How Accurate are the USDA's Baseline Projections?

Siddhartha Bora (The Ohio State University),
Ani Katchova (The Ohio State University),
Todd Kuethe (Purdue University)

NC-1177 Meeting
October 22, 2020

Background

- USDA's statistical agencies such as NASS and ERS provide forecasts of agricultural production, prices, trade, uses, inventories, and farm income.
- USDA forecasts such as Farm Income Forecasts and WASDE forecasts are eagerly awaited by stakeholders.
- A number of previous studies suggest that many USDA forecasts are biased and/or inefficient.
- Bora, Katchova, and Kuethe (2020) show that asymmetric loss functions can rationalize these USDA forecasts.

USDA Long-Run Agricultural Baseline Projections

- The USDA baseline projections describe the factors influencing agricultural markets for the next decade, and include projections of commodity prices, production, global agricultural trade and farm income.
- A conditional scenario that serves as a basis for comparison of alternative policies, and for analysing market developments, subject to assumptions.
- Important for understanding the status of the economy several quarters or years from the current year.
- Used for formulating policy such as preparing the President's budget and program allocations.
- As the projections are available for next 10 years, a relevant question is how long these projections stay informative.

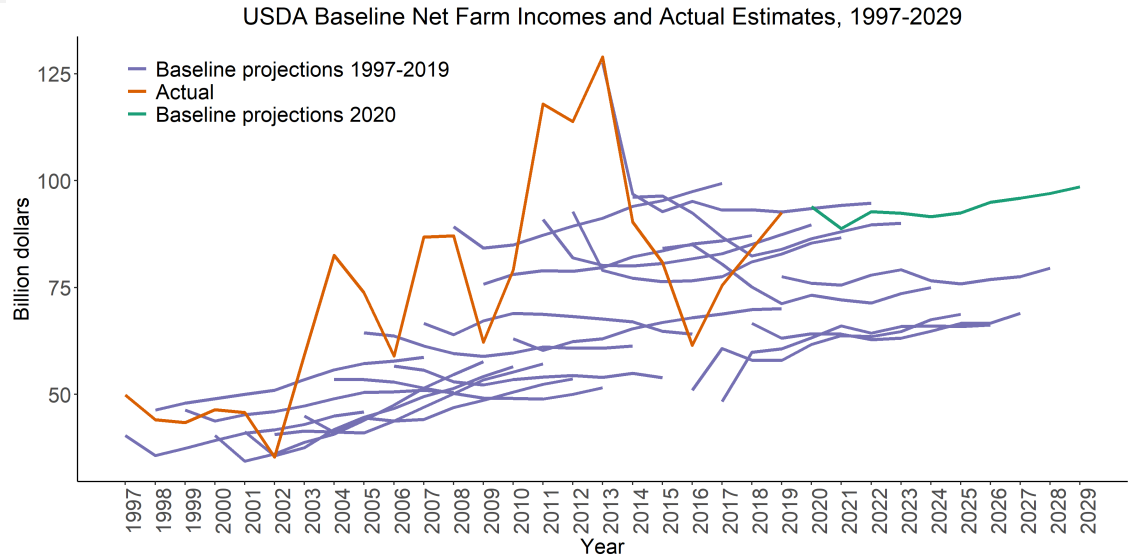
Evaluating USDA Baseline Projections

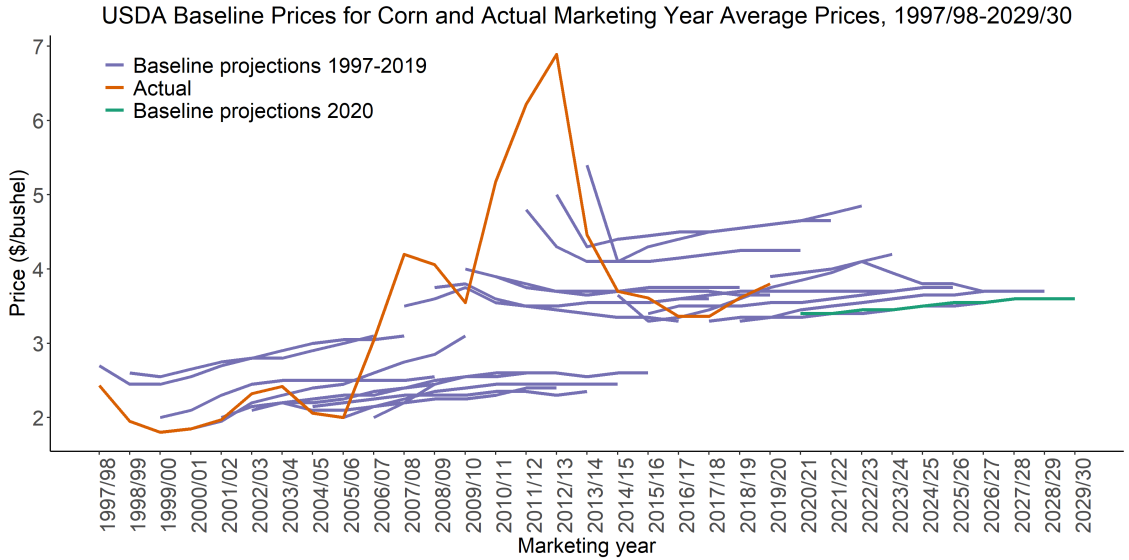
- Previous Literature:
  - Despite their importance in shaping agricultural policy, USDA's baseline projections have not been rigorously evaluated in the literature.
  - Traditional tests of assessing predictive content usually compares the actual forecast with some "naive" forecasts.
  - Path forecasts are when the forecaster predicts a set of values for multiple horizons. Forecast accuracy tests of path forecasts take into account the dynamics along the forecast path across the horizon, and typically follow a joint approach (Martinez, 2020).
- Our study:
  - We use a novel method that compares mean-squared forecast error with the unconditional variance of the target variable.
  - This method sidesteps the requirement of a "naive benchmark," as only the forecasts and the actual values are need for evaluation.
  - We calculate the largest informative forecast horizon for key baseline variables.
- We find that:
  - The baseline projections are barely informative beyond a year or two years ahead.

Data and Notations

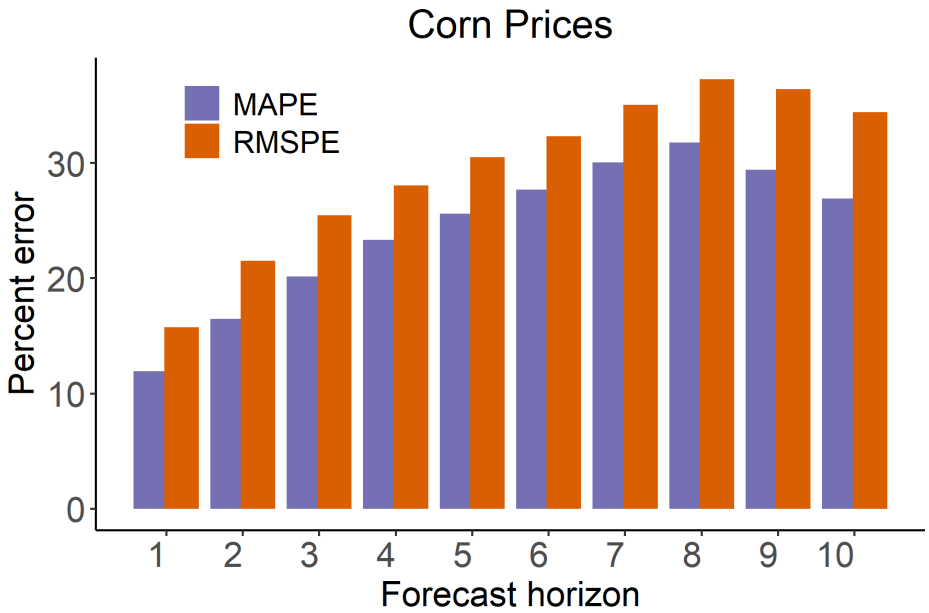
- USDA baseline projections are available since 1997, and updated in February of each year.
- Data set includes the actual estimates for the previous year, nowcast (the projection for the current year), and projections for the next 10 years.
- Our analysis is focused on two main projections tables:
  - farm sector income, and
  - prices of three commodities: corn, soybeans, and wheat.
- Definitions:
  - $\{Y_{1+h}, Y_{2+h}, \ldots, Y_{T+h}\}$ denote the set of T observed actual values corresponding to model forecasts $\hat{Y}_{t+h|t}$, $t = 1, 2, \ldots T$ at time t (h=forecast horizon).
  - $\{\hat{Y}_{1+h}, \hat{Y}_{2+h}, \ldots, \hat{Y}_{T+h}\}$ denote the set of h-step ahead forecasts.

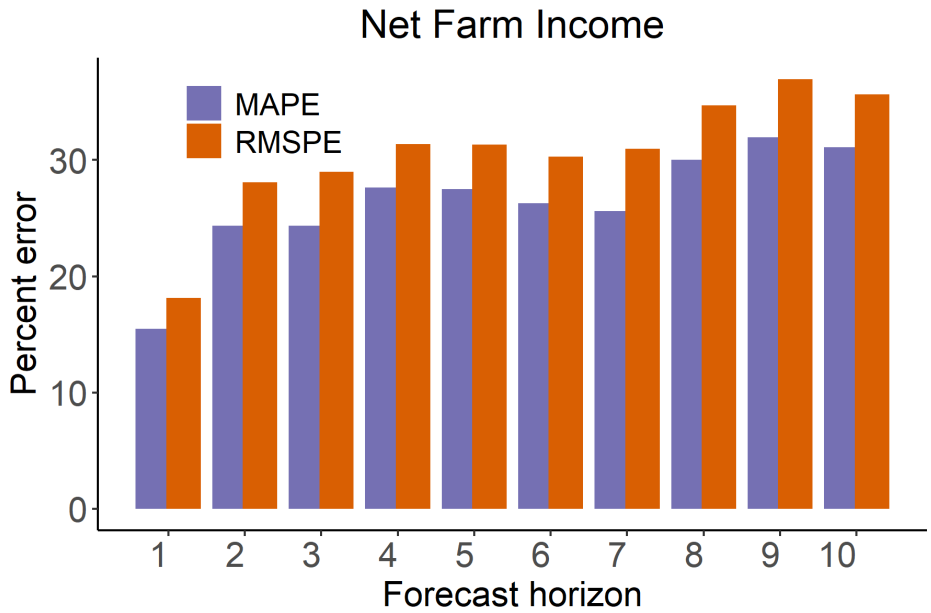
Baseline projections for net farm income and corn price





MAPE and RMSPE by horizon





Testing for Predictive Content in the Baseline Projections

- Set Up:
  - Actual values Y_t are generated by a stationary process $\{Y_t\}$.
  - Forecasts $\hat{Y}_{t+h|t}$, $t = 1, 2, \ldots T$ are realizations of a forecast generating process $\{Y^{\theta}_{t+h|t}\}$, (θ is a model parameter).
- Hypothesis testing about whether the baseline projections at horizon h are informative (Breitung and Knüppel, 2020):

$$H_0 : E(Y_{t+h} - \hat{Y}_{t+h|t})^2 \geq E(Y_{t+h} - \mu)^2, \text{ for } h > h^* \text{ and } t \in \{1, \ldots, T\} \quad (1)$$

$$H_1 : E(Y_{t+h} - \hat{Y}_{t+h|t})^2 < E(Y_{t+h} - \mu)^2 \quad (2)$$

where, $\mu = E(Y_t)$ is the unconditional mean.

- Hypothesis testing about whether the conditional mean is constant within the sample:

$$H_0 : E(Y_{t+h}|I_t) = \mu_{h,t} = \mu, \text{ for } h > h^* \text{ and } t \in \{1, \ldots, T\} \quad (3)$$

where, I_t is the information set at time t.

- The maximum informative forecast horizon is $h^* = h_{min} - 1$ where h_{min} is the smallest horizon for which the null hypothesis is not rejected.

Testing for Predictive Content in the Baseline Projections

- These hypotheses can be tested in a Mincer-Zarnowitz regression:

$$Y_{t+h} = \alpha_h + \beta_h \hat{Y}_{t+h|t} + \nu_{t+h} \tag{4}$$

- No information hypothesis: $H_0 : \beta_h \geq 0.5$, $H_1 : \beta_h < 0.5$.
- Constant mean hypothesis: $H_0 : \beta_h = 0$.
- LM test statistic is constructed as,

$$\tau_a = \frac{1}{\hat{\omega}_a \sqrt{T}} \sum_{t=1}^{n} a_t$$

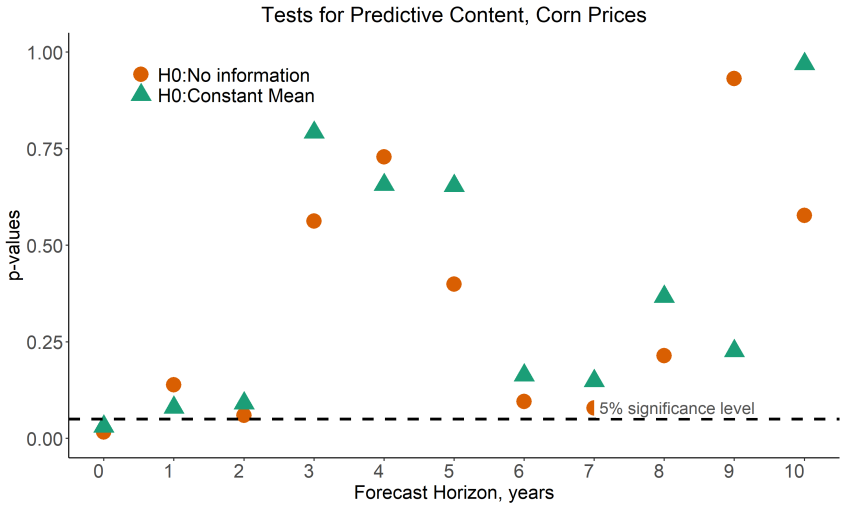
$$a_t = \left[Y_{t+h} - \overline{Y}^h - 0.5(\hat{Y}_{t+h|t} - \overline{\hat{Y}_h}) \right] (\hat{Y}_{t+h|t} - \overline{\hat{Y}_h}) \text{ for } H_0 : \beta_h = 0.5$$

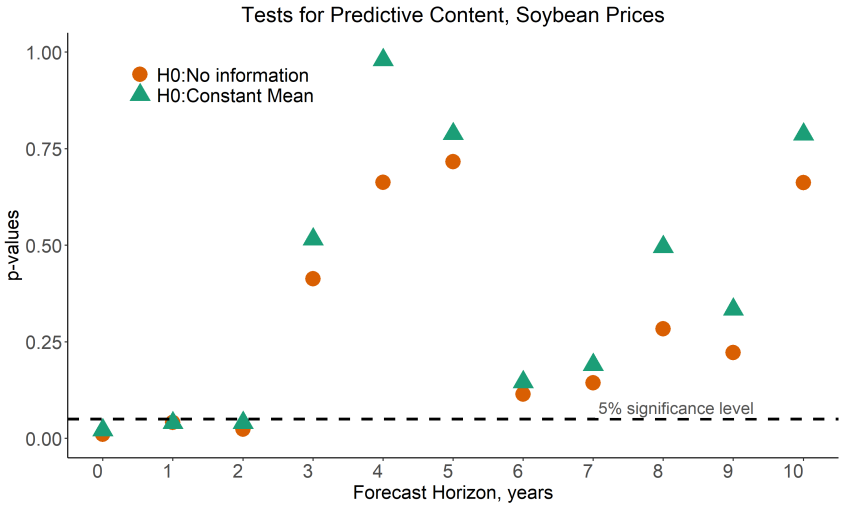
$$a_t = (Y_{t+h} - \overline{Y}^h)(\hat{Y}_{t+h|t} - \overline{\hat{Y}_h}) \text{ for } H_0 : \beta_h = 0$$

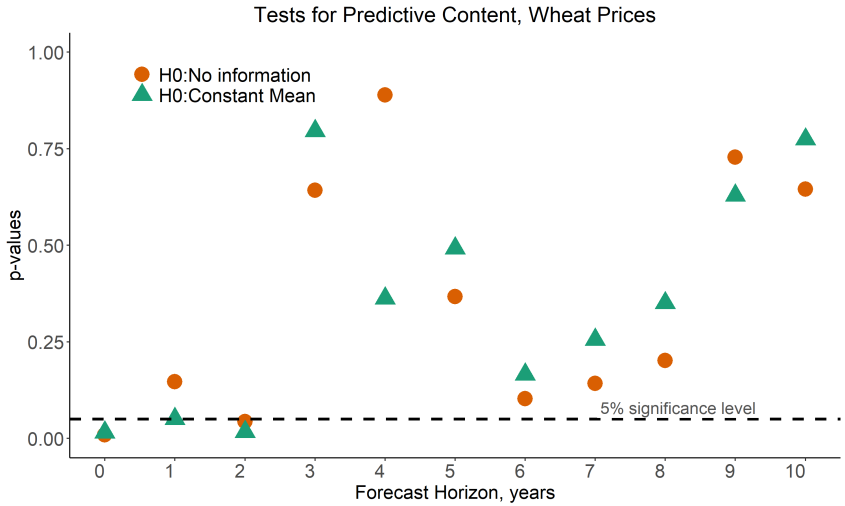
where, $\hat{\omega}^2$ is a consistent estimator of the long-run variance of a_t.

- The LM-statistic has a standard normal distribution.

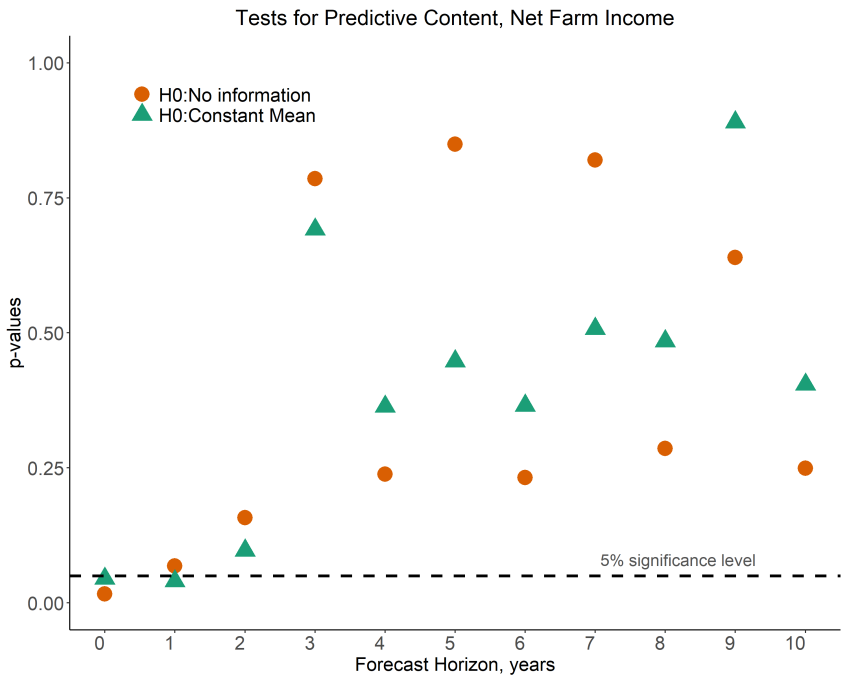
Predictive Content for Commodity Price Projections

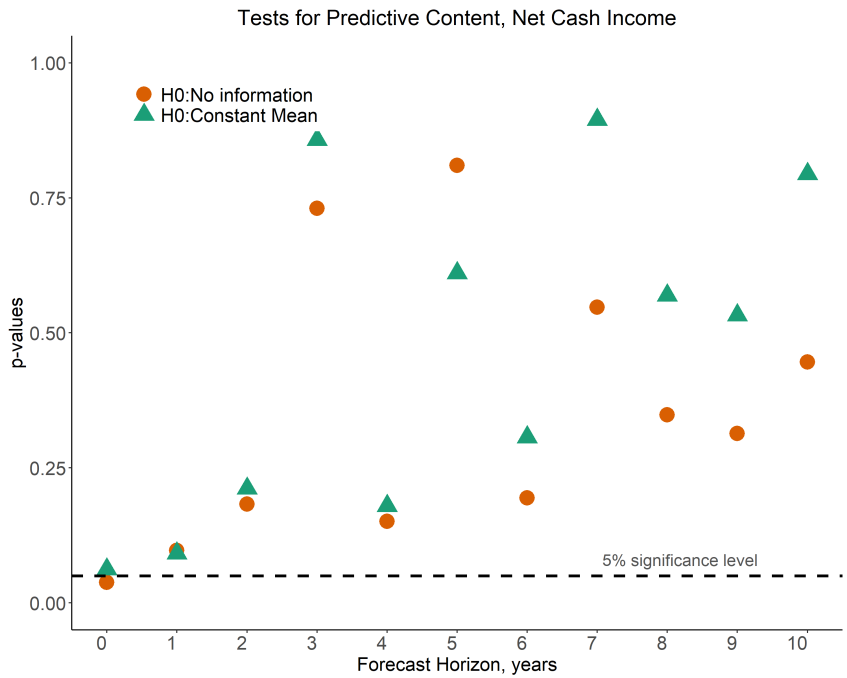






Predictive Content for Farm Income





Maximum Informative Forecast Horizon

Table 1: Maximum informative forecast horizons for selected variables

	Variable	No information test	Constant Mean test
Corn	Price	0	0
	Yield		
	Acreage		
Soybean	Price	2	2
	Yield	0	
	Acreage	0	
Wheat	Price	0	0
	Yield	0	0
	Acreage	1	1
Farm Income	Net farm income	0	1
	Net cash income	0	
	Crop receipts	0	
	Livestock receipts	0	0
	Govt. payments	0	
	Cash expenses	0	

Notes: h* calculated at 5% significance level

Conclusions and Implications

- The information content of the USDA baseline projections diminishes after the nowcast of the current year.
- Even in the best case scenarios, the projections remain informative only up to two to three years ahead and become uninformative after that.
- These results inform USDA forecasters who create these baseline projections.

Questions?

Please email Siddhartha Bora at bora.19@osu.edu.

References

Bora, S.S., A.L. Katchova, and T.H. Kuethe. 2020. "The Rationality of USDA Forecasts under Multivariate Asymmetric Loss." *American Journal of Agricultural Economics* n/a.

Breitung, J., and M. Knüppel. 2020. "How far can we forecast? Statistical tests of the predictive content." Bundesbank discussion paper, Frankfurt a. M.

Martinez, A. 2020. "Testing for Differences in Path Forecast Accuracy." Working paper, Office of Macroeconomic Analysis, U.S. Department of the Treasury, January.