# A Comprehensive Evaluation of USDA Cotton Forecasts

by
Olga Isengildina-Massa, Stephen MacDonald,
and Ran Xie

Suggested citation format:

Isengildina-Massa, O., S. MacDonald, and R. Xie. 2011. “A Comprehensive Evaluation of USDA Cotton Forecasts.” Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, MO. [http://www.farmdoc.illinois.edu/nccc134].

**A Comprehensive Evaluation of USDA Cotton Forecasts.**

by

Olga Isengildina-Massa, Stephen MacDonald and Ran Xie[1]

*Paper presented at the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management*
*St. Louis, Missouri, April 18-19, 2011*



[1] Assistant Professor, in the Department of Applied Economics and Statistics at Clemson University, Senior Economist, ERS, USDA, and Graduate student in the Department of Applied Economics and Statistics at Clemson University, respectively. 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. 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

# A Comprehensive Evaluation of USDA Cotton Forecasts

*This study provides a comprehensive examination of accuracy and efficiency of all USDA cotton supply and demand estimates for the U.S. (including unpublished price forecasts), China and rest of the world (ROW) over1985/86 through 2009/10. Our findings show that USDA overestimated China's exports and underestimated China's domestic use and ROW imports. Based on correlation of forecast errors with levels, estimates of U.S. domestic use, ending stocks and China's exports were too extreme while forecasts of China's ending stocks and ROW production and exports were too conservative. Correlations with past errors suggest that USDA tends to repeat errors in ROW production forecasts and overcorrect errors in ROW exports forecasts. Significant positive correlation between subsequent revisions indicating forecast "smoothing" was detected in the U.S. production, domestic use, exports and ending stocks forecasts, China's imports, domestic use and exports forecasts and the ROW production and domestic use forecasts. While China's ending stocks and production forecasts significantly improved over time, (unpublished) U.S. price forecasts became worse. Based on correlations of errors we conclude that better forecasts of U.S. ending stocks and domestic use forecasts, China's imports and ROW ending stocks and exports forecasts are essential for improving U.S. cotton price forecasts.*

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

## Introduction

It is a commonly held belief of agricultural market participants and analysts that USDA forecasts function as the "benchmark" to which other private and public forecasts are compared. The dominant role of USDA forecasts 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. Because of their importance, there is a vast body of literature devoted to analyzing accuracy and efficiency of USDA forecasts (e.g., Irwin, Gerlow and Liu, 1994; Bailey and Brorsen, 1998; Sanders and Manfredo, 2002; Isengildina, Irwin, and Good, 2004; Isengildina, Irwin, and Good, 2006b). These studies focus on production and price forecasts of major U.S. commodities, such as corn, soybeans, wheat, hogs and cattle. Other major commodities, such as cotton, received relatively little attention. Only a few studies have concentrated on a subset of USDA forecasts for cotton (MacDonald, 2002) or included examination of cotton in studies of USDA export forecasts for a number of commodities (MacDonald, 1999 and MacDonald 2005).

Another limitation of the previous literature is that the accuracy of USDA forecasts other than production and price has largely been overlooked. Previous forecast evaluation literature has largely focused on accuracy of production (e.g., Gunnelson, Dobson and Pamperin, 1972; Thomson, 1974; Isengildina, Irwin, and Good, 2006b) and price (e.g., Marquardt and McGann, 1977; Just and Rausser, 1981; Irwin, Gerlow and Liu, 1994; Sanders and Manfredo, 2002; Egelkraut et al., 2003; Isengildina, Irwin, and Good, 2004) forecasts. The importance of price forecasts is obvious, given the role price expectations play in decisions on resource allocation. Production forecasts are important because they are a major determinant of future supply. The accuracy of most other USDA forecasts describing supply and demand forces has been overlooked in the previous literature. To the best of our knowledge, only one previous study (Botto et.al., 2006) investigated the accuracy of all forecasts for U.S. corn and soybeans published within WASDE (World Agricultural Supply and Demand Estimates) reports over 1980/81-2003/04 marketing years.

Knowledge of supply and demand forecast accuracy is important because these categories serve as building blocks for price forecast accuracy. Furthermore, supply and demand estimates are published within a set of other forecasts in USDA's WASDE reports that have been shown to affect the markets (e.g., Colling and Irwin, 1990; Fortenbery and Sumner, 1993; Baur and Orazem, 1994; Isengildina, Irwin, and Good, 2006a). Even less is known about the accuracy and efficiency of WASDE forecasts of the world and foreign supply and demand categories that may affect U.S. markets through trade. Cotton presents a good opportunity to study these linkages as it is one of the most trade dependent U.S. commodities.

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 WASDE upland cotton forecasts. This study will concentrate on USDA cotton forecasts as little is known about their accuracy. In fact, the USDA was legally prohibited from forecasting cotton prices from 1929 to 2008. Although cotton price forecasts were not published, USDA's Interagency Commodity Estimates Committee for cotton calculated unpublished price forecasts each month. The accuracy of these unpublished forecasts should be evaluated as USDA moves forward with its cotton price forecasting mission. Cotton also presents a challenge of being one of the most trade dependent U.S. commodities. Therefore, its evaluation will require us to look beyond U.S. WASDE categories, which has not been done before. This study will use data from monthly WASDE balance sheets for cotton for the U.S, China and the World over 1985/1986 through 2009/2010 including unpublished price forecasts.

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) bias, 2) efficiency with respect to forecast levels, 3) efficiency with respect to past errors, 4) efficiency in forecast revisions, 5) forecast improvement over the forecasting cycle, 6) forecast improvement over the study period, 7) investigation of how errors in ending stocks forecasts originate from errors in other balance sheet categories and whether U.S. cotton price errors are correlated with other U.S. as well as foreign balance sheet errors.

## Data

WASDE reports are released by the USDA usually between the 9$^{th}$ and the 12$^{th}$ 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 cotton). The first forecast for a marketing year is released in the May preceding the U.S. marketing year. USDA forecasts for China were historically not released until the July preceding the U.S. marketing year. Estimates for the United States are largely finalized 18 months later, by November after the marketing year (Figure 1) except for the U.S. production forecasts which are finalized by May (month 13 of the forecasting cycle). A similar 19 month forecasting cycle is used in this study for non-U.S. estimates as well, with China production forecasting cycle ending by July (month 15).[2]

USDA WASDE forecasts are fixed-event forecasts because the series of forecasts is related to the same terminal event ( $y_t^I$ ), where $I$ is the release month of the final estimate for the category for the $t^{th}$ marketing year. The forecast of the terminal event for month $i$ is denoted as $y_t^i$ , where $i$=1, ..., $I$, $I$=19, and $t$=1985/86, …, 2009/10. Thus, each subsequent forecast is essentially an update of the previous forecast as it describes the same terminal event. The WASDE forecasting cycle generates 18 updates for each forecasted variable except production (12-14 updates) within each marketing year.

WASDE forecasts for the U.S. and the world follow a balance sheet approach to account for supply and utilization (see Vogel and Bange (1999) for a detailed description of the USDA crop forecast generation process). The major components of the balance sheet are beginning stocks, production and imports on the supply side and domestic use, exports and ending stocks reflecting utilization. The balance sheet approach means that individual estimates are cross checked against each other, across commodities and countries. For example, "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, p. 10). WASDE price estimates describe marketing year average prices received by farmers, which is an average of monthly prices weighed by the amounts marketed at these prices. While the price forecasts have been published in the form of an interval since 2008, unpublished price forecasts were calculated as a point estimate. To overcome this inconsistency and keep the analysis consistent across all forecasts, midpoints of the published price forecast intervals were used in this study.

The focus of this study is monthly WASDE forecasts for the U.S., China, and World Upland Cotton for the marketing years of 1985/86 through 2009/10. To separate the influence of the U.S. and China forecast errors on the World forecasts, the" Rest of the World" (ROW) forecasts were calculated by subtracting U.S. and China values from the World values. Table 1 shows the descriptive statistics for the final estimates ($I$=19) of the supply and demand categories for these regions. The means for various categories reported in this table show that both China and the U.S. are major cotton producers that jointly produce over 43% of cotton in the world. China is also a major consumer of cotton with the growing textile sector supported by domestic production and supplemented by imports. The demands of China's textile sector are also facilitated by relatively high levels of stocks.

The U.S. cotton industry is characterized by a shrinking textile industry and essentially no raw cotton imports. Changes in international cotton trade that occurred in the mid-1990s resulted in the growth of U.S. cotton exports and the decline in domestic use and stocks. The U.S. cotton price averaged about 56.75 cents/lb during the period of this study. Similar to other major U.S. commodities, cotton price has been supported by the farm programs prior to 1985, but has become more market oriented since. Due to the increased export orientation of the U.S. cotton

[2] Even though trade and consumption revisions for non-U.S. estimates sometimes occurred later in the forecasting cycle, in aggregate the revisions were small.

industry, the price of U.S. cotton is becoming increasingly affected by international market forces (Isengildina and MacDonald, 2009). Positive skewness and kurtosis values near or above 1 in China's import and export forecasts indicate the frequent presence of mostly positive outliers in the data. These outliers illustrate occurrences such as sharp changes in imports as Chinese imports tripled from 2004 to 2005.

The standard deviations and the coefficients of variation indicate absolute and relative variability in the forecasted categories. For example, Table 1 shows that China has some of the largest values of coefficients of variation in all categories (except production). These large values imply that China's supply and demand categories are very volatile and difficult to forecast. In the US, the coefficients of variation for exports and ending stocks were nearly twice as large as those of the other US categories (about 41% and 38%, respectively), indicating higher volatility and potential challenges in forecasting these categories. Similarly, the coefficient of variation for ROW ending stocks at 26% echoes the pattern observed in the US, suggesting that ending stocks are more difficult to forecast than other supply and demand categories.

Descriptive statistics of the forecast errors shown in table 1 demonstrate that in absolute terms errors were the largest for the categories that are most difficult to forecast, China's exports, imports and ending stocks, and U.S. exports and ending stocks. Huge maximum and minimum errors in China's imports, exports and endings stocks forecasts illustrate challenges with obtaining reliable data from China. As described by Skelly, Colby and Johnson (2010), "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." Since NDRC estimates were deemed more realistic, USDA switched to forecasting NDRC rather than NBS 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).

## Forecast Evaluation Framework

To evaluate the accuracy and efficiency of WASDE cotton forecasts it is necessary to establish a set of testable properties for an optimal forecast as suggested in recent evaluation studies (e.g., Sanders and Manfredo, 2003; Timmerman, 2006). Following Timmerman (2006), this study assumes that the objective function is of the mean squared error (*MSE*) type so the forecasts minimize a symmetric, quadratic loss function. The properties of WASDE forecasts are investigated using error and revision analysis. For each category, monthly announcement and marketing year forecast errors and revisions were calculated in percentage terms to take into account changes in forecast levels:

$$e_t^i = 100 * \ln\left( y_t^I \big/ y_t^i \right); \quad i=1, \ldots, I\text{-}1; \quad t=1985/86,\ldots,2009/10 \tag{1}$$

$$r_t^i = 100 \times \ln\left( y_t^i \big/ y_t^{i-1} \right); \quad i=2, \ldots, I; \qquad t=1985/86, \ldots, 2009/10$$

where $e_t^i$ corresponds to the error, $r_t^i$ is the revision for a given report month $i$, and marketing year $t$. As defined earlier, $y_t^i$ is the forecast for marketing year $t$ released in month $i$ and $y_t^I$ corresponds to the final estimate for marketing year $t$, $I$=19 for cotton.

The fundamental measures of optimal forecasts are bias and efficiency (Diebold and Lopez, 1998). The test of bias can be performed using a regression:

(2) $$e^i_t = \alpha_0 + \varepsilon^i_t \qquad i=1, \ldots, I-1; \; t=1985/86, \ldots, 2009/10$$

The null hypothesis for an unbiased forecast is $\alpha_0 = 0$. If $\alpha_0 > 0$, forecasts are consistently underestimating the final estimate. If $\alpha_0 < 0$, forecasts are consistently overestimating the final estimate.

Weak efficiency tests evaluate whether forecast errors are orthogonal to forecasts themselves as well as to prior forecast errors (Nordhaus, 1987). Following Pons (2000) and Sanders and Manfredo (2002, 2003) weak efficiency with respect to forecast levels is tested using the following regression:

(3) $$e^i_t = \alpha_1 + \beta_1 f^i_t + \varepsilon^i_t \qquad i=1, \ldots, I-1; \; t=1985/86, \ldots, 2009/10,$$

where forecast levels are measured as a percent change from the previous year's levels[3]:

$$f_t^i = 100 * \ln\left( y_t^i \big/ y_{t-1}^I \right) \quad i=1, \ldots, I-1; \; t=1985/86, \ldots, 2009/10.$$

The null hypotheses for efficient forecasts is $\beta_1 = 0$. If $\beta_1 > 0$ in equation (3), that means that when forecasts become larger, so do the positive errors (greater underestimation) while the negative errors become smaller (smaller overestimation). If $\beta_1 < 0$, it implies that larger forecasts are correlated with smaller positive errors (smaller underestimation) and larger negative errors (larger overestimation).

Forecast efficiency with respect to past errors is measured as:

(4) $$e^i_t = \alpha_2 + \beta_2 e_{t-1}^{\,i} + \varepsilon^i_t \qquad i=1, \ldots, I-1; \; t=1985/86, \ldots, 2009/10.$$

Note that for fixed event forecasts, the forecast error for the previous event (marketing year) should be used for this test. The null hypotheses for efficient forecasts is $\beta_2 = 0$. If $\beta_2 \neq 0$ in

[3] This transformation was necessary to be consistent with our measurement of percent forecast errors. When forecast errors are measured as $e_t^i = 100 * \ln\left( y_t^I \big/ y_t^i \right) = 100 * \left( \ln\left( y_t^I \big/ y_{t-1}^I \right) - \ln\left( y_t^i \big/ y_{t-1}^I \right) \right)$, $\ln\left( y_t^I \big/ y_{t-1}^I \right)$ represents the final estimate and $\ln\left( y_t^i \big/ y_{t-1}^I \right)$ denotes the forecasted value.

equation (4), there is a systematic component in forecast errors that can be predicted using past errors.

Furthermore, 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. Following Isengildina, Irwin, and Good (2006b), this property can be tested formally using the following regressions:

$$(5) \qquad r^{i}_{t} = \gamma\, r^{i-1}_{t} + \varepsilon^{i}_{t} \qquad i=1, \ldots, I\text{-}1;\ t=1985/86,\ldots,2009/10,$$

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

Another property of an optimal forecast is that the variance of the forecast error should decline as more information becomes available. This means that as $t$ increases, variance of forecast errors should decrease. This can be tested through a variance ratio test or (more appropriately given the small sample size here) by considering patterns in the variance of forecast errors associated with different forecast horizons.[4]

Following the methodology used by Bailey and Brorsen (1998), Sanders and Manfredo (2003) suggested a test that evaluates forecast improvement over time by focusing on the changes in the absolute value of the forecast error:

$$(6) \qquad \left| e^{i}_{t} \right| = \alpha_t + \beta_3 T + \varepsilon^{i}_{t} \quad i=1, \ldots, I\text{-}1;\ t=1985/86,\ldots,2009/10,$$

where T is a linear time trend. The null hypothesis is $\beta_3 = 0$, which indicates no systematic change in the size of forecast error. If $\beta > 0$, the forecasts become worse over time as evidenced by larger errors. If $\beta < 0$, the forecasts improve over time as demonstrated by smaller errors.

If the forecasts are interrelated (as is the case with WASDE forecasts since they are constructed using a balance sheet approach), errors in individual categories may contribute to errors in aggregate categories, such as ending stocks and price. Botto et al., (2006) examined such relationship by regressing errors in corn and soybean ending stocks and price forecasts against errors in individual balance sheet categories. For cotton WASDE forecasts, we hypothesize that errors in ending stocks forecasts will be caused by errors in production, imports, domestic use, and exports forecasts from the same balance sheet. Due to heavy trade dependency of the U.S. cotton, errors in the U.S. price forecasts may be correlated with errors in U.S., as well as foreign (China and ROW) balance sheet categories. To avoid the effects of collinearity between

[4] Alternatively, changes in forecast error over time and across the forecasting horizon can be investigated using a framework developed by Bailey and Brorsen (1998).

independent variables, and the adding up condition for the ending stocks equation[5] these relationships are evaluated in this study using Pearson correlation coefficients.

As described in the data section, WASDE forecasts are fixed event forecasts. Because of this nature, all monthly forecasts for the same category within the same marketing year are related to the same terminal event (annual estimate for the category) and hence are correlated. Therefore, forecast evaluation regressions described in equations 2 - 6 were initially evaluated separately for each forecast month across marketing years.[6] This approach allows observing changes in forecast accuracy and efficiency across the forecasting cycle.

A more general analysis involved simultaneous estimation of monthly equations while allowing coefficients to change linearly over time and within the forecasting cycle. This was accomplished by pooling the data across all marketing years and forecast months in a panel where forecast months were treated as cross sections. Panel least squares method with correlated and group-wise heteroskedastic panels was applied to estimate the parameters and to the calculation of standard errors and covariances using the White cross-section method. 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. This estimator is robust to cross-equation correlation as well as different error variances in each cross-section (Wooldridge, 2002, p 148-153 and Arellano, 1987).

To allow coefficients to change within the forecasting cycle, a trend variable "Month" was added. To prevent the effect of the trend variable on hypothesis testing from equations 2-6, the variable was coded such that its mean was always equal zero. For example, for 17 months of the forecasting cycle, the first forecast month (May prior to harvest) was coded as -8, this value increased by one each subsequent month and was equal 8 in month 17 (September after harvest). Similar approach was used to allow coefficients to change over time by using a trend variable "Year" with a zero mean. These month and year trends were included in panel estimation of equations 2 – 6. Additionally, a dummy variable with the values of 1 for all forecast months in 2006 and beginning of 2007 (through September, $i$=5) was added to the evaluation of China's forecasts to take into account the impact of the change in information source used to provide these forecasts.

## Results

The results of the test of bias shown in table 2 suggest that bias was present in China's export and domestic use forecasts as well as ROW import and ending stock forecasts. The bias in China's export forecasts was particularly astounding: it averaged 28% overestimation within our sample. The error decreased over the forecasting cycle (from $May_t$ to $November_{t+1}$ for each marketing year) by about 2.6% a month. The coefficient on the Month variable can be interpreted in combination with the constant relative to the reference point of forecast month 10

[5] Within the balance sheet, ending stocks = total supply – total use, must be maintained.

[6] Results not presented here due to space considerations, but available from authors upon request.

(January) for the U.S. and 11 (February) for China and ROW (when Month=0). It means that the overestimation in China's export forecasts increased by 2.6% each month from 28% in months prior February and decreased by 2.6% in the months following February over the forecasting cycle. Thus, the overestimation in the first (July) forecast for China's exports averaged as much as 46%. This value should be interpreted relative to the average size of these forecasts of 0.72 million bales, thus it indicates about 0.33 million bales average overestimation. This pattern illustrates the impact of changes in Chinese trade policies on the ability of USDA to forecast them. Underestimation in China's domestic use forecasts was about 1.52% or 416 thousand bales on average within our sample. While underestimation in China import forecasts was not significantly different from zero, a coefficient on the Year trend indicates a significant reduction in the level of underestimation and an increasing tendency for overestimation that started around 2002/03. A similar pattern was observed in the U.S. domestic use forecasts, where the tendency for overestimation started in 1997/98. This pattern illustrates difficulties USDA had with forecasting structural changes and increased export orientation in the U.S. cotton industry. A dummy variable was included in the tests of the China production and ending stock forecasts to account for USDA's shift of sources for its China production data in October 2007. Ending stocks is to a large extent forecasted as a residual, so USDA's estimates for ending stocks adjusted with its estimates for production.

Bias in ROW import and ending stocks forecasts was not accompanied by any monthly or yearly patterns. The coefficients indicate that USDA underestimated ROW imports by 1.4 % or 347 thousand bales based on the 24.77 million bales average forecast level and ending stocks by 2.71 percent or 607.6 thousand bales based on 22.42 million bale average value. No bias was found in the U.S. cotton forecasts. The only significant coefficient across the U.S. categories was the Year trend for domestic use forecasts over time similar to what we observed for China's import forecasts. In the case of U.S. domestic use, while the mean error of USDA's forecasts over the 1985-2009 period is not statistically different from zero, the forecasts do have a form of systemic bias. The forecasts typically underestimated future domestic use during the initial years of the study period, and overestimated it by the end. The implication is that USDA might be expected to overestimate U.S. domestic use of cotton in future years, despite the lack of statistically significant bias for the overall historical sample.

Results of the forecast efficiency tests reported in table 3 indicate that errors of U.S. domestic use and ending stocks forecasts, China's exports and ending stocks forecasts and ROW production and exports forecasts were significantly correlated with forecast levels. For U.S. domestic use, ending stocks and China's exports the observed correlation was negative indicating that larger forecasts were associated with negative errors (overestimation) while smaller forecasts were associated with positive errors (underestimation). Thus, our findings suggest that USDA forecasts of the affected categories were probably too extreme. For example, for each 10% increase in U.S. domestic use forecasts from the previous year's level, we expect forecast error to grow by -3.3%. Hence the forecast should be scaled down by that amount. Another way to interpret these estimates is to recall that equation (3) is derived from the traditional Mincer-Zarnowitz test of forecast efficiency:

$$
\begin{aligned}
&e_i^t = (y_t^I - y_t^i) = \alpha_1 + \beta y_t^i + \varepsilon_t^i = \alpha_1 + (\gamma - 1) y_t^i + \varepsilon_t^i = \alpha_1 + (\gamma y_t^i - y_t^i) + \varepsilon_t^i \\
&y_t^I = \alpha_1 + \gamma y_t^i + \varepsilon_t^i
\end{aligned} \tag{7}
$$

Thus, forecasts with significant coefficients in equation (7) can be made more accurate by adjusting them with $\gamma=1+\beta$ and $\alpha$. For example, the China export forecasts would be improved on average by multiplying them by 0.72, and subtracting 31.3 percentage points.

A negative coefficient on the year trend in the U.S. domestic use equation also suggests a growing tendency for overestimation over time, as highlighted in the bias test above. This finding likely illustrates inability of USDA to accurately predict the speed of the shrinkage in the US textile sector in the recent decades. While the forecasts of China's exports have a similar pattern of being too extreme, a significant coefficient on the month variable indicates that the tendency to overestimate declines during the forecasting cycle. China's cotton exports have tended to decline over time and have been negligible in recent years—comparable to U.S. imports. China was typically a net importer during the period studied and exports often required policy-based assistance, which has not been forthcoming in recent years. On the other hand, China's ending stocks were typically underestimated. Table 3 indicates that forecasts adjusted by multiplying by 1.18 and adding 6.61 percentage points would be more accurate on average than those published by USDA. Correlations with forecast levels in the ROW categories were some of the weakest: 0.13% for production and 0.12% for exports forecasts, indicating that these forecasts were slightly too conservative.

Table 4 presents the results of the test of efficiency with respect to past errors. The only evidence of inefficiency was found in ROW forecasts where forecast errors of production and exports correlated with respective errors from the previous marketing year. Positive correlation in ROW production forecast errors suggests a tendency to repeat past errors. Thus, if last year USDA underestimated ROW production by 10%, one can expect that this year's forecast will underestimate it by 2.9% and can be adjusted by adding 2.9% to the published value accordingly. This finding is consistent with the evidence of positive correlation with previous forecast errors in USDA livestock price and production forecasts documented by Sanders and Manfredo (2003, 2002). As Sunders and Manfredo argue, the tendency to repeat past errors may reflect difficulties with modeling structural changes (Sanders and Manfredo, 2003). Negative correlation observed in ROW exports forecasts suggests the tendency to overcorrect errors from the previous year. Thus, if USDA underestimated ROW export forecast last year by 10%, one can expect this year's forecast to overestimate exports by 1.8% and adjust it accordingly.

Correlation in forecast revisions demonstrated in table 5 can also be used to predict errors in forecasts. Significant positive correlation between subsequent revisions was detected in the U.S. production, domestic use, exports and ending stocks forecasts, China's imports, domestic use and exports forecasts and the ROW production and domestic use forecasts. This pattern 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, 2006). For example, if U.S. cotton production forecast is revised up by 10% in September ($i$=5), one can expect the following October ($i$=6) forecast to be 2.6% higher than September. The magnitude of smoothing was between 20 and 33 percent for the US and ROW and between 12 and 16 percent for China forecasts. A significant positive coefficient on the month variable indicates that the correlation becomes stronger over the forecasting cycle. This pattern was detected in U.S. exports and ROW domestic use forecasts. For example, U.S.

export forecast revisions in 1997 had about 2.66% correlation with previous revisions early in the forecasting cycle (*i*=3) and about 3.61% correlation with previous revisions toward the end of the forecasting cycle (*i*=17). Note that the degree of smoothing also depends on the magnitude of revision in the previous month so as revisions become smaller during the forecasting cycle so does the predictable component of the next forecast. A negative coefficient on the year variable indicates that the correlation became weaker during the study period. Thus, the correlation in U.S. domestic use forecast revisions weakened by about 4 percentage points a year so that the correlation between January(*i*=9) and February(*i*=10) revisions changed from about 3.82% in the beginning of the sample (*t*=1985/86) to about 2.87% at the end of the sample (*t*=2009/10). These time effects were much stronger in China's import forecasts where the correlation between January and February revisions changed from about 6.02% in the beginning of the sample to about -2.7% at the end of the sample.

Changes in the standard deviation of forecast errors illustrated in figure 2 exhibit a pattern of decreasing variance as more information becomes available, thus satisfying this requirement for an optimal forecast. This pattern is violated only in the Chinese import forecasts where the standard deviation in increasing between months 3 and 4 of the forecasting cycle and decreasing thereafter, but the difference between the months' 3 and 4 variances are not statistically significant. This figure also shows that U.S. ending stocks and exports forecasts had the highest variability among the U.S. categories illustrating the challenges with forecasting these categories likely caused by structural changes that resulted in increased export orientation in the US cotton industry. The standard deviation of errors of China's exports and imports forecasts was more than double that of any other forecasts. This pattern illustrates a great deal of uncertainty and variability in these categories. Ending stocks forecasts were the third most variable category on the Chinese balance sheet, but their standard deviation was comparable to that of the U.S. ending stocks. Interestingly, the variation in the percent errors for the rest of the world forecasts was much lower than that of China's and the U.S. categories illustrating that the U.S. and China have the largest impact on the uncertainty in the world cotton market.

The results of the test of forecast improvement shown in table 6 illustrate the average magnitude of forecast errors (in month 10 for the U.S. and 11 for China and ROW of 1997/98, when both month and year equal 0) in the constant. Consistent with the pattern observed in figure 2, the largest forecast errors were observed in China's exports, imports and ending stocks forecasts, 54.32%, 34.26%, and 18.9% respectively. The largest errors in the U.S. balance sheet were associated with ending stocks and export forecasts, 17.61% and 9.27%, respectively. The magnitude of forecast errors in the ROW categories was relatively smaller, with the largest errors associated with ending stocks forecasts at 6.25%. The coefficient on the month variable describes how forecast errors decreased throughout the forecasting cycle. For example, a -2.09 coefficient for the U.S. ending stocks indicates that the magnitude of forecast error became smaller by 2.09 percent each month during the forecasting cycle. When this coefficient and the constant are interpreted relative to the reference point of 0 (January, *i*=10 for the U.S.), it indicates that the average error was 17.61% in January and it decreased by 2.09% in each following month, but increased by 2.9% in each preceding month of the forecasting cycle, thus bringing the average May forecast error to 34.3%. The coefficient on the year variable tests the hypothesis of whether the magnitude of forecast error changed over time. We found significant evidence of forecast improvement in China's ending stocks and production forecasts. China's endings stocks forecast errors became 0.75% smaller each year during our study period. The improvement in China's production forecasts measured at 0.23% a year. On the other hand, we found that the U.S. price forecasts became significantly worse with the error increasing by 0.16% a year over 1985/86 through 2009/10 marketing year. Note, however, that this finding refers to unpublished forecasts.

The final property of forecasts investigated in this study was correlation in errors of related categories. The premise behind this analysis is to investigate to what extent errors in individual categories contribute to errors in aggregate categories, such as ending stocks and price. Figure 3 shows correlations between ending stocks errors with individual categories from the same balance sheet.[7] This graph demonstrates that U.S. ending stock errors are strongly negatively correlated with errors in exports forecasts. This correlation remains significant at 10% level through month 15 of the forecasting cycle. Correlation between ending stocks errors and production errors is around 40% and significant through month 6 of the forecasting cycle, but declining thereafter. Thus the uncertainty regarding the size of the U.S. cotton crop (and its impact on the ending stocks forecasts) begins to resolve only after October forecast. Domestic use forecast errors had a relatively modest contribution to ending stocks forecast errors in the U.S. cotton balance sheet.

Figure 3 also demonstrates that China's endings stocks forecast errors were mostly due to errors in production forecasts with correlations ranging from 0.5 to 0.7 through most of the forecasting cycle. Domestic use forecast errors are significantly correlated with ending stocks forecast errors through month 11 and again in the last two months of the forecasting cycle. Thus, the final refinements in China's forecasts are largely in the production and domestic use categories. Trade categories had relatively little impact on China's ending stocks forecast errors throughout the forecasting cycle. These findings suggest that better information about China's cotton production will have the largest impact on the improvement of China's ending stocks forecasts.

The bottom panel of figure 3 shows that ROW ending stocks forecast errors are also mostly driven by errors in production forecasts with correlations ranging between 0.4 to 0.8 through month 16 of the forecasting cycle. World exports forecast errors also had a significant impact on ending stocks errors between month 9 and 13 of the forecasting cycle. Domestic use and import forecast errors had relatively little impact on ROW ending stocks. Thus, ROW ending stocks forecasts will benefit most from better information on world production and exports.

Figure 4 shows that errors in the U.S. average price forecasts errors were driven by errors in U.S. ending stocks and domestic use forecasts, China's imports forecasts and ROW ending stocks and exports forecasts. Most of these relationships were significant only in the first nine months of the forecasting cycle, through January (approximately through the middle of the U.S. marketing year). Once uncertainty in these forecasts becomes partially resolved post January, their impact on U.S. price forecast errors becomes insignificant (with the exception of U.S. ending stocks forecast errors in the last few months of the forecasting cycle). This finding suggests that USDA

[7] Due to the balance sheet nature of these forecasts, where ending stocks are calculated as the difference between total supply and total use, the errors ending stocks forecasts are due entirely to errors in individual categories.

may be able to improve U.S. price forecasts by improving their forecasts of U.S. ending stocks and domestic use forecasts, China's imports and ROW ending stocks and exports forecasts. The improvement will likely involve better access to data and data analysis as well as correction for inefficiencies in some of these forecasts revealed in this study.

## Summary and Conclusions

Given the limited knowledge about the accuracy of the USDA cotton forecasts, the goal of this study was to provide a comprehensive examination of accuracy and efficiency of all supply and demand categories of these forecasts. This study included data from monthly WASDE balance sheets for upland cotton for the U.S., China and the World over 1985/1986 through 2008/2009 including unpublished price forecasts. Within our study period the U.S. and China's cotton sectors underwent through structural changes in the 1990s with China growing its textile sector and the U.S. shrinking its domestic use and becoming mostly export oriented. Based on the coefficients of variation, most Chinese supply and demand categories as well as U.S. exports and ending stocks and ROW ending stocks were the most difficult to forecast. Positively skewed and leptokurtic distributions of trade categories for China indicate the frequent presence of positive outliers, which further illustrate challenges in forecasting these categories.

Our analysis of forecast bias revealed that USDA overestimated China's exports by an average of 28% or 202 thousand bales and underestimated China's domestic use by an average of 1.52% or 416 thousand bales during our study period. USDA underestimated ROW imports by 1.4 % or 347 thousand bales and ending stocks by 2.71 percent or 607.6 thousand bales. While no statistically significant bias was found in China's imports and US. domestic use forecasts, there is evidence of the tendency to overestimate these forecasts increasing over time.

Efficiency tests revealed that errors of U.S. domestic use and ending stocks forecasts, China's exports and ending stocks forecasts and ROW production and exports forecasts were significantly correlated with forecast levels. For U.S. domestic use, ending stocks and China's exports the observed correlation was negative suggesting that these forecasts were too extreme. For China's ending stocks and ROW production and exports forecasts, the correlation was positive indicating that these forecasts were too conservative.

Efficiency with respect to past errors was not rejected in all but ROW forecasts of production and exports. Positive correlation in ROW production forecast errors suggests a tendency to repeat past errors. On the other hand, negative correlation observed in ROW exports forecasts suggests the tendency to overcorrect errors from the previous year. Thus, if USDA underestimated ROW export forecast last year by 10%, one can expect this year's forecast to overestimate exports by 1.8% and adjust it accordingly.

Efficiency with respect to forecast revisions was rejected in multiple cases in USDA cotton forecasts. Significant positive correlation between subsequent revisions was detected in the U.S. production, domestic use, exports and ending stocks forecasts, China's imports, domestic use and exports forecasts and the ROW production and domestic use forecasts. This pattern 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, 2006). Thus, information on previous revisions can be used to improve forecasts.

Forecast optimality was not rejected with respect to decreasing error variance over the forecasting cycle in all cases except for early China's imports forecasts. While errors decreased over the forecasting cycle, they did not necessarily become smaller during the study period. Our analysis revealed that only China's ending stocks and production forecasts significantly improved over time. China's endings stocks forecast errors became 0.75% smaller each year during our study period. The improvement in China's production forecasts measured at 0.23% a year. On the other hand, we found that the U.S. price forecasts became significantly worse with the error increasing by 0.16% a year over 1985/86 through 2009/10 marketing year. Note, however, that this finding refers to unpublished forecasts.

Investigation of correlation in errors of related categories was conducted to figure out whether errors in supply and demand categories cause errors in aggregate categories such as ending stocks and price. Our findings suggest that U.S. ending stocks forecast errors are mostly caused by errors in U.S. exports and production forecasts. Better information about China's production and domestic use will have the largest impact on the improvement of China's ending stocks forecasts. ROW ending stocks forecasts will benefit most from better information on world production and exports. We also found that U.S. cotton price forecast errors are driven by factors reaching far beyond the domestic balance sheet. Better forecasts of U.S. ending stocks and domestic use forecasts, China's imports and ROW ending stocks and exports forecasts are essential for improving U.S. cotton price forecasts.

The results of this study highlight various inefficiencies of USDA cotton forecasts. A lot of these inefficiencies are likely due to challenges in forecasting supply and demand factors for a very dynamic industry undergoing structural changes and faced with data quality and availability issues from foreign countries. Regardless of these limitations USDA is providing a very important service to information-starved cotton market. In this sense this study will echo multiple previous authors that argue that USDA provide valuable information to market participants that can be used to assist and improve private forecasts and enhance welfare by reducing price uncertainty (e.g., Sanders and Manfredo, 2003; Isengildina, Irwin and Good, 2006a) However, it would be constructive if USDA would use the results of this study to "learn from their mistakes." An on-going forecast quality analysis similar to the one presented in this study would allow USDA to identify problem areas in their forecasting procedures and to address them in a timely manner in order to ensure the highest quality of information that they provide.

## References

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

Bailey, D.V., and B.W. Brorsen. "Trends in the Accuracy of USDA Production Forecasts for Beef and Pork." *Journal of Agricultural and Resource Economics* 23 (1998): 515–26.

Baur, R.F., and P.F. Orazem. "The Rationality and Price Effects of USDA Forecasts of Oranges." *Journal of Finance* 49(1994): 681-696.

Botto, A.C., Isengildina, O., S.H. Irwin, D.L. Good. "Accuracy Trends and Sources of Forecast Errors in WASDE Balance Sheet Categories for Corn and Soybeans." Selected Paper presented at AAEA 2006 Annual Meetings, Long Beach, California, July 14-17, 2006.

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

Diebold, F.X., and J.A. Lopez. "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, 1998.

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

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

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

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

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

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

Isengildina, O., S.H. Irwin, and D.L. Good. "The Value of USDA Situation and Outlook Information in Livestock and Hogs Markets." *Journal of Agricultural and Resource Economics*, 31(2) (2006 a): 262-282.

Isengildina, O., S.H. Irwin, D.L. Good. "Are Revisions to USDA Crop Production Forecasts Smoothed?" *American Journal of Agricultural Economics*, 88(4) (2006 b): 1091-1104.

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

MacDonald, Stephen, “A Preliminary Evaluation of USDA’s Export Forecasts,” The 10th Federal Forecasters Conference 1999: Papers and Proceedings and Selected Papers of 19th International Symposium on Forecasting, U.S. Department of Education Office of Educational Research and Improvement, pp. 305-316.

MacDonald, Stephen, “Rational Commodity Forecasts: Improving USDA’s Cotton Analysis,” 12th Federal Forecasters Conference—Papers and Proceedings. 2002, Veterans Health Administration, U.S. Department of Veterans Affairs, pp. 355-360.

MacDonald, Stephen, “A Comparison of USDA’s Agricultural Export Forecasts with ARIMA-based Forecasts,” 14th Federal Forecasters Conference—Papers and Proceedings, 2005, Veteran’s Health Administration, U.S. Department of Veterans Affairs, pp. 155-161.

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

Nordhaus, W.D. “Forecasting Efficiency: Concepts and Applications.” *Review of Economics and Statistics* 69(1987):667-674.

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): 114–28.

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

Skelly, C., H. Colby, and J. Johnson. “Developments in the Estimation of China’s Cotton Supply and Demand. Proceedings of the 2010 Beltwide Cotton Conferences, New Orleans, Louisiana, January 4-7, 2010, pp. 414-421.

Timmerman, Allan. “An Evaluation of the World Economic Outlook Forecasts.” International Monetary Fund Working Paper. March 2006.

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

Vogel, F.A., and G.A. Bange. 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.

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

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

| | Category | Production (M.Bal) | Imports (M.Bal) | Domestic Use (M.Bal) | Exports (M.Bal) | Ending Stocks (M.Bal) | Price (¢/lb) |
|---|---|---|---|---|---|---|---|
| US | Mean/a | 16.85 | | 7.98 | 8.92 | 5.23 | 56.72 |
| | Std Deviation/a | 3.24 | | 2.35 | 3.62 | 1.99 | 10.03 |
| | Coeff Variation/a | 19.23 | | 29.46 | 40.58 | 37.95 | 17.68 |
| | Skewness/a | -0.04 | | -0.23 | 0.66 | 0.64 | -0.22 |
| | Kurtosis/a | -0.39 | | -1.04 | -0.43 | -0.39 | -0.34 |
| | Mean Error/b | 0.48 | | 0.48 | 0.30 | -3.66 | -1.45 |
| | Mean Absolute Error/b | 4.67 | | 4.77 | 9.27 | 17.60 | 5.69 |
| | Std Dev of Error/b | 7.06 | | 6.80 | 15.11 | 24.44 | 8.90 |
| | Max error/b | 27.84 | | 24.69 | 29.48 | 72.31 | 20.39 |
| | Min error/b | -19.76 | | -20.35 | -93.65 | -80.35 | -45.56 |
| China | Mean/a | 22.96 | 3.97 | 27.40 | 0.72 | 12.22 | |
| | Std Deviation/a | 5.56 | 5.07 | 11.30 | 0.74 | 4.66 | |
| | Coeff Variation/a | 24.24 | 127.74 | 41.23 | 101.74 | 38.14 | |
| | Skewness/a | 1.13 | 1.55 | 1.22 | 1.17 | 0.20 | |
| | Kurtosis/a | 0.15 | 1.40 | -0.10 | 0.71 | -0.45 | |
| | Mean Error/c | 1.39 | 5.60 | 1.50 | -28.02 | 3.78 | |
| | Mean Absolute Error/c | 5.89 | 35.77 | 4.10 | 54.02 | 19.54 | |
| | Std Dev of Error/c | 8.68 | 59.45 | 5.47 | 83.00 | 26.63 | |
| | Max error/c | 25.53 | 409.43 | 18.23 | 128.09 | 81.59 | |
| | Min error/c | -23.36 | -120.40 | -18.23 | -299.57 | -85.91 | |
| ROW | Mean/a | 52.45 | 24.77 | 58.44 | 18.85 | 22.42 | |
| | Std Deviation/a | 6.56 | 2.01 | 4.67 | 3.11 | 5.89 | |
| | Coeff Variation/a | 12.50 | 8.13 | 7.99 | 16.50 | 26.26 | |
| | Skewness/a | 0.82 | -0.39 | 0.39 | 0.65 | 0.97 | |
| | Kurtosis/a | -0.22 | 0.10 | -0.50 | -0.01 | -0.42 | |
| | Mean Error/c | 0.21 | 1.40 | 0.11 | 0.73 | 2.71 | |
| | Mean Absolute Error/c | 2.34 | 3.46 | 1.60 | 4.60 | 6.25 | |
| | Std Dev of Error/c | 3.69 | 5.01 | 2.20 | 7.46 | 8.15 | |
| | Max error/c | 18.92 | 18.09 | 6.84 | 23.14 | 30.75 | |
| | Min error/c | -11.97 | -14.37 | -7.69 | -43.39 | -33.03 | |

Notes: a/Calculated using the final November($I$ =19) estimate for each category. Errors are calculated as the natural logarithm of the ratio of the forecast in month $I$ =19 to the forecast in month $i$ , times 100. b/ Calculated for May($i$ =1) through September(i=17) forecasts for US except production, which use May($i$ =1) through March($i$ =11) forecasts. c/Calculated for July($i$ =3) through September($i$ =17) forecasts for China and ROW except China production, which use July($i$ =3) through July($i$ =15) forecasts.

**Table 2. Test of Bias for WASDE Cotton Forecasts, 1985/86-2009/10 Marketing Years.**

| | | Production | Imports | Domestic Use | Exports | Ending Stocks | Average Price |
|---|---|---|---|---|---|---|---|
| **U.S.** | | | | | | | |
| | C | 0.48 | | 0.49 | 0.30 | -3.67 | -1.45 |
| | Month | -0.14 | | 0.01 | 0.05 | 0.31 | 0.22 |
| | Year | -0.01 | | -0.50 *** | 0.40 | 0.19 | -0.21 |
| | N | 275 | | 425 | 425 | 425 | 425 |
| **China** | | | | | | | |
| | C | 0.75 | 5.39 | 1.52 * | -28.19 ** | 2.97 | |
| | Month | 0.03 | -0.21 | -0.15 | 2.61 * | -0.48 | |
| | Year | 0.28 | -1.85 ** | -0.04 | -0.80 | 0.95 | |
| | 2006 | 12.87 *** | | | | 17.52 *** | |
| | N | 325 | 374 | 375 | 375 | 375 | |
| **ROW** | | | | | | | |
| | C | 0.21 | 1.40 * | 0.11 | 0.73 | 2.71 ** | |
| | Month | 0.11 | -0.06 | 0.06 | 0.06 | -0.27 | |
| | Year | -0.05 | -0.10 | -0.03 | -0.17 | -0.18 | |
| | N | 375 | 375 | 375 | 375 | 375 | |

Notes: Panel Least Squares regressions use data for May($i$ =1) through September($i$ =17) forecasts for US except production, which use May($i$ =1) through March($i$ =11) forecasts. ROW and China regressions use data for July($i$ =3) through September($i$ =17) forecasts except China production, which use July($i$ =3) through July($i$ =15) forecasts. Standard errors are calculated using White cross-section correction. Single, double, and triple asterisks (*, **, ***) denote statistical significance at 10%, 5%, and 1% levels, respectively.

**Table 3. Test of Efficiency with respect to Forecast Levels for WASDE Cotton Forecasts, 1985/86-2009/10 Marketing Years.**

| | Production | Imports | Domestic Use | Exports | Ending Stocks | Average Price |
|---|---|---|---|---|---|---|
| **U.S.** | | | | | | |
| C | 0.67 | | -1.17 | 1.98 | -3.45 | -1.83 |
| Forecast Level | 0.02 | | -0.33 *** | -0.02 | -0.20 ** | 0.01 |
| Month | -0.16 | | 0.08 | -0.21 | 0.25 | 0.28 |
| Year | -0.05 | | -0.64 *** | 0.01 | 0.13 | -0.18 |
| N | 253 | | 391 | 391 | 391 | 391 |
| **China** | | | | | | |
| C | 1.69 | 5.35 | 1.32 | -31.31 ** | 6.61 * | |
| Forecast Level | -0.05 | 0.01 | -0.07 | -0.28 * | 0.18 * | |
| Month | -0.11 | -0.41 | -0.07 | 2.21 * | -1.07 * | |
| Year | 0.05 | -2.42 | 0.09 | 0.45 | 0.42 | |
| 2006 | 14.45 *** | | | | 21.23 *** | |
| N | 299 | 344 | 345 | 345 | 345 | |
| **ROW** | | | | | | |
| C | -0.12 | 0.73 | -0.13 | -0.14 | 2.70 ** | |
| Forecast Level | 0.13 *** | 0.04 | -0.03 | 0.12 *** | 0.08 | |
| Month | 0.15 | -0.02 | 0.08 | 0.16 | -0.31 | |
| Year | 0.01 | 0.07 | 0.04 | -0.04 | -0.21 | |
| N | 345 | 345 | 345 | 345 | 345 | |

Notes: Panel Least Squares regressions use data for May($i$ =1) through September($i$ =17) forecasts for US except production, which use May($i$ =1) through March($i$ =11) forecasts. ROW and China regressions use data for July($i$ =3) through September($i$ =17) forecasts except China production, which use July($i$ =3) through July($i$ =15) forecasts. Standard errors are calculated using White cross-section correction. Single, double, and triple asterisks (*, **, ***) denote statistical significance at 10%, 5%, and 1% levels, respectively.

**Table 4. Test of Efficiency with respect to Past Errors for WASDE Cotton Forecasts, 1985/86-2009/10 Marketing Years.**

| | Production | Imports | Domestic Use | Exports | Ending Stocks | Average Price |
|---|---|---|---|---|---|---|
| **U.S.** | | | | | | |
| C | 0.57 | | -0.12 | 1.94 | -3.47 | -1.99 * |
| Lagged error | 0.16 | | -0.06 | 0.02 | 0.09 | -0.10 |
| Month | -0.14 | | 0.12 | -0.21 | 0.30 | 0.31 |
| Year | -0.08 | | -0.50 *** | 0.01 | 0.21 | -0.22 |
| N | 253 | | 391 | 391 | 391 | 391 |
| **China** | | | | | | |
| C | 1.66 | 4.90 | 0.99 | -24.86 ** | 6.48 | |
| Lagged error | 0.02 | 0.13 | 0.09 | 0.31 | -0.15 | |
| Month | -0.12 | -0.39 | -0.07 | 2.23 | -0.97 * | |
| Year | 0.06 | -2.09 * | 0.07 | 0.28 | 0.56 | |
| 2006 | 13.81 *** | | | | 22.63 *** | |
| N | 299 | 343 | 345 | 345 | 345 | |
| **ROW** | | | | | | |
| C | 0.00 | 0.80 | -0.16 | 0.14 | 2.49 ** | |
| Lagged error | 0.29 * | -0.08 | 0.02 | -0.18 * | 0.05 | |
| Month | 0.09 | -0.03 | 0.08 | 0.17 | -0.27 | |
| Year | -0.01 | 0.06 | 0.04 | -0.07 | -0.21 | |
| N | 345 | 345 | 345 | 345 | 345 | |

Notes: Panel Least Squares regressions use data for May($i$ =1) through September($i$ =17) forecasts for US except production, which use May($i$ =1) through March($i$ =11) forecasts. ROW and China regressions use data for July($i$ =3) through September($i$ =17) forecasts except China production, which use July($i$ =3) through July($i$ =15) forecasts. Standard errors are calculated using White cross-section correction. Single, double, and triple asterisks (*, **, ***) denote statistical significance at 10%, 5%, and 1% levels, respectively.

**Table 5. Test of Independence of Forecast Revisions for WASDE Cotton Forecasts, 1985/86-2009/10 Marketing Years.**

| | Production | Imports | Domestic Use | Exports | Ending Stocks | Average Price |
|---|---|---|---|---|---|---|
| **U.S.** | | | | | | |
| Lagged Revision | 0.26 *** | | 0.33 *** | 0.31 *** | 0.20 *** | 0.10 |
| Month | 0.00 | | 0.01 | 0.07 * | -0.10 | 0.04 |
| Year | 0.00 | | -0.04 *** | 0.03 | 0.05 | -0.04 |
| N | 225 | | 375 | 375 | 375 | 375 |
| **China** | | | | | | |
| Lagged Revision | 0.03 | 0.13 ** | 0.12 ** | 0.16 *** | -0.03 | |
| Month | 0.09 | -0.39 | 0.02 | -0.08 | 0.07 | |
| Year | 0.05 * | -0.36 ** | 0.00 | -0.05 | 0.13 ** | |
| N | 275 | 325 | 325 | 325 | 325 | |
| **ROW** | | | | | | |
| Lagged Revision | 0.32 *** | 0.09 | 0.28 *** | 0.15 | 0.09 | |
| Month | 0.02 | 0.06 ** | 0.03 *** | 0.04 * | -0.06 * | |
| Year | 0.00 | -0.01 | 0.00 | -0.02 | -0.01 | |
| N | 325 | 325 | 325 | 325 | 325 | |

Notes: Panel Least Squares regressions use data for May($i$ =1) through August($i$ =16) forecasts for US except production, which use May($i$ =1) through March($i$ =11) forecasts. ROW and China regressions use data for July($i$ =3) through August($i$ =16) forecasts except China production, which use July($i$ =3) through July($i$ =15) forecasts. Standard errors are calculated using White cross-section correction. Single, double, and triple asterisks (*, **, ***) denote statistical significance at 10%, 5%, and 1% levels, respectively.

**Table 6. Test of Forecast Improvement for WASDE Cotton Forecasts, 1985/86-2009/10 Marketing Years.**

| | Production | Imports | Domestic Use | Exports | Ending Stocks | Average Price |
|---|---|---|---|---|---|---|
| **U.S.** | | | | | | |
| C | 4.67 *** | | 4.77 *** | 9.27 *** | 17.61 *** | 5.69 *** |
| Month | -1.11 *** | | -0.52 *** | -1.22 *** | -2.09 *** | -0.83 *** |
| Year | 0.04 | | -0.02 | -0.11 | 0.25 | 0.16 ** |
| N | 275 | | 425 | 425 | 425 | 425 |
| **China** | | | | | | |
| C | 5.30 *** | 34.26 *** | 4.13 *** | 54.32 *** | 18.90 *** | |
| Month | -0.95 *** | -4.53 *** | -0.38 *** | -4.53 *** | -1.91 *** | |
| Year | -0.23 *** | -0.80 | 0.09 | 0.62 | -0.75 ** | |
| 2006 | 12.13 *** | | | | 15.89 *** | |
| N | 325 | | | | 375 | |
| **ROW** | | | | | | |
| C | 2.34 *** | 3.46 *** | 1.60 *** | 4.60 *** | 6.25 *** | |
| Month | -0.31 *** | -0.32 *** | -0.20 *** | -0.55 *** | -0.54 *** | |
| Year | 0.02 | -0.09 | -0.01 | 0.00 | 0.07 | |
| N | 375 | 375 | 375 | 375 | 375 | |

Notes: Panel Least Squares regressions use data for May($i$ =1) through September($i$ =17) forecasts for US except production, which use May($i$ =1) through March($i$ =11) forecasts. ROW and China regressions use data for July($i$ =3) through September($i$ =17) forecasts except China production, which use July($i$ =3) through July($i$ =15) forecasts. Standard errors are calculated using White cross-section correction. Single, double, and triple asterisks (*, **, ***) denote statistical significance at 10%, 5%, and 1% levels, respectively.

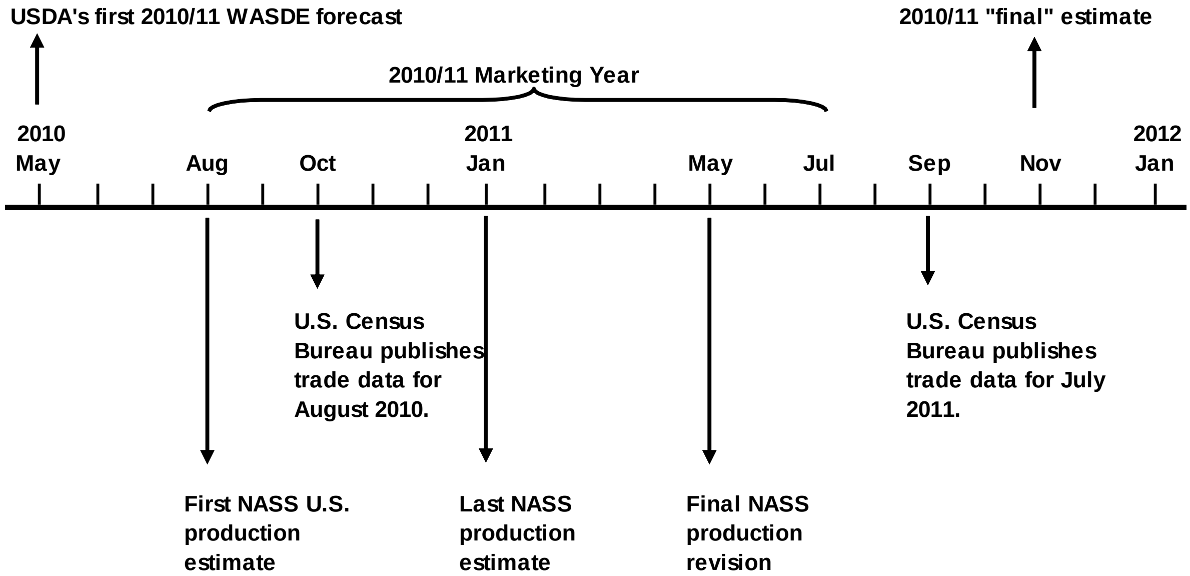
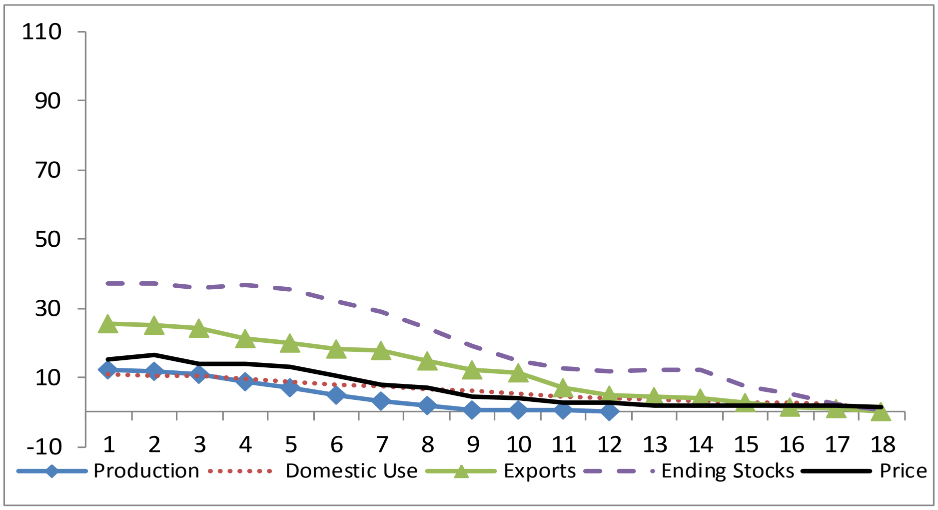
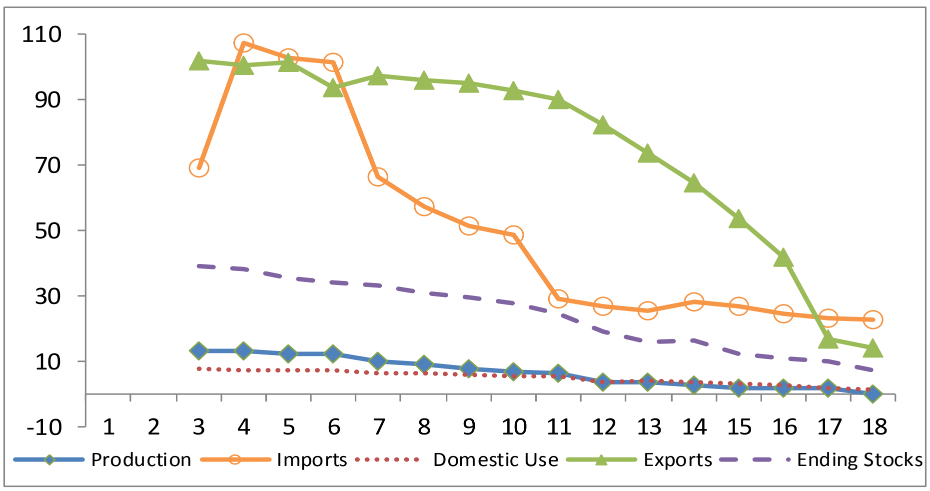



**Figure 1. WASDE Forecasting Cycle for Cotton Relative to the 2010/11 U.S. Marketing Year**

**Panel 1: U.S.**




**Panel 2: China**




**Panel 3: Rest of the World**

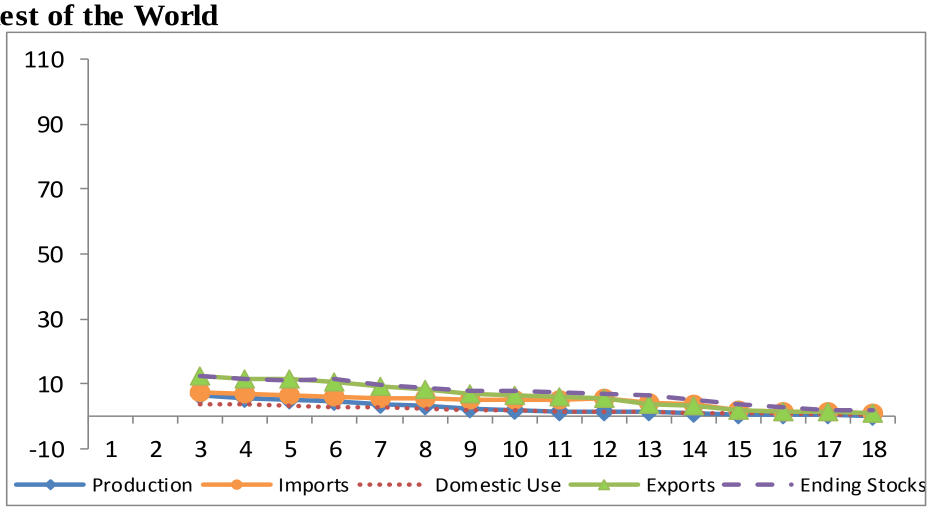
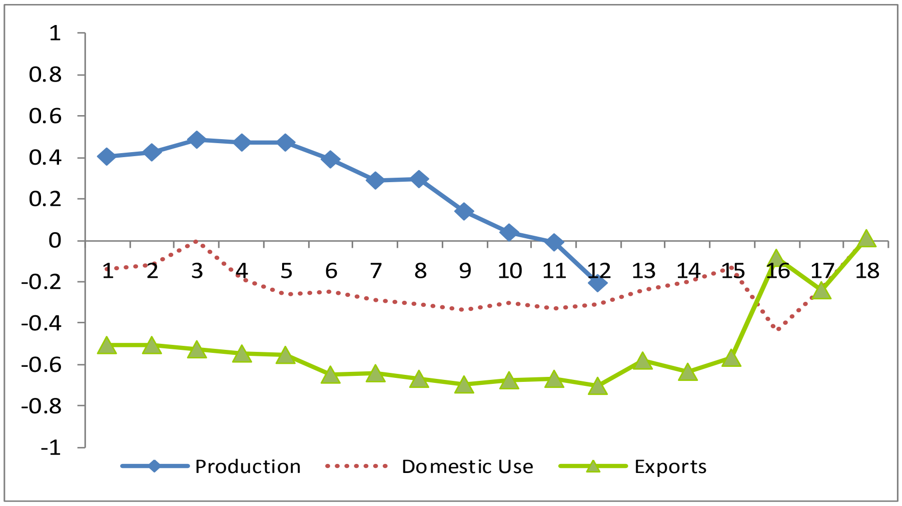
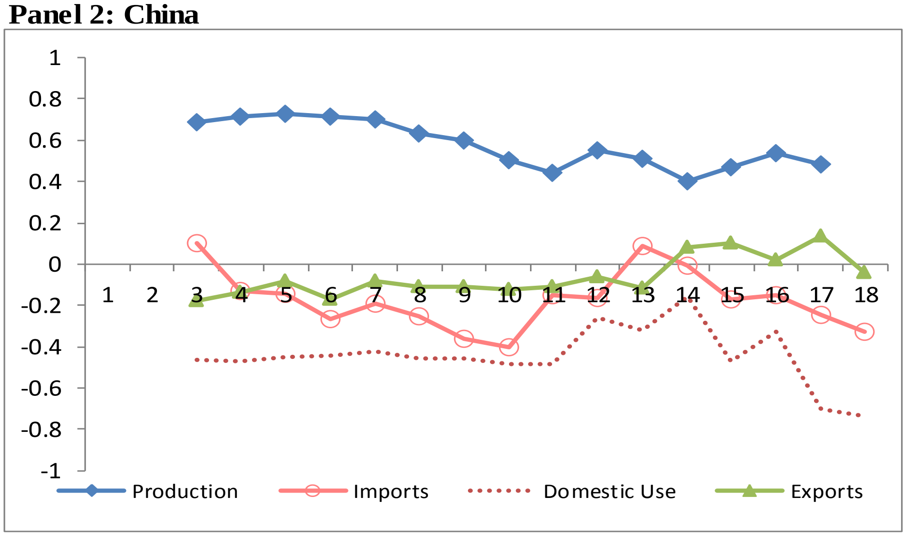



**Figure 2. Changes in the Standard Deviation of Percent Forecast Errors across the 19 months of the forecasting cycle.**

**Panel 1: U.S.**




**Panel 2: China**




**Panel 3: Rest of the World**

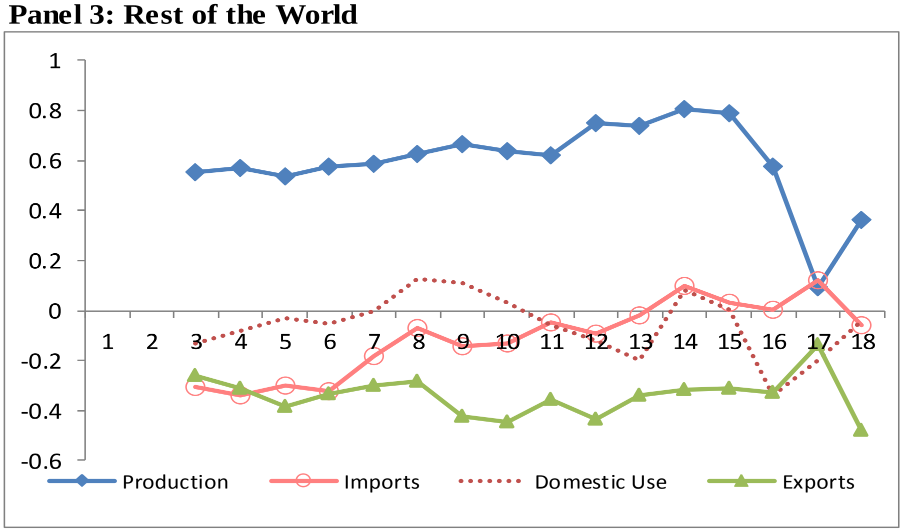
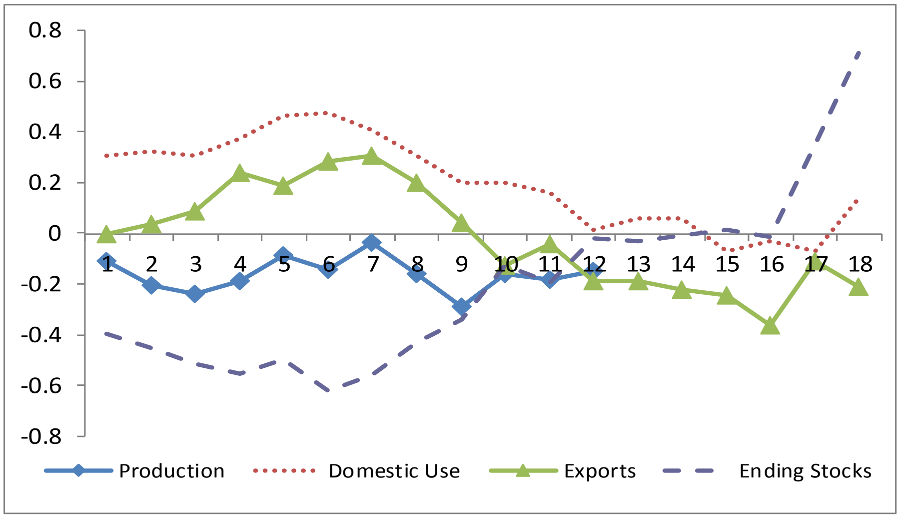
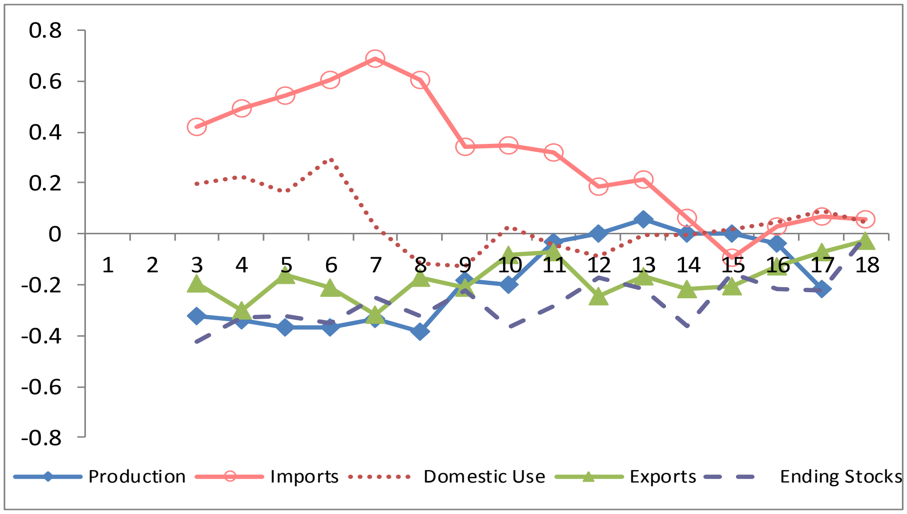



**Figure 3. Correlation of Percent Errors in Ending Stocks Forecasts with Percent Errors in Other Forecasts.**

**Panel 1: U.S.**




**Panel 2: China**




**Panel 3: Rest of the World**

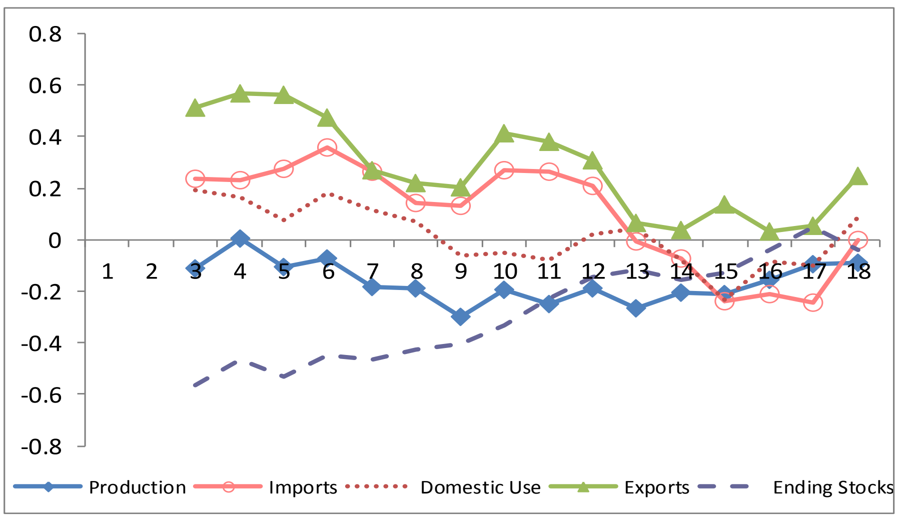



**Figure 4. Correlation of Percent Errors in U.S. Price Forecasts with Percent Errors in Other Forecasts.**