A Comparison of USDA's Agricultural Export Forecasts with ARIMA-based Forecasts

MacDonald, Stephen

U.S. Department of Agriculture, Economic Research Service

2005

A Comparison of USDA's Agricultural Export Forecasts with ARIMA-based Forecasts

Stephen MacDonald
U.S. Department of Agriculture, Economic Research Service

This study indicates that a number of USDA forecasts lack information that is readily available from monthly U.S. export data. This is determined by comparing the accuracy USDA's FY2001-04 forecasts with forecasts based on trends for each commodity. ARIMA models utilizing the monthly data available at the time each USDA forecast was published were estimated. Out of 24 separate commodity forecasts examined, USDA forecasts were superior to ARIMA forecast in only 9 cases. ARIMA forecasts were superior in 11 cases, and there was no difference in 4 cases.

Introduction

U.S. agricultural export forecasts are one subset of the voluminous information USDA provides on agriculture. From one perspective, these forecasts are but one small facet of a broad, integrated program of analysis. In addition to indicating developments in U.S. exports, published trade forecasts serve as useful indicators of USDA's perspective on current developments in global commodity markets. Similarly, the process of developing these forecasts may have positive externalities for other USDA priorities, both analytical and with respect to policy.

From another perspective, forecasting U.S. trade might be considered a diversion of resources that USDA could apply directly to other priorities. Published trade forecasts are only useful if they contain information not already published elsewhere. If USDA's published forecasts are no more accurate than forecasts anyone could develop from already published data, then USDA could increase public welfare by focusing on other priorities.

A balanced view of USDA's trade forecasting may lie between these two perspectives. Only a small percentage of the resources USDA devotes to U.S. export forecasting are devoted exclusively to this process, so the gains from eliminating the task may be small. Also, knowledge about trade clearly strengthens USDA's efforts on other commodity topics. However, if USDA's forecasts are no more accurate than those of an easily updatable model based on trends in monthly data, then rationality suggests there may be circumstances when using the model is preferred to USDA's more extensive efforts. At the very least, it suggests that the output of the model should be added to the information set available to USDA forecasters.

Methods and Data

This study compares the accuracy of USDA's fiscal year export value forecasts for FY 2001-04 with forecasts based on trends in each commodity's monthly exports. USDA's forecasts are published quarterly in the *Outlook for U.S. Agricultural Trade*, The trend forecasts were produced with ARIMA models utilizing the monthly data available at the time each USDA forecast was published. The models were specified and estimated with the Tramo/Seats software developed by the Bank of Spain. This software was incorporated by Eurostat into a software package, Demetra, which was this study's interface for Tramo/Seats.

For a given fiscal year (October-September), USDA forecasts U.S. agricultural export value by commodity 5 times. The first forecast is published in August, before the fiscal year begins, and updates are published in November, February, May and the following August. As an illustration, Table 1 compares the ARIMA forecasts of U.S. cotton export value with those USDA published each November during FY 2001-04.

International agricultural trade is in a constant state of flux. Economic development around the world has induced significant structural change for consumers, producers, and traders. For economists, structural change necessitates newly specified or estimated models. For forecasters, structural change means reorienting toward new countries, to different segments of the supply chain, or to different portions of the marketing year. This study's efforts were confined to the last 4 years to limit the impact of inevitable changes in world markets on the validity of its conclusions.

Historical data for the ARIMA modeling was downloaded from the Foreign Agricultural Service's website. In order to ensure a sufficient number of observations for ARIMA modeling, June 2000 was the end point of the oldest data set, and August 2000 was the date of the earliest USDA forecast analyzed.

Results

Table 1 shows that USDA's cotton forecast error each November during FY2001-04 was smaller than what an ARIMA-based forecast would have produced. Note that the software chooses 3 different ARIMA specifications over the 4 years studied, perhaps indicative of market volatility that hinders the accuracy of ARIMA forecasting. The ARIMA model's error in FY 2002 is extraordinarily large, and the ARIMA methodology's root mean squared error (RMSE) during these 4 years is substantially higher than USDA's as a result. The ratio between the RMSE of the ARIMA methodology and the RMSE of USDA's forecasts is 3.2 / 0.2 = 13.7. This was the highest ratio for any commodity for any of the forecast update months (Table 2). In Table 2, ARIMA modeling is less accurate than USDA if the ratio is greater than 1. This ratio provides a simple indication of relative performance.

ARIMA forecasts of aggregate commodities, like Grains and feeds, are the sum of forecasts by ARIMA models for each component of the aggregate. This includes a forecast of residuals for aggregate groupings. USDA's published forecast for Total U.S. agricultural exports is essentially a sum of its published forecasts of each component of agricultural trade. However, USDA's forecast of Grains and feeds exports (for example) is larger than its published forecasts of specific categories of grains and feeds. Therefore, there is an implied forecast of the remaining products. Table 2 indicates that USDA's RMSE has been at least twice as large as the error that ARIMA-based forecasting would have realized for this grains and feeds residual.

While cotton and the residual category for grains and feeds have relative RMSEs that clearly indicate the dominance of ARIMA or USDA forecasting, these are atypical. For most forecasts, the ratio is much closer to 1.0. To formally compare the accuracy of the methodologies, two statistics were calculated. A general measure of forecast accuracy has been developed by Diebold and Marino, the Morgan-Granger-Newbold test (Table 3). A sign test was also used to determine if the frequency of a given forecast's greater relative accuracy was significant during 2001-04 (Table 4).

While some patterns are apparent in Tables 3 and 4, one further step was taken to make these patterns clearer. A summary statistic was created for each commodity. To create this statistic, USDA's forecast of each commodity was assigned a score based on its performance in the MGN and sign tests. For each update month for which USDA's 2001-04 performance was significantly better (at least 10 percent significance) than the ARIMA-methodology's for a given test, a score of 1 was assigned. If the significance of USDA's dominance was 1 percent or better, a score of 2 was assigned. On the other hand, when ARIMA forecasting dominated, the scores were -1 and -2, respectively. A commodity's composite score is the sum of its MGN and sign test scores over the 5 update months. Conceivably, a commodity's score could be as high as 20 or as low as -20.

Table 5 ranks the scores in ascending order, and negative scores are more common than positive scores. The scores range from 8 to -9. Soybeans, soybean meal, and cotton have the best scores, while a number of high-value products and rice have the worst. Interestingly, the accuracy of USDA's estimates for Horticultural products in total is lower than for virtually all the forecasts of the components of the total. This isn't the case for any of the other aggregates: Grains and feeds, Oilseeds and products, and Livestock and products.

These results have an important implication for USDA's forecasting of U.S. total agricultural export value. A combination of USDA and ARIMA forecasts is more accurate than either alone. A forecast of total U.S. agricultural export value can be created by using ARIMA forecasts for all commodities with a composite score of -3 or below and USDA forecasts for all other commodities. This combined forecast would have been more accurate than USDA's forecast in 3 out of the last 4 years (except for the initial August release, for which there was a tie in one year). This frequency of dominance is not statistically significant. However, the MGN test indicates the combined forecast was significantly more accurate (at the 1 percent or better level) in November and February over 2001-04. This comparison does not take into account USDA's practice of revising commodity forecasts to create a total export value

forecast that rounds to the nearest $500 million, but perhaps the rationale and implications of that practice bear examination.

Conclusions

Benchmarked against ARIMA-based forecasting, USDA's quarterly U.S. export value forecasts are more often dominated than dominant. As measured by a composite score, USDA forecasts were superior to ARIMA forecast in only 9 out of the 24 separate commodities examined. ARIMA forecasts were superior in 11 cases, and there was no difference in 4 cases. While it should be noted that ARIMA forecasts did not dominate a majority of the commodities, the onus is probably on USDA to dominate the ARIMA forecasts, which it has often failed to do.

USDA does not devote equivalent resources to each commodity's forecast. Some receive a great deal of attention, some very little. All USDA forecasts are approved by the World Agricultural Outlook Board (WAOB). Examination of the WAOB's publication, the *World Agricultural Supply and Demand Estimates* (*WASDE*), reveals different levels of detail for different commodities. Variations in the levels of detail correspond to variations in the intensity of USDA's forecasting efforts. Variations in USDA's accuracy also correspond to these variations in forecasting efforts.

Complete supply and demand estimates for the United States and other major producers, consumers, importers, and exporters comprise the greatest level of detail any commodity receives in the *WASDE*. The next level of detail is to provide supply and demand forecasts for only the United States. In each case, these forecasts are produced by interagency committees that meet monthly, reviewing developments in U.S. and world markets (see Vogel and Bange for discussion).

The commodities with the greatest level of detail in the *WASDE* can be grouped into a "high attention" category :

> Wheat, corn, rice, soybeans, soybean meal, soybean oil, and cotton.

The commodities with only U.S. supply and demand tables in the *WASDE* can be grouped into a "medium attention" category:

> Sugar, beef, pork, broilers, turkeys, eggs, and milk.

However, the majority of the 24 commodities examined in this study are not included in the *WASDE*. The interagency committees overseeing these forecasts meet less frequently, and the supply and demand estimates USDA provides for these commodities include only a small number of countries outside the United States (Table 6). These other commodities can be grouped into a "low attention" category.

Averaging the composite scores of commodities in the "high attention" category (8 forecasts) gives an average of 1.9, indicating the superiority of the USDA forecasts. The average of commodities in the "medium attention" category (3 forecasts) is -1.0, and the "low attention" category's average is -1.3 (13 forecasts).

The implication is that, for the majority of commodities included in USDA's quarterly *Outlook for U.S. Agricultural Trade*, publication of ARIMA-based forecasts of U.S. exports would be an improvement from previous efforts in terms of accuracy. These forecasts are primarily in the "low attention" category.

At the very least, the information embodied in ARIMA-based forecasts would make a useful contribution to USDA's analysis of these commodities. This also holds for some of the commodities already receiving a significant amount of forecasting resources. A forecast can only be considered rational if it embodies all information available when the forecast is developed. Software is now available that can readily provide this information, offering a viable opportunity to improve USDA's accuracy.

This study indicates that a number of USDA forecasts lack information that is readily available from monthly U.S. export data. The appropriate response to this challenge would vary by commodity and would be best implemented by specialists concentrating on these commodities. The "high attention" commodities, on average, have forecasts superior to ARIMA-based forecasts, appropriately enough. The advantages of adding such trend analysis to forecasters' information set are not immediately obvious. However, as circumstances change, it is appropriate to consider all options as any forecasting institution reviews its changing mix of priorities and resources.

References

Diebold, F.X, and R.S. Mariano (1995), Comparing predictive accuracy, *Journal of Business and Economic Statistics*, 13(3): 253-263.

DEMETRA 2.0, Service Pack 1 - Release (May 2002) http://forum.europa.eu.int/irc/dsis/eurosam/info/data/demetra.htm

U.S. Department of Agriculture, *Outlook for U.S. Agricultural Trade*, various issues 2000-2004.

U.S. Department of Agriculture, *World Agricultural Supply and Demand Estimates*, various issues.

Vogel, Frederic, and Gerald Bange (1999). *Understanding USDA Crop Forecasts*, Office of the Chief Economist, U.S. Department of Agriculture. Miscellaneous Publication No. 1554.