over time as demonstrated by smaller errors.

Weak efficiency tests evaluate whether forecast errors are orthogonal to forecasts themselves and to prior forecast errors (Nordhaus, 1987). This study tests weak efficiency with respect to forecast levels using the basic approach of (Pons, 2000) and (Sanders and Manfredo, 2002, 2003), with a trend term:[4]

$$(6) \qquad e_t^i = \gamma_0 + \gamma_1 f_t^i + \gamma_2 T + \varepsilon_t^i \qquad \text{where} \quad i = 1,\ldots,I-1; t = 1,\ldots,25,$$

[4] This equation can be easily traced back to the standard Mincer-Zarnowitz formulation:

1. $f_t^I = \beta_0 + \beta_1 f_t^i + \varepsilon_t^i$

The null hypothesis for efficient forecasts is $\gamma_1 = 0$. If $\gamma_1 > 0$, the absolute value of the forecast is smaller than the actual realization. Since the forecasts here are stated in terms of change from the previous year, $\gamma_1 > 0$ means that change is underestimated, in either direction. If $\gamma_1 < 0$, the change is overestimated.

Forecast efficiency with respect to past errors is measured as:

(7) $$e_t^i = \rho_0 + \rho_1 e_{t-1}^i + \rho_2 T + \varepsilon_t^i \qquad \text{where} \quad i = 1,\ldots,I-1; t = 2,\ldots,25,$$

For fixed event forecasts, the forecast error for the previous event (marketing year) should be used for this test. The null hypothesis for efficient forecasts is $\rho_1 = 0$. If $\rho_1 \neq 0$, there is a systematic component in forecast errors that can be predicted using past errors.

Weak form efficiency of fixed-event forecasts implies independence of forecast revisions (Nordhaus, 1987). According to Nordhaus, if forecasts are weak form efficient, revisions should follow a random walk. This property is tested using the approach outlined in Isengildina, Irwin, and Good (2006a), with a time trend:

(8) $$r_t^i = \lambda_1 r_t^{i-1} + \lambda_2 T + \varepsilon_t^i \qquad \text{where} \quad i = 2,\ldots,I; t = 1,\ldots,25,$$

For ($i = 3$), λ_1 represents the slope coefficient of all October revisions made from 1985/86 to 2009/10 regressed against the September revisions ($i - 1 = 2$) for the same respective years. The null hypothesis for efficiency in forecast revisions is $\lambda_1 = 0$. If $\lambda_1 > 0$, the forecasts are "smoothed," as they are partially based on the previous revision. If $\lambda_1 < 0$, the forecasts are "jumpy," as they tend to over-adjust the previous revision.

All equations were estimated for each of stages 1-4 using data pooled across all marketing years and the 3 to 4 months of the forecasting cycle contained in each stage. A dummy variable with the value of 1 for all forecast months in 2006 and the beginning of 2007 (through September, $i = 5$) was added to the evaluation of China's production and ending stock forecasts to account for the USDA's October 2007 shift from forecasting production as estimated by China's National Bureau of Statistics (NBS) to targeting the estimate from the National Development and Reform Commission (NDRC). Regressions were estimated using panel least squares method with White cross-section correction in standard error calculation. The White cross-section method treats the pooled regression as a multivariate regression (with an equation for each cross-section) and computes White-type robust standard errors for the system of equations. The coefficient covariance matrix is estimated as:

(9) $$\left(\frac{N^*}{N^* - K^*}\right)\left(\sum_t X_t' X_t\right)^{-1}\left(\sum_t X_t' \hat{\varepsilon}_t \hat{\varepsilon}_t' X_t\right)\left(\sum_t X_t' X_t\right)^{-1}$$

where N^* is the number of cross sections (months of the forecasting cycle), K^* is the number of estimated parameters and X_t is a vector of regressors. This estimator is robust to contemporaneous correlation as well as different error variances in each cross-section (Wooldridge, 2002, p. 148-153).

Forecast Optimality Evaluation Results

Table 3 shows the results of the test of bias in equation (4) , which suggest that on average none of the U.S. cotton forecasts were biased. However, the U.S. domestic use forecast errors have a strong negative time trend, demonstrating a tendency for underestimation prior to 1997 and overestimation beginning in 1997/98. This finding demonstrates difficulties with forecasting structural changes in the domestic textile industry; in particular, the USDA underestimated the speed of contraction in the U.S. textile sector beginning in the late 1990s. While the bias in China's export forecasts was very large in percentage terms–averaging 42% overestimation in the first stage–its importance is

2. $f_t^I - f_t^i = \beta_0 + \beta_1 f_t^i - f_t^i + \varepsilon_t^i = \beta_0 + (\beta_1 - 1) f_t^i + \varepsilon_t^i$,
which restates the test in terms of the forecasted variable (f_t^i) rather than forecast error ($f_t^I - f_t^i$).

Table 3. Test of Bias for WASDE Cotton Forecasts, 1985/86-2009/10 Marketing Years[a]

Country/ Stage of forecasting cycle	Coeff.	Production	Imports	Domestic Use	Exports	Ending Stocks	Average Price	N
U. S.								
Stage 1	α_0	1.31		0.66	−1.35	−3.59	−4.21	75
	α_1	−0.03		−1.00***	0.72	0.76	−0.49	
Stage 2	α_0	0.49		0.01	0.95	−5.09	−1.82	75
	α_1	0.02		−0.82***	0.55	0.62	−0.38	
Stage 3	α_0	−0.01		0.51	1.15	−6.73	−0.49	100
	α_1	−0.02		−0.49***	0.43	0.06	−0.16	
Stage 4	α_0			0.75	0.48	−2.12	−0.89**	100
	α_1			−0.22**	0.22*	−0.19	−0.06	
China								
Stage 1-2	α_0	−0.06	6.70	2.16	−41.86**	4.14		100
	α_1	0.48	−2.58	−0.07	−0.70	1.70		
	α_2	14.65***				18.39*		
Stage 3	α_0	1.84	−0.22	2.05*	−30.66	6.59		100
	α_1	0.30	−0.73*	−0.07	−1.18	0.90		
	α_2	10.56***				7.62		
Stage 4	α_0	0.66	−9.04	1.10	−25.39*	0.99		100
	α_1	0.13*	1.48	0.02	−0.74	0.77		
	α_2	12.09***				18.83***		

Notes: e is forecast error, T is a trend, and 2006 is a dummy variable. Stage 1 includes forecasts released in months 1-3, Stage 2 includes months 4-6, Stage 3 includes months 7-10, Stage 4 includes months 11-14 of the forecasting cycle. Regressions estimated using panel least squares with White heteroscedasticity correction. N is the number of observations. Single, double, and triple asterisks (*, **, ***) denote statistical significance at 10%, 5%, and 1% levels, respectively.

[a] $e_t^i = \alpha_0 + \alpha_1 T + \alpha_2 2006 + \varepsilon_t^i$

likely limited due to generally small forecast levels (0.72 million bales) during the study period. Overestimation in China's export forecasts is similar to the pattern previously found in broiler prices by Sanders and Manfredo (2003). More important was a downward bias of 2.05% in the forecasts of China's consumption during the third stage. Since China is the world's largest consumer of cotton, this finding suggests that third stage China consumption forecasts have been on average about 562,000 bales below actual. This tendency to underestimate China's consumption is reminiscent of previous evidence of underestimation in forecasts of U.S. beef, pork, and broiler production (Bailey and Brorsen, 1998; Sanders and Manfredo, 2002).

Table 4 illustrates forecast improvement test results derived from equation (5), which demonstrate significant evidence of forecast improvement in China's ending stocks, production, and import forecasts. China's ending stocks forecast errors became 1.58% and 1.31% smaller each year during the first/second and third stage of the forecasting cycle, respectively. China's import forecast errors were reduced by 4.51% and 3.75% each year in the first/second and the third stage, respectively. The improvement in China's production forecasts measured 0.38% a year in the first/second and third stage. These findings are likely associated with better data regarding the Chinese cotton industry that have become available toward the end of the study period. Forecast improvement results are similar to Bailey and Brorsen's (1998) findings in beef and pork production forecasts and Sanders and Manfredo's (2003) findings in broiler price forecasts. In contrast, U.S. price forecasts became significantly worse, with error increasing by 0.53% per year in the first stage and 0.08% per year in the fourth stage. This finding likely reflects challenges with forecasting price in an increasingly internationally-oriented market (Isengildina and MacDonald, 2009). Note, however, that this finding refers to unpublished forecasts. U.S. ending stock forecasts in stage 3

Table 4. Test of Forecast Improvement for WASDE Cotton Forecasts, 1985/86-2009/10 Marketing Years[a]

Country/ Stage of forecasting cycle	Coeff.	Production	Imports	Domestic Use	Exports	Ending Stocks	Average Price	N
U. S.								
Stage 1	β_0	9.51***		8.93***	17.26***	30.11***	12.08***	
	β_1	0.00		−0.11	−0.27	−0.05	0.53**	75
Stage 2	β_0	5.66***		6.68***	14.75***	28.43***	10.04***	
	β_1	0.11		−0.05	−0.20	0.66	0.09	75
Stage 3	β_0	1.32***		4.60***	10.61***	19.63***	4.73***	
	β_1	0.02		0.00	−0.16	0.51**	0.06	100
Stage 4	β_0			2.89***	3.82***	8.69***	1.89***	
	β_1			−0.03	0.05	0.03	0.08***	100
China								
Stage 1-2	β_0	9.53***	77.05***	6.00***	72.51***	27.74***		
	β_1	−0.38***	−4.51*	0.15	1.08	−1.58**		100
	β_2	13.21***				25.65***		
Stage 3	β_0	5.95***	52.93***	4.90***	66.56***	23.04***		
	β_1	−0.38**	−3.75**	0.06	1.16	−1.31***		100
	β_2	12.54***				11.07***		
Stage 4	β_0	1.58***	27.36***	3.25***	49.60***	14.20***		
	β_1	−0.03	−2.33	0.06	0.04	−0.16		100
	β_2	12.53***				13.97***		

Notes: e is forecast error, T is a trend, and 2006 is a dummy variable. Stage 1 includes forecasts released in months 1-3, Stage 2 includes months 4-6, Stage 3 includes months 7-10, Stage 4 includes months 11-14 of the forecasting cycle. Regressions estimated using panel least squares with White heteroscedasticity correction. N is the number of observations. Single, double, and triple asterisks (*, **, ***) denote statistical significance at 10%, 5%, and 1% levels, respectively.

[a] $|e_t^i| = \beta_0 + \beta_1 T + \beta_2 2006 + \varepsilon_t^i$

also saw a significant increase in absolute error over time. Since β_0 in equation (5) equals the forecast mean absolute percent error (MAPE), table 4 shows that MAPE improves (declines) as the forecasting cycle progresses from stages 1 to 4.

Table 5 reports results of the forecast efficiency tests from equation (6). Errors of early U.S. domestic use and ending stocks forecasts and China's exports forecasts were significantly negatively correlated with forecast levels, indicating that the USDA tends to overestimate change in these variables early in the forecasting cycle. For example, for each 10% change from the previous year's forecast in U.S. ending stocks, we expect first stage forecast error to grow by 3.2%; hence the forecast should be adjusted by that amount. Based on the γ_1 coefficient of -0.32 for stage 2 ending stocks forecasts, the forecasts should be scaled down by a factor of 0.68 $(1 + \gamma_1)$. In other words, if USDA forecasts ending stocks to rise or fall by 10% in their May-July forecasts, a better estimate of the actual change is 6.8%. A similar pattern of overestimation of change was previously found in livestock production forecasts by Sanders and Manfredo (2002). The opposite pattern was observed in stage 4 U.S. price and China ending stocks forecasts, where positive correlation with forecast levels signals underestimation of change. Thus, U.S. price and China ending stocks forecasts should be scaled up by 4% and 29% $(1 + \gamma_1)$, respectively. Note that the magnitude of forecast levels should be taken into account in interpreting the relative impact of this inefficiency. For example, since average forecasts for stage 2 U.S. ending stocks are 5.23 million bales, a 3.2% overestimation is 170,000 bales, which is greater in absolute terms than the 4.6% overestimation in China's stage 1 average export forecasts of 0.78 million bales (equivalent to 36,000 bales). The magnitude of average annual variation is also a factor to be taken into account when ranking the impact of inefficiency in

Table 5. Estimated γ_1 for the Test of Efficiency with Respect to Forecast Levels for WASDE Cotton Forecasts, 1985/86-2009/10 Marketing Years[a]

Country/ Stage of forecasting cycle	Production	Imports	Domestic Use	Exports	Ending Stocks	Average Price	N
U. S.							
Stage 1	0.04		−0.66**	0.24	−0.16	0.08	75
Stage 2	0.06		−0.41***	0.14	−0.32**	−0.12	75
Stage 3	0.03		−0.21	0.09	−0.10	0.01	100
Stage 4			−0.10	0.01	−0.07	0.04**	100
China							
Stage 1-2	−0.06	−0.16	−0.11	−0.46**	0.18		100
Stage 3	0.10	0.13	0.10	−0.31*[1]	0.35		100
Stage 4	0.03	0.06	−0.05	−0.24	0.29***		100

Notes: e is forecast error, T is a trend, and 2006 is a dummy variable. Stage 1 includes forecasts released in months 1-3, Stage 2 includes months 4-6, Stage 3 includes months 7-10, Stage 4 includes months 11-14 of the forecasting cycle. Regressions estimated using panel least squares with White heteroscedasticity correction. N is the number of observations. Single, double, and triple asterisks (*, **, ***) denote statistical significance at 10%, 5%, and 1% levels, respectively.

[a] $e_t^i = \gamma_0 + \gamma_1 f_t^i + \gamma_2 T + \gamma_3 2006 + \varepsilon_t^i$

different variables. Thus, the challenges associated with predicting the rate of change were the most pronounced in the U.S. and China ending stocks forecasts.

Table 6 shows the pattern of positive annual serial correlation of errors (equation 7) for several China forecasts: stage 3-4 production, stage 1-2 domestic use, and stage 1 and 4 exports.[5] This finding suggests a tendency to repeat previous errors. Thus, a 10% error in stage 4 production forecasts last year will be followed by a 2.2% error in the same direction this year. In other words, if 2008 stage 4 China's production forecast error was 2.48%, the production forecast during the same period for 2009 should be adjusted by subtracting 0.55% (0.22 times 2.48). A similar tendency was observed in U.S. livestock price forecasts (Sanders and Manfredo, 2003) and broiler production forecasts (Sanders and Manfredo, 2002). On the other hand, negative correlation of errors was observed in stage 1 U.S. price forecasts, stage 3 U.S. domestic use and China's imports forecasts, and stage 4 China's domestic use forecasts. This pattern suggests a tendency to offset previous years' errors rather than repeat them. Thus, a 10% overestimation in a stage 1 U.S. price forecast will be followed by a 3% underestimation next year, and the forecast should be adjusted accordingly.

Table 7 demonstrates the most frequently observed evidence of inefficiency in U.S. forecasts: positive correlation in forecast revisions, which was present in forecasts in all stages and for all variables but price. This pattern was also observed in early forecasts of China's production, imports, and domestic use and suggests that the affected forecasts are "smoothed," since the new information is not fully incorporated in the forecasts as it becomes available and is carried over into subsequent revisions (Nordhaus, 1987; Isengildina, Irwin, and Good, 2006a). This finding implies that revisions in one month help predict revisions in the subsequent month. Most estimated values of λ_1 in equation (8) range from about 0.23 to 0.37, indicating that a 10% revision to the forecast in month i is associated with a 2.3 to 3.7% revision in the subsequent month ($i + 1$). For example, an April 2009 (stage 4) 6.45% downward revision from March U.S. ending stocks forecast represented an actual change of 8.45% ($6.45 + 6.45 \times 0.31$), if the expected component in the next revision (6.45×0.31) is also taken into account. These findings resemble the evidence of smoothing in corn and soybean production forecasts shown by Isengildina, Irwin, and Good (2006a).

[5] In addition to this test, the Ljung-Box Q-statistic was calculated for each variable's forecasts for each month i across years t out to 5 lags. U.S. domestic use stands out in this test with rejections of the null hypothesis of white noise errors with at least 0-2% significance for the first 5 months of forecasts, and 10% in the 6[th] month. The U.S. use errors follow an AR(3) process, an initial sign of inefficiency. For other variables, virtually every forecast error satisfied the conditions for white noise.

Table 6. Estimated ρ_1 for the Test of Efficiency with Respect to Past Errors for WASDE Cotton Forecasts, 1985/86-2009/10 Marketing Years[a]

Country/ Stage of forecasting cycle	Production	Imports	Domestic Use	Exports	Ending Stocks	Average Price	N
U. S.							
Stage 1	0.14		−0.16	−0.04	−0.05	−0.31*	72
Stage 2	0.11		−0.16	0.00	0.08	−0.10	72
Stage 3	0.11		−0.36*	0.07	0.10	0.24	96
Stage 4			−0.32	−0.06	−0.19	−0.03	96
China							
Stage 1-2	−0.02	0.12	0.25**	0.38***	−0.11		96
Stage 3	0.16*	−0.29***	0.08	0.16	−0.14		96
Stage 4	0.22***	0.04	−0.31***	0.48***	−0.12		96

Notes: e is forecast error, T is a trend, and 2006 is a dummy variable. Stage 1 includes forecasts released in months 1-3, Stage 2 includes months 4-6, Stage 3 includes months 7-10, Stage 4 includes months 11-14 of the forecasting cycle. Regressions estimated using panel least squares with White heteroscedasticity correction. N is the number of observations. Single, double, and triple asterisks (*, **, ***) denote statistical significance at 10%, 5%, and 1% levels, respectively.
[a] $e_t^i = \rho_0 + \rho_1 e_{t-1}^i + \rho_2 T + \rho_3 2006 + \varepsilon_t^i$

Table 7. Estimated λ_1 for the Test of Independence of Forecast Revisions for WASDE Cotton Forecasts, 1985/86-2009/10 Marketing Years[a]

Country/ Stage of forecasting cycle	Production	Imports	Domestic Use	Exports	Ending Stocks	Average Price	N
U. S.							
Stage 1-2	0.23**		0.36***	0.25***	0.12*	0.11	100
Stage 3	0.34***		0.37**	0.35***	0.29**	0.11	100
Stage 4			0.27***	0.37***	0.31***	0.03	100
China							
Stage 1-2	0.31***	0.11	0.32**	−0.03	0.03		50
Stage 3	0.02	0.11**	0.37***	0.02	−0.08		100
Stage 4	0.01	0.09	0.05	0.16	0.04		100

Notes: r is forecast revision and T is a trend. Stage 1 includes forecasts released in months 1-3, Stage 2 includes months 4-6, Stage 3 includes months 7-10, Stage 4 includes months 11-14 of the forecasting cycle. Regressions estimated using panel least squares with White heteroscedasticity correction. N is the number of observations. Single, double, and triple asterisks (*, **, ***) denote statistical significance at 10%, 5%, and 1% levels, respectively.
[a] $r_t^i = \lambda_1 r_t^{i-1} + \lambda_2 T + \varepsilon_t^i$

By including all supply and demand categories from the U.S. and China's balance sheets, the above results demonstrate the relative performance of the forecasts across categories. In this respect, the USDA appears most successful in forecasting production in both the United States and China. U.S. exports were also fairly well forecast: while MAPE was higher than for the production forecasts, inefficiency was limited to smoothing. Ending stocks forecasts in both regions failed to detect the direction of change early in the season. U.S. ending stocks forecasts were smoothed and showed a tendency to overestimate change. Chinese ending stocks forecasts were more successful and showed significant improvement over time with inefficiency limited to conservative forecasts of change. Larger challenges were found in the domestic use forecasts, with evidence of underestimation in the early part of the sample and overestimation in the later part of the study period; smoothing and correlation with past errors were found in both U.S. and China forecasts and overestimation of change was seen in the U.S. forecasts.

Summary and Conclusions

The goal of this study was to provide a comprehensive examination of the accuracy and efficiency of all supply and demand categories of USDA cotton forecasts using data from monthly WASDE balance sheets for U.S. and Chinese cotton from 1985/86 through 2009/10, including unpublished price forecasts. Global cotton markets underwent considerable structural change during our study period, with China realizing significant growth in its textile sector and the U.S. shrinking its domestic use and becoming mostly export oriented.

Overall, results demonstrate that USDA cotton forecasts were not optimal during the period of study. Forecast smoothing was the most widespread and persistent type of inefficiency, observed in most U.S. variables at every stage of the forecasting cycle. Correlation with past errors was more common in China forecasts and indicated both the tendency to repeat past errors in most cases as well as to offset them in some cases. Some evidence of a tendency to overestimate growth was also found in several U.S and China cotton forecasts. Bias was uncommon and limited to several cases of overestimation in China's exports and U.S. price and underestimation of China's domestic use. While forecasts of China's imports and endings stocks have improved, U.S. price and ending stock forecast errors became larger toward the end of the study period.

Separating the forecasting cycle into several stages allowed us to examine how forecast performance progressed during the forecasting cycle. While the size of the forecast errors declined as more information became available during the forecasting cycle, inefficiencies found in stage 4 highlight challenges associated with rationally incorporating this information into forecasts. This evidence suggests that the USDA needs to re-examine the procedures it uses to analyze information.

Published USDA forecasts are always the Department's best estimate of expected future realizations of the variables, but this study shows that these forecasts have several limitations in terms of both accuracy and efficiency. Over the years the USDA has periodically made an effort to improve its forecasting capability, including a series of annual interagency conferences from the mid-1990's to 2002. In recent years, the USDA's Economic Research Service has pursued additional funding to upgrade its market analysis program and has occasionally been successful (Office of Management and Budget, 2008; Allred, Gouge, and Maw, 2008; Office of Management and Budget, 2009). However, it appears that there is no continuous process to monitor and adjust forecasts based on their performance. This study provides background for developing such a system.

Another issue is access to information. While the same analysts are involved in providing forecasts for both the United States and China, the U.S. data collection system is one of the best in the world with respect to both resources and transparency of the process. In contrast, access to information from China has historically been difficult to obtain, despite marked improvements in recent years. This issue is particularly important in an environment where G-20 countries are working to set up an agricultural market information system (AMIS) in response to growing price volatility that will improve access to data on production and stocks of most food commodities (Boschat and Moffett, 2011). The differences in mean absolute errors between the U.S. and China forecasts and the relative prevalence of serially correlated forecast errors for the China forecasts shown in this study indicate that such efforts may significantly improve trade-related forecasts.

The findings of this study can be used by market participants to help interpret USDA information. If market participants are fully aware of the flaws and inefficiencies in USDA forecasts and adjust for them in their decision making process, limited or no economic losses would result (Orazem and Falk, 1989). The degree to which market participants anticipate and adjust information contained in USDA forecasts is outside the scope of this study and presents an interesting area for future research.

[Received July 2011; final revision received October 2011.]

References

Allred, E., E. Gouge, and I. Maw. "R&D in the U.S. Department of Agriculture." In Intersociety Working Group, ed., *AAAS Report XXXIII Research And Development FY 2009,* American Association for the Advancement of Science, 2008.

Bailey, D. and B. W. Brorsen. "Trends in the Accuracy of USDA Production Forecasts for Beef and Pork." *Journal of Agricultural and Resource Economics* 23(1998):515–525.

Boschat, N. and S. Moffett. "G-20 Tackles Food Production." 2011.

Botto, A. C., O. Isengildina, S. H. Irwin, and D. L. Good. "Accuracy Trends and Sources of Forecast Errors in WASDE Balance Sheet Categories for Corn and Soybeans." 2006. Paper presented at AAEA 2006 Annual Meetings, Long Beach, CA.

Clements, M. P. "Evaluating the Rationality of Fixed-Event Forecasts." *Journal of Forecasting* 16(1997):225–239.

Diebold, F. X. and J. A. Lopez. "Forecasting Evaluation and Combination." In G. S. Maddala and C. R. Rao, eds., *Handbook of Statistics 14: Statistical Methods in Finance,* Amsterdam: North-Holland, 1998, 241–268.

Elliott, G. and A. Timmermann. "Economic Forecasting." *Journal of Economic Literature* 46(2008):3–56.

Fortenbery, T. R. and D. A. Summer. "The Effects of USDA Reports in Futures and Options Markets." *The Journal of Futures Markets* 13(1993):157–173.

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

Henriksson, R. D. and R. C. Merton. "On Market Timing and Investment Performance. II. Statistical Procedures for Evaluating Forecasting Skills." *Journal of Business* 54(1981):513–533.

Irwin, S. H., M. E. Gerlow, and T. R. Liu. "The Forecasting Performance of Livestock Futures Prices: A Comparison to USDA Expert Predictions." *Journal of Futures Markets* 14(1994):861–875.

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

———. "Are Revisions to USDA Crop Production Forecasts Smoothed?" *American Journal of Agricultural Economics* 88(2006a):1091–1104.

———. "The Value of USDA Situation and Outlook Information in Livestock and Hogs Markets." *Journal of Agricultural and Resource Economics* 31(2006b):262–282.

Isengildina, O. and S. MacDonald. "Cotton Prices and the World Cotton Market: Forecasting and Structural Change." 2009. Economic Research Report Number 80, U.S. Department of Agriculture, Economic Research Service.

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

Leuthold, R. M. "On the Use of Theil's Inequality Coefficients." *American Journal of Agricultural Economics* 57(1975):344–346.

MacDonald, S. "Rational Commodity Forecasts: Improving USDA's Cotton Analysis." In *12th Federal Forecasters Conference–Papers and Proceedings,* Veteran's Health Administration, U.S. Department of Veterans Affairs, 2002, 355–360.

———. "A Comparison of USDA's Agricultural Export Forecasts with ARIMA-Based Forecasts." In *14th Federal Forecasters Conference–Papers and Proceedings,* Veteran's Health Administration, U.S. Department of Veterans Affairs, 2005, 155–161.

McIntosh, C. S. and J. H. Dorfman. "Qualitative Forecast Evaluation: A Comparison of Two Performance Measures." *American Journal of Agricultural Economics* 74(1992):209–214.

Mincer, J. A. and V. Zarnowitz. "The Evaluation of Economic Forecasts." In J. A. Mincer, ed., *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance,* Washington, DC: National Bureau of Economic Research, 1969, 1–46.

Nordhaus, W. D. "Forecasting Efficiency: Concepts and Applications." *The Review of Economics and Statistics* 69(1987):667–674.

Office of Management and Budget. *Statistical Programs of the United States Government Fiscal Year 2009.* 2008.

———. *Statistical Programs of the United States Government Fiscal Year 2010.* 2009.

Orazem, P. and B. Falk. "Measuring Market Responses to Error-Ridden Government Announcements." *Quarterly Review of Economics and Business* 29(1989):41–55.

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

Sanders, D. R. and M. R. Manfredo. "USDA Production Forecasts for Pork, Beef, and Broilers: An Evaluation." *Journal of Agricultural and Resource Economics* 27(2002):114–128.

———. "USDA Livestock Price Forecasts: A Comprehensive Evaluation." *Journal of Agricultural and Resource Economics* 28(2003):316–334.

Schnader, M. H. and H. O. Stekler. "Evaluating Predictions of Change." *The Journal of Business* 63(1990):99–107.

Skelly, C., H. Colby, and J. Johnson. "Developments in the Estimation of China's Cotton Supply and Demand." In *Proceedings of the 2010 Beltwide Cotton Conferences,* New Orleans, LA, 2010, 414–421.

Theil, H. *Economic Forecasts and Policy*. Amsterdam: North-Holland Publishing Co., 1958.

U.S. Department of Agriculture, Office of the Chief Economist. "World Agricultural Supply and Demand Estimates." 1985-2010. Monthly report, USDA, Washington, DC. Various issues. Available online at http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1194.

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

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