The Accuracy and Informativeness of Agricultural Baselines

Ani Katchova

The Ohio State University

Sabbatical/visiting scholar at ERS for Ani Katchova
Duration: August 16, 2021 to May 15, 2022
...
Cooperative Agreement between USDA-ERS and Ohio State University
Cooperator: Ani Katchova, The Ohio State University
Graduate Research Assistants: Siddhartha Bora, Rabail Chandio, Kexin Ding, Xiaoyi Fang
Duration: August 16, 2021 to August 31, 2023

Project Overview

Projects for the sabbatical and cooperative agreement

- Accuracy and informativeness of ag baselines - Bora, Katchova, and Kuethe (AJAE second revise and resubmit)
- Revisions in ag baselines - Ding and Katchova
- Herding in the USDA baselines - Chandio and Katchova
- Baselines using Deep Learning - Bora and Katchova
- Comparison of USDA and OECD baselines - Fang and Katchova
- 5 outreach reports on ag baselines
- papers and reports posted on the OSU website: https://aede.osu.edu/our-people/ani-katchova

Accuracy and Informativeness of Ag Baselines - Bora, Katchova, and Kuethe

- USDA's statistical agencies such as NASS and ERS provide forecasts of agricultural production, prices, trade, use, inventories, and farm income.
- Stakeholders eagerly await USDA forecasts such as Farm Income Forecasts and WASDE forecasts.
- Previous studies suggest that many USDA forecasts are biased and/or inefficient.
- Ag baseline projections provide information about the farm sector for the next decade.
- Importance of the long-run information as the economic recovery from the pandemic continues.
- Implications for the Farm Bills which usually run at 5-year cycles.

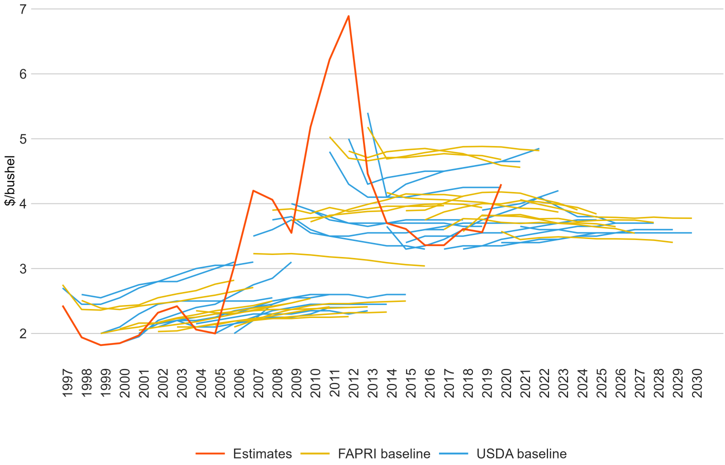
Agricultural Baselines

- The 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.
- Important for understanding the status of the economy several quarters or years from the current year.
- Informative for formulating policy such as preparing the President's budget and program allocations.
- Produced by the USDA Interagency Agricultural Projections Committee, comprising experts from 10 USDA agencies and offices in February every year.
- A composite of model results and judgment-based analysis.
- Food & Agricultural Policy Research Institute (FAPRI), University of Missouri is another source of agricultural baseline projections for the US.

Evaluating Baseline Projections

- Previous Literature:
  - Despite their importance in shaping agricultural policy, USDA's baseline projections have not been rigorously evaluated in the literature.
  - Boussios, Skorbiansky, and MacLachlan (2021) evaluated the projection errors for harvested area.
  - Kuethe, Bora, and Katchova (2022) show that short-term projections in the USDA baseline report contain information that might improve ERS farm income forecasts released in February.
- Our study:
  - Examines accuracy and bias in USDA and FAPRI baseline.
  - Tests whether the baseline projections contain any useful information beyond a certain year into the future, and determines the maximum informative projection horizon.
  - Compares USDA and FAPRI baselines by taking into account all the projection horizons together.

Baseline projections for corn farm price vs realized values



Data and Notations

- Data include USDA and FAPRI baselines from 1997 to 2020.
- The baseline reports contain estimates for the previous year(s), and projections for the next 10 years including the current year.
- We examine two main projections tables:
  - Harvested acres, farm price, and yield of three major commodities: corn, soybeans, and wheat.
  - Net cash income and its components: crop receipts, livestock receipts, direct government payments, farm-related cash income, and cash expenses.

Data and Notations

- We use natural logarithms of the variables for our analysis: $\hat{y}^i_{t+h|t} = \ln(\hat{Y}^i_{t+h|t})$, where $\hat{Y}^i_{t+h|t}$ is the projection made in year t for year $t+h$ by the agencies $i = \{USDA, FAPRI\}$.
- Similarly, we use log transforms of the realized values: $y_t = \ln(Y_t)$, where Y_t is the realized value for year t.
- The projection horizon h can take values between 0 and 9, where $h = 0$ stands for current year.

Accuracy

- For each variable, the percent projection error at horizon h is defined as: $e^i_{t+h|t} = 100 \times (Y_{t+h} - \hat{Y}^i_{t+h|t})/Y_{t+h}$, where $i = \{FAPRI, USDA\}$
- Mean absolute percent error (MAPE) and root mean squared percent error (RMSPE) are defined as,

$$\text{MAPE}^i_h = \frac{1}{T}\sum_{t=1}^{T}|e^i_{t+h|t}| \tag{1}$$

$$\text{RMSPE}^i_h = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(e^i_{t+h|t})^2} \tag{2}$$

Average Percent Errors

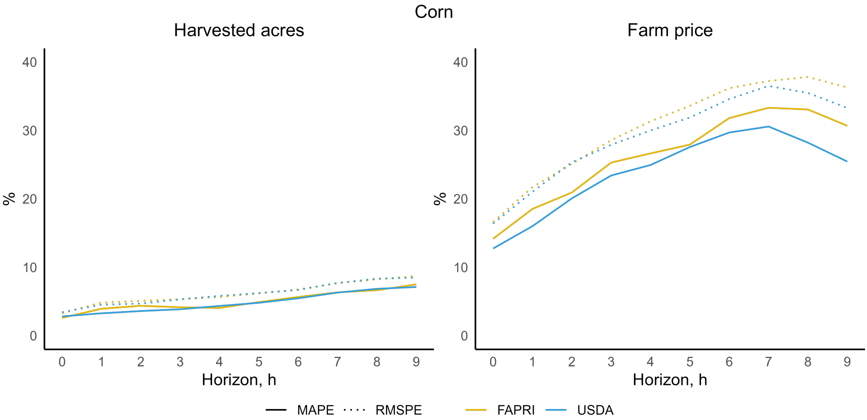


Figure 1: MAPE and RMSPE of corn harvested acres and average farm price

Average Percent Errors

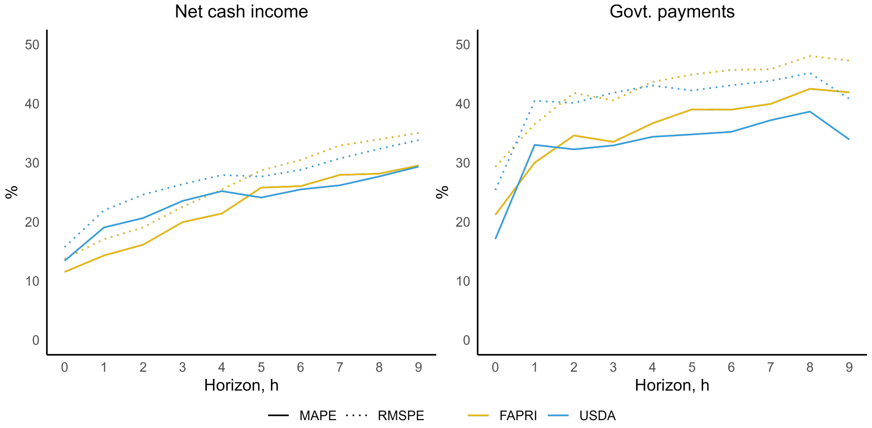


Figure 2: MAPE and RMSPE of net cash income and government payments

Bias

- Bias: For each series of projections from both agencies, we run the following regression (Holden and Peel, 1990) for $h \in \{0, 1, \ldots 9\}$:

$$y_{t+h} - \hat{y}^i_{t+h|t} = \alpha^i_h + \varepsilon^i_{t+h}. \tag{3}$$

- Estimated using Ordinary least squares (OLS) with an HAC standard error (Newey and West, 1987).

Bias

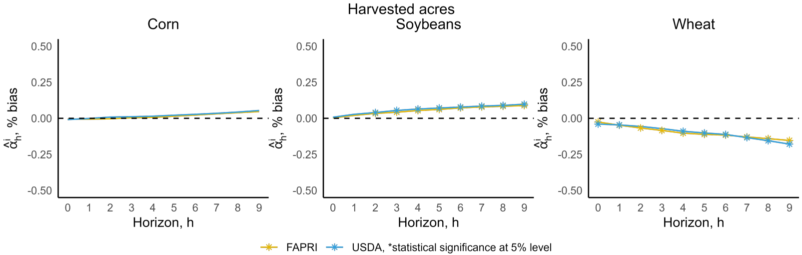


Figure 3: Bias in projections of harvested acres of corn, soybeans and wheat

- Corn harvested acres are unbiased, soybean harvested acres show downward bias, and wheat harvested acres show upward bias, similar to Boussios, Skorbiansky, and MacLachlan (2021).

Bias

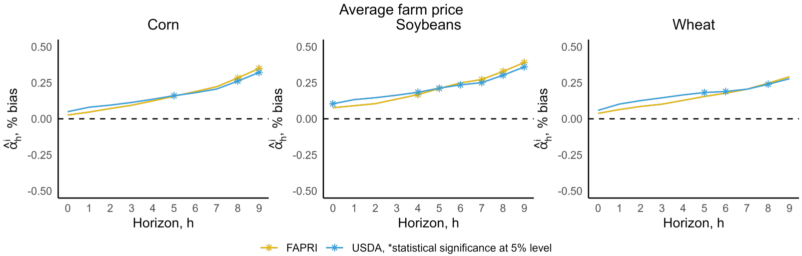


Figure 4: Bias in projections of average farm price of corn, soybeans and wheat

- Bias in average farm price at longer horizons.

Bias

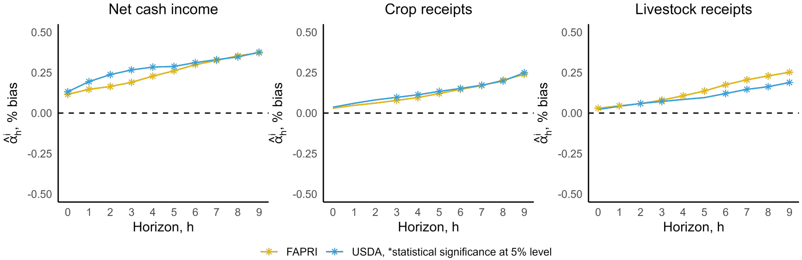


Figure 5: Bias in projections of net cash income components

- Net cash income bias increase with horizon.
- Net cash income components also show bias, and of higher magnitude at larger horizons.

Tests for predictive content in the baseline projections

- The tests for informativeness compares the mean-squared prediction error of the projections to the unconditional variance of the target variable (Breitung and Knüppel, 2021).
- We test the following hypothesis,

$$H_0 : E(y_{t+h} - \hat{y}_{t+h|t})^2 \geq E(y_{t+h} - \mu)^2, \text{ for } h > h^* \text{ and } t \in \{1, \dots, T\} \tag{4}$$

$$H_1 : E(y_{t+h} - \hat{y}_{t+h|t})^2 < E(y_{t+h} - \mu)^2 \tag{5}$$

  where, $\mu = E(y_t)$ is the unconditional mean.
- The null hypothesis, termed as *no information* hypothesis, states that there exists a maximum forecast horizon h^* beyond which the process y_t is unpredictable with respect to the information set I_t.
- If the projection $\hat{y}_{t+h}$ is identical to the conditional mean $\mu_{h,t}$, then the *no information* hypothesis is equivalent to the hypothesis that the conditional expectation is constant within the evaluation sample (*constant mean* hypothesis).

$$H_0 : E(\hat{y}_{t+h}|I_t) = \mu_{h,t} = \mu, \text{ for } h > h^* \text{ and } t \in \{1, \dots, T\} \tag{6}$$

Tests for predictive content in the baseline projections

- 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.
- The *no information* hypothesis is equivalent to testing the null hypothesis $\beta \leq 0.5$ against the alternative $\beta > 0.5$ in the regression,

$$y_{t+h} = \alpha_h + \beta_h \hat{y}_{t+h|t} + \nu_{t+h} \tag{7}$$

- The *constant mean* hypothesis is equivalent to testing $\beta = 0$ in the same regression.
- The tests of the parameters β_h can be performed using a HAC t-statistic.

Tests for predictive content, average farm price

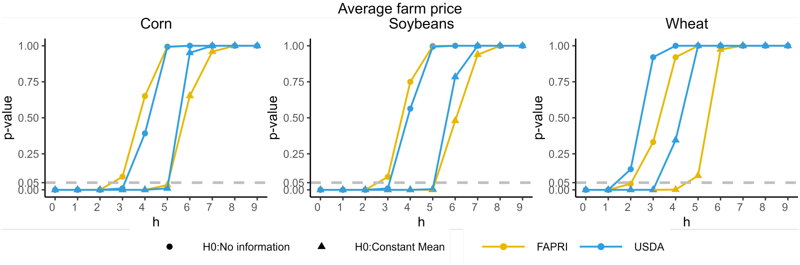


Figure 6: Tests for predictive content for farm prices of corn, soybeans and wheat

- As per *no information test*, corn farm price projections stay informative upto $h = 2$ for FAPRI and $h = 3$ for USDA. Same for soybeans.
- For wheat farm price projections, $h^* = 2$ and $h^* = 1$ for FAPRI and USDA respectively.

Tests for predictive content, net cash income

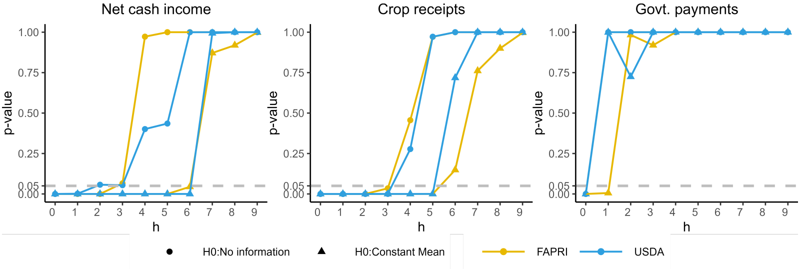


Figure 7: Tests for predictive content for net cash income components

- Net cash income projections stay informative upto $h = 2$ for FAPRI and $h = 1$ for USDA.
- Crop receipts projections stays informative upto $h = 3$ for FAPRI and $h = 3$ for USDA.
- Government payments are difficult to predict beyond current year.

Maximum informative projection horizons

Maximum informative projection horizons, h^*

	H0:No information		H0: Constant mean	
	FAPRI	USDA	FAPRI	USDA
Corn				
Harvested acres	5	5	7	7
Farm price	2	3	5	5
Yield	6	9	9	9
Soybean				
Harvested acres	3	1	9	8
Farm price	2	3	5	5
Yield	9	9	9	9
Wheat				
Harvested acres	3	3	9	9
Farm price	2	1	4	3
Yield	9	8	8	8
Farm income				
Net cash income	2	1	6	6
Crop receipts	3	3	5	5
Livestock receipts	4	3	7	7
Govt. payments	0	0	1	0
Farm-related income	7	8	7	8
Cash expenses	4	4	7	7

Multi-horizon comparison of the baselines

- Traditional Diebold-Mariano (DM) type of tests examines the expected loss differential between two projection series according to a loss function (Diebold and Mariano, 1995).
- For multiple horizon projections such as the baseline, these tests might give inconsistent results across horizons.
- We use the tests of multi-horizon *superior predictive ability* proposed by Quaedvlieg (2021) which jointly consider all horizons of the entire projection path.
- For models $i \in \{USDA, FAPRI\}$, we denote the vector of projections by, $\hat{\boldsymbol{y}}_{i,t} = [\hat{y}^1_{i,t}, \hat{y}^2_{i,t}, \ldots, \hat{y}^H_{i,t}]$, where $\hat{y}^h_{i,t}$ is model i's projection of $\boldsymbol{y_t}$ based on the information set at forecast horizon h.
- A loss function $\boldsymbol{L}_{i,t} = L(\boldsymbol{y_t}, \hat{\boldsymbol{y}}_{i,t})$ maps the projection errors into an H-dimensional vector.
- The loss differential between the USDA and FAPRI is also an H-dimensional vector,

$$\boldsymbol{d}_t = \boldsymbol{L}_{USDA,t} - \boldsymbol{L}_{FAPRI,t} \tag{8}$$

Multi-horizon comparison of the baselines

- The comparison of the two projections are based on the mean loss differential between them, $\boldsymbol{\mu} = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} E(\boldsymbol{d}_t)$, which again is an H-dimensional vector.
- DM-type tests compare the two models by testing the null hypothesis that the mean loss differential is zero ($H_0 : \mu^h = 0$) using a standard t-test, separately for each horizon.
- Uniform Superior Predictive Ability (uSPA): uSPA of the FAPRI model requires that it is better than USDA model at every projection horizon. Define,

$$\mu^{Unif} = \min_h \mu^h \tag{9}$$

  equivalent to testing the null hypothesis, $H_{0,uSPA} : \mu^{Unif} <= 0$ against the alternative, $\mu^{Unif} > 0$.
- Average Superior Predictive Ability (aSPA): Based on a weighted loss differential, which allows losses at different horizon compensate each other. Define,

$$\mu^{Avg} = \boldsymbol{w}'\boldsymbol{\mu} = \sum_h w_h \mu^h \tag{10}$$

  and test the null hypothesis, $H_{0,aSPA} : \mu_{ij}^{Avg} <= 0$ against alternative $\mu_{ij}^{Avg} > 0$.

Multi-horizon comparison of the baselines

- We use two different weighing schemes: equal weights, and weighted by the variance of the loss differential for each horizon.
- The test-statistic are constructed for uSPA and aSPA as,

$$t_{uSPA} = \min_h \frac{\sqrt{T}\bar{d}^h}{\hat{\omega}} \tag{11}$$

$$t_{aSPA} = \frac{\sqrt{T}\bar{d}^h}{\hat{\zeta}} \tag{12}$$

- The variances $\hat{\omega}$ and $\hat{\zeta}$ are estimated using an HAC-type estimator (Newey and West, 1987).

Multi-horizon comparison of USDA and FAPRI

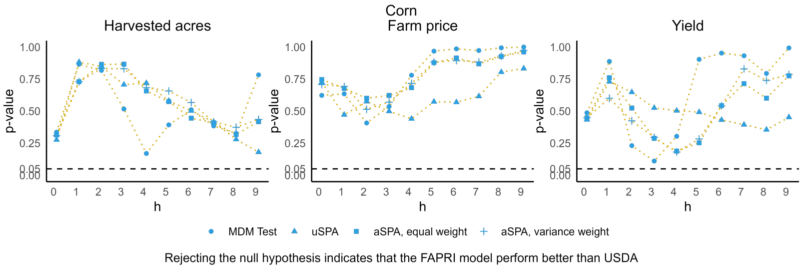


Figure 8: Tests of SPA of FAPRI over USDA for corn projections

- For corn variables, the models do not outperform one another.
- Similar results for soybeans and wheat.

Multi-horizon comparison of USDA and FAPRI

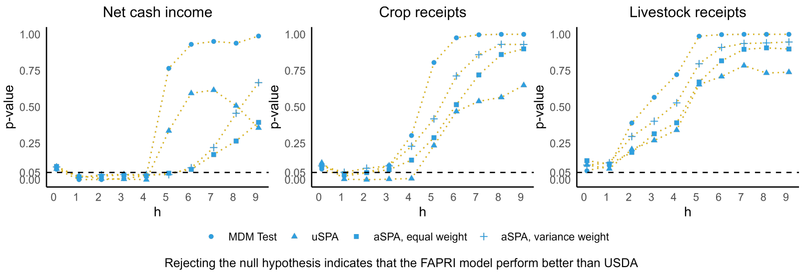


Figure 9: Tests of SPA of FAPRI over USDA for net cash income projections

- The FAPRI model performs better for net cash income projections at shorter projection horizons, but not when the entire projection path is considered.

Revisions in Agricultural Baselines - Ding and Katchova

- Objective
  - To evaluate whether revisions of the baseline projections improve the accuracy of the projections.
- Methods
  - Following Nordhaus (1987), a s-step (s equals the difference between horizons for the same-year projection) revision in the baselines is defined as $R^i_{h|t} = \hat{Y}^i_{h|t} - \hat{Y}^i_{h+s|t}, h = \{0, ..., H_t - 1\}, \text{where} s = \{1, ..., 9\}$. Within the study period from 1997 to 2020: there have been 171 1-step revisions for the commodity projections and 170 1-step revisions for the farm income projections produced by USDA and FAPRI.
  - The mean projection, $\hat{y}_{h|t}$, is defined as $\frac{1}{2}(\ln \hat{Y}^{USDA}_{h|t} + \ln \hat{Y}^{FAPRI}_{h|t})$, the realized value as $y_{h|t} = \ln Y_{h|t}$, and the s-step revision as $r^s_{h|t} = \hat{y}_{h|t} - \hat{y}_{h+s|t}$.
  - The absolute percent errors (APE) is defined as $\text{APE}^i_{h|t} = |100 \times \frac{Y_t - \hat{Y}^i_{h|t}}{Y_t}|$.

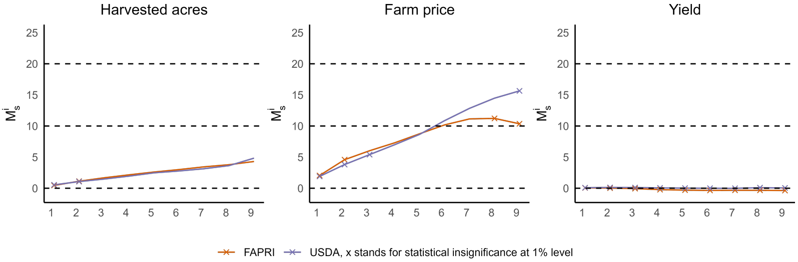


Figure 10: Differences in Absolute Errors of Revised Baseline Projections of corn by revision step *s*, 1997–2020

- For harvested acres, the reductions in errors made by 1-step and 2-step revisions for all crops for both USDA and FAPRI are insignificant. For farm price, the reduction from small-step revisions is insignificant. No significant reduction in errors for all-step revisions for yield.

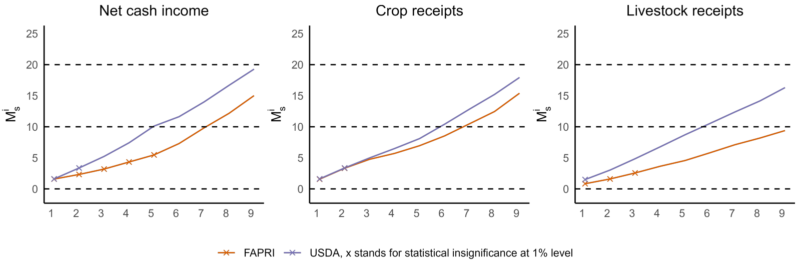


Figure 11: Differences in Absolute Errors of Revised Baseline Projections of net cash income, crop receipts, and livestock receipts by revision step *s*, 1997–2020

- For farm income projections, the revisions made by FAPRI show less significant reduction in projection errors.

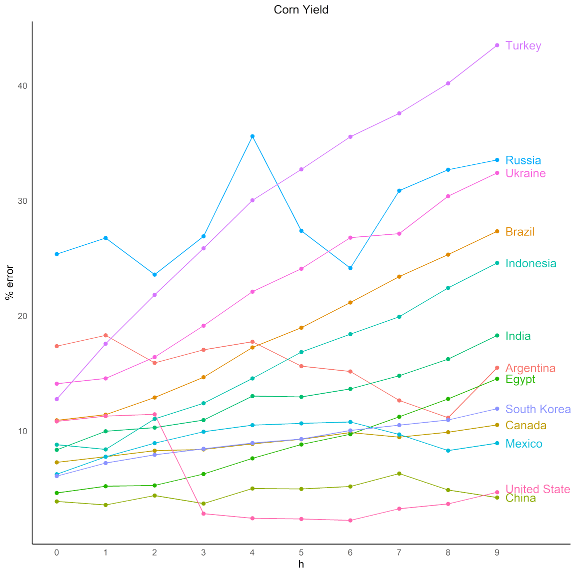
Herding in the USDA Baselines - Chandio and Katchova

- **Objective**
  - To examine whether the baseline projections are grouped together for certain crops across different countries (i.e. herding behavior), producing similar projection trends, and whether that contributes to bias.
- **Data**
  - Yield, area harvested, ending stocks, imports, exports, and total consumption for three major commodities: corn, soybeans, and wheat from USDA International Baseline Projections since 2002.
- **Methods**
  - To assess accuracy, we use root mean squared percentage error:

$$RMSPE_{rcvh} = \left(\frac{1}{T} \sum_{t} (100(\hat{Y}_{rcvth} - Y_{rcvth})/Y_{rcvth})^2 \right) \qquad (13)$$

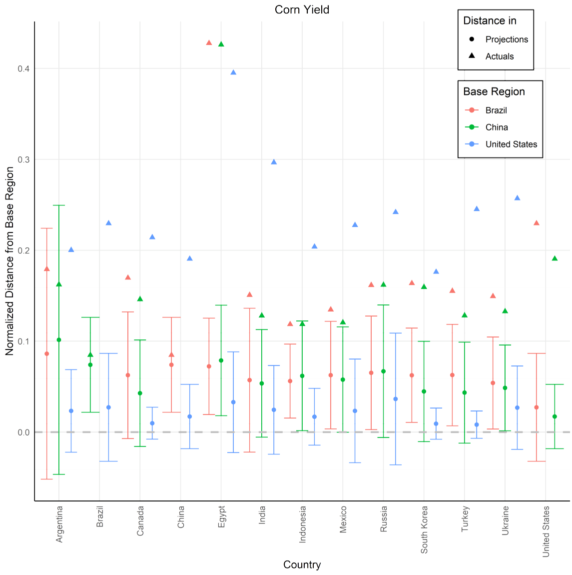
  - Y_{rcvth} is the actual value realized for the projection $\hat{Y}_{rcvth}$
  - We use the dynamic time warping (DTW) algorithm to compute the distances between all available country pairs for each crop-variable-year-horizon and determine whether herding occurs.

Figure 12: RMSPE for corn yield projections of various countries



- Corn yield projections errors for the US remain substantially low for all horizons
- Projections for other countries show higher bias, which increases for longer horizons

Figure 13: Dynamic Time Warping distance between corn yield projections of various countries from US, China, and Brazil



- When comparing projections for other countries to the US, all confidence intervals overlap 0.
- That is, projections for all countries show herding behavior when compared to the US.
- For the countries where realized values are not herded, this increases bias.
- When mapping a relationship between projections correlation with US and bias, we find a positive association for multiple crop-variable combinations.

Baselines using Deep Learning - Bora and Katchova

- What is the issue?
  - Previous studies show that many variables in the baselines are biased.
  - As prediction error increases with horizon, and the predictive content diminishes, and the projections stop being informative.
  - Current baseline models do not utilize information efficiently.
  - Baseline projections process is time-consuming.
- What did the study find?
  - This study compares the performance of various deep learning methods against USDA baseline and a naïve benchmark.
  - Findings suggest that while current baselines perform well in shorter horizons, the deep learning methods perform well in longer horizons.
- How was the study conducted?
  - Deep neural networks were trained using past history of commodity indicators.
  - Performance of the deep neural networks were compared with USDA baseline and naïve benchmark on a test sample.

Deep Learning Methods

Data

- Harvested area and yield for three major commodities: corn, soybeans, and wheat from NASS Quickstats API since 1960.
- USDA baseline projections since 1997.

Prediction problem

- Supervised learning problem where a set of input features $\boldsymbol{X}$ are mapped to an output variable $\boldsymbol{y}$.
- For year t, $\boldsymbol{X_t}$ consists of lagged features while $\boldsymbol{y_t}$ consists of future values starting with year t.
- Due to small sample size we limit our study to upto $H = 5$ forecast horizons.
- $\{\boldsymbol{X}, \boldsymbol{y}\}$ available between 1965–2017. Last $N = 12$ years used as test sample.

Methods

- Naïve no-change benchmark
- USDA baseline
- Long Short-Term Memory Recurrent neural networks (LSTM-RNN)
- Encoder Decoder LSTM
- Convolutional neural networks(CNN)-LSTM

Evaluation Criteria

- Root mean squared error (RMSE)
- Mean absolute percent error (MAPE)

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(A_i - F_i)^2}$$

$$MAPE = \frac{100}{N}\sum_{i=1}^{N}\left|\frac{A_i - F_i}{A_i}\right|$$

A=actual, F=forecast

Deep Learning Results

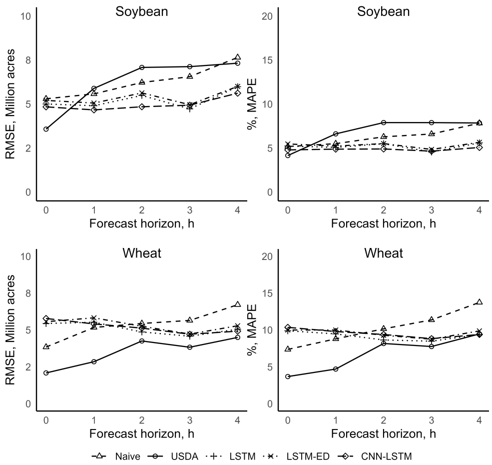


Figure 14: Comparison of accuracy of different methods for harvested area projections

- For soybean harvested area, deep learning methods perform better than USDA baselines for $h > 0$.
- For wheat harvested area, USDA baselines do well at shorter horizons, but deep learning methods improve by $h = 4$.
- USDA baselines perform better for corn harvested area, and yield of the three crops.
- CNN-LSTM shows most promise among the three deep learning methods.

Summary and Conclusions

- Accuracy: Projection error increases with horizon. Bias in net cash income components is consistent bias shown for USDA net cash income forecasts (Bora, Katchova, and Kuethe, 2021).
- Informativeness: the predictive content of the baseline projections start diminishing 4-5 years from current year.
- Multi-horizon comparisons: Except for net cash income projections at shorter horizons, neither USDA nor FAPRI projections outperform each other.
- Baseline revisions: Reductions in projection errors for small-step revisions are insignificant.
- Herding in baselines: Projections of all countries show herding behavior when compared to U.S. which increases bias.
- Deep learning methods perform better than USDA baselines for longer horizons than 4 years.

Implications

- Our findings are relevant for future revisions of the USDA baseline models and processes.
- Underlines importance of stochastic baselines and conditional scenarios.
- Implications for various market participants who use these projections.
- Implications for potential extension of projection horizon for climate change applications.

THANK YOU

For any questions: Ani Katchova, katchova.1@osu.edu

Many thanks to ERS-MTED for providing me with a sabbatical and a cooperative agreement to work on these projects.

References

Bora, S.S., A.L. Katchova, and T.H. Kuethe. 2021. "The Rationality of USDA Forecasts under Multivariate Asymmetric Loss." *American Journal of Agricultural Economics* 103:1006–1033.

Boussios, D., S.R. Skorbiansky, and M. MacLachlan. 2021. "Evaluating U.S. Department of Agriculture's Long-Term Forecasts for U.S. Harvested Area." Economic Research Report No. 285, Economic Research Service, USDA, Feb.

Breitung, J., and M. Knüppel. 2021. "How far can we forecast? Statistical tests of the predictive content." *Journal of Applied Econometrics* 36:369–392.

Diebold, F.X., and R.S. Mariano. 1995. "Comparing predictive accuracy." *Journal of Business & Economic Statistics* 13:253–263.

Holden, K., and D.A. Peel. 1990. "On testing for unbiasedness and efficiency of forecasts." *The Manchester School of Economic & Social Studies* 58:120–127.

Kuethe, T.H., S.S. Bora, and A.L. Katchova. 2022. "Improving ERS's Net Cash Income Forecasts using USDA Baseline Projections." *Journal of Agricultural and Resource Economics*, pp. .

Newey, W.K., and K.D. West. 1987. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation." *Econometrica* 55:703–708.

Quaedvlieg, R. 2021. "Multi-Horizon Forecast Comparison." *Journal of Business & Economic Statistics* 39:40–53.