




RESEARCH ARTICLE

Futures-Based Forecasts of Cotton Prices: Beyond Historical Averages

Armine Poghosyan , Olga Isengildina-Massa  and Shamar L. Stewart 

Department of Agricultural and Applied Economics, College of Agriculture and Life Sciences, Virginia Tech, Blacksburg, VA, USA

Corresponding author: Armine Poghosyan; Email: armine@vt.edu

Abstract

This study explores several alternative specifications of futures-based forecasting models to improve existing approaches constrained by restrictive assumptions and limited information sets. In lieu of historical averages, our approaches use rolling regressions and include current market information reflected in the deviation of the current basis from its historical average. To mitigate potential challenges arising from nonstationarity and structural changes in the relationship between farm and futures prices, we employ a 5-year rolling estimation window. We find that a rolling regression approach offers significant improvements (as evidenced by our Modified Diebold–Mariano test) in the accuracy and information content of forecasts of cotton season-average prices (SAPs) mostly at short forecast horizons.

Keywords: Basis deviation; commodity price forecasting; forecast evaluation; futures; season-average price (SAP)

JEL classifications: C22; Q02; Q11

1. Introduction

Volatile market environments make commodity price forecasting a challenging and essential task for many economic agents. For companies, reliable price forecasts can help make informed long-term investment decisions and predict budgets and profits (Kwas and Rubaszek, 2021). Accurate price projections also significantly impact the macroeconomic performance of countries whose revenues are predominantly dependent on commodity exports (Bowman and Husain, 2004). Therefore, many agricultural economists and government agencies strive to provide reliable commodity price projections and continuously seek ways to enhance and improve these forecasts.

One of the most prominent examples of such forecasts is season-average price (SAP) projections for all major crops published by the US Department of Agriculture (USDA) in the monthly World Agricultural Supply and Demand Estimates (WASDE) reports. These reports are an essential information tool for producers, traders, stakeholders, and other government and private agencies to guide their production, investment, and trading decisions (Fiene, 2023; Zoller, 2021). For example, the May WASDE report is particularly important for farmers as it provides the first estimates/projections for the new marketing year's supply and demand, while the August and September WASDE reports provide the first production estimates and forecasts for the spring-planted crops (CME Group, 2024). Previous studies (Adjemian, 2012; Isengildina-Massa et al., 2008) have demonstrated that WASDE reports offer valuable insights to market participants, and the announcements related to these reports can have a lasting effect on market prices. Along with supply and demand estimates, WASDE reports also forecast the SAP (known as the

marketing year average price), which represents the average price received by US producers throughout the marketing year¹ for all grades and qualities of the crop. The SAP is a key parameter used to assess the financial health of the US farm sector and determine commodity program payments. For example, under the Agricultural Improvement Act of 2018, the SAP is a key input for calculating the Price Loss Coverage (PLC) and Agricultural Risk Coverage (ARC) payment rates for program commodities (Hoffman and Meyer, 2018).

WASDE SAP forecasts are a result of an extensive projection process that “involves the interaction of expert judgment, econometric price forecasting models, futures prices, market information, and in-depth research by USDA analysts on key domestic and international issues”.² Because of its potential impact on the market, the report is prepared under strict confidentiality, not known to any entities (except the analysts preparing the report) until its release time through a “lock-up” procedure (Hoffman et al., 2015). While most of the components of the WASDE SAP forecasts are not publicly available, the futures-based price estimates for corn, soybeans, wheat, and cotton are well documented and updated every month.³ These models serve as key inputs for WASDE SAP forecasts.

Numerous previous studies demonstrate that futures-based forecasts of commodity prices perform well relative to time-series and judgmental price forecasts, especially at shorter forecasting horizons (see Etienne et al., 2019; Figuerola-Ferretti et al., 2021; Hoffman et al., 2015; Just and Rausser, 1981; Kastens et al., 1998; Manfredo and Sanders, 2004; Rausser and Just, 1979; Roache and Reichsfeld, 2011, among others). The advantages of futures-based price forecasts relative to time-series alternatives stem from their forward-looking nature (Adjemian et al., 2020), as futures prices of commodities with liquid derivative markets reflect expectations regarding market conditions at contract maturities. The forward-pricing role of the futures markets for storable commodities goes back to the seminal work of Working (1948), which was later extended by Black (1976) and Peck (1985). These studies established that futures prices are an unbiased predictor of the cash price for a given delivery location and time period when the futures markets are efficient (Fama, 1970, 1991). Furthermore, some argued that futures-based forecasts would beat projections based on structural or econometric models using publicly available data (e.g., Just and Rausser, 1981; Tomek, 1997; Zulauf et al., 1999). However, Chavas et al. (1983) and McKenzie and Holt (2002) suggest that when government programs and risk premiums are factored into futures pricing, futures prices alone might not be a reliable predictor of cash prices. Hence, further enhancements in futures-based price projections are necessary.

Early SAP forecasting models used by USDA were developed by Hoffman (Hoffman et al., 2007, 2015; Hoffman, 2005; Hoffman and Meyer, 2018) and involved adding a 5-year moving average of basis (the difference between cash and futures prices) to the futures price. Hoffman et al. (2007, 2015) found that reducing historical context to 5-year moving averages addresses the impact of structural changes and produces a reliable basis forecast for corn, soybeans, and wheat. To increase the precision of basis forecast for cotton, Hoffman and Meyer (2018) used several specifications of simple (e.g., 5-year) and Olympic moving averages and found that a 7-year Olympic average (which excludes the maximum and minimum of the selected seven values) for basis forecast leads to more precise projections of cotton prices. More recently Etienne et al. (2019, 2023) developed an alternative futures-based procedure to forecast the SAP for the US corn. The authors employed regression models of corn cash prices on the nearby futures contract prices and the previous month's cash prices. They used projected marketing weights for each month to

¹A marketing year is a period of 1 year designated for reporting and/or analysis of production, marketing, and disposition of a commodity.

²Further information on how WASDE is prepared and organized can be found here: <https://www.usda.gov/oce/commodity/wasde>.

³Further information on these SAP forecasts can be found here: <https://www.ers.usda.gov/data-products/season-average-price-forecasts/>.

generate SAP forecasts but were able to achieve only marginal accuracy improvements relative to Hoffman *et al.* (2007) approach for most of the forecast periods considered.

Given the importance of the SAP forecasts and the limitations of the currently available models, our study aims to explore several alternative specifications that may enhance the accuracy of the SAP projections for cotton. Our focus on cotton prices is motivated by the fact that it is significantly under-researched (relative to corn and soybeans) despite being one of the most critical and volatile global commodities. Beyond producing textiles, cotton is also used in animal feed, packaging, and biofuels (Makhula, 2021). Projecting cotton prices is particularly complex as they are sensitive to natural and political factors (Balseca and Bennion, 2011; Janzen *et al.*, 2018; Shoulberg, 2021). However, between 1929 and 2008, the USDA was prohibited from publishing cotton prices (Townsend, 1989). This long-standing prohibition began following congressional hearings that investigated a September 1927 USDA cotton price forecast that was believed to have triggered a market selloff. During the prohibition period, the SAP forecasts of cotton prices were produced for internal use only and were not included in WASDE reports. Therefore, it is not surprising that little is known about SAP cotton forecasts.

In this study, we develop several alternative model specifications for SAP forecasts for US cotton. One potential area of improvement of Hoffman's method is a more flexible adjustment for structural change using a rolling regression approach. Additionally, several studies demonstrated that the accuracy of basis forecasts can be improved by incorporating current market information. One such measure is the difference between the current basis and its historic average (see Hayenga and Jiang, 1997; Liu *et al.*, 1994; Taylor *et al.*, 2006; Tomek, 1980; Tonsor *et al.*, 2004, for instance). Similarly, we examine whether incorporating current market information (via a *basis deviation* term) improves the accuracy of futures-based forecasts of cotton prices. Therefore, the first model modifies Hoffman's benchmark by including a "basis deviation term" (computed as the difference between the current basis and its historical average) to capture the new market information. The second and third models utilize time-series regression models of cotton farm prices on futures prices with and without the basis deviation term.

To assess potential improvements, we evaluate each model's out-of-sample performance relative to Hoffman and Meyer (2018) benchmark model and the WASDE forecasts. Since the WASDE SAP forecasts are available from 2008 onward, our evaluation period is 2008–2023. We examine the projections using standard measures of accuracy such as mean absolute percentage error (MAPE) and root mean square percentage error (RMSPE). We also investigate systematic biases (the tendency to consistently over- or under-estimate the actual prices) of our forecasts using a two-tailed *t*-test on the mean percentage errors (MPE). We find over-prediction in Hoffman's benchmark model for all forecast horizons and under-prediction in the regression-based models at short-to-mid (3–6 months) horizons. At the same time, we find that the proposed regression-based forecasting approaches perform well relative to Hoffman's benchmark approach and produce forecasts with lower errors.

Next, using the Modified Diebold–Mariano (MDM) test, we explore whether the forecasts from alternative model specifications, Hoffman's benchmark approach, and WASDE projections significantly differ from each other. The results show that the regression model without the basis deviation term yields statistically similar results to that of WASDE projection for the short and long forecast horizons while producing significantly more accurate results at the short-term (5–7 months) projection periods when compared to Hoffman's approach.

Finally, we test the predictive content of the alternative forecasts using an encompassing test and find that our proposed model forecasts contain additional information missing in Hoffman's benchmark model at every forecast horizon. Our proposed model forecasts also contain incremental information missing from WASDE forecasts both at the short (3–7 months) and the long-term (12–14 months) projection horizons. Therefore, the projected cotton prices from our proposed regression model are less biased, more accurate, and informative relative to the currently used procedures.

2. Data

By definition, SAP is an average of monthly farm prices weighted by historical monthly marketings. These data are collected by the USDA's National Agricultural Statistics Service (NASS) via a monthly survey of US cotton buyers (merchandisers, mills, and others) to obtain information on the quantity and price of upland cotton purchased directly from US farmers during a given month. The monthly farm price is derived by dividing the total cost (purchase price times quantity) by the total quantity purchased (farm sales or marketings) (USDA-NASS, 2011). Monthly upland cotton prices received by US producers are published in Agricultural Prices reports and available from the Quick Stats database with a 2-month lag. For each marketing year (August–July), this information is finalized by October after the conclusion of the marketing year, when the final cotton SAP received by farmers is published. USDA projections of cotton SAP, similar to other commodities included in WASDE forecasts, begin in May, prior to the start of the marketing year. Table 1 illustrates the SAP forecasting cycle for upland cotton relative to marketing year and crop seasons. Our focus in this study is on the forecasts released during the marketing year, as highlighted in Table 1, as they are considered most relevant for decision-making. While the official USDA SAP forecasts published in the WASDE reports described above are based on multiple sources, Etienne et al. (2023) and Hoffman and Meyer (2018) show that futures-based forecasts rely on futures prices for the nearby contract months, past monthly and season-average farm prices, and past monthly marketing weights (see Figure 1).

Table 1. Marketing year, crop season, and forecasting cycle for 2016/17 cotton season-average price

| Month/year | Crop season | Marketing year | WASDE forecast cycle | Months ahead | Futures contract maturity |
|------------|--|----------------|----------------------|--------------|---------------------------|
| May 2016 | Growing | | Forecast 1 | 17 | |
| Jun 2016 | | | Forecast 2 | 16 | |
| Jul 2016 | | | Forecast 3 | 15 | |
| Aug 2016 | Harvest | Month 1 | Forecast 4 | 14 | Oct 2016 |
| Sept 2016 | | Month 2 | Forecast 5 | 13 | Oct 2016 |
| Oct 2016 | | Month 3 | Forecast 6 | 12 | Dec 2016 |
| Nov 2016 | | Month 4 | Forecast 7 | 11 | Dec 2016 |
| Dec 2016 | | Month 5 | Forecast 8 | 10 | Mar 2017 |
| Jan 2017 | | Month 6 | Forecast 9 | 9 | Mar 2017 |
| Feb 2017 | Post-harvest | Month 7 | Forecast 10 | 8 | Mar 2017 |
| Mar 2017 | | Month 8 | Forecast 11 | 7 | May 2017 |
| Apr 2017 | | Month 9 | Forecast 12 | 6 | May 2017 |
| May 2017 | | Month 10 | Forecast 13 | 5 | Jul 2017 |
| Jun 2017 | | Month 11 | Forecast 14 | 4 | Jul 2017 |
| Jul 2017 | Month 12 | Forecast 15 | 3 | Oct 2017 | |
| Aug 2017 | | | Forecast 16 | 2 | |
| Sept 2017 | | | Forecast 17 | 1 | |
| Oct 2017 | Season-average price published for marketing year 2016/17 | | | | |

This table, adapted from Etienne et al. (2023), illustrates the WASDE SAP forecasting cycle, marketing year, and crop season. WASDE begins the SAP forecast three months before the start of the marketing year and continues it until the actual SAP publication date in October. Our analysis focuses on the marketing year, highlighted in the rectangular box in this table.

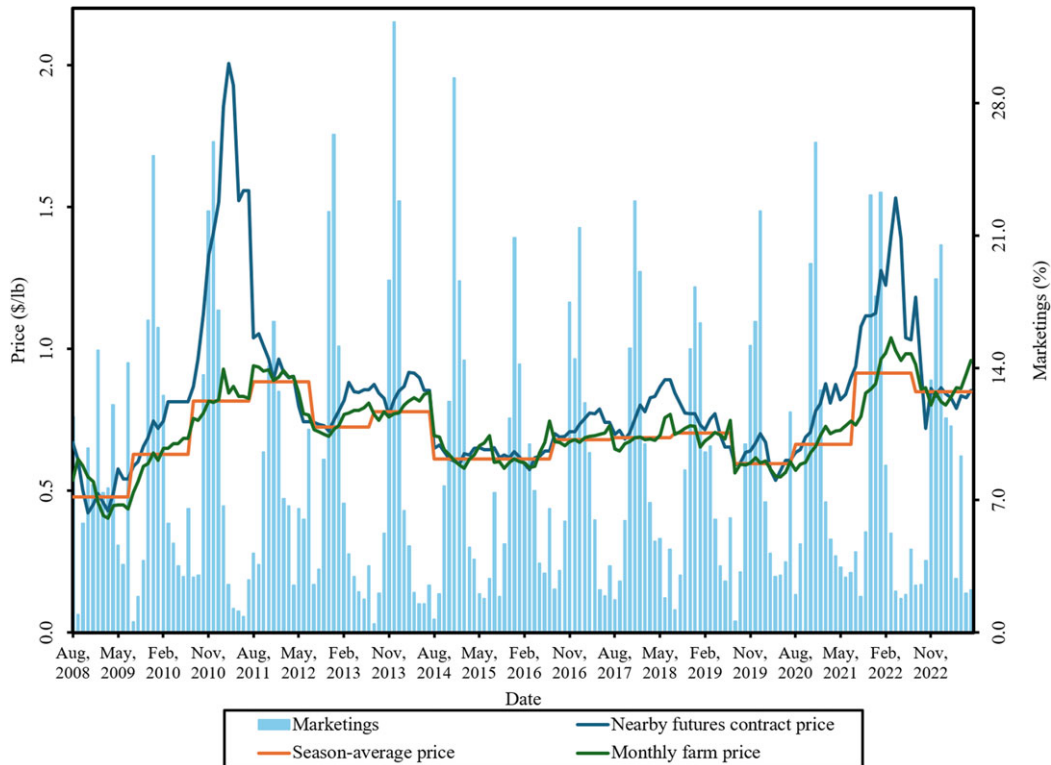


Figure 1. Overview of actual cotton prices, nearby futures contract prices, season-average prices, and marketings for the August 2008–December 2023 period.

Note: This figure closely resembles the one by Hoffman and Meyer (2018). It presents the essential components used in developing the season-average price projections for the Hoffman model. The variable marketings (presented on the secondary axis) represent the historical marketing weights. The nearby futures contract prices represent the prices of the futures contracts closest to expiration (as detailed in Table 1). Season-average prices are computed as averages of monthly farm prices weighted by the marketings. The monthly farm price is derived by dividing the total cost (purchase price times quantity) by the total quantity purchased (farm sales or marketings).

We collected monthly average farm prices received (hereafter referred to as *farm prices*) for cotton from January 2000 to December 2023 from the Quick Stats Database. Monthly average closing futures prices for the Intercontinental Exchange (ICE) cotton contract were obtained from Quandl. The sixth column of Table 1 shows the selection of the nearby futures contract for each spot price month, k . Since these futures contracts expire within the first two weeks of the contract expiration month, we switch to the next-to-expire futures contract at the start of the expiration month.

Changes in the monthly upland US cotton prices between 2000 and 2023 demonstrate significant volatility during the first decade of the 2000s. Much of these observed variations in cotton prices resulted from changes in international trade and China's growing demand for boosting its textile industry (Isengildina-Massa and MacDonald, 2009). For example, in 2011, cotton prices reached their historical high due to net supply shortage (Janzen *et al.*, 2018, Robinson, 2011a; 2011b). Unfavorable weather conditions in the major cotton-growing countries such as the United States, China, Pakistan, and India, combined with a prolonged multi-year reduction in worldwide cotton reserves, sparked a wave of panic buying in the farm market, resulting in significant shortages across the global market (Robinson, 2011a; Shoulberg, 2021). Concurrently, the Indian government banned cotton exports, further limiting the world cotton supply (Balseca and Bennion, 2011). These unanticipated shocks in the cotton market make price

forecasting challenging as historical patterns and relationships do not always hold (Isengildina-Massa and MacDonald, 2013).

Panel A of Table 2 presents the descriptive statistics of the cotton farm and futures prices and the basis. The farm price of cotton ranged from 27 cents to one dollar, with an average of 63 cents per pound. While the average price of the futures contract was comparable to the farm prices, the futures prices exhibited more considerable variation. The normalized results from the coefficient of variation reveal that the futures markets exhibited, on average, 10 percentage points higher variability than the cash markets. These wide price gaps and higher volatility in the futures market often explain the dispersion dynamics and swings in computed basis— the difference between farm and nearby futures prices.

Table 2. Data description

| Panel A. Summary statistics | | | |
|---------------------------------|--------------------|-----------------------------|-----------|
| Statistic | Farm price (\$/lb) | Futures price in spot month | Basis |
| Mean | 0.629 | 0.718 | -0.088 |
| Standard deviation | 0.165 | 0.253 | 0.152 |
| Minimum | 0.267 | 0.314 | -1.163 |
| Maximum | 1.040 | 2.007 | 0.144 |
| Coefficient of variation (%) | 26.171 | 35.279 | -172.561 |
| Panel B. Stationarity analysis | | | |
| Stationarity test | Farm price (\$/lb) | Futures price in spot month | Basis |
| Augmented Dickey-Fuller | -2.104 | -2.432 | -4.176*** |
| Augmented Dickey-Fuller (trend) | -2.716 | -2.701 | -4.170*** |
| Augmented Dickey-Fuller (drift) | -2.104** | -2.432*** | -4.176*** |
| Phillips-Perron | -2.258 | -3.010** | -4.593*** |
| Phillips-Perron (trend) | -3.012 | -3.384* | -4.588*** |

This table provides descriptive statistics for farm prices, futures prices, and the basis for cotton, where the basis represents the difference between farm and futures prices (Farm prices minus Futures prices). The total number of observations in the full dataset is 288. The coefficient of variation, given in percentages, is the ratio of the standard deviation to the mean. The p-values of the null hypothesis of nonstationarity are shown in parentheses. Statistical significance is indicated with asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

When developing futures-based forecasts, the most common approach (e.g., Hoffman et al., 2007, 2015; Hoffman, 2005; Hoffman and Meyer, 2018) is to use the basis instead of the price level as it is believed to be more stable and, consequently, much easier to predict. However, this assumption does not always hold, as evidenced in Figure 2 and our discussion above regarding Panel A of Table 2. To formally test this hypothesis, we use the Bai-Perron test to check for the presence of structural breaks in the basis. The advantage of the Bai-Perron test is its ability to identify multiple unknown breaks. Consistent with visual assessments of Figure 2, the Bai-Perron test suggests December 2008 and March 2012 as potential breakpoints in the basis series. We also tested farm prices for the presence of structural breaks. We found breakpoints in March 2003, April 2007, July 2010, and July 2014, as shown in Figure 2.⁴ These findings indicate that structural

⁴We also formally tested our farm and futures price series for the presence of seasonality and did not detect any seasonal patterns.

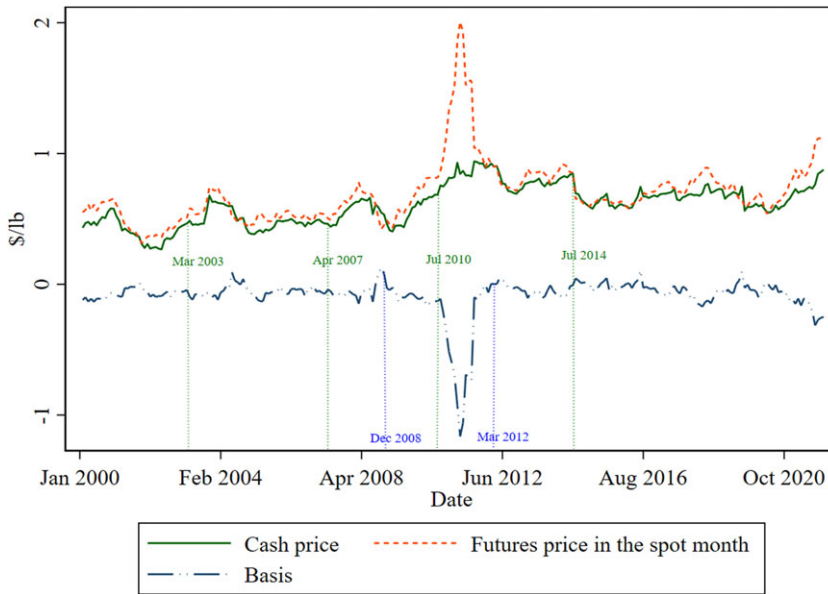


Figure 2. Farm prices, futures prices, and basis, January 2000–December 2023.

Note: This figure presents the time plot of the spot and futures price of cotton as well as the basis. Our basis series is computed as the difference between the spot price and the price of the nearby futures contract. The dates highlighted on the graph represent the breakpoints detected in farm prices (presented in green) and basis (presented in blue) according to the Bai–Perron test for unknown structural breaks.

changes in the basis have to be taken into account when using these futures-based forecasting models.

In the presence of structural breaks, unit-root tests have to be interpreted with caution. Panel B of Table 2 demonstrates that cotton farm and futures prices (but not basis) were nonstationary using traditional unit-root tests. On the other hand, the results of the Johansen and Engel–Granger tests show that the spot and futures prices were cointegrated. This suggests that both price series comove and the difference between them is stable in the long run (which is confirmed by the stationarity of basis). Cointegration allows us to analyze the relationship between spot and futures prices without the need for differencing to render both price series stationary. Hence, the main issue that needs to be considered for developing a reliable forecasting model is the presence of structural breaks. Potential ways of addressing the issue of structural breaks include using Olympic moving averages (for example, Hoffman and Meyer, 2018), using rolling regressions (discussed in Section 3.1), and/or incorporating current market information using a basis deviation term (discussed in Section 3.2).

3. Forecasting approaches

SAP projections are calculated using monthly price forecasts, observed monthly farm prices (when they become available), and historical monthly marketing weights (Hoffman and Meyer, 2018). Thus, the biggest challenge is in forecasting monthly farm prices, which is the main objective of our paper. The underlying premise of futures-based forecasting methods is that futures prices capture the expectations of cash market conditions at contract maturity. This notion is based on market efficiency, which stipulates that futures prices are unbiased predictors of farm prices at delivery location and time period (Fama, 1970, 1991) and by the convergence of futures and farm prices at contract expiration. Following Algeri and Kalkuhl (2014) and Bowman and Husain

(2004), the fundamental relationship between farm and futures prices is described as follows:

$$S_{k,t} = \beta_0 + \beta_1 \mathcal{F}_{h,t,i} + \varepsilon_{t,i}; \quad \varepsilon \sim N(0, \sigma) \tag{1}$$

where S is the farm price of cotton in calendar month k , year $t = 2000, \dots, 2023$; F is the price of the futures contract expiring in month h , as described in Table 1, and forecast horizon $i = 1, \dots, 12$; and ε is the error term.

The forecasts are made at various horizons, i , describing the forecast lead time relative to farm price, as illustrated in Table 1.

3.1. Benchmark model

The most current and widely used futures-based forecasting model for cotton was developed by Hoffman and Meyer (2018).⁵ This model uses a 7-year Olympic moving average (by excluding the maximum and minimum for computing the average) of nearby basis to overcome nonstationarity and structural break. Thus, a relationship in Equation 1 is simplified to:

$$\widehat{S}_{k,t} = \overline{\mathcal{B}}_{k,t} + \mathcal{F}_{h,t,i} \quad t = 2000, \dots, 2023 \quad i = 1, \dots, 12 \quad k = \text{Jan}, \dots, \text{Dec} \tag{2}$$

$$\overline{\mathcal{B}}_{k,t} = \frac{1}{5} \left[\sum_{j=1}^7 \mathcal{B}_{k,(t-12;j)} - \max \left(\sum_{j=1}^7 \mathcal{B}_{k,(t-12;j)} \right) - \min \left(\sum_{j=1}^7 \mathcal{B}_{k,(t-12;j)} \right) \right]$$

where \widehat{S} is the predicted spot price, $\overline{\mathcal{B}}_{k,t}$ is the predicted basis, calculated as a 7-year Olympic average of nearby basis, ($\mathcal{B}_{k,t} = S_{k,t} - \mathcal{F}_{h,t,0}$), observed during spot month. This model assumes that β_0 from Equation 1 can be approximated by a short-term historical basis, and $\beta_1 = 1$. The benefit of this model is that it is simple, easy to use, and relies on readily available information. This model tends to perform well when its assumptions in Equation 1 are not violated, that is, β_0 (represented by $\overline{\mathcal{B}}_{k,t}$) is stable, and β_1 is close to 1. However, when the basis deviates from its historical average (as shown in Figure 2) and/or β_1 diverges from 1, this method becomes less accurate.

3.2. Alternative model 1: Moving average model with basis deviation term

Several previous studies (Etienne et al., 2023; Liu et al., 1994; Taylor et al., 2006; Tomek, 1980; Tonsor et al., 2004, among others) showed that spot price projection can be enhanced by incorporating current market information into the forecasts. For example, Taylor et al. (2006) demonstrated that the basis deviation from its 5-year moving average may be helpful to account for variability left in the residuals from the benchmark model. The basis deviation term is calculated as a difference between the current month’s nearby basis and its historical moving average:

$$BDEV_{k,t} = \mathcal{B}_{k,t} - \overline{\mathcal{B}}_{k,t} \quad t = 2000, \dots, 2023 \quad k = \text{Jan}, \dots, \text{Dec} \tag{3}$$

where $\mathcal{B}_{k,t}$ is the nearby basis, and $\overline{\mathcal{B}}_{k,t}$ is the predicted basis, calculated as a 7-year Olympic average of nearby basis observed during spot month (refer to Equation 2). For example, in December 2000, the current basis includes the December 2000 farm price and the futures price of the nearest-to-maturity March 2001 contract, and a 7-year Olympic moving average is calculated as an average of these differences over the last seven observations as per Equation 2.

Our first modification is to include the basis deviation term directly into the benchmark model as follows:

⁵Similar models for corn, soybeans, and wheat were developed by Hoffman (2005) and Hoffman et al. (2007).

$$\widehat{S}_{k,t} = \mathcal{F}_{h,t,i} + \bar{B}_{k-2,t} + BDEV_{k-2,t} \quad t = 2000, \dots, 2023 \quad i = 1, \dots, 12 \quad k = \text{Jan}, \dots, \text{Dec} \quad (4)$$

To incorporate the most recent market data into the model, in Equation 4, the \bar{B} and $BDEV$ terms are lagged for two months. This is done because the USDA reports farm prices with a 2-month lag (i.e., June prices are published in August).

3.3. Alternative model 2: OLS model with nearby contract prices

The next alternative specification utilizes rolling regression analysis to develop a forecast based on Equation 1. The main difference between this specification and the benchmark approach is the use of regression analysis instead of moving averages (similar to Etienne *et al.*, 2023). This relaxes the potentially restrictive assumptions related to the model coefficients, as discussed in Section 3.1.

Another notable difference between this specification and the benchmark model lies in the data used for basis prediction. In the benchmark approach, the measure of basis expectation does not change over various forecasting horizons as it is always calculated as a moving average of nearby basis during the spot month. In our approach, the estimated coefficient of the intercept, β_0 , represents the difference between the farm price and its nearby futures price at a selected horizon. For instance, a forecast of August 2016 spot prices ($k = 8$, $t = 2016$) generated in January 2016 uses the October 2016 futures contract prices observed in January 2016 (see Table 1). Here, the estimated coefficient of the intercept, β_0 , will represent the difference between the farm price and its nearby futures price observed 7 months prior (January 2016), when the forecast is made.

To develop spot price forecasts, we begin by estimating Equation 1 with a 5-year rolling window approach using the data starting in January 2000 to December 2004 ($n = 60$) to compute out-of-sample forecasts for January 2005. This method is similar to Hoffman's benchmark model technique for addressing the issues of non-stationarity and structural breaks by focusing on the short-time period (discussed in Section 3.1). Hence, our approach of relaxing these assumptions and allowing the coefficients to vary by moving to regression analysis may help improve these forecasts.

3.4. Alternative model 3: OLS model with basis deviation term

Another alternative specification combines the new information given by the basis deviation term and the regression approach in Equation 1, as discussed in the previous section, to capture situations when the basis becomes unstable. This leads to estimating the following regression:

$$S_{k,t} = \beta_0 + \beta_1 \mathcal{F}_{h,t,i} + \beta_2 BDEV_{k-2,t} + v_{t,i} \quad t = 2000, \dots, 2023 \quad i = 1, \dots, 12 \quad k = \text{Jan}, \dots, \text{Dec} \quad (5)$$

This model is used to generate out-of-sample forecasts of farm prices using a 5-year rolling window procedure described in Section 3.3.

3.5. SAP forecast

After forecasting monthly cotton prices using the alternative model specifications, we combine those forecasts with actual farm prices received (when available) to generate projections for the SAP. We follow the procedure outlined in Hoffman and Meyer (2018) to use the monthly forecast of farm prices, the actual realized spot prices, and marketing year weights assigned to each month within the marketing year to generate SAP forecasts (see Table 3).

To illustrate this process, consider a scenario where we need to generate the SAP forecast for the 2016/17 marketing year. Our forecasting process begins in August 2016, before any actual prices for that marketing year are available. The USDA publishes the first observed monthly farm price with a 2-month lag. Thus, the August 2016 price will not be released until October 2016. As a result, the initial SAP forecast for 2016/17, generated in August 2016, must rely entirely on futures

Table 3. Season-average price forecasts by forecast months and marketing year months

| WASDE forecast month | Marketing year months | | | | | | | | | | | |
|----------------------|-----------------------|----------|---------|---------|---------|---------|---------|---------|---------|----------|----------|----------|
| | Aug (1) | Sept (2) | Oct (3) | Nov (4) | Dec (5) | Jan (6) | Feb (7) | Mar (8) | Apr (9) | May (10) | Jun (11) | Jul (12) |
| August (4) | | | | | | | | | | | | |
| September (5) | | | | | | | | | | | | |
| October (6) | | | | | | | | | | | | |
| November (7) | | | | | | | | | | | | |
| December (8) | | | | | | | | | | | | |
| January (9) | | | | | | | | | | | | |
| February (10) | | | | | | | | | | | | |
| March (11) | | | | | | | | | | | | |
| April (12) | | | | | | | | | | | | |
| May (13) | | | | | | | | | | | | |
| June (14) | | | | | | | | | | | | |
| July (15) | | | | | | | | | | | | |

This table, adapted from Hoffman and Meyer (2018), presents the forecasting procedure we follow to generate SAP. As discussed in Section 3.5, the SAP forecast for August and September months fully relies on futures prices since the USDA publishes the first observed monthly farm prices with a 2-month lag. Once the actual August 2016 prices become available in October 2016, we generate a composite forecast by combining the observed cash prices with futures-based forecasts for the remaining ten months of the marketing year. From October 2016 onward, as each subsequent month's actual price is published, we replace the corresponding futures-based forecast with the realized data point, thereby enhancing the accuracy of our SAP projection.

prices using either the Hoffman or an alternative model specification. The September 2016 forecast will also depend solely on futures prices.

Once the actual August 2016 farm price becomes available in October 2016, we generate a composite forecast by combining this observed farm price with futures-based forecasts for the remaining ten months of the marketing year. From October 2016 onward, as each subsequent month's actual price is published, we replace the corresponding futures-based forecast with the realized data point, thereby enhancing the accuracy of our SAP projection. This procedure is schematically presented in Table 3.

To obtain a weighted forecast for each month, we use the 7-year Olympic average of marketing weights for that particular month within the marketing year.⁶ Then, we multiply the forecasted prices by these assigned weights and compute the SAP by summing the weighted price projections.

4. Model selection procedures

Since these models are designed for forecasting, we select the best model based on its out-of-sample forecasting performance in three steps. First, we assess the accuracy and potential bias of each price forecast across different projection horizons. Second, we determine whether the differences in predictive accuracy of the benchmark and alternative models are statistically

⁶Following Hoffman and Meyer (2018), we use the 7-year Olympic average of marketing weights, obtained from the USDA (<https://www.ers.usda.gov/data-products/season-average-price-forecasts/>). By maintaining consistent weights both for alternative model specifications and Hoffman approach, we ensure that any differences in forecast performance are attributable to the models themselves rather than variations in the marketing weights. This consistency supports the validity of our comparison and reinforces our findings.

significant. Finally, we compare the predictive content of price forecasts generated by the proposed alternative forecast approaches with that of the benchmark model using the encompassing test.

4.1. Forecast accuracy and bias

We evaluate the accuracy of out-of-sample forecasts using traditional evaluation criteria described in this section. First, following MacDonald and Isengildina-Massa (2012) and Isengildina-Massa et al. (2021), we define the prediction error in percentage terms, denoted by $pe_{k,t}$:

$$pe_{k,t} = (\ln S_{k,t} - \ln \widehat{S}_{k,t}) * 100 \quad t = 2005, \dots, 2023 \quad k = \text{Jan}, \dots, \text{Dec} \quad (6)$$

where $\ln S$ and $\ln \widehat{S}$ are the logarithmic transformations of the actual and predicted spot prices, respectively.

We use the two most common measures of forecast error size – MAPE and root mean squared percentage error (RMSPE) defined as:

$$MAPE = \frac{1}{n} \sum_{t=1}^n |pe_{k,t}| \quad t = 2005, \dots, 2023 \quad k = \text{Jan}, \dots, \text{Dec} \quad (7)$$

$$RMSPE = \sqrt{\frac{1}{n} \sum_{t=1}^n pe_{k,t}^2} \quad t = 2005, \dots, 2023 \quad k = \text{Jan}, \dots, \text{Dec} \quad (8)$$

A significant advantage of using MAPE measure is that it is relatively unaffected by occasional large prediction errors. In contrast, RMSPE is scale-independent, which ensures consistent and reliable performance evaluation across different measurement scales (Hyndman et al., 2006). When comparing the performance of the alternative model with that of the benchmark model, smaller values of both MAPE and RMSPE measures indicate more accurate and precise projections.

While the MAPE and RMSPE measures describe the magnitude of forecast errors, they do not say much about the presence of bias or systematic errors in the generated price projections. Assuming that forecasters try to minimize a symmetric loss function, we assess the presence and direction of bias, using the MPE criterion:

$$MPE = \frac{1}{n} \sum_{t=1}^n pe_{k,t} \quad t = 2005, \dots, 2023 \quad k = \text{Jan}, \dots, \text{Dec} \quad (9)$$

We apply a two-tailed t -test to the MPE to assess the statistical significance of potential forecast biases. The null hypothesis of the t -test is that the average percentage error is not significantly different from 0. Rejection of the null hypothesis will indicate whether the forecasts under- or over-predict the observed spot prices. Negative and significant MPE values will indicate that the model over-predicts the actual price. In contrast, positive and significant MPE measures indicate an under-prediction of spot prices.⁷

⁷This approach of assessing the presence and direction of bias in the price projection is tangential to the regression-based method developed by Holden and Peel (1990). The Holden and Peel test statistics can be obtained by regressing the forecast errors on a constant, which accommodates testing whether the forecast errors are mean-zero. This approach is conceptually similar to conducting a t -test on MPE statistics. The major difference between these approaches is that the Holden and Peel test uses a regression model with heteroskedasticity and autocorrelation consistent standard errors (Newey and West, 1986). We find that the results and conclusions of the Holden and Peel test approach are consistent with our t -test procedure.

4.2. Test for differences between benchmark and alternative forecasts

Following previous studies (see Manfredo and Sanders, 2004; Colino et al., 2011, for instance), we test for statistically significant differences in the forecast accuracy of the benchmark and alternative forecasts across different months using the MDM test developed by Harvey et al. (1997):

$$MDM_k = \frac{\bar{d}}{\sqrt{\frac{1}{n}(\delta_0 + 2 \sum_{q=1}^{t-1} \delta_q)}} \cdot \sqrt{\frac{n + 1 - 2h + \frac{h(h-1)}{n}}{n}} \tag{10}$$

$$\delta_0 = \frac{1}{n} \sum_{t=1}^n (d_t - \bar{d})^2; \quad \delta_q = \frac{1}{n} \sum_{t=q+1}^n (d_t - \bar{d})(d_{t-q} - \bar{d}) \quad t = 2005, \dots, 2023$$

where h is the forecast horizon, $d_t = |pe_{1,k,t}| - |pe_{2,k,t}|$, $pe_{1,k,t}$ is the percentage error from the benchmark model, $pe_{2,k,t}$ is the percentage error from the competing/alternative model specification. \bar{d} is the average difference across the forecasts for each forecast period, δ_0 is the variance of d_t , while δ_q is the q -th order auto-covariance term.

The null hypothesis of the MDM test is $H_0 : \xi = \mathbf{0}$, where $\xi = E[V(pe_{1,k,t}) - V(pe_{2,k,t})]$, and $V(pe_{.k,t}) = |pe_{.k,t}|$ is a loss function when RMSPE is used as a measure of forecast accuracy (Franses, 2016; Hyndman et al., 2006). This tests whether the forecast performance of the benchmark and alternative model specifications are significantly different from each other. Rejection of the null hypothesis indicates that the difference in the size of errors between the two models is significantly different from zero. We also confirm our results with one-tailed tests.

4.3. Test for forecast encompassing

Sometimes forecasts with larger errors may contain incremental information that is missing in the alternative forecasts (Granger and Newbold, 1973). Harvey et al. (1998) developed the encompassing test to evaluate whether (i) a given forecast incorporates the information content present in another projection, or (ii) the information in the benchmark model outweighs that of the alternative model, or vice versa. In effect, this test operates on the premise that if one forecast fully encompasses the other– the information content of the superior forecast would be so comprehensive that the inferior forecast would add little to no valuable insight to a combined forecast. Essentially, the contribution of the inferior forecast to the combined forecast would be minimal, with its optimal weight approaching zero.

If our forecasts at time t conditional on the full information set Ω_t , $y_{t|\Omega_t}^{(a)}$, encompasses the benchmark $y_{t|\Omega_t}^{(b)}$, the difference between both forecasts ought not to be a significant predictor of our forecast errors. This notion yields the regression-based test for estimating the optimal weights for each component of the composite forecast as follows:

$$\underbrace{y_t - y_{t|\Omega_t}^{(a)}}_{pe_{2,k,t}} = \underbrace{\phi(y_{t|\Omega_t}^{(b)} - y_{t|\Omega_t}^{(a)})}_{pe_{2,k,t} - pe_{1,k,t}} + \xi_t \tag{11}$$

$$\Downarrow$$

$$y_t = \phi y_{t|\Omega_t}^{(b)} + (1 - \phi)y_{t|\Omega_t}^{(a)} + \xi_t$$

where $pe_{1,k,t}$ is the percentage error from the benchmark model, and $pe_{2,k,t}$ is the percentage error from our alternative model specifications.

We estimate the regression in Equation 11 and test the null hypothesis of $H_0 : \phi = 0$ using the two-tailed t test. The null hypothesis assumes that the benchmark forecast encompasses the competing forecast. Consistent with the second equation in 11, rejection of the null suggests that a combination of the two forecasts (a composite forecast) would yield a more informative forecast and, simultaneously, lower prediction errors. Put simply, the alternative forecast contains

incremental information not otherwise present in the benchmark model forecasts. Alternatively, failure to reject the null hypothesis implies that the benchmark model forecast fully encompasses the information content of the alternative model projections and further implies that the combined forecast would be composed entirely of the projection from the benchmark model.

5. Forecast performance results

To evaluate the out-of-sample performance of benchmark and alternative model specifications, we begin the analysis by estimating Equations 1, 2, 4, and 5 with data from January 2000 to December 2004 ($n = 60$) to obtain forecasts for January 2005. We extend this process by “sliding” our estimation window forward, dropping the oldest observation and including the most recent one for each subsequent month until the end of the sample. These rolling window forecasts were generated for the entire out-of-sample period from January 2005 to December 2023 ($n = 228$). Once we obtain these monthly price forecasts, we identify which ones need to be considered as inputs in the SAP forecast, as shown in Table 3. After filling in Table 3 with the appropriate futures-based monthly forecasts and available observed farm prices received by farmers, we apply the 7-year Olympic average marketing weights computed from NASS data to these prices and compute the sum of the weighted monthly farm prices, that is, the SAP. Using Equation 6, we calculate percent errors by comparing these SAP forecasts from WASDE and alternative futures models to actual prices. We then use these percentage errors to calculate the statistical bias and accuracy measures (discussed in Section 4.1) and evaluate the out-of-sample predictive performance of each forecasts. While the discussion here focuses on SAP forecasts, additional details for each monthly forecast as well as sensitivity analysis are presented in the appendix.

5.1. Bias

The tests of bias, conducted based on the MPE statistic (see Table 4 and Figure 3), show that WASDE forecasts had positive forecast errors indicating a tendency to under-predict cotton prices during November and December (harvest period). On the other hand, Hoffman’s model had negative forecast errors, over-predicting the SAPs in each forecast month. The alternative models proposed in this study provide a significant improvement in terms of bias over Hoffman’s model as Models 1 and 2 produce unbiased forecasts at all forecast horizons, and Model 3 generates unbiased forecasts through March (with under-prediction in April–July, postharvest). Moreover, the magnitude of any bias estimates throughout Table 4 is extremely small, indicating that while there are detectable biases in some forecasts, they are not substantial in absolute terms. This suggests that the improvements achieved by the alternative models, while statistically significant, result in only minor adjustments to the forecast values. Consequently, even the biased forecasts, such as those from Hoffman model, are relatively close to the actual SAPs.

5.2. Accuracy

To evaluate the accuracy of our models with respect to WASDE projections, we rely on the most commonly used accuracy measures of MAPE and RMSPE. Table 5 presents MAPE statistics for each forecast month. Our findings demonstrate that both Hoffman’s and Models 2 and 3 offer accuracy improvements to WASDE forecasts in August and September during growing period, but only Model 2 offers additional accuracy improvements in May–July after harvest. Furthermore, Model 2 outperforms the Hoffman model in 10 out of the 12 forecast month, while the other models provide more modest accuracy improvements. Figure 4 shows the differences in MAPEs between WASDE and alternative forecasts and indicates accuracy improvements where these differences are negative. This figure demonstrates that accuracy improvements in August and September were fairly substantial, while the ones later in the forecasting cycle (May–July) rather minor. This figure may also be interpreted relative to the size

Table 4. Mean percentage error (MPE) for season-average price forecast

| Forecast month | Model 1 | Model 2 | Model 3 | Hoffman model | WASDE model |
|----------------|---------|---------|----------|---------------|-------------|
| August | 0.058 | 0.015 | 0.035 | -0.093*** | 0.004 |
| September | 0.058 | 0.020 | 0.040 | -0.079*** | 0.013 |
| October | 0.033 | 0.020 | 0.041 | -0.068** | 0.027 |
| November | 0.015 | 0.011 | 0.038 | -0.073** | 0.021* |
| December | -0.024 | -0.002 | 0.026 | -0.092*** | 0.016* |
| January | -0.025 | -0.014 | 0.012 | -0.098*** | 0.008 |
| February | -0.027 | -0.012 | 0.015 | -0.082** | -0.004 |
| March | -0.013 | -0.009 | 0.016 | -0.061** | 0.001 |
| April | -0.011 | -0.007 | 0.019** | -0.050** | 0.003 |
| May | 0.002 | -0.003 | 0.023*** | -0.030** | 0.002 |
| June | 0.006 | 0.000 | 0.023*** | -0.019* | 0.004 |
| July | -0.004 | -0.002 | 0.020** | -0.015* | 0.005 |

Note 1: Model 1 represents the moving average approach with basis deviation term.
 Note 2: Model 2 represents the regression model with nearby futures prices.
 Note 3: Model 3 represents the regression model with basis deviation term.
 Note 4: The null hypothesis states that the MPEs are not significantly different from 0.
 Note 5: Statistical significance of a two-tailed test is indicated with the asterisk: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

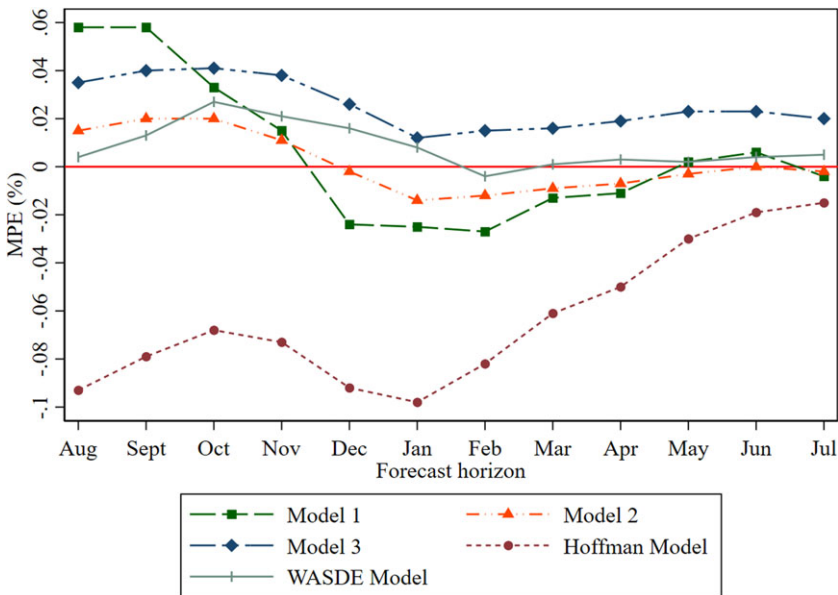


Figure 3. Mean percentage error (MPE) across SAP forecast months.
 Notes: This figure presents the accuracy statistics of the MPE criterion (detailed in Equation 9) across models and horizons. For each forecast month, a two-tailed *t*-test is conducted to assess bias in MPEs and determine if they significantly differ from zero. Model details are provided in Section 3: Model 1 denotes a moving average approach with a basis deviation term, Model 2 presents a regression model with nearby futures prices, and Model 3 represents a regression model with a basis deviation term.

Table 5. Mean absolute percentage error (MAPE) for season-average price forecast

| Forecast month | Model 1 | Model 2 | Model 3 | Hoffman model | WASDE model |
|----------------|---------------------|---------------------|---------------------|--------------------|-------------|
| August | 0.168 | 0.095 ^{††} | 0.092 ^{††} | 0.099 [†] | 0.104 |
| September | 0.164 | 0.080 [†] | 0.089 [†] | 0.079 [†] | 0.092 |
| October | 0.091 | 0.079 [†] | 0.085 [†] | 0.089 | 0.072 |
| November | 0.070 [†] | 0.088 | 0.099 | 0.081 | 0.040 |
| December | 0.053 [†] | 0.078 [†] | 0.085 [†] | 0.094 | 0.029 |
| January | 0.048 [†] | 0.065 [†] | 0.056 [†] | 0.098 | 0.023 |
| February | 0.045 [†] | 0.051 [†] | 0.047 [†] | 0.087 | 0.021 |
| March | 0.032 [†] | 0.030 [†] | 0.035 [†] | 0.069 | 0.016 |
| April | 0.023 [†] | 0.024 [†] | 0.028 [†] | 0.056 | 0.016 |
| May | 0.022 [†] | 0.017 ^{††} | 0.030 | 0.035 | 0.017 |
| June | 0.018 [†] | 0.014 ^{††} | 0.029 | 0.024 | 0.015 |
| July | 0.013 ^{††} | 0.011 ^{††} | 0.028 | 0.019 | 0.014 |

Values marked with † indicate that a specific model performs better in terms of lower forecast errors compared to the WASDE projections. Values marked with † denote that a specific model shows lower forecast errors compared to the Hoffman model projections. Model 1 represents a moving average approach with a basis deviation term. Model 2 represents a regression model using nearby futures prices. Model 3 denotes a regression model with a basis deviation term.

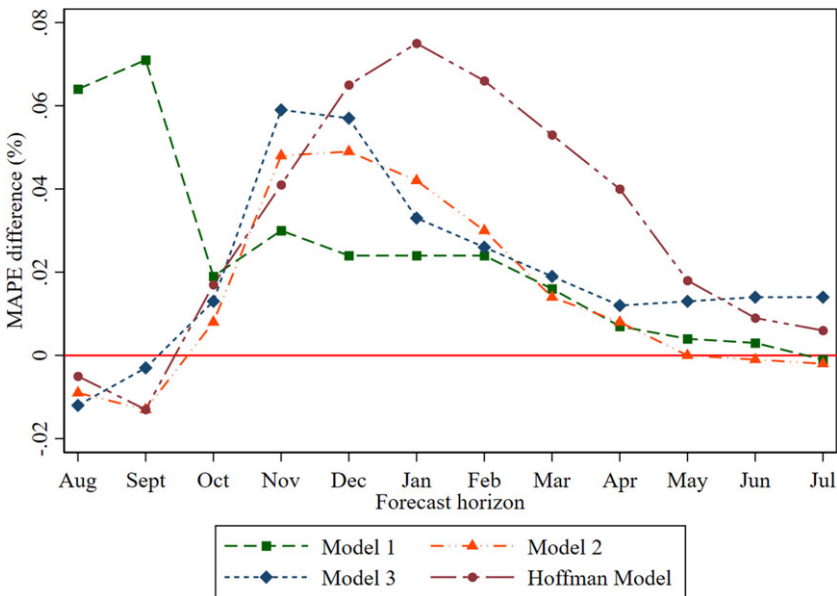


Figure 4. Mean absolute percentage error (MAPE) differences across SAP forecast months.

Notes: This figure illustrates the differences in MAPE statistics between forecasts from alternative models and WASDE projections across forecast months. Negative (positive) MAPE differences for each forecast month indicate that the alternative model produces lower (higher) MAPE values, indicating more (less) accurate forecasts compared to the WASDE model. Model details from Section 3 are as follows: Model 1 denotes a moving average approach with a basis deviation term, Model 2 presents a regression model with nearby futures prices, and Model 3 represents a regression model with a basis deviation term.

Table 6. Root mean squared percentage error (RMSPE) for season-average price forecast

| Forecast month | Model 1 | Model 2 | Model 3 | Hoffman model | WASDE model |
|----------------|---------------------|---------------------|---------------------|---------------|-------------|
| August | 0.311 | 0.111 ^{†‡} | 0.107 ^{†‡} | 0.147 | 0.129 |
| September | 0.290 | 0.098 ^{†‡} | 0.099 ^{†‡} | 0.119 | 0.114 |
| October | 0.158 | 0.103 [†] | 0.103 [†] | 0.115 | 0.088 |
| November | 0.093 [†] | 0.107 [†] | 0.116 [†] | 0.127 | 0.050 |
| December | 0.070 [†] | 0.099 [†] | 0.114 [†] | 0.144 | 0.034 |
| January | 0.068 [†] | 0.079 [†] | 0.079 [†] | 0.146 | 0.031 |
| February | 0.072 [†] | 0.067 [†] | 0.066 [†] | 0.143 | 0.025 |
| March | 0.055 [†] | 0.044 [†] | 0.048 [†] | 0.120 | 0.019 |
| April | 0.030 [†] | 0.030 [†] | 0.035 [†] | 0.089 | 0.021 |
| May | 0.035 [†] | 0.022 ^{†‡} | 0.036 [†] | 0.053 | 0.022 |
| June | 0.023 [†] | 0.018 ^{†‡} | 0.036 [†] | 0.039 | 0.020 |
| July | 0.016 ^{†‡} | 0.015 ^{†‡} | 0.033 | 0.030 | 0.020 |

Values marked with † indicate that a specific model performs better in terms of lower forecast errors compared to the WASDE projections. Values marked with ‡ denote that a specific model shows lower forecast errors compared to the Hoffman model projections. To determine whether the improvements are statistically significant, we use a Modified Diebold–Mariano test (refer to section 5.3). Model 1 denotes a moving average approach with a basis deviation term. Model 2 represents a regression model using nearby futures prices. Model 3 denotes a regression model with a basis deviation term.

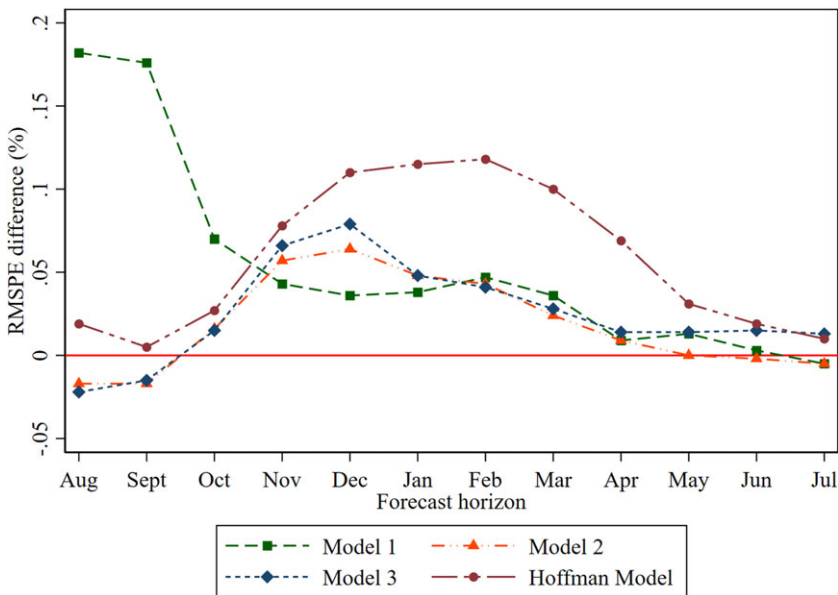


Figure 5. Root mean squared percentage error (RMSPE) difference across SAP forecast months.

Notes: This figure illustrates the differences in RMSPE statistics between forecasts from alternative models and WASDE projections across forecast months. Negative (positive) RMSPE differences for each forecast month indicate that the alternative model produces lower (higher) RMSPE values, indicating more (less) accurate forecasts compared to the WASDE model. Detailed in Section 3, Model 1 denotes a moving average approach with a basis deviation term, Model 2 presents a regression model with nearby futures prices, and Model 3 represents a regression model with a basis deviation term.

of Hoffman’s model forecasts and indicates very substantial improvements in accuracy for all of the proposed models from December through May.

These findings are largely confirmed in RMSPE analysis which is more sensitive to large errors, as shown in Table 6 and Figure 5). The main difference is that there is no longer evidence of Hoffman’s forecasts being more accurate than WASDE, while the accuracy enhancements from Model 2 are confirmed for August–September and May–June. Using this metric, Model 2 consistently outperforms Hoffman’s model’s accuracy and Models 1 and 3 also offer consistent accuracy improvements. Figure 5 shows the differences in RMSPEs between WASDE and alternative forecasts, indicating an accuracy improvement during the forecast months where these differences are negative. The results are consistent with those from MAPEs. Among the three proposed models, Model 2 stands out by demonstrating the most accuracy improvements compared to WASDE, particularly in August and September, with relatively minor improvements from May to July months.

5.3. MDM test

The MDM test results in Table 7 indicate that WASDE forecasts had significantly lower errors compared to Hoffman’s and our proposed models, particularly during the harvest and postharvest months (starting from November forward). In addition, for the months where our regression approaches outperformed WASDE in terms of lower MAPE and RMSPE (Tables 5 and 6), the MDM tests indicate that these improvements were not statistically significant. However, compared to the Hoffman model, Models 1 and 2 produce significantly lower forecast errors during March–May months, while Model 3 generates lower forecast errors only for the January and February forecast months. Overall, these findings indicate that our proposed model specifications offer significant improvements in forecast accuracy over the Hoffman’s model, demonstrating their potential for more accurate forecasting during certain periods.

Table 7. Modified Diebold–Mariano test

| Forecast month | Benchmark: WASDE model | | | | Benchmark: Hoffman model | | |
|----------------|------------------------|-----------|-----------|---------------|--------------------------|---------|---------|
| | Model 1 | Model 2 | Model 3 | Hoffman Model | Model 1 | Model 2 | Model 3 |
| August | −0.924 | 0.913 | 1.044 | −0.054 | −0.929 | 0.623 | 0.664 |
| September | −1.081 | 0.942 | 0.609 | 0.451 | −1.315 | 0.427 | 0.136 |
| October | −0.493 | 0.095 | −0.120 | −0.565 | −0.124 | 0.638 | 0.411 |
| November | −1.540 | −2.607** | −3.177*** | −1.466 | 0.517 | −0.388 | −0.787 |
| December | −1.815* | −2.713** | −2.730** | −1.974* | 1.186 | 0.361 | 0.034 |
| January | −1.959* | −3.209*** | −2.138** | −2.651** | 1.340 | 1.409 | 1.855* |
| February | −1.836* | −1.935* | −1.699 | −2.074* | 1.719 | 1.655 | 1.774* |
| March | −1.115 | −1.284 | −1.855* | −1.960* | 2.332** | 2.006* | 1.681 |
| April | −1.292 | −1.825* | −2.193** | −2.169** | 2.067* | 2.058* | 1.499 |
| May | −0.894 | −0.357 | −2.393** | −1.789* | 2.360** | 1.847* | 0.586 |
| June | −0.648 | 0.031 | −2.286** | −1.190 | 1.149 | 1.426 | −0.311 |
| July | 0.159 | 0.440 | −2.503** | −0.903 | 1.376 | 1.644 | −0.977 |

Note 1: Model 1 represents the moving average approach with basis deviation term.

Note 2: Model 2 represents the regression model with nearby futures prices.

Note 3: Model 3 represents the regression model with basis deviation term.

Note 4: The null hypothesis for the two-tailed Modified Diebold–Mariano test states that the two forecasts have similar forecast accuracy.

Note 5: Statistical significance of a two-tailed test is indicated with the asterisk: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.4. Encompassing

The encompassing tests indicate that our proposed models incorporate additional information absent from the WASDE projections and the Hoffman model. Although the MDM test did not show a significant enhancement in forecast accuracy compared to the WASDE benchmark, the encompassing test revealed that our models contained incremental information missing from the WASDE forecasts during months when our models exhibited lower forecast errors. Furthermore, the encompassing test results (see Table 8) demonstrate that Models 2 and 3 encompass additional information absent from the Hoffman model across all forecast months, while Model 1 incorporates incremental information over the Hoffman model primarily during the harvest and postharvest periods, starting from November onward. These findings suggest that our proposed models offer additional information over the Hoffman approach and may result in more accurate combination forecasts.

Table 8. Encompassing test

| Forecast month | Benchmark: WASDE model | | | | Benchmark: Hoffman model | | |
|----------------|------------------------|----------|----------|---------------|--------------------------|----------|----------|
| | Model 1 | Model 2 | Model 3 | Hoffman Model | Model 1 | Model 2 | Model 3 |
| August | 0.080 | 0.729** | 0.787** | 0.300 | 0.171 | 0.693*** | 0.687*** |
| September | 0.132 | 0.618*** | 0.611*** | 0.470** | 0.180** | 0.602*** | 0.586*** |
| October | 0.290*** | 0.426*** | 0.419*** | 0.390*** | 0.235 | 0.570*** | 0.556*** |
| November | 0.193 | 0.172 | 0.125 | 0.177** | 0.742*** | 0.631*** | 0.560*** |
| December | 0.231** | 0.089 | 0.040 | 0.087 | 0.975*** | 0.821*** | 0.688*** |
| January | 0.074 | 0.040 | 0.016 | 0.028 | 1.042*** | 0.961*** | 0.901*** |
| February | -0.024 | 0.066 | 0.069 | 0.005 | 1.367*** | 1.133*** | 1.008*** |
| March | 0.102 | 0.175* | 0.115 | 0.019 | 1.563*** | 1.223*** | 1.032*** |
| April | 0.174 | 0.289* | 0.136 | -0.014 | 1.209*** | 1.149*** | 0.875*** |
| May | 0.240 | 0.498** | 0.180 | 0.028 | 0.652*** | 1.034*** | 0.653*** |
| June | 0.416** | 0.595*** | 0.097 | 0.128 | 0.697*** | 1.015*** | 0.546*** |
| July | 0.682*** | 0.731*** | 0.069 | 0.252 | 1.110*** | 1.172*** | 0.440** |

Note 1: Model 1 represents the moving average approach with basis deviation term.

Note 2: Model 2 represents the regression model with nearby futures prices.

Note 3: Model 3 represents the regression model with basis deviation term.

Note 4: The null hypothesis for the encompassing test states that the benchmark model forecast contains all information present in alternative model forecasts.

Note 5: Statistical significance of a two-tailed test is indicated with the asterisk: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Overall, our results highlight the superiority of Model 2 over the Hoffman model, demonstrating its ability to significantly enhance forecast accuracy throughout the marketing year and producing more informative price projections. While Model 2 does not outperform the WASDE benchmark, on average, it offers notable improvements in accuracy during certain months.

6. Summary and conclusions

Accurate price forecasts are valuable and important for producers, processors, and policymakers. This study explored several avenues for improving cotton SAP forecasting model proposed by Hoffman and Meyer (2018), which is currently used as an information source for the USDA's WASDE outlook program. Three alternative model specifications were developed based on using

a less restrictive regression approach and including a basis deviation term that captures current market information. The first model included a basis deviation from its 5-year average to Hoffman's model, which is used as a benchmark. The second model utilized rolling OLS regressions of the spot on the nearby futures prices. The third model combined the rolling regression approach and basis deviation term.

Hoffman's benchmark model and the three alternative specifications were used to generate monthly out-of-sample forecasts of cotton prices, which were later used to compute the SAPs. The forecasts from alternative models were compared to WASDE and Hoffman's benchmark approaches over 2008–2023 using standard accuracy measures such as MPE, MAPE, and RMPSE, and forecast evaluation criteria such as MDM and encompassing tests. We found that our second model offered the largest improvements to Hoffman's benchmark both in terms of accuracy and encompassing. These improvements are similar to the findings of Etienne *et al.* (2023) for corn forecasts. Thus, our findings offer enhancements to current USDA forecasting procedures without increasing computational costs. This will directly benefit farmers, policymakers, and exporters who use USDA projections as an information source for directing their future activities.

The findings of this paper may also be used by all market participants interested in obtaining accurate and reliable commodity price projections. Previous studies (for example Colino and Irwin, 2010; Hoffman *et al.*, 2015; Irwin *et al.*, 1994; Irwin and Good, 2015; Manfredo and Sanders, 2004) demonstrate that futures-based forecasts provide a viable alternative to USDA forecasts. Since many public and private companies use commodities for diversifying their investment portfolios, they may find it fruitful to adopt some of the forecasting techniques discussed in this paper to form expectations regarding future price movements.

While this study offers enhancements to current USDA cotton forecasting procedures, there are several avenues for future research that could further improve forecast accuracy. First, extending the analysis to longer forecast horizons beyond the marketing year months examined here could allow a more direct comparison with WASDE projections. Second, our use of monthly average futures prices might be limiting the models' responsiveness to current market conditions; incorporating more recent futures price information, such as month-ending prices, could enhance forecast accuracy. Third, examining the structure of forecast errors could reveal patterns that might be exploited to improve subsequent forecasts. Fourth, future research could explore incorporating the lagged basis term instead of the lagged basis deviation to improve the forecast accuracy. Fifth, forecasts for the final month of the crop year used futures prices from the next crop year and an alternative approach for the last month might prove more accurate. Additionally, applying and evaluating the proposed modeling approaches on other agricultural commodities would assess their generalizability across different markets. Furthermore, examining the inclusion of additional explanatory variables beyond the basis deviation term, such as inventory levels, production-related factors, and other current market information, could potentially capture more sources of price variation and enhance predictive power. Future work could also explore advanced time-series techniques, econometric specifications, or machine learning methods as alternatives to the regression-based models used here. We also recognize that the 7-year Olympic moving average used here might be longer than that typically used for other commodities. A shorter moving average might be considered in future research. Lastly, periodically re-estimating the models with updated data would help maintain forecast accuracy over time as market conditions evolve.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/aae.2024.36>.

Data availability statement. The data that support the findings of this study are openly available in the USDA National Agricultural Statistics Service Quick Stats, and Quandl.

Author contribution. Conceptualization: A.P. and O.I.M.; Methodology: A.P., O.I.M., and S.L.S.; Formal Analysis: A.P. and O.I.M.; Data Curation: A.P. and O.I.M.; Writing – Original Draft: A.P.; Writing – Review and Editing: A.P., O.I.M., and S.L.S.; Project Supervision, A.P., O.I.M., and S.L.S.

Funding statement. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interests. The authors declare none.

Declaration of AI. No generative AI was used in this work.

References

- Adjemian, M.K. "Quantifying the wasde announcement effect." *American Journal of Agricultural Economics* 94,1(2012):238–56.
- Adjemian, M.K., V.G. Bruno, and M.A. Robe. "Incorporating uncertainty into USDA commodity price forecasts." *American Journal of Agricultural Economics* 102,2(2020):696–712.
- Algieri, B., and M. Kalkuhl. (2014). Back to the futures: An assessment of commodity market efficiency and forecast error drivers. Available at SSRN. 2509820.
- Balseca, E., and E. Bennion. "Focus on prices and spending. Import and export prices: Second quarter 2011." *US Bureau of Labor Statistics* 2,5(2011):1–3.
- Black, F. "The pricing of commodity contracts." *Journal of Financial Economics* 3,1-2(1976):167–79.
- Bowman, C., and A. Husain. *Forecasting commodity prices: Futures versus judgment*. 2004. IMF Working Paper No. 04/41. Available at SSRN: = <https://ssrn.com/abstract=878864> I got this information from: = https://papers.ssrn.com/sol3/papers.cfm?abstract_id=878864.
- Chavas, J.-P., R.D. Pope, and R.S. Kao. "An analysis of the role of futures prices, cash prices and government programs in acreage response." *Western Journal of Agricultural Economics* 8,1(1983):27–33.
- CME Group. (2024). Understanding major usda reports for grains and oilseed markets. Internet site: <https://www.cmegroup.com/articles/2024/understanding-major-usda-reports.html>.
- Colino, E.V., and S.H. Irwin. "Outlook vs. futures: Three decades of evidence in hog and cattle markets." *American Journal of Agricultural Economics* 92,1(2010):1–15.
- Colino, E.V., S.H. Irwin, P. Garcia. "Improving the accuracy of outlook price forecasts." *Agricultural Economics* 42,3(2011):357–71.
- Etienne, X., S. Farhangdoost, and L. Hoffman. (2019). An alternative method to forecast season-average price for US corn, Available at SSRN, 3306590.
- Etienne, X.L., S. Farhangdoost, L.A. Hoffman, and B.D. Adam. "Forecasting the US season-average farm price of corn: Derivation of an alternative futures-based forecasting model." *Journal of Commodity Markets* 30(2023):100333.
- Fama, E.F. "Session topic: Stock market price behavior." *The Journal of Finance* 25,2(1970):383–417.
- Fama, E.F. "Time, salary, and incentive payoffs in labor contracts." *Journal of Labor Economics* 9,1(1991):25–44.
- Fiene, D. (2023). Importance of the wasde report for grain speculation. Internet site: <https://www.premiercooperative.com/news/co-op-news/wasde-report>.
- Figuerola-Ferretti, I., A. Rodriguez, and E. Schwartz. "Oil price analysts' forecasts." *Journal of Futures Markets* 41,9(2021):1351–74.
- Franses, P.H. "A note on the mean absolute scaled error." *International Journal of Forecasting* 32,1(2016):20–2.
- Granger, C.W., and P. Newbold. "Some comments on the evaluation of economic forecasts." *Applied Economics* 5,1(1973):35–47.
- Harvey, D., S. Leybourne, and P. Newbold. "Testing the equality of prediction mean squared errors." *International Journal of Forecasting* 13,2(1997):281–91.
- Harvey, D.I., S.J. Leybourne, and P. Newbold. "Tests for forecast encompassing." *Journal of Business & Economic Statistics* 16,2(1998):254–9.
- Hayenga, M.L., and B. Jiang. *Corn and Soybean Basis Behavior and Forecasting: Fundamental and Alternative Approaches*. Iowa State University, Department of Economics, 1997, Technical report.
- Hoffman, L., and L. Meyer. "Forecasting the US Season-Average Farm Price of Upland Cotton: Derivation of a Futures Price Forecasting Model." *Electronic Outlook Report from the Economic Research Service (CWS-18I-01)*, 2018.
- Hoffman, L.A. "Forecasting the counter-cyclical payment rate for US corn: An application of the futures price forecasting model." Economic Research Service Electronic Outlook Report No. FDS-05a-01, U.S. Department of Agriculture. (2005). Available at <https://www.ers.usda.gov/publications/pub-details/?pubid=36456>.
- Hoffman, L.A., X.L. Etienne, S.H. Irwin, E.V. Colino, and J.I. Toasa. "Forecast performance of WASDE price projections for US corn." *Agricultural Economics* 46,S1(2015):157–71.
- Hoffman, L.A., S.H. Irwin, and J. Toasa. (2007). Forecast performance of futures price models for corn, soybeans, and wheat. (381-2016-22161), 32.
- Holden, K., D.A. Peel. "On testing for unbiasedness and efficiency of forecasts." *The Manchester School of Economic & Social Studies* 58,2(1990):120–7.
- Hyndman, R.J., et al. "Another look at forecast-accuracy metrics for intermittent demand." *Foresight: The International Journal of Applied Forecasting* 4,4(2006):43–6.

- Irwin, S., and D. Good. "Long-term corn, soybeans, and wheat price forecasts and the farm bill program choice." *Farmdoc Daily* 5,20(2015). Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign.
- 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* 14,7(1994):861–75.
- Isengildina-Massa, O., S.H. Irwin, D.L. Good, and J.K. Gomez. "Impact of WASDE reports on implied volatility in corn and soybean markets." *Agribusiness: An International Journal* 24,4(2008):473–90.
- Isengildina-Massa, O., B. Karali, T.H. Kueth, A.L. Katchova. "Joint Evaluation of the System of USDA's Farm Income Forecasts." *Applied Economic Perspectives and Policy* 43,3(2021):1140–60.
- Isengildina-Massa, O., and S. MacDonald. (2009). U.S. cotton prices and the world cotton market: Forecasting and structural change. (1477-2016-121197), 26.
- Isengildina-Massa, O., and S. MacDonald. "Forecasting US Cotton Prices in a Changing Market." *Advances in Business and Management Forecasting*. Emerald Group Publishing Limited, 2013.
- Janzen, J.P., A. Smith, and C.A. Carter. "Commodity price comovement and financial speculation: The case of cotton." *American Journal of Agricultural Economics* 100,1(2018):264–85.
- Just, R.E., and G.C. Rausser. "Commodity price forecasting with large-scale econometric models and the futures market." *American Journal of Agricultural Economics* 63,2(1981):197–208.
- Kastens, T.L., R. Jones, and T.C. Schroeder. "Futures-based price forecasts for agricultural producers and businesses." *Journal of Agricultural and Resource Economics* 23,1(1998):294–307.
- Kwas, M., and M. Rubaszek. "Forecasting commodity prices: Looking for a benchmark." *Forecasting* 3,2(2021):447–59.
- Liu, S.-M., B.B. Wade, C.M. Oellermann, and P.L. Farris. "Forecasting the nearby basis of live cattle." *The Journal of Futures Markets* 14,3(1994):259–273.
- MacDonald, S., and O. Isengildina-Massa. (2012). Information rigidity and correcting inefficiency in USDA's commodity forecasts. Technical report.
- Makhula, A.T. (2021). 33 uses of cotton. Internet site: <https://engineeringtextile.com/fiber/169-uses-of-cotton>.
- Manfredo, M.R., and D.R. Sanders. "The value of public price forecasts: Additional evidence in the live hog market." *Journal of Agribusiness* 2016(2004):119–31.
- McKenzie, A.M., and M.T. Holt. "Market efficiency in agricultural futures markets." *Applied Economics* 34,12(2002):1519–32.
- Newey, W.K., and K.D. West. (1986). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix.
- Peck, A.E. "The economic role of traditional commodity futures markets." *Futures Markets: Their Economic Role*, 1–81.
- Rausser, G.C., and R.E. Just. *Agricultural Commodity Price Forecasting Accuracy: Futures Markets Versus Commercial Econometric Models*. Berkeley: UC, Department of Agricultural and Resource Economics, 1979, Technical report.
- Roache S. K. and D.A. Reichsfeld. *Do Commodity Futures Help Forecast Spot Prices?*. 2011. IMF Working Paper No. 11/254, Available at SSRN: <https://ssrn.com/abstract=1956401> This information is taken from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1956401.
- Robinson, J. "Case Study of December 2010 ICE Cotton Futures and Options." *Texas A & M AgriLife Extension*, 2011a. <https://cottonmarketing.tamu.edu/marketing-strategies/december-2010/>
- Robinson, J. "Case Study of December 2011 ICE Cotton Futures and Options." *Texas A & M AgriLife Extension*, 2011b. <https://cottonmarketing.tamu.edu/marketing-strategies/december-2011/>
- Shoulberg, W. "What the spike in cotton prices to a 10-year high means." *Forbes*, 2021.
- Taylor, M.R., K.C. Dhuyvetter, and T.L. Kastens. "Forecasting crop basis using historical averages supplemented with current market information." *Journal of Agricultural and Resource Economics* 31,3(2006):549–67.
- Tomek, W.G. "An analysis of the futures-cash price basis for live beef cattle: comment." *North Central Journal of Agricultural Economics* 2,1(1980):81.
- Tomek, W.G. "Commodity futures prices as forecasts." *Applied Economic Perspectives and Policy* 19,1(1997):23–44.
- Tonsor, G.T., K.C. Dhuyvetter, and J.R. Mintert. "Improving cattle basis forecasting." *Journal of Agricultural and Resource Economics* 29, 2(2004):228–41.
- Townsend, T.P. (1989). This was a high-powered instrument that sent its projectile to the vitals of the industry: why usda does not forecast cotton prices. In: *Proceedings-Beltwide Cotton Production Research Conferences (USA)*, 374.
- USDA-NASS. (2011). Season-average price forecasts. Internet site: <http://www.ers.usda.gov/Data/PriceForecast/>.
- Working, H. "Theory of the inverse carrying charge in futures markets." *Journal of Farm Economics* 30,1(1948):1–28.
- Zoller, C. (2021). What is the waste report and why is it important? Internet site: <https://bpb-us-w2.wpmucdn.com/u.osu.edu/dist/9/29991/files/2021/07/OAM-WASDE-July-2021.pdf>.
- Zulauf, C.R., S.H. Irwin, J.E. Ropp, and A.J. Sberna. "A reappraisal of the forecasting performance of corn and soybean new crop futures." *Journal of Futures Markets* 19,5(1999):603–18.

Cite this article: Poghosyan, A., O. Isengildina-Massa, and S.L. Stewart (2025). "Futures-Based Forecasts of Cotton Prices: Beyond Historical Averages." *Journal of Agricultural and Applied Economics* 57, 114–134. <https://doi.org/10.1017/aae.2024.36>