Evaluating USDA Agricultural Forecasts

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Siddhartha S. Bora, M.S.

Graduate Program in Department of Agricultural, Environmental, and Development Economics

The Ohio State University

2022

Dissertation Committee:

Ani L. Katchova, Ph.D., Advisor

Wuyang Hu, Ph.D.

Brian E. Roe, Ph.D.

Todd H. Kuethe, Ph.D.

Abstract

The timely availability of accurate forecasts plays a vital role in informing decisions by farm sector stakeholders. In this dissertation, I evaluate the rationality, accuracy, and informativeness of a range of agricultural forecasts and projections published by the United States Department of Agriculture (USDA) and other agencies and examine ways to improve them. The findings have implications for future revisions of the forecasting processes and for policymakers, agricultural businesses, and other stakeholders who use these forecasts.

In Chapter 1, I show that some of the reported biases and inefficiencies in USDA forecasts may be due to an asymmetric loss of the forecaster. Many previous studies suggest that many USDA forecasts are biased and/or inefficient. These findings, however, may be the result of the assumed loss function of USDA forecasters. I test the rationality of the USDA net cash income forecasts and the World Agricultural Supply and Demand Estimates (WASDE) production and price forecasts between 1988-2018 using a flexible multivariate loss function that allows for asymmetric loss and non-separable forecast errors. My results provide robust evidence that USDA forecasters are rational expected loss minimizers yet demonstrate a tendency to place a greater weight on under- or over-prediction. As a result, this study provides an alternate interpretation of previous findings of forecast irrationality.

Agricultural baselines play an important role in shaping agricultural policy by providing information about the farm sector for a ten-year horizon, yet these projections have not been rigorously evaluated. In Chapter 2, I evaluate the accuracy and informativeness of two widely used baselines for the US farm sector published by the USDA and the Food and Agricultural Policy Research Institute (FAPRI) in three steps. First, I examine the average percent errors of the projections and perform tests of bias. Second, I use a novel testing framework based on the encompassing principle to test the predictive content of the projections for each horizon, determining the longest informative projection horizon. Third, I compare the USDA and FAPRI baseline projections using a multi-horizon framework that considers all projection horizons jointly. I find that prediction error and bias increase with the horizon's length. The predictive content of the baselines projections for most variables diminishes after 4-5 years. The multi-horizon comparison suggests that neither USDA nor FAPRI projections have uniform or average superior predictive ability over the other for most variables.

Multi-step forecasts about commodity market indicators play an important role in informing policy and investment decisions by governments and market participants. In Chapter 3, I examine whether the accuracy of long-term forecasts can be improved using deep learning models. I first formulate a supervised learning problem and set benchmarks for forecast accuracy. I train a set of deep neural networks on a training sample and measure their performance against the benchmark model on a test sample using a walk-forward validation strategy. I find that while the USDA baseline projections perform better for the shorter horizon, the performance of the deep neural networks improves for the longer forecast horizons.

Acknowledgments

I would like to thank everyone who helped me with my research throughout graduate school. First, I would like to thank my advisor Dr. Ani Katchova for the constant support that I could rely on during my doctoral studies. As a member of the Farm Income Enhancement Program led by Dr. Katchova, I had the opportunity to share the findings of my research with farmers, the media, and other stakeholders. Her supervision helped me maintain an organized work style and contribute efficiently to the team. She always had high expectations for me, keeping me motivated and elevating the quality of my research. I am grateful for her immense contribution to my growth as a researcher.

I thank all of my committee members for their help in improving the quality of my research. I extend my special thanks to Dr. Todd Kuethe, from whom I learned a lot about the art of writing. Dr. Brian Roe has always been generous with his insightful comments, and my work benefitted greatly from the hallway conversations with him. I thank Dr. Wuyang Hu for his valuable feedback on my research and his wholehearted support. I am thankful to Dr. Abdoul Sam for his comments on my research during the candidacy examination.

I want to express my gratitude to the present and the past members of the Farm Income Enhancement Program, who had often been the first to comment on the early stages of my research. I am proud of the shared knowledge I helped create as part of this team. I extend my special thanks to Dr. Robert Dinterman, who greatly influenced my analysis workflow. I thank the participants of the conferences and seminars where I presented my research for their helpful comments on my work. The feedback from the participants of the doctoral seminar in my department was instrumental in improving my research.

I would like to thank my family for being there for me as a source of emotional support. I would especially like to thank my classmates for their friendship and support throughout graduate school. I thank Dr. Zoë Plakias for always providing advice on research and professional development. I thank many others in the department, including Dr. Leah Bevis, Dr. Sathya Gopalakrishnan, and Dr. Allen Klaiber, for their support during various stages of graduate school.

Vita

2011 .... B. Tech., Electronics and Communications Engineering, National Institute of Technology (NIT), Warangal, India

2015 .... Post Graduate Diploma in Rural Management, Institute of Rural Management Anand (IRMA), India

2019 .... M.S. in Agricultural, Environmental and Development Economics, The Ohio State University.

2017-present .... Graduate Research Associate, Agricultural, Environmental, and Development Economics, The Ohio State University, Columbus, Ohio

Publications

Research Publications

Kuethe, T.H, S.S. Bora, and A.L. Katchova. "Improving ERS Net Cash Income Forecasts using USDA Agricultural Baseline Projections." *Journal of Agricultural and Resource Economics*, Volume 47, Issue 2, May 2022, Pages 246-261 [link]

Bora, S.S., A.L. Katchova, and T.H. Kuethe. "The Rationality of USDA Forecasts under Multivariate Asymmetric Loss." *American Journal of Agricultural Economics* (2021), 103, 3:1006-1033. [link

]

Fields of Study

Major Field: Agricultural, Environmental and Development Economics

Table of Contents

Page

List of Tables

Table **Page**

List of Figures

Chapter 1: The Rationality of USDA Forecasts under Multivariate Asymmetric Loss

> "The U.S. Agriculture Department said on Wednesday it had pulled all staff from an annual crop tour after an employee was threatened, and three sources said the threat of violence was made during a phone call from an angry farmer."
>
> Huffstutter and Polansek (2019)

1.1 Introduction

The Federal government's statistical agencies and programs generate a large volume of data that "the public, businesses, and governments need to make informed decisions" (Office of Management and Budget, 2020, pp. 3). The U.S. Department of Agriculture (USDA) plays an important role within the Federal government's statistical agencies and programs. The USDA is responsible for producing a variety of principal federal economic indicators. Further, the USDA is home to two of the Federal government's thirteen principal statistical agencies, the National Agricultural Statistics Service (NASS) and the Economic Research Service (ERS), which accounted for approximately 8% of the $3.2 billion appropriated to primary statistical agencies in fiscal year 2017 (Office of Management and Budget, 2020). In order to provide timely information for decision-makers across the agricultural sector, the USDA's statistical agencies and programs provide forecasts of agricultural production, prices, trade, uses, inventories, and farm income. The existing literature, however, suggests that many USDA forecasts are not rational (Bailey & Brorsen, 1998; Isengildina, Irwin, & Good, 2006; Isengildina-Massa, MacDonald, & Xie, 2012; Kuethe, Hubbs, & Sanders, 2018; Sanders & Manfredo, 2003; Xiao, Hart, & Lence, 2017). The overwhelming evidence of irrationality may lead many forecast users to question the usefulness of USDA forecasts.

In this study, we examine the degree to which prior findings of irrationality in USDA forecasts can be attributed to assumptions researchers make about the costs of forecast errors, by examining the forecaster's loss function. The conventional practice is (i) to assume *a priori* that USDA forecasts are generated to minimize a mean squared error (MSE) loss function and (ii) to test the weak form conditions of rationality under MSE loss.[1] Many of the weak form conditions of forecast rationality, however, do not hold under other loss functions (Patton & Timmermann, 2007). As a result, rejections of forecast rationality may be due to misspecified loss functions, rather than lack of rationality (Elliott, Timmermann, & Komunjer, 2005). Our empirical approach, by contrast, only assumes that the forecaster's loss function belongs to a flexible class of loss functions, of which MSE is a special case. We then back out the parameters of the loss function that are consistent with the observed forecasts.

As Auffhammer (2007) argues, a forecast is only optimal for a particular forecast user when his or her loss function matches that of the forecast producer. As

[1]Notable exceptions include evaluations of USDA's interval forecasts of commodity prices, including Sanders and Manfredo (2003), Isengildina, Irwin, and Good (2004), and Isengildina-Massa and Sharp (2012). These studies examine the proportion of actual market prices that fall in the forecasted range, or "hit rate," of various USDA interval forecasts. In a spirit similar to the current study, Isengildina et al. (2004) examine the degree to which inaccuracy of USDA's interval forecasts of corn and soybean price forecasts can be attributed to inefficient use of information or the utility function of the forecasters. The authors find evidence of the latter.

a result, we empirically estimate the loss functions associated with two of USDA's most prominent forecasts from 1988 through 2018. First, we examine USDA's forecasts of annual net farm income and its components. The USDA's farm income forecasts are among the department's most cited statistics (McGath et al., 2009a). They are closely monitored by various farm sector stakeholders, including farm input and machinery suppliers, lenders, other farm related industries, and state and local governments (Dubman, McElroy, & Dodson, 1993). In addition, USDA's farm income forecasts are frequently cited in farm policy debates in Congress and are inputs in numerous other statistical models, such as the estimation of gross domestic production (McGath et al., 2009a). Second, we examine a set of production and price forecasts from the World Agricultural Supply and Demand Estimates (WASDE) for three major commodities: corn, soybeans, and wheat. WASDE is an important source of information for commodity production, consumption, trade, and prices. The existing literature provides robust evidence that WASDE releases move commodity markets (Adjemian, 2012; Adjemian & Irwin, 2018; Dorfman & Karali, 2015; Fortenbery & Sumner, 1993; Isengildina-Massa, Cao, et al., 2020; Isengildina-Massa, Irwin, Good, & Gomez, 2008; Karali, Isengildina-Massa, Irwin, Adjemian, & Johansson, 2019; McKenzie, 2008; Sumner & Mueller, 1989). Further, the USDA production forecasts were subject to intense scrutiny during the recent 2019 growing season. USDA's corn acreage and yield forecasts were even believed to motivate threats of violence against USDA employees by disgruntled farmers (Huffstutter & Polansek, 2019).

This study makes a number of important contributions to the literature. First, our empirical approach relaxes the assumptions of MSE loss. We build on the generalized method of moments (GMM) framework of Elliott et al. (2005) that jointly estimates the forecaster's loss function and tests for rationality of the forecasts. Elliott et al.'s method is flexible as it allows for several parameterizations, as well as asymmetry in the loss function due to differential costs of over- and under-prediction. As detailed in the next section, there are a number of reasons to believe that the bias and inefficiencies documented in USDA forecasts may be the direct result of asymmetric costs of over- and under-prediction.

Second, we evaluate forecasts in a multivariate framework. The overwhelming majority of prior evaluations of USDA forecasts test rationality on each variable independently, even though the forecasts are released as joint forecasts of several variables.[2] This practice implicitly assumes that the marginal costs for forecast errors in one variable are independent of the costs for other variables or that there are no interactions between variables. Thus, previous researchers implicitly assume separable loss functions. We instead apply the estimation procedure of Komunjer and Owyang (2012) that generalizes the approach of Elliott et al. (2005) to a multivariate setting with non-separable loss. Both net farm income and WASDE are based on accounting equations, so the forecast errors of each variable likely depend on the errors of other variables. For example, within the farm income forecast, the costs associated with over-predicting cash receipts are compounded when jointly under-predicting cash expenses, or within WASDE, the costs associated with over-predicting yields are compounded when jointly over-predicting acreage.

Third, while our results are generally consistent with previous findings, our analysis yields an alternative interpretation of the results of prior research. Batchelor (2007)

[2]Isengildina-Massa, Karali, Kuethe, and Katchova (2020) is one recent exception.

draws a distinction between a rational *forecaster* and a "technically" rational *forecast*. A rational forecaster uses all available information when constructing a forecast, as in Muth (1960), and a forecast is "technically" rational when it is unbiased and efficient, as in Diebold and Lopez (1996). Batchelor (2007) identifies three possible explanations why rational forecasters may publish "technically" irrational forecasts. One, the forecaster may lack the skill to use information efficiently and learn from forecasting errors. Two, the forecaster may have the skill to use information efficiently, but the forecaster's information set is insufficient. Three, the forecaster may have skill and sufficient data but responds to incentives to make an optimistic or pessimistic forecast, namely, the forecaster responds to asymmetries in the consequences of over- versus under-prediction. We find robust evidence that USDA forecasts are generated to minimize an asymmetric loss function. While many USDA forecasts are technically irrational under traditional tests (biased and/or inefficient), this study suggests that USDA forecasters are rational expected loss minimizers under asymmetric loss. As Keane and Runkle (1990, pp. 719) state, "if forecasters have differential costs of over- and under-prediction, it could be rational for them to produce biased forecasts. If we were to find that forecasts are biased, it could still be claimed that forecasters were rational if it could be shown that they had such differential costs."

1.2 Background

As Elliott and Timmermann (2008, pp. 8) suggest, "no forecast is going to always be correct, so a specification of how costly different mistakes are is needed to guide the procedure." The loss function is a mathematical representation of the costs associated with forecast errors, and forecasters generate projections that minimize the expected loss function. Elliott and Timmermann (2008) argue that the loss function interpretation of forecast evaluation is valid when (i) forecasters care about the accuracy of their forecasts, and (ii) forecasters can adjust their forecasts in a way that incorporates any costs associated with forecast errors. The two most common loss functions employed in the existing literature are mean squared error (MSE) and mean absolute error (MAE).

Forecast evaluation under MSE is particularly popular because rationality implies a number of properties that can be easily tested empirically. A rational forecast under MSE loss is unbiased, forecast errors are serially uncorrelated, and the unconditional variance of the forecast error is a non-decreasing function of the forecast horizon (Diebold & Lopez, 1996). However, many forecasts fail to meet these conditions. Traditional tests of forecast rationality under MSE loss suffer from the "joint hypothesis problem," as rejection of rationality stems from either irrationality of the forecasts or a misspecified test (Fritsche, Pierdzioch, Rülke, & Stadtmann, 2015). Thus, traditional rationality tests may be misspecified with respect to the assumed loss function of the forecaster.

The loss function is sometimes referred to as the utility function of the forecast producer. Kahneman and Tversky (1973) argue that forecasters may intentionally produce technically irrational forecasts because of behavioral biases in information processing. For example, when forecasters have a large utility of a positive outcome, they may assign greater probability weights to some values out of anticipation, hope, or greed (Weber, 1994). Similarly, when forecasters have a large disutility of a negative outcome, they may assign greater probability weights to some values out of fear of the negative consequences associated with underestimating probability (Weber, 1994).

The asymmetries in probability weights mirrors the asymmetric reaction to gains and losses (Kahneman & Tversky, 1979). Asymmetries in the consequences of over- or under-prediction of uncertain quantities are frequently referred to as *asymmetric loss functions*. Weber (1994) demonstrates that asymmetric loss functions can be derived from either an expected utility or a rank-dependent utility framework.

West, Edison, and Cho (1993) argue that an asymmetric loss function is a natural candidate to evaluate forecasts when one seeks to emulate a utility-function-based approach to forecast evaluation. Under asymmetric loss functions, the optimal forecast is the conditional mean (MSE) or median (MAE) plus an optimal bias term (Christoffersen & Diebold, 1997; Granger, 1969, 1999; Zellner, 1986b). The size of the optimal bias will depend on the parameters of the loss function (Granger, 1969). In addition, the forecast errors will not be orthogonal to variables in the forecaster's information set (Batchelor & Peel, 1998). Patton and Timmermann (2007) assert that optimal forecasts are only unbiased when they meet the "double symmetry" condition, in which both the variable forecasted and the forecaster's loss function are distributed symmetrically.

In addition to internal asymmetric costs, such as anticipation or fear, forecasters may face external asymmetric consequences for forecast errors (Weber, 1994). For example, Laster, Bennett, and Geoum (1999) show that when forecasters are rewarded based on both accuracy and their ability to generate publicity, their efforts to attract publicity may compromise forecast accuracy. In an experimental setting, Maddox and Bohil (1998) show that people react in the appropriate direction to asymmetric payoff functions, but they are often too conservative in their reactions. In addition to publicity, forecasters may derive some benefit from cultivating a reputation as optimists or pessimists (Batchelor & Dua, 1990). For example, when forecasters rely on others for information, optimism may help to build relationships with information providers (Francis, Hanna, & Philbrick, 1997; Francis & Philbrick, 1993; Lim, 2001). Similarly, forecasters may alter their predictions to make their forecasts more attractive to particular client groups or forecast users (Batchelor, 2007). While USDA forecasters may be subject to limited internal asymmetric costs, such as anticipation or fear, they may be subject to external asymmetric consequences for forecast errors, given their reliance on information from farmers and other agricultural sector professionals, who are also USDA forecast users.

Asymmetric loss functions carry important consequences for fixed event forecasts, such as USDA's farm income or WASDE forecasts. Kahneman and Tversky (1973) argue that forecasters may overweight their own past forecasts and under-react to new information. Thus, any bias in initial forecasts will propagate forward. Batchelor (2007) demonstrates that bias in initial forecasts will also propagate forward if forecasters face penalties for forecast revision or are rewarded for consistency. Given that the optimal forecast under asymmetric loss includes an optimal bias, any asymmetric loss early in the forecast process may carry forward throughout later forecast revisions.

An asymmetric loss function in any one forecast may also carry important consequences for later forecasts of other economic variables. When forecasts are made sequentially by different agents, each published forecast becomes part of the information set of the next forecaster. Graham (1999) demonstrates that this process of "information cascades" may lead to herding when later forecasts are biased towards early forecasts.[3] As previously stated, USDA produces a variety of forecasts, and any asymmetries in one USDA forecast may carry over to later USDA forecasts through information cascades. For example, if WASDE forecasts project a significant decline in the production of a particular commodity, this information will likely be incorporated in the USDA farm income forecasts.

Finally, the existing literature offers several explanations as to why forecasts produced by government agencies, such as USDA, may be generated under asymmetric loss. A number of previous studies suggest that government agency forecasts tend to be conservative or cautious (Capistrán, 2008; Caunedo, Dicecio, Komunjer, & Owyang, 2018; Ellison & Sargent, 2012). Government forecasters may be cautious because stability is a crucial economic policy goal (Capistrán, 2008), policy-making may require "worst case scenario" forecasts (Ellison & Sargent, 2012), or over-predicting prosperity may be worse for policymakers than under-predicting (Caunedo et al., 2018). In addition, government forecasts may also be used to stimulate some private sector response (Estrin & Holmes, 1990). For example, Beaudry and Willems (2018) demonstrate that overly optimistic GDP growth forecasts triggers public and private debt accumulation. Finally, it has been argued that government agency forecasts may be used as an instrument to justify a particular policy response (Frankel, 2011; Jonung & Larch, 2006) or to put the incumbent party in a favorable light (Ulan, Dewald, & Bullard, 1995).

As previously stated, traditional forecast evaluation methods test the weak form properties of rationality under an assumed loss function, such as MSE or MAE, yet

[3]Fritsche et al. (2015), conversely, identify an "anti-herding" behavior in professional forecasters where later forecasters strategically differentiate their forecasts from those previously published.

whether one can conclude that bias or inefficiency represents irrationality requires knowledge of the shape of the forecaster's loss function (Keane & Runkle, 1998). A number of studies develop alternative loss functions that account for asymmetry (Batchelor & Peel, 1998; Christoffersen & Diebold, 1996; Granger & Pesaran, 2000; Ulu, 2013; Varian, 1975; Zellner, 1986a).

Elliott et al. (2005), in contrast, develop an alternative forecast evaluation framework that maintains the assumption of loss minimizing behavior and estimates the shape of the loss function, or class of loss functions, that are consistent with the observed forecasts. The method is flexible as it allows for several alternative parameterizations of the loss function, with symmetry as a special case. Elliott et al.'s (2005) method jointly estimates the asymmetry parameters of the forecaster's loss function and tests for rationality of the forecasts. Krüger and LeCrone (2019) show that this method has a high power and is robust to fat tails, serial correlation, and outliers. The method has been used to evaluate forecasts of a number of economic variables by professional forecasters (Aretz, Bartram, & Pope, 2011; Christodoulakis, 2020; Fritsche et al., 2015; Mamatzakis & Koutsomanoli-Filippaki, 2014; Pierdzioch, Reid, & Gupta, 2016; Pierdzioch, Rülke, & Stadtmann, 2013; Tsuchiya, 2016a, 2016b), government agencies (Auffhammer, 2007; Giovannelli & Pericoli, 2020; Krol, 2013; Tsuchiya, 2016a), international organizations (Christodoulakis & Mamatzakis, 2008; Giovannelli & Pericoli, 2020; Tsuchiya, 2016a), and central banks (Ahn & Tsuchiya, 2019; Baghestani, 2013; Capistrán, 2008; Caunedo, Dicecio, Komunjer, & Owyang, 2020; Pierdzioch, Rülke, & Stadtmann, 2015). These studies overwhelmingly suggest that forecasts that are biased or inefficient under MSE loss are rational under asymmetric loss.

1.3 Description of USDA Forecasts

1.3.1 Farm Income Forecasts

Since 1910, the USDA has produced annual estimates of net farm income, a measure of the return to farm operators for their labor, capital, and management after all production expenses are deducted (Lucier, Chesley, & Ahearn, 1986). USDA's official farm income estimates are produced with a significant time lag. They are typically released in August following the reference year. In order to provide more timely information, the USDA produces a series of forecasts each year. The forecasts relate to a calendar year and are typically released in February, August, November, and the following February.[4] The August forecast coincides with the release of the official estimates of the prior year, and the last forecast in February coincides with the release of the first forecast of the new calendar year.

USDA's farm income accounts include a variety of income and wealth measures. Our analysis examines the vector of forecasts related to net cash income. Net cash income (NCI) is a measure of farm-sector earnings, including cash receipts from farming, farm-related income, and government payments less cash expenses. It is calculated using the accounting equation:

$$\begin{aligned}\text{Net Cash Income (NCI)} = &\ \text{Crop Receipts (CR)} + \text{Livestock Receipts (LR)} \\ &+ \text{Direct Government Payments (GP)} \\ &+ \text{Cash Farm-Related Income (FRI)} \\ &- \text{Cash Expenses (EXP)}.\end{aligned} \tag{1.1}$$

[4]USDA may further revise the estimates in subsequent releases, mainly to correct errors or to incorporate information that was not available earlier. However, to maintain consistency, we consider the first official estimates as the realized values throughout the study, following Kuethe et al. (2018) and Isengildina-Massa, Karali, et al. (2020).

Crop receipts include cash receipts from eight major crops, and livestock receipts include four categories (meat animals, dairy products/milk, poultry and eggs, and miscellaneous livestock). Direct government payments represent funds that the Federal Government pays to farmers and ranchers who produce program commodities, participate in resource conservation, and receive compensation for natural disasters. Cash farm-related income includes income from items such as recreational activities, custom work, machine hire, forest products, and other farm sources. The first four items are added up to calculate gross cash income, after which cash production expenses are subtracted to arrive at net cash income. Since cash farm-related income forecasts are not available for most of the years and since cash farm-related income contributes less than 10% to gross cash income, we exclude this variable from our analysis.

McGath et al. (2009a) document the economic model and estimation procedure for each component. The forecast procedure relies on data obtained from a variety of sources, including WASDE, Agricultural Resource Management Survey (ARMS), and NASS. While the data sources remain constant throughout the farm income forecast process, the timing of the release of the forecast revisions is selected to reflect changes in information. For example, the August revision reflects updates in crop production estimates and cash receipts from the USDA's survey-based production and yield estimates, and the November revision reflects updated crop production and harvest information. As time progresses, many of the forecasted values are substituted with the official estimates. For example, by February of next year, the WASDE acreage and yield values are final estimates, however, prices for the marketing year remain forecasts (McGath et al., 2009a).

Kuethe et al. (2018) previously found that USDA's bottom-line net farm income forecasts are biased and inefficient. Specifically, initial forecasts systematically under-predict realized values, and later forecast revisions over-react to new information.[5] Isengildina-Massa, Karali, et al. (2020) extend the work of Kuethe et al. (2018) by examining the vector of net cash income and its components, including crop receipts, livestock receipts, government payments, gross cash income, and cash expenditures. Isengildina-Massa, Karali, et al. (2020) find a similar downward bias in initial net cash income forecasts, which mainly stems from forecasts of crop receipts.

1.3.2 WASDE Forecasts

We also examine USDA production forecasts for area harvested ($Acreage$), yield per harvested acre ($Yield$), and average farm price ($Price$) for three major commodities: corn, soybeans, and wheat. The forecasts were obtained from USDA's WASDE. WASDE is coordinated by the World Agricultural Outlook Board (WAOB) and relies on data and expertise from a variety of USDA agencies including NASS, ERS, Farm Service Agency (FSA), Agricultural Marketing Service (AMS), and Foreign Agricultural Service (FAS). A detailed description of WASDE's balance sheet approach to crop forecast generation is provided by Vogel and Bange (1999).

WASDE forecasts and estimates are produced for marketing year averages. For corn and soybeans, the marketing year is defined as September through August of the following calendar year, and, for wheat, the marketing year is defined as June through May of the following calendar year. WASDE forecasts are released by USDA between the 9th and 12th of each month. The first marketing year forecasts for

[5]Kuethe et al. (2018) examine bottom-line *net farm income* which includes non-cash income and expenses. The differences between the two measures is documented in McGath et al. (2009a).

corn, soybeans, and wheat are released in May. For corn and soybeans, the acreage and yield estimates are finalized in December, and for wheat, the acreage and yield estimates are finalized in September. Season average prices are finalized in November of the following calendar year for corn and soybeans and in September of the following calendar year for wheat.

Isengildina-Massa, Karali, and Irwin (2013) document the bias and inefficiency of WASDE price and production forecasts for corn, soybeans, and wheat. The authors find strong evidence that forecast errors are affected by behavioral and macroeconomic factors. Xiao et al. (2017) similarly show that WASDE forecasts of ending stocks for the same three commodities are inefficient and conservative. Isengildina et al. (2006) also identify informational inefficiencies in NASS corn and soybean production forecasts. Despite the prior findings of bias and inefficiency, Hoffman, Etienne, Irwin, Colino, and Toasa (2015a) find that WASDE projections of season-average corn price provide useful information to the market. The evidence of irrationality for other commodities, however is mixed. Isengildina-Massa et al. (2012) find evidence of bias and inefficiency in WASDE cotton forecasts, yet Lewis and Manfredo (2012) fail to reject the rationality of WASDE sugar production and consumption forecasts.

1.3.3 Preliminary Analysis

Following Isengildina-Massa et al. (2013), forecasts are expressed as percent changes from the previous year to avoid the impact of changing forecast levels over the study period. The percent change is calculated as: $\mathbf{f}_{t,h} = 100 * \ln(\mathbf{F}_{t,h}/\mathbf{F}_{t-1,h})$, where $\mathbf{F}_{t,h}$ is the forecast level for a reference year t and at horizon h months before the final estimate and $\mathbf{F}_{t-1,h}$ is the forecast from the previous year $t-1$ for the same time horizon h. The forecasts for the previous year are replaced with the final estimates after they are released (in August for farm income, November for corn and soybeans, and September for wheat). The percentage forecast errors at each horizon h are defined as the difference between the estimate and the h-horizon forecast, $\mathbf{e}_{t,h} = \mathbf{f}_{t,0} - \mathbf{f}_{t,h}$. For the farm income forecast, $\mathbf{f}_{t,h}$ represents the vector of net cash income and its components, where the horizon h is February, August, November, and next February. The j-th component of the vector of h-horizon forecasts is ${f^j}_{t,h}$, where $j \in \{NCI, CR, LR, GP, EXP\}$. The WASDE forecasts are similarly expressed with h representing each month of May through December for corn and soybeans and May through September for wheat. For all three commodities, $j \in \{Acreage, Yield, Price\}$. The forecast errors are stationary for all forecast series and horizons, as verified using standard tests of stationarity (Dickey & Fuller, 1979; Said & Dickey, 1984).

Figures 1.1 and 1.2 depict the average annual forecast errors for net cash income and its components and for the WASDE acreage, price, and yield, respectively. Net cash income forecast errors exhibit large variation over time, even though errors in cash expenses tend to offset errors in receipts and government payments. Notably, the 2007-2008 period of sharply increasing crop prices resulted in some of the largest forecast errors for crop prices, but not for farm income, acreage, or yield.

Tables 1.1 - 1.4 summarize the forecast errors for each variable and horizon h over the period 1988-2018. The summaries include three common measures of forecast accuracy: the mean absolute percent error $MAPE(|e|) = \frac{1}{T}\sum_{t=1}^{T}|e_t|$, the root mean square percent error $RMSPE(\sqrt{e^2}) = (\frac{1}{T}\sum_{t=1}^{T}e_t^2)^{\frac{1}{2}}$, and the mean percent error $MPE(e) = \frac{1}{T}\sum_{t=1}^{T}e_t$. An important implication of forecast rationality is that the forecast should become more accurate as the forecast horizon shortens (Patton & Timmermann, 2007). As shown in tables 1.1 - 1.4, the $MAPE$, $RMSPE$, and MPE for each variable generally decrease over the forecasting horizon, with few exceptions. In addition, the last two columns of each table report the t-statistic and p-value for the bias test developed by Holden and Peel (1990a). The test statistic is calculated by regressing the forecast errors on a constant. As expected, table 1.1 suggests that net cash income and crop receipts are biased. We also find some evidence of bias in the forecasts of corn acreage (table 1.2), soybean prices (table 1.3), and wheat acreage and yield (table 1.4). In terms of the magnitude of the bias, the $MPEs$ for 12 out of 20 farm income forecasts and its components, 7 out of 24 for corn, 10 out of 24 for soybeans, and 8 out of 15 for wheat forecasts are above 1% in absolute values, which may be considered economically significant. Net cash income has the largest bias when compared to its components which is expected since the bias in its components is additive. For the WASDE forecasts, the bias for the price forecasts is much larger when compared to the bias for the yield and acreage forecasts. In addition, forecasts of farm income and its components have larger bias than that of WASDE production and price forecasts. Lastly, in cases where bias exceeds 1% in absolute value, it is generally positive, which indicates that the forecasts generally under-predict realized values. These results are consistent with our later findings that forecasters have a greater cost of over-prediction than under-prediction.

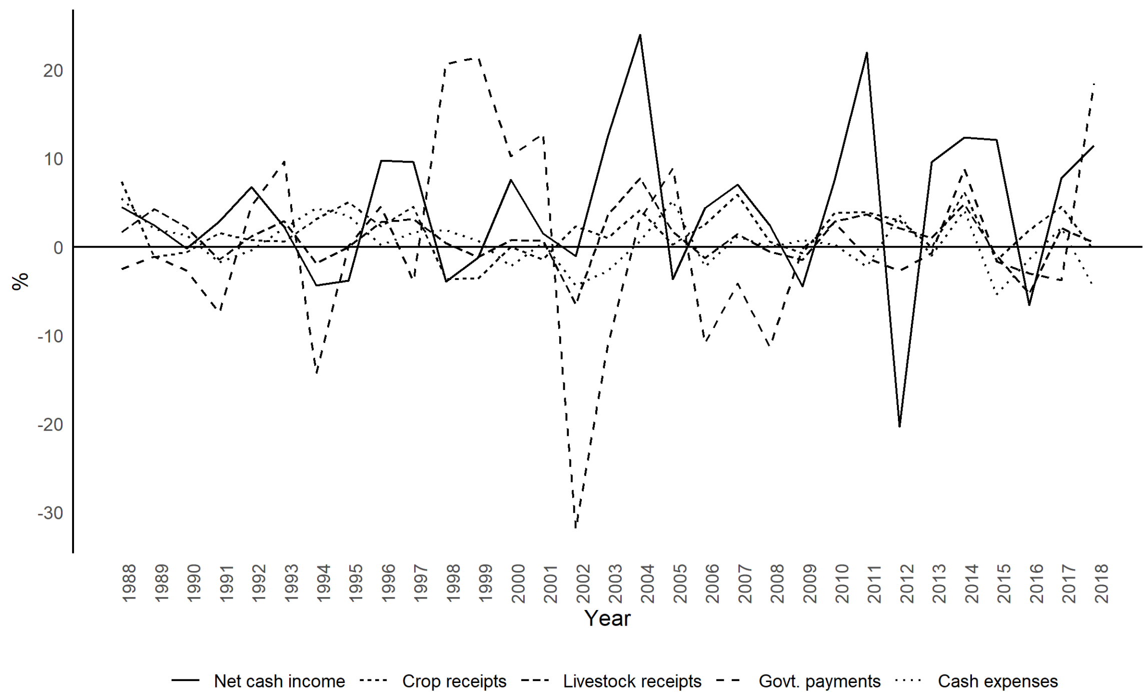


Figure 1.1: Average annual errors of net cash income forecasts and its components, 1988-2018

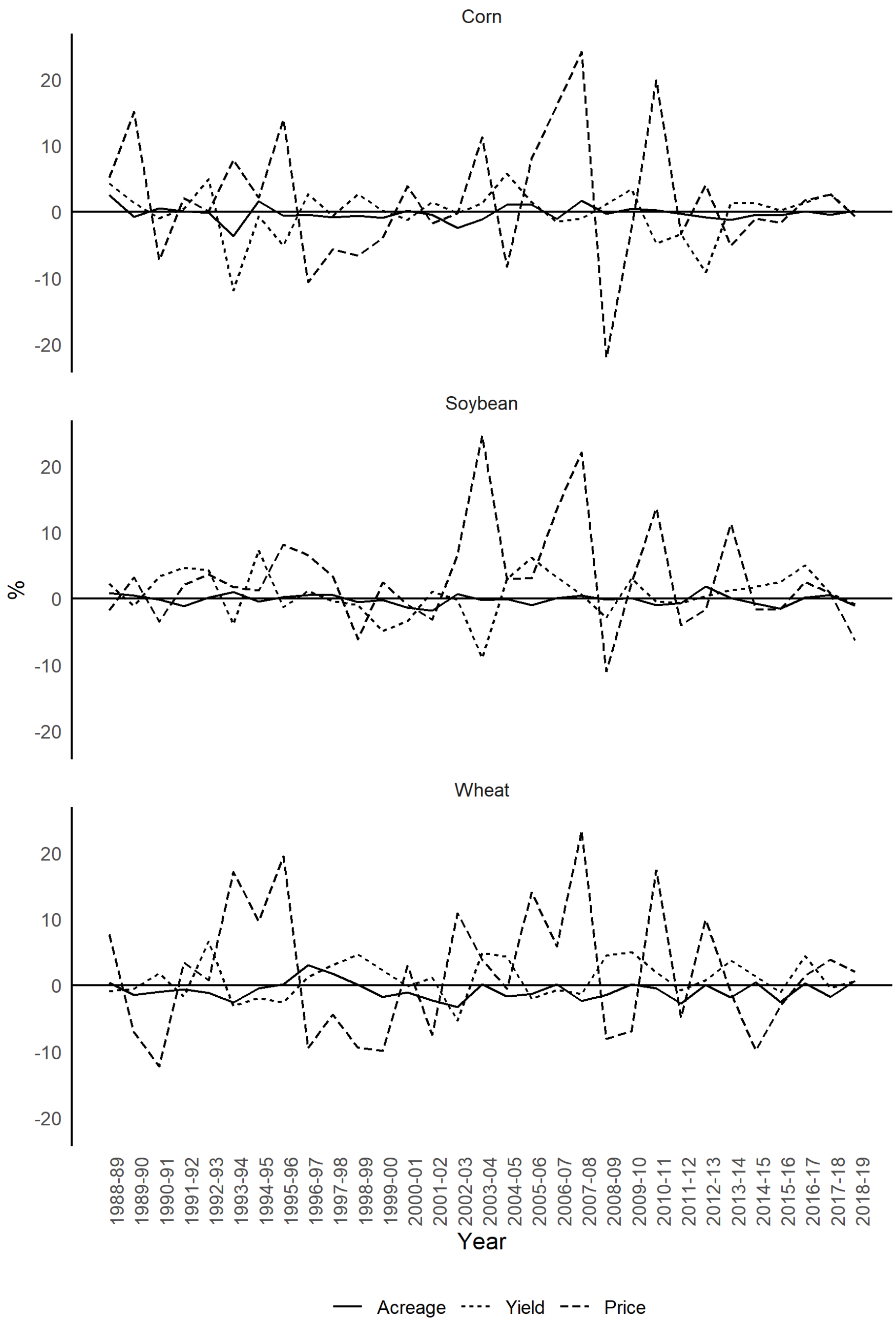


Figure 1.2: Average annual errors of WASDE forecasts for acreage, yield, and price, 1988-2018

Table 1.1: MAPE, RMSPE and MPE of USDA net cash income forecasts, 1988-2018

Variable	Forecast	MAPE ($\|e\|$)	RMSPE ($\sqrt{e^2}$)	MPE (e)	Bias (t-statistic)	p-value
Net cash income	February	13.177	16.001	8.153	3.225	0.003
	August	7.620	11.196	3.742	1.999	0.055
	November	5.817	8.208	2.373	1.742	0.092
	Next February	5.861	8.305	2.584	1.972	0.058
Crop receipts	February	4.314	5.339	3.048	3.459	0.002
	August	2.943	3.724	1.761	2.999	0.006
	November	2.389	3.026	1.018	2.076	0.047
	Next February	1.863	2.332	0.813	2.051	0.049
Livestock receipts	February	6.278	7.832	3.147	2.358	0.025
	August	2.261	2.778	0.415	0.739	0.466
	November	1.482	1.888	0.280	0.791	0.435
	Next February	1.076	1.529	-0.101	-0.363	0.719
Govt. payments	February	19.311	27.815	7.564	1.209	0.236
	August	11.374	18.518	-0.066	-0.017	0.986
	November	7.425	11.427	-4.967	-2.643	0.013
	Next February	4.934	7.100	-1.162	-0.876	0.388
Cash expenses	February	3.551	4.307	1.805	2.659	0.012
	August	2.683	3.282	0.470	0.734	0.469
	November	1.960	2.436	-0.375	-0.852	0.401
	Next February	1.926	2.665	-0.251	-0.575	0.570

Table 1.2: RMSPE, MAPE and MPE of WASDE corn forecasts, 1988/89-2018/19

Variable	Forecast	MAPE ($\|e\|$)	RMSPE ($\sqrt{e^2}$)	MPE (e)	Bias (t-statistic)	p-value
Acreage	May	2.043	2.807	-0.729	-1.383	0.179
	June	1.802	2.542	-0.341	-0.732	0.471
	July	1.145	1.651	-0.755	-2.891	0.008
	August	0.745	0.961	-0.339	-2.367	0.025
	September	0.763	0.961	-0.195	-1.310	0.200
	October	0.604	0.827	-0.041	-0.267	0.791
	November	0.534	0.737	0.047	0.350	0.729
	December	0.534	0.737	0.047	0.350	0.729
Price	May	13.255	17.153	5.467	1.900	0.067
	June	13.194	16.932	3.612	1.265	0.216
	July	10.962	14.161	2.202	0.934	0.358
	August	10.335	13.374	-0.245	-0.101	0.921
	September	9.456	12.079	0.720	0.326	0.746
	October	7.641	9.363	1.171	0.662	0.513
	November	4.853	6.599	1.083	0.929	0.360
	December	3.722	5.126	1.100	1.239	0.225
Yield	May	6.543	12.256	-3.427	-1.122	0.273
	June	5.204	8.382	-0.685	-0.411	0.685
	July	4.554	6.329	-0.156	-0.127	0.900
	August	3.687	4.943	0.736	1.052	0.302
	September	3.459	4.380	0.942	1.542	0.134
	October	2.028	2.758	0.333	0.907	0.372
	November	0.771	0.952	-0.058	-0.351	0.728
	December	0.805	1.028	-0.023	-0.128	0.899

Notes: (a) For acreage and yield, the May, June and July forecasts were available for the years 1993/94-2018/19 while the August forecasts were missing for the year 1988/89. (b) Other forecasts were available for 1988/89-2018/19.

Table 1.3: RMSPE, MAPE and MPE of WASDE soybean forecasts, 1988-2018

Variable	Forecast	MAPE ($\|e\|$)	RMSPE ($\sqrt{e^2}$)	MPE (e)	Bias (t-statistic)	p-value
Acreage	May	1.600	1.913	0.104	0.252	0.803
	June	1.411	1.752	-0.085	-0.244	0.809
	July	0.996	1.165	-0.485	-2.150	0.041
	August	0.838	1.057	-0.254	-1.332	0.193
	September	0.852	1.073	-0.239	-1.257	0.218
	October	0.621	0.850	-0.047	-0.288	0.775
	November	0.620	0.846	0.028	0.167	0.869
	December	0.620	0.846	0.028	0.167	0.869
Price	May	9.934	13.762	6.025	2.841	0.008
	June	9.865	13.647	4.788	2.192	0.036
	July	9.508	13.506	4.317	2.003	0.054
	August	8.961	12.243	1.992	1.071	0.293
	September	7.533	10.578	1.108	0.618	0.541
	October	6.086	7.997	2.260	1.443	0.159
	November	4.416	6.412	2.286	2.217	0.034
	December	3.643	5.348	1.733	1.957	0.060
Yield	May	5.136	6.666	-0.173	-0.121	0.905
	June	5.115	6.659	-0.088	-0.061	0.952
	July	4.816	6.275	0.705	0.548	0.589
	August	4.330	5.510	1.533	1.568	0.127
	September	3.908	4.757	1.538	1.791	0.083
	October	1.977	2.528	0.951	2.207	0.035
	November	0.930	1.110	0.128	0.598	0.554
	December	0.930	1.110	0.128	0.598	0.554

Notes: (a) For acreage and yield, the May, June and July forecasts were available for the years 1993/94-2018/19. (b) Other forecasts were available for 1988/89-2018/19.

Table 1.4: MAPE, RMSPE and MPE of WASDE wheat forecasts, 1988-2018

Variable	Forecast	MAPE ($\lvert e \rvert$)	RMSPE ($\sqrt{e^2}$)	MPE (e)	Bias (t-statistic)	p-value
Acreage	May	2.158	2.700	-0.407	-0.763	0.452
	June	2.142	2.663	-0.301	-0.535	0.598
	July	1.272	1.625	-1.145	-5.660	0.000
	August	1.056	1.351	-0.903	-4.839	0.000
	September	1.056	1.351	-0.903	-4.839	0.000
Price	May	11.840	14.839	2.583	0.935	0.357
	June	11.083	13.653	2.220	0.886	0.383
	July	8.570	10.794	2.709	1.401	0.172
	August	6.435	7.866	1.169	0.797	0.432
	September	4.812	5.774	1.056	1.004	0.324
Yield	May	4.759	5.923	2.149	1.943	0.063
	June	3.957	4.749	1.985	2.302	0.030
	July	2.470	3.358	0.797	1.591	0.123
	August	1.674	2.173	0.259	0.736	0.468
	September	1.230	1.607	0.186	0.638	0.528

Notes: (a) For acreage and yield, the May and June forecasts were available for the years 1993/94-2018/19 while the July forecasts were missing for the year 1988/89. (b) Other forecasts were available for 1988/89-2018/19.

1.4 Methodology

Traditional forecast evaluation assumes that the forecaster's objective is to minimize the univariate mean square error (MSE) loss function:

$$L(f^j_{t,0}, f^j_{t,h}) = (f^j_{t,0} - f^j_{t,h})^2 \tag{1.2}$$

where $L(\cdot)$ is the loss function, $f^j_{t,0}$ is the realized value of variable j for period t, and $f^j_{t,h}$ is the forecast of $f^j_{t,0}$ conducted at a horizon of h months ahead of the realized value.

As previously stated, Elliott et al. (2005) develop a method for testing forecast rationality under a flexible class of asymmetric loss functions which nest MSE loss as a special case. Their generalized method of moments (GMM) approach jointly estimates the asymmetry parameters of the loss function and tests for rationality. In this study, we follow Komunjer and Owyang (2012), who develop a generalized version of the Elliott et al. (2005) approach that examines multivariate forecasts and allows for non-separable loss. We define the multivariate loss function $L_p(\boldsymbol{\tau}, \mathbf{e})$ as,

$$L_p(\boldsymbol{\tau}, \mathbf{e}) = (||\mathbf{e}||_p + \boldsymbol{\tau}'\mathbf{e})\, ||\mathbf{e}||_p^{p-1}, \tag{1.3}$$

where $1 \leq p < \infty$, $1/p + 1/q = 1$, $\mathbf{e} \in \mathbb{R}^n$ and $\boldsymbol{\tau} \in \mathfrak{B}^n_q = \{\mathbf{u} \in \mathbb{R}^n : ||\mathbf{u}||_q < 1\}$. The vector $\mathbf{e}$ comprises the forecast errors of n variables. The asymmetry parameter $\boldsymbol{\tau}$ determines the relative losses due to positive and negative errors for each component of the error vector. The scalar p determines the shape of the loss function.

The loss function (1.3) is flexible as it can accommodate a wide variety of loss functions by varying the shape parameter p. The loss function allows for asymmetry and non-separability for a value of $p \geq 1$, and nests many well-known loss functions, such MSE and MAE loss (Komunjer & Owyang, 2012). Figure 1.3 shows several examples of special cases of univariate loss functions with shape parameters $p = 1$ and $p = 2$. Panel (a) depicts several linear loss functions, such as the symmetric absolute deviation loss (MAE) and the asymmetric lin-lin loss functions where $p = 1$. Panel (b) depicts several quadratic loss functions, such as the symmetric squared loss (MSE) and asymmetric quad-quad loss functions where $p = 2$.

The magnitude of the asymmetry parameter $\boldsymbol{\tau}$ indicates the direction and degree of asymmetry in the loss function. In equation (1.3), the sign of the forecast error of a variable enters the loss function only if the asymmetry parameter for that variable is non-zero, $\tau^j \neq 0$. The univariate version of the multivariate loss function in equation (1.3) for $p = 2$ is given by,

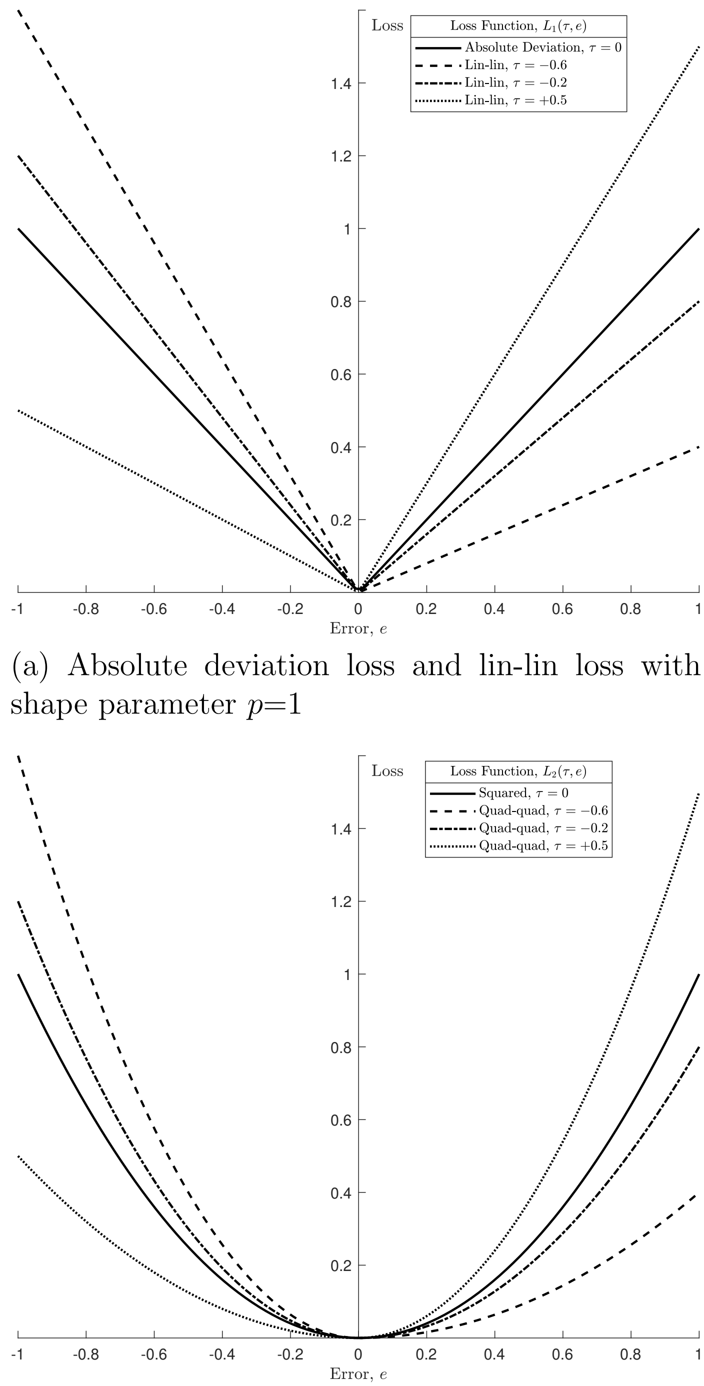
$$L_2(\tau^j, e^j) = (e^j)^2 + \tau^j sgn(e^j)(e^j)^2, \tag{1.4}$$

where j is the variable, e^j is the forecast error for variable j, and $sgn(e^j)$ is the sign of e^j.

We define the relative loss of over-prediction as the ratio of the loss due to over-prediction and the loss due to under-prediction of the same magnitude (negative and positive forecast errors of the same magnitude). From equation (1.4), the relative loss of over-prediction for variable j can be expressed as:

$$\frac{L_2(\tau^j, -|e^j|)}{L_2(\tau^j, |e^j|)} = \frac{1-\tau^j}{1+\tau^j}. \tag{1.5}$$

If $\tau^j = 0$, the costs of over-prediction and under-prediction are the same, and the loss function is symmetric in variable j. A negative value of the asymmetry parameter suggests that the relative loss of over-prediction is greater than one, suggesting over-predictions are costlier than under-predictions. Figure 1.3 shows how the sign and



(a) Absolute deviation loss and lin-lin loss with shape parameter p=1

(b) Squared loss and quad-quad loss with shape parameter p=2

Figure 1.3: Special cases of univariate loss functions

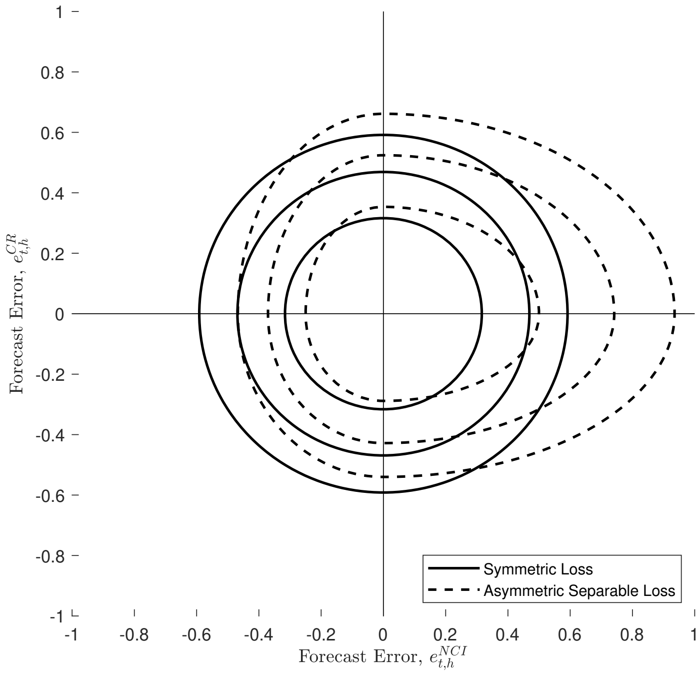
magnitude of the asymmetry parameter influence the univariate lin-lin and quad-quad loss functions.

Under the multivariate loss function (1.3), losses due to errors in the components of the vector are additively non-separable. If $\boldsymbol{\tau} \neq \mathbf{0}$, the sum of univariate losses does not equal the multivariate loss, i.e. $\sum_j L_2(\tau^j, e^j) \neq L_2(\boldsymbol{\tau}, \boldsymbol{e})$. The non-separability stems from the second term in equation (1.3) which represents the interaction between the forecast errors of the components that contribute toward the multivariate loss. When the forecaster's loss is symmetric in all components ($\boldsymbol{\tau} = \mathbf{0}$), this term disappears and the multivariate loss function becomes additively separable, i.e., $\sum_j L_2(0, e^j) = L_2(\mathbf{0}, \boldsymbol{e})$. Figure 1.4 shows the isoloss contours of a separable loss (i.e. the sum of univariate losses) and a non-separable multivariate loss against the symmetric loss using a bivariate loss function as an example. For symmetric loss, the isoloss contours are circular. In the case of separable loss and non-separable loss, the isoloss contours are distorted, and they have different shapes.

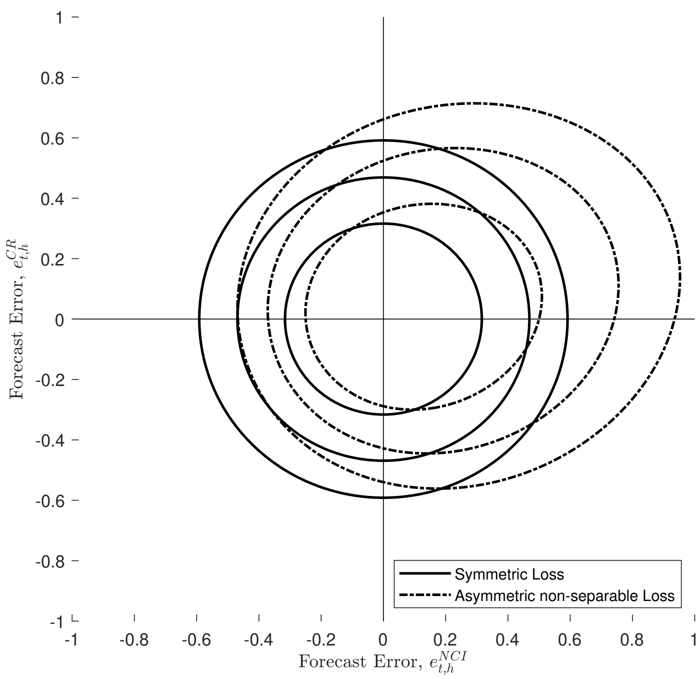
1.4.1 Estimation Procedure

We observe the multivariate forecasts $\mathbf{f}_{t,h}$, realized values $\mathbf{f}_{t,0}$, and a set of d instruments $\mathbf{x}_{t-1,h}$, which are a subset of the forecasters' information set and include the lagged forecasts of the same horizon, for P periods.[6] We assume that the forecaster minimizes an expected loss when constructing the forecasts and that the loss function belongs to the general class of loss functions (1.3). Using this information, we seek to estimate the asymmetry parameter $\boldsymbol{\tau}_h$ for each forecast horizon that is consistent with the characteristics of $\mathbf{f}_{t,h}$.

[6]The instrumental variables must be stationary with a full rank covariance matrix and satisfy the standard exclusion restrictions, as outlined in Komunjer and Owyang (2012), Appendix pp. 1078-1080.



(a) Asymmetric separable loss, $(\tau_h^{NCI}, \tau_h^{CR}) = (-0.6, -0.2)$, $p = 2$



(b) Asymmetric non-separable loss, $(\tau_h^{NCI}, \tau_h^{CR}) = (-0.6, -0.2)$, $p = 2$

Figure 1.4: Isoloss contours in a bivariate case

We follow Komunjer and Owyang (2012) and use a GMM-based strategy to estimate the asymmetry parameters. The procedure requires two assumptions. First, we assume that the shape parameter p is given. Second, we assume that the forecaster uses a rolling window of information to construct the forecasts. For example, to construct the forecast for the first period under our study, the forecaster uses information from the previous R periods. The information window is then rolled forward to construct all forecasts until the P^{th} period. Under these assumptions, the GMM estimator of the asymmetry parameter is given by,

$$\hat{\boldsymbol{\tau}}_h = \underset{\boldsymbol{\tau} \in \mathfrak{B}_q^n}{\arg\min} \left[P^{-1} \sum_{t=1}^{P} \mathbf{g}_p(\boldsymbol{\tau}; \mathbf{e}_{t,h}, \mathbf{x}_{t-1,h}) \right]' \times \hat{\mathbf{S}}^{-1} \times \left[P^{-1} \sum_{t=1}^{P} \mathbf{g}_p(\boldsymbol{\tau}; \mathbf{e}_{t,h}, \mathbf{x}_{t-1,h}) \right] \quad (1.6)$$

where,

$$\mathbf{g}_p(\boldsymbol{\tau}; \mathbf{e}_{t,h}, \mathbf{x}_{t-1,h}) = p\nu(\mathbf{e}_{t,h}) + \boldsymbol{\tau}||\mathbf{e}_{t,h}||_p^{p-1} + (p-1)\boldsymbol{\tau}'\mathbf{e}_{t,h}||\mathbf{e}_{t,h}||_p^{-1}\nu(\mathbf{e}_{t,h}) \otimes \mathbf{x}_{t,h}$$

$$\nu(\mathbf{e}_{t,h}) = (sgn(e_{t,h}^{j_1})|e_{t,h}^{j_1}|^{p-1}, \ldots, sgn(e_{t,h}^{j_n})|e_{t,h}^{j_n}|^{p-1})$$

The optimal weight matrix $\hat{S}^{-1}$ is iteratively determined during the GMM estimation using the equation,

$$\hat{S}(\tilde{\boldsymbol{\tau}}) = P^{-1} \sum_{t=1}^{P} \mathbf{g}_p(\tilde{\boldsymbol{\tau}}; \mathbf{e}_{t,h}, \mathbf{x}_{t-1,h}) \mathbf{g}_p(\tilde{\boldsymbol{\tau}}; \mathbf{e}_{t,h}, \mathbf{x}_{t-1,h})' \quad (1.7)$$

For a given shape parameter p, Komunjer and Owyang (2012) outline the conditions on the observed errors so that the GMM estimate of the asymmetry parameter is asymptotically normal (see Theorem 3 pp. 1072). Komunjer and Owyang (2012) also construct a J-statistic with $d > 1$ instruments to test the rationality of the multivariate forecasts.

$$\hat{J}_h = \left[P^{-1} \sum_{t=1}^{P} \mathbf{g}_p(\hat{\boldsymbol{\tau}}_h; \mathbf{e}_{t,h}, \mathbf{x}_{t-1,h}) \right]' \times \hat{\mathbf{S}}^{-1} \times \left[P^{-1} \sum_{t=1}^{P} \mathbf{g}_p(\hat{\boldsymbol{\tau}}_h; \mathbf{e}_{t,h}, \mathbf{x}_{t-1,h}) \right] \sim \chi^2_{n(d-1)} \quad (1.8)$$

A failure to reject the null hypothesis of rationality would suggest that, for a given set of instruments, there exists some value of the asymmetry parameter for which the forecasts are rational. Komunjer and Owyang (2012) further provide Monte Carlo evidence that the GMM estimation under non-separable loss yields consistent estimates of the asymmetry parameter even when the components of the vector $\mathbf{f}_{t,h}$ are highly correlated. If the loss function is misspecified as separable, it would likely produce biased estimates. Moreover, many equations employed in USDA's forecast models use time-lagged information from several sources as inputs (Dubman et al., 1993; McGath et al., 2009a; Vogel & Bange, 1999). Therefore, it is reasonable to assume that the forecaster uses a rolling window of information to generate the forecasts. These advantages make the GMM approach well-suited for evaluating the USDA forecasts.

1.4.2 Robustness Checks

Our estimation procedure requires two important assumptions: the instrumental variable set $\mathbf{x}_{t-1,h}$ and the shape parameter p. To ensure that our results are not overly influenced by these choices, we offer two important robustness checks. First, following Elliott et al. (2005), our preferred specification uses an instrument set consisting of a constant and one year lagged forecasts of a single variable of the same horizon (net cash income for net cash income forecasts and average farm prices for WASDE production forecasts). To ensure that our results are robust to the choice of instruments, we compute the asymmetry parameters using several sets of alternative instruments that are plausibly part of the forecaster's information set. Second, in our preferred specification, the shape parameter of the loss function was fixed at $p = 2$, which corresponds to the well-known quadratic loss in the univariate case. Komunjer and Owyang (2012) have shown that different values of the shape parameter p result in consistent estimates of the asymmetry parameter. Yet, to ensure that our results are robust to the choice of p, we estimate the model under different choices for the shape parameter: $p = 1.5, 2, 2.5$. Finally, our preferred specification estimates a single asymmetry parameter over the observation period 1988 – 2018. There may be some concern as to whether the asymmetry parameters are stable over time. The forecast performance may change due to changes in forecasting procedures over time or unexpected shocks, such as price disturbances, that may affect the loss function parameters. Isengildina-Massa et al. (2013), for example, showed that structural changes in the commodity markets during the mid 2000s accounted for the largest increase in errors in several WASDE forecasts for corn, soybeans, and wheat. Previous research suggests that, if the underlying data generating process of a variable is not stable, it is rational for error-minimizing forecasters to make serially correlated forecast revisions and systematic forecast errors as they learn about changes in the process driving the target variable (Batchelor, 2007; Muth, 1960). Further, Isengildina-Massa et al. (2013) demonstrate that WASDE forecast errors grew during periods of economic growth. Higgins and Mishra (2014) show that when forecasters are concerned with missing turning points, the forecasts are biased downward during expansions and biased upward during recessions.

To test for the stability of the estimated loss function parameters, we use an out-of-sample technique of detecting forecast breakdowns proposed by Giacomini and Rossi (2009) which has been used in similar econometric settings (Christodoulakis, 2020; Mamatzakis & Koutsomanoli-Filippaki, 2014; Mamatzakis & Tsionas, 2015).

Following Isengildina-Massa et al. (2013), we specifically examine the potential for forecast breakdowns during the 2007-2008 commodity price boom.

Giacomini and Rossi (2009) defines forecast breakdown as a situation where the out-of-sample performance of a forecast model is significantly inferior to its in-sample performance. The method involves dividing the sample period P into in-sample and out-of-sample windows of length m and n, where $P = m+n$. Then a "surprise loss" is calculated as the difference between the out-of-sample loss and the average in-sample loss:

$$SL_{t+1}(\hat{\boldsymbol{\tau}}_t) = L_{t+1}(\hat{\boldsymbol{\tau}}_t) - \overline{L_t}(\hat{\boldsymbol{\tau}}_t), \quad \text{for} \quad t = m, \ldots, (P-1) \tag{1.9}$$

The average in-sample loss $\overline{L_t}(\hat{\boldsymbol{\tau}}_t)$ is computed by first estimating the asymmetry parameters $\hat{\boldsymbol{\tau}}_{\boldsymbol{t}}$ for the in-sample window. Then the out-of-sample loss $L_{t+1}(\hat{\boldsymbol{\tau}}_t)$ is calculated using the in-sample asymmetry parameter estimates. The test is based on the hypothesis that in the absence of forecast breakdowns, the out-of-sample mean of the surprise losses should be zero.

$$H_0 : E\left(n^{-1}\Sigma_{t=m}^{P-1} SL_{t+1}(\boldsymbol{\tau}_t)\right) = 0 \tag{1.10}$$

The test statistic is calculated using a Newey-West standard error as,

$$t_{m,n,1} = n^{1/2} \frac{\overline{SL}_{m,n}}{\hat{\sigma}_{m,n}} \tag{1.11}$$

If the null hypothesis is rejected, a forecast breakdown is detected. We use three different forecasting schemes which follow different assumptions about the data generating process, (a) a fixed scheme with in-sample window $t = 1, \ldots, m$ for all t; (b) a rolling forecasting scheme with in-sample window $t = t - m + 1, \ldots, t$ at time t; and (c) a recursive forecasting scheme with in-sample window $t = 1, \ldots, t$ at time t.

The forecast breakdown test is performed for each of the three forecasting schemes by using the period before 2007 as the in-sample window.

1.5 Results

The estimated asymmetry parameters for our preferred specification are presented in table 1.5 and figure 1.5 for the USDA net cash income forecasts and in table 1.6 and figure 1.6 for the WASDE production and price forecasts for corn, soybean, and wheat. Following Komunjer and Owyang (2012) and Caunedo et al. (2018), we use two instruments to avoid size distortions in the J-test. The instrument sets used for farm income forecasts consist of a constant and one year lagged forecasts of net cash income. For the WASDE forecasts, the instruments include a constant and one year lagged forecasts of the average farm price. In each case, the shape parameter was fixed at $p = 2$.

The results for the J-test for rationality under $\hat{\tau}$ are presented in the bottom two rows of tables 1.5 and 1.6. The results show that the null hypothesis of rationality could not be rejected for any of the forecasts at the 5% significance level, for both separable and non-separable losses. This suggests that the USDA net cash income and WASDE forecasts are rationalizable under asymmetric loss.

Table 1.5 shows the GMM estimates of the asymmetry parameters for each component of the net cash income forecast under separable and non-separable loss, along with their standard errors. Figure 1.5 graphically presents the asymmetry parameters for net cash income and its components with 95% confidence intervals under non-separable loss. While the magnitude of τ^j is generally larger in absolute terms for separable loss than for non-separable loss, the sign and significance are mostly consistent under separable and non-separable loss. For example, the asymmetry estimate for the February (18-month-ahead) forecast of net cash income is –0.624 assuming separable loss, but only –0.458 under non-separable loss. The estimates of the asymmetry parameters for crop and livestock receipts and cash expenses are much closer to symmetric under non-separable loss, yet the asymmetry parameters are still significantly different from zero. In contrast, the estimates are markedly asymmetric under separable loss. These results are consistent with Komunjer and Owyang (2012), who demonstrate that rationality could be achieved with smaller degree of asymmetry under non-separable loss relative to separable loss. The pattern also follows the empirical findings of Caunedo et al. (2018), who show that Federal Reserve's forecasts of growth, inflation, and unemployment are asymmetric, yet the degree of asymmetry is less under non-separable loss. For example, the asymmetry parameter for growth and unemployment were -0.30 and 0.32 under separable loss and -0.29 and 0.03 under non-separable loss.

Given that USDA's net cash income forecasts are constructed using the accounting identity (1.1), we focus our discussion on the results under non-separable loss which is a more realistic assumption. The relative losses associated with the asymmetry parameters are calculated using equation (1.5) and presented in figure 1.7. The estimates of the asymmetry parameter of net cash income are negative and significant, which suggests that USDA has 2.7 times higher costs associated with over-predicting 1% in the February net cash income forecast than under-predicting it by 1%. This finding provides an alternative explanation of the bias findings in table 1.1 and of previous studies' findings that the February (18-month-ahead) forecasts tend to under-predict net cash income and net farm income (Isengildina-Massa, Karali, et al., 2020; Kuethe

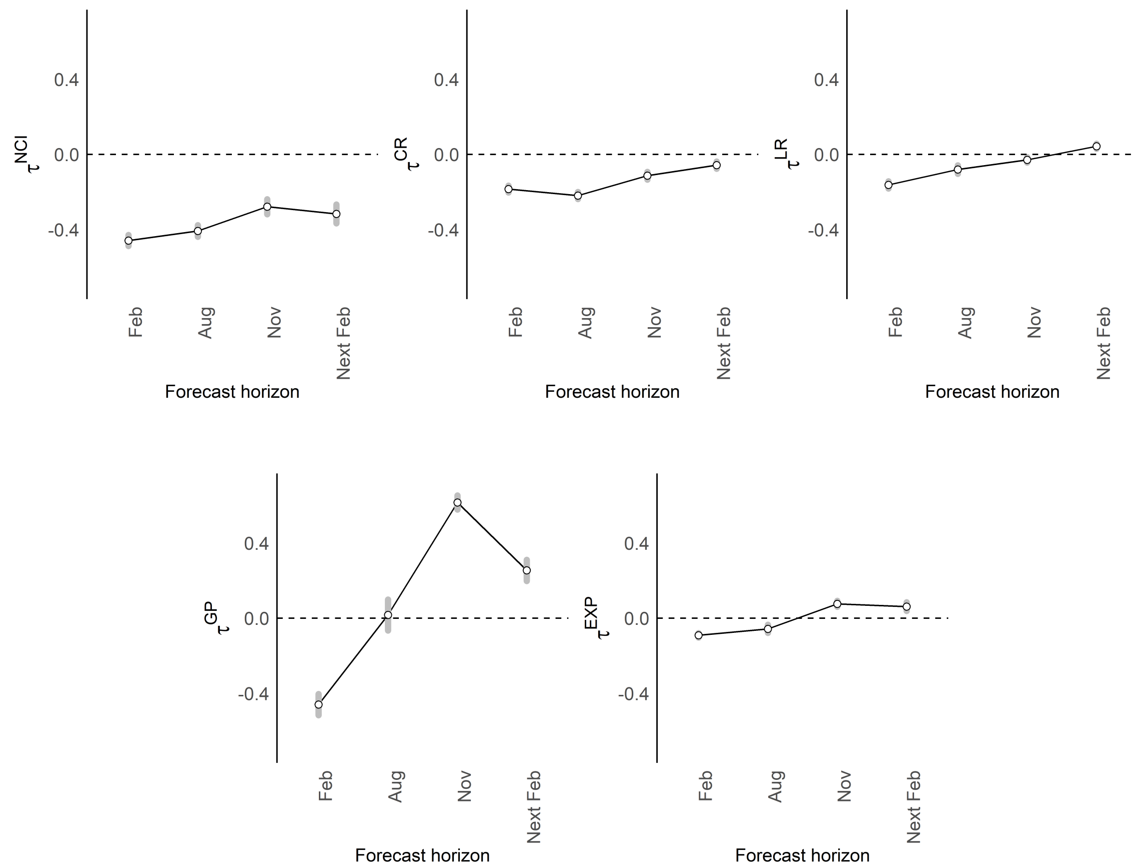


Figure 1.5: Asymmetry parameter estimates under non-separable loss for net cash income forecasts with 95% confidence intervals

Note: (a) The instrument set consists of a constant and one year lagged forecasts of net cash income. (b) Estimates are plotted for non-separable loss with shape parameter $p = 2$

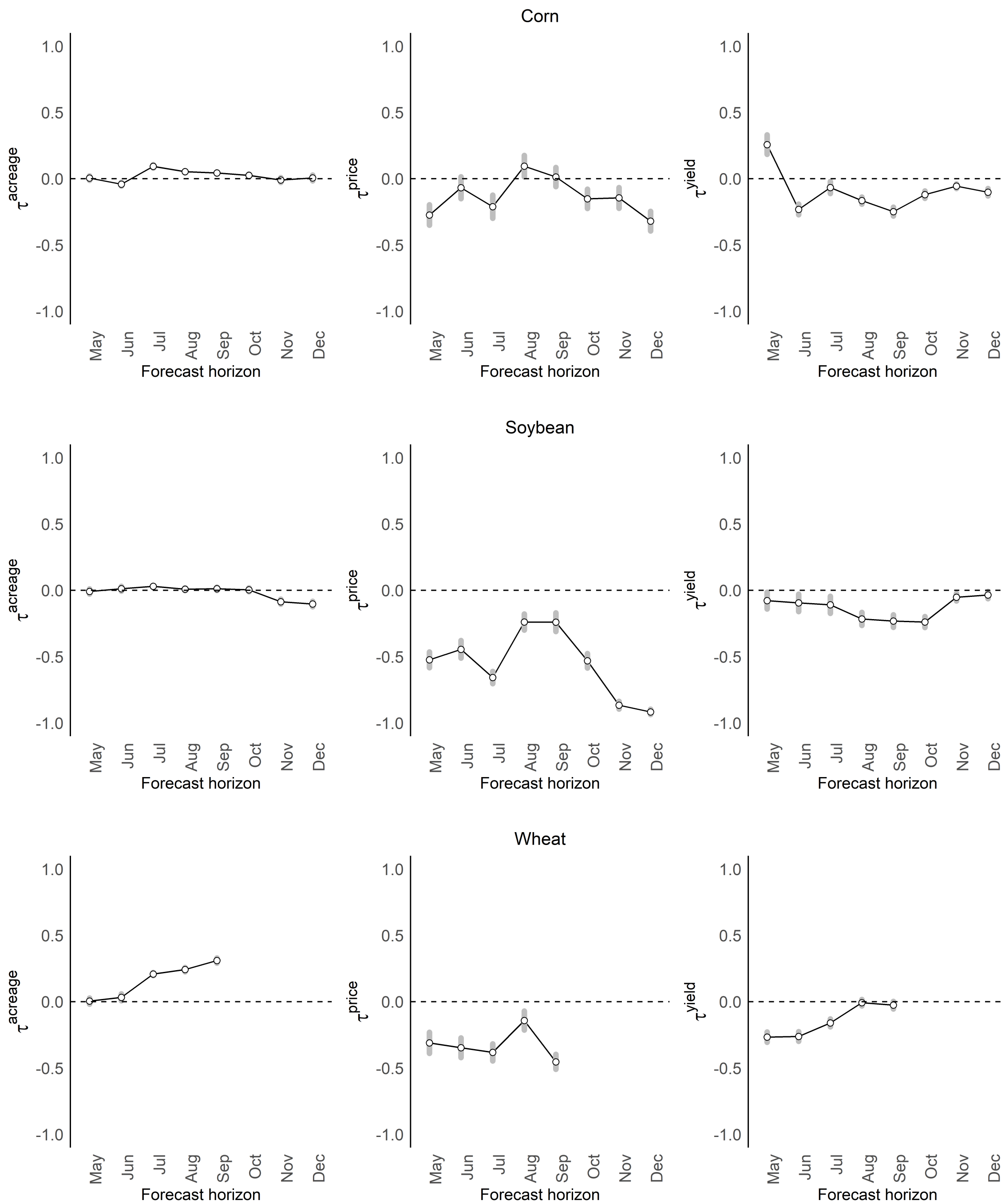


Figure 1.6: Asymmetry parameter estimates under non-separable loss for WASDE forecasts with 95% confidence intervals

Note: (a) The instrument set consists of a constant and one year lagged forecasts of price. (b) Estimates are plotted for non-separable loss with shape parameter $p = 2$

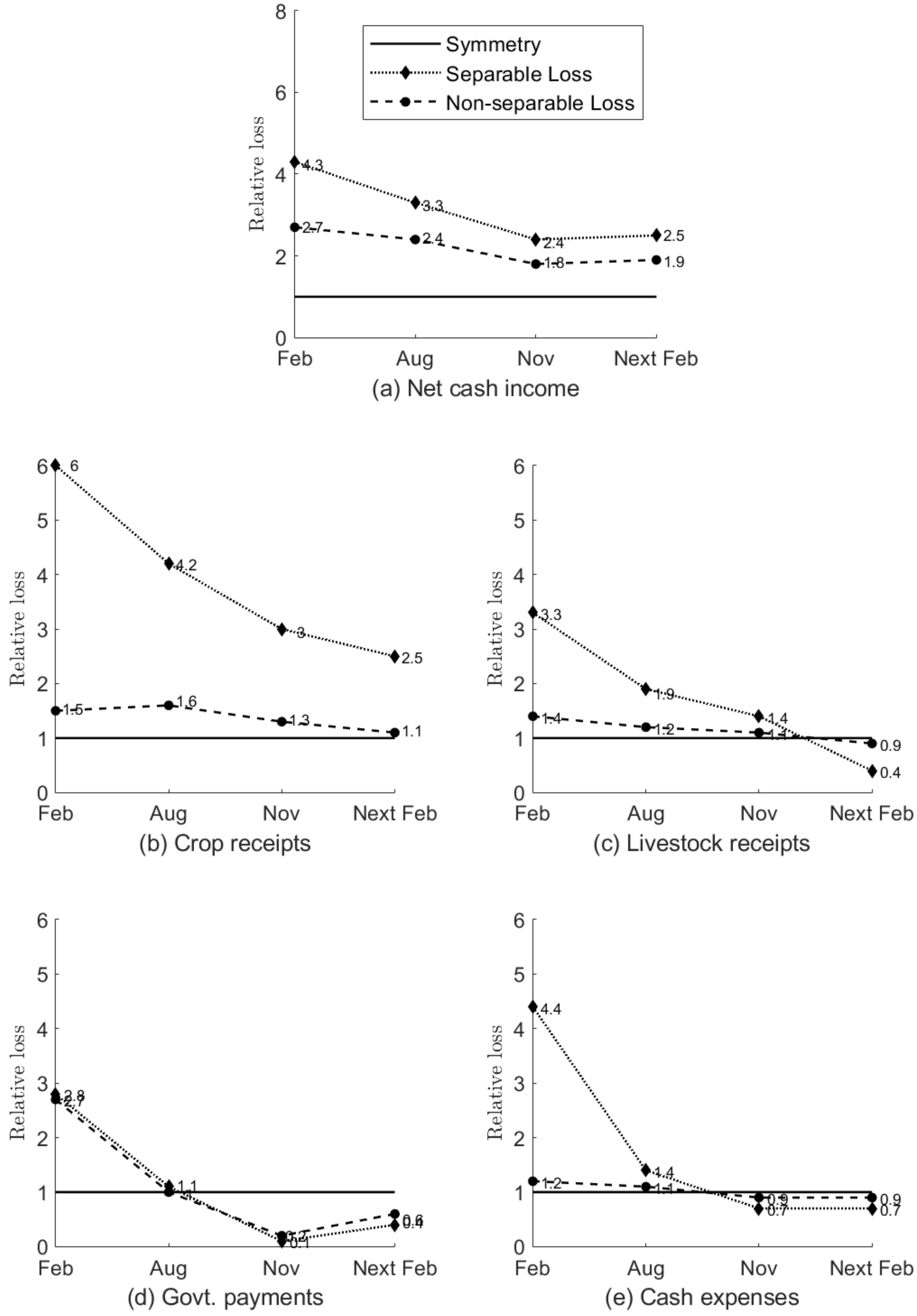


Figure 1.7: Relative losses for net cash income and its components, $p = 2$
Note: The instrument set consists of a constant and one year lagged forecasts of net cash income.

et al., 2018). Following Granger (1969), USDA's initial forecasts appear biased because because over-predictions are costlier, and therefore, the forecasts are rationally conservative. Among the components of net cash income, the asymmetry parameter estimate is significant and negative, particularly for the February forecast of government payments. On the other hand, the asymmetry parameter estimates for crop and livestock receipts and production expenses are closer to zero (symmetry) even though they are significant in most cases. These asymmetry parameters suggest that USDA has 1.2 to 1.5 times higher costs associated with over-predicting crop and livestock receipts and cash expenses.

The estimates of asymmetry parameters generally move closer to symmetry as the terminal event of releasing the USDA official estimates approaches, with the exception of government payments (figure 1.5). One implication of forecast rationality is that the forecast error should be a weakly non-decreasing function of the forecast horizon (Patton & Timmermann, 2007), or alternatively, that the forecasts become more accurate as the forecast horizon reduces from say 18 months ahead to 6 months ahead of the final estimate. As a result, a smaller degree of asymmetry is required to rationalize the forecasts as the horizon reduces. Direct government payments, however, show an interesting pattern across the forecast horizon. The asymmetry parameter is negative for the February (18-month-ahead) forecast of government payments, suggesting over-predictions by 1% percent are 2.8 times costlier than under-predictions by 1%. However, for the November (9-month-ahead) forecasts of government payments, and to some extent, for the next February (6-month-ahead) forecasts, the asymmetry estimate is positive, suggesting under-predictions are costlier. That is, before production, at the beginning of the calendar year USDA does not want to over-predict government outlays. However, in November, after the growing season, the USDA does not want to under-predict government program payments to farmers. This behavior, while curious, is consistent with the bias in USDA forecasts of government payments previously reported by Isengildina-Massa, Karali, et al. (2020) and shown in table 1.1.

Several of the findings described above for the USDA farm income forecasts also hold for the WASDE production and price forecasts for corn, soybeans, and wheat. The estimates of the asymmetry parameters for the WASDE acreage, yield, and price forecasts are presented in table 1.6 and graphically presented in figure 1.6 under non-separable loss, while the separable loss results are presented in table 1.7. Most of the asymmetry parameters for acreage are positive or not significant, with some negative asymmetry parameters particularly for the November and December forecasts for soybean acreage. The asymmetry parameters for acreage are relatively small in magnitude, suggesting that although USDA tends to over- or under-predict some acreage forecasts, the relative costs of over-predicting are not very high. The asymmetry parameters for price, on the other hand, are mostly negative and large in magnitude. Further, the asymmetry does not appear to decrease over the forecast horizon. In the case of corn and soybeans, the asymmetry parameters are highest during planting and harvest. For example, the asymmetry parameter of –0.524 for the soybean price forecast in May shows that the relative cost of over-predicting soybean price by 1% is 3.2 times higher than the cost of under-predicting it by 1%. Similarly, the asymmetry parameter estimates for yield are mostly negative and significant, but they generally move toward more symmetry over the time horizon, particularly during the last couple of months before harvest. These findings closely correspond to the bias results shown in tables 1.2 – 1.4. Overall, these findings show that the relative costs of over-predicting prices and yields are higher, while the relative costs of under-predicting acreage are generally higher, particularly for wheat.

Beyond our preferred specification, we also show that our findings are robust to both the choice of shape parameter and the instrument sets used in the estimation procedure. Through Monte Carlo simulation, Komunjer and Owyang (2012) show that different values of the shape parameter p result in consistent estimates of the asymmetry parameter. As a robustness check, we obtain similar results using different shape parameters. In figure 1.8, we plot non-separable asymmetry estimates with shape parameter values $p = 1.5$ and $p = 2.5$ along with our preferred specification of $p = 2$. In both cases, the asymmetry parameter estimates have the same sign as reported in the main results, and the magnitudes are similar, except for government payments in the August forecasts. These estimates show that our main results are not driven by the shape of the loss function, and the presence of asymmetry cannot be ruled out under alternative specifications. As an additional robustness check, we hold the shape parameter at $p = 2$ but vary the instruments sets. These estimates are plotted in figure 1.9. The results show that the estimates of the asymmetry parameters are similar to those reported in table 1.5, both in terms of sign and magnitude, except for the August forecast of government payments.

The results of the structural breakdown test of Giacomini and Rossi (2009) are presented in table 1.8 for the net cash income forecasts and in table 1.9 for the WASDE acreage, yield, and price forecasts. Using the fixed, rolling, and recursive forecasting schemes to test for structural breaks before and after the 2007 commodity price spikes, the test statistics are not significant at the 5% level, with the exception

Table 1.5: Estimates of asymmetry parameters and rationality tests for net cash income forecasts, 1988-2018

Asymmetry Parameters	Separable Loss				Non-separable Loss			
	Feb	Aug	Nov	Next Feb	Feb	Aug	Nov	Next Feb
Net cash income, τ^{NCI}	-0.624**	-0.530**	-0.403**	-0.421**	-0.458**	-0.408**	-0.278**	-0.317**
	(0.024)	(0.029)	(0.038)	(0.038)	(0.015)	(0.016)	(0.020)	(0.026)
Crop receipts, τ^{CR}	-0.714**	-0.613**	-0.501**	-0.432**	-0.185**	-0.219**	-0.113**	-0.057**
	(0.030)	(0.025)	(0.030)	(0.033)	(0.009)	(0.009)	(0.010)	(0.010)
Livestock receipts, τ^{LR}	-0.532**	-0.309**	-0.151**	0.477**	-0.162**	-0.081**	-0.029**	0.043**
	(0.031)	(0.044)	(0.041)	(0.035)	(0.010)	(0.011)	(0.008)	(0.007)
Govt. payments, τ^{GP}	-0.468**	-0.066	0.822**	0.397**	-0.460**	0.017	0.615**	0.254**
	(0.041)	(0.055)	(0.017)	(0.045)	(0.029)	(0.042)	(0.019)	(0.029)
Cash expenses, τ^{EXP}	-0.627**	-0.182**	0.193**	0.148**	-0.091**	-0.057**	0.077**	0.061**
	(0.026)	(0.040)	(0.039)	(0.040)	(0.006)	(0.011)	(0.008)	(0.012)
J-statistic	4.074	1.715	4.146	7.276	3.263	4.164	3.034	6.098
p-value	0.539	0.887	0.529	0.201	0.660	0.526	0.695	0.297

Notes: (a) The numbers are estimates of asymmetry parameters, $(\tau^{NCI}, \tau^{CR}, \tau^{LR}, \tau^{GP}, \tau^{EXP})'$, and standard errors (SE) are reported in parentheses. (b) Instruments used are a constant and one year lagged forecasts of net cash income. (c) Number of periods, P=31. (d) (**) denotes significant at 5%. (e) p-values of the J-test correspond to a χ^2 distribution with 5 degrees of freedom (f) Shape parameter, $p = 2$.

of the November and December soybean forecasts. Our interpretation is that even though crop prices sharply increased in 2007 resulting in high price forecast errors, the forecast errors for farm income and crop production were not the highest in 2007 as compared to the rest of the years in our sample (figures 1.1 and 1.2). Even though we do not find evidence of structural breakdown before and after 2007 when considering the vector of forecasts for net cash income and WASDE production and prices, our test does not preclude the possibility that individual asymmetry parameters for specific components at specific time horizons may differ across sub-periods.

1.6 Conclusions

Previous studies suggest that many of the forecasts generated by the USDA are technically irrational (biased and/or inefficient). A rejection of rationality, however,

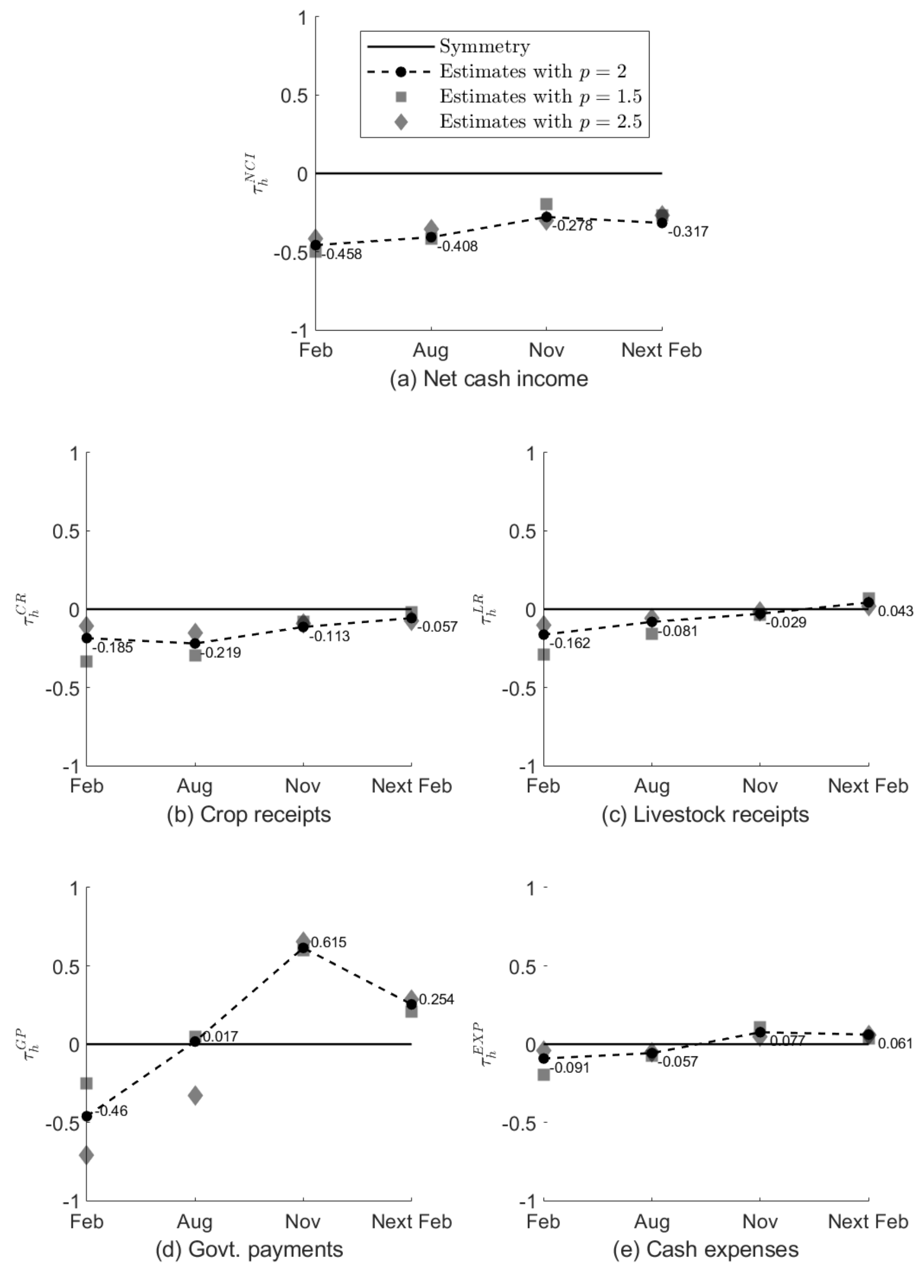


Figure 1.8: Robustness of net cash income asymmetry parameters to choice of shape parameter

Note: (a) Non-separable loss function with shape parameter, $p = 2$. (b) The instrument set consists of a constant and one year lagged forecasts of net cash income (same as Table 1.5).

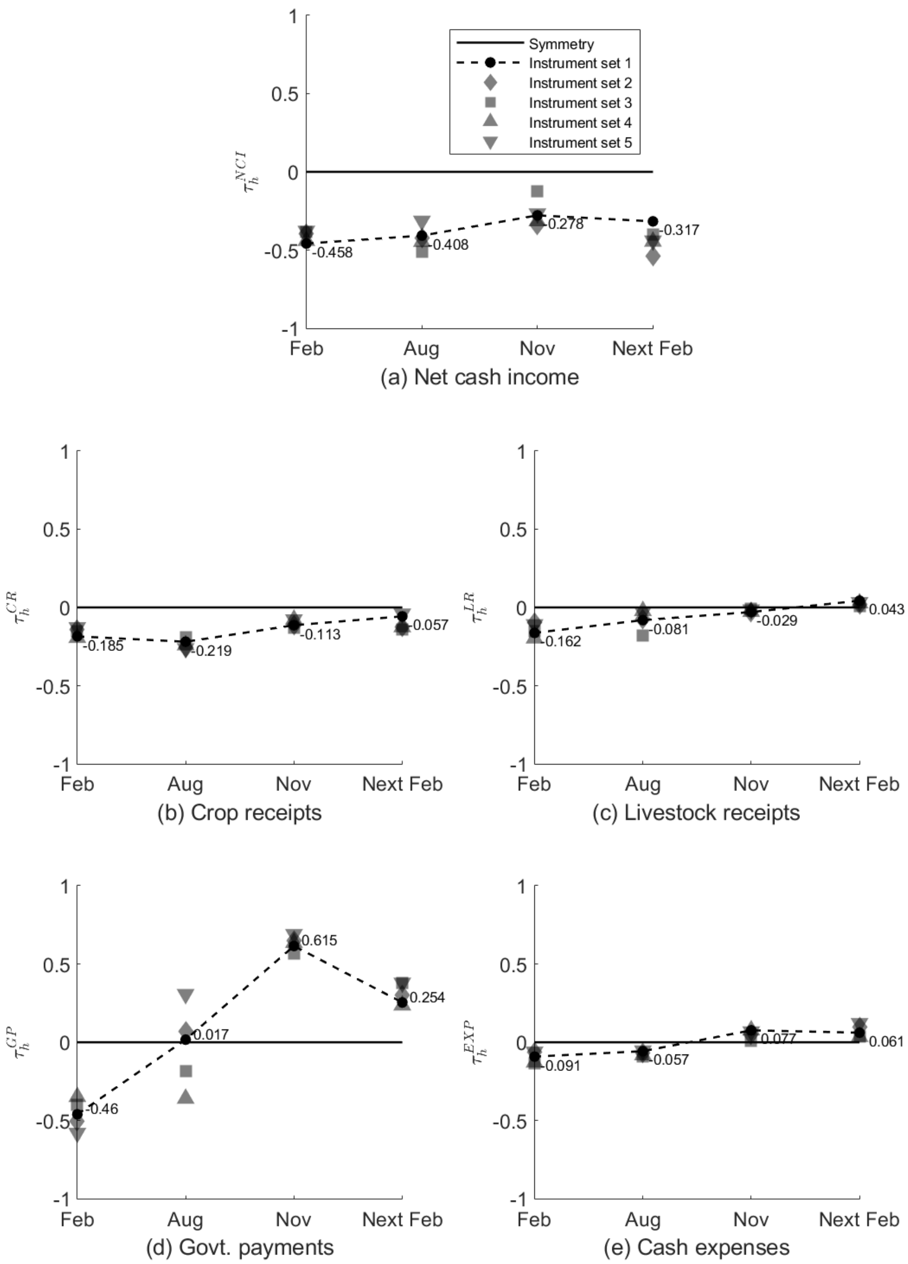


Figure 1.9: Robustness of net cash income asymmetry parameters to choice of instruments

Note: (a) Non-separable loss function with shape parameter, $p = 2$. (b) The instrument consists of a constant and one year lagged forecasts of 1) net cash income(same as Table 1.5), 2) crop receipts, 3) livestock receipts, 4) government payments, and 5) net cash income and crop receipts.

may be the result of either the forecaster's inefficient use of information or a misspecification of the forecaster's loss function (Elliott et al., 2005). In this study, we jointly estimate the parameters of USDA forecasters' loss function and test for rationality for two important sets of USDA forecasts: net cash income and WASDE production and price forecasts for corn, soybeans, and wheat. Following Komunjer and Owyang (2012), the loss function is estimated under a flexible multivariate loss function that allows for asymmetry and non-separability in the forecast errors.

Our analysis yields two important findings. First, we find evidence of asymmetric loss in both net cash income and WASDE forecasts. These results are consistent with previous findings of bias and inefficiency (Isengildina-Massa et al., 2013; Isengildina-Massa, Karali, et al., 2020; Kuethe et al., 2018), yet our empirical approach provides an alternate interpretation of these results. For example, Isengildina-Massa, Karali, et al. (2020) find that initial USDA net cash income forecasts are downward biased. Our results suggest that a 1% over-prediction of the initial net cash income is 2.7 times as costly as an under-prediction of the same percent. Thus, USDA is averse to over-predicting net cash income at the early stages of the forecasting process. We similarly find that USDA has a higher cost over-predicting both price and yield for corn, soybeans, and wheat. Second, we find that under asymmetric loss, the USDA forecasters are rational expected loss minimizers. Economic theory provides a variety of internal and external costs that may lead otherwise rational forecasters to release "technically irrational" forecasts (Batchelor, 2007; Weber, 1994).

Our findings are important for a variety of USDA forecast users, including farmers, lenders, agricultural business leaders, and agricultural policymakers. As Auffhammer (2007) argues, a forecast is only optimal for a particular forecast user when his or her loss function matches that of the forecast producer. Accurately describing the loss function of USDA forecasters is therefore an important first step in forecast evaluation. Given the important role that USDA's farm income forecasts play in farm policy debates and WASDE's influence in commodity markets, the USDA should consider the internal and external forces that influence the cost of forecast errors. Merola and Pérez (2013) argue that government forecasting processes can be improved by increasing (i) transparency on data reporting, (ii) accountability of forecasters, and (iii) *ex ante* incentives to release unbiased forecasts. Previous research suggests that biased public forecasts can influence private decision making. For example, Beaudry and Willems (2018) demonstrate that over-optimistic GDP growth forecasts leads to higher public and private debt accumulation and later recessions. Thus, our findings may also help inform future revisions of USDA forecast models and procedures.

Table 1.6: Estimates of asymmetry parameters and rationality tests for WASDE forecasts under non-separable loss, 1988/89-2018/19

Asymmetry Parameters	May	June	July	August	September	October	November	December
Corn								
$\tau^{acreage}$	0.005	-0.041**	0.093**	0.053**	0.043**	0.025**	-0.010	0.006
	(0.009)	(0.008)	(0.007)	(0.005)	(0.005)	(0.006)	(0.009)	(0.011)
τ^{price}	-0.274**	-0.068	-0.211**	0.095**	0.013	-0.152**	-0.145**	-0.321**
	(0.039)	(0.042)	(0.044)	(0.041)	(0.037)	(0.037)	(0.040)	(0.038)
τ^{yield}	0.257**	-0.232**	-0.066**	-0.165**	-0.248**	-0.120**	-0.057**	-0.102**
	(0.038)	(0.020)	(0.023)	(0.014)	(0.017)	(0.014)	(0.009)	(0.014)
J-statistic	4.591	4.542	0.258	1.570	4.380	4.215	4.342	5.356
p-value	0.204	0.209	0.968	0.666	0.223	0.239	0.227	0.148
Soybean								
$\tau^{acreage}$	-0.009	0.012	0.030**	0.009	0.012	0.004	-0.086**	-0.103**
	(0.010)	(0.009)	(0.004)	(0.005)	(0.006)	(0.008)	(0.010)	(0.011)
τ^{price}	-0.524**	-0.445**	-0.657**	-0.239**	-0.240**	-0.532**	-0.866**	-0.917**
	(0.031)	(0.034)	(0.023)	(0.031)	(0.036)	(0.028)	(0.015)	(0.010)
τ^{yield}	-0.078**	-0.095**	-0.109**	-0.216**	-0.232**	-0.239**	-0.052**	-0.036**
	(0.032)	(0.033)	(0.033)	(0.025)	(0.024)	(0.021)	(0.014)	(0.014)
J-statistic	1.345	1.487	4.213	3.249	3.144	1.617	3.340	3.736
p-value	0.718	0.685	0.239	0.355	0.370	0.655	0.342	0.291
Wheat								
$\tau^{acreage}$	0.004	0.033**	0.208**	0.242**	0.310**	-	-	-
	(0.012)	(0.013)	(0.006)	(0.008)	(0.010)	-	-	-
τ^{price}	-0.310**	-0.347**	-0.382**	-0.142**	-0.453**	-	-	-
	(0.040)	(0.038)	(0.033)	(0.037)	(0.029)	-	-	-
τ^{yield}	-0.267**	-0.263**	-0.160**	-0.007	-0.026	-	-	-
	(0.020)	(0.019)	(0.015)	(0.012)	(0.014)	-	-	-
J-statistic	1.862	2.943	1.374	3.789	3.916	-	-	-
p-value	0.602	0.400	0.712	0.285	0.271	-	-	-

Notes: (a) The numbers are estimates of asymmetry parameters, $(\tau^{acreage}, \tau^{yield}, \tau^{price})'$, and standard errors (SE) are reported in parentheses. (b) Instruments used are a constant and one year lagged forecasts of average farm price. (c) (**) denotes significant at 5%. (d) p-values of the J-test correspond to a χ^2 distribution with 5 degrees of freedom. (e) Non-separable loss function with shape parameter, $p = 2$. (f) Corn estimates are conducted for the period 1994/95-2018/19 for the May, June and July forecasts, and 1990/91-2018/19 for the August forecasts. (g) Soybean estimates are conducted for the period 1994/95-2018/19 for the May, June and July forecasts. (h) Wheat estimates are conducted for the period 1994/95-2018/19 for the May and June forecasts, and 1990/91-2018/19 for the July forecasts.

Table 1.7: Estimates of asymmetry parameters for WASDE forecasts under separable loss, 1988/89-2018/19

Asymmetry Parameters	May	June	July	August	September	October	November	December
Corn								
$\tau_h^{acreage}$	0.416**	-0.007	0.630**	0.430**	0.265**	0.013	-0.085	-0.081
	(0.042)	(0.046)	(0.034)	(0.034)	(0.035)	(0.043)	(0.045)	(0.045)
τ_h^{price}	-0.303**	-0.123**	-0.200**	0.088	-0.105**	-0.182**	-0.229**	-0.307**
	(0.046)	(0.047)	(0.049)	(0.043)	(0.039)	(0.038)	(0.041)	(0.040)
τ_h^{yield}	0.477**	-0.337**	-0.136**	-0.386**	-0.627**	-0.416**	0.095**	0.091**
	(0.053)	(0.053)	(0.056)	(0.032)	(0.025)	(0.033)	(0.037)	(0.039)
J-statistic	4.071	4.455	0.798	2.440	5.208	5.193	4.421	4.120
p-value	0.254	0.216	0.850	0.486	0.157	0.158	0.219	0.249
Soybean								
$\tau^{acreage}$	-0.038	0.092	0.828**	0.326**	0.274**	0.052	-0.143**	-0.115**
	(0.050)	(0.049)	(0.027)	(0.039)	(0.039)	(0.047)	(0.048)	(0.048)
τ^{price}	-0.571**	-0.463**	-0.506**	-0.194**	-0.159**	-0.495**	-0.818**	-0.918**
	(0.032)	(0.036)	(0.033)	(0.035)	(0.042)	(0.033)	(0.020)	(0.010)
τ^{yield}	-0.048	-0.065	-0.118	-0.365**	-0.396**	-0.478**	-0.134**	-0.141**
	(0.058)	(0.059)	(0.059)	(0.039)	(0.033)	(0.034)	(0.040)	(0.040)
J-statistic	2.256	2.129	4.014	4.018	2.904	1.653	6.722	8.331
p-value	0.521	0.546	0.260	0.260	0.407	0.647	0.081	0.040
Wheat								
$\tau^{acreage}$	0.163**	0.185**	0.913**	0.869**	0.855**	-	-	-
	(0.048)	(0.047)	(0.009)	(0.011)	(0.012)	-	-	-
τ^{price}	-0.232**	-0.230**	-0.343**	-0.227**	-0.449**	-	-	-
	(0.048)	(0.045)	(0.036)	(0.037)	(0.031)	-	-	-
τ^{yield}	-0.553**	-0.609**	-0.348**	-0.062	-0.012	-	-	-
	(0.038)	(0.033)	(0.036)	(0.036)	(0.038)	-	-	-
J-statistic	1.222	1.341	0.564	5.560	5.074	-	-	-
p-value	0.748	0.719	0.905	0.135	0.166	-	-	-

Notes: (a) Shape parameter, $p = 2$. (b) The numbers are estimates of asymmetry parameters, $(\tau^{acreage}, \tau^{price}, \tau^{yield})'$, and standard errors (SE) are reported in parentheses. (c) Instrument set consists of a constant and one year lagged forecasts of average farm price. (d) (**) denotes significant at 5%. (e) p-values of the J-test correspond to a χ^2 distribution with 5 degrees of freedom. (f) Corn estimates are conducted for the period 1994/95-2018/19 for the May, June and July forecasts, and 1990/91-2018/19 for the August forecasts. (g) Soybean estimates are conducted for the period 1994/95-2018/19 for the May, June and July forecasts. (h) Wheat estimates are conducted for the period 1994/95-2018/19 for the May and June forecasts, and 1990/91-2018/19 for the July forecasts.

Table 1.8: Tests of forecast breakdown for net cash income forecasts, 1988-2018

Forecast Horizon	Forecasting scheme					
	Fixed		Rolling		Recursive	
	t-statistic	p-value	t-statistic	p-value	t-statistic	p-value
February	-1.499	0.134	-0.468	0.640	-0.764	0.445
August	-0.885	0.376	-1.071	0.284	-0.887	0.375
November	0.030	0.976	0.022	0.983	0.135	0.893
Next February	0.299	0.765	0.428	0.668	0.389	0.698

Note: (a) T-test statistics and corresponding p-values of structural break test using surprise loss are reported. Rejection of the null hypothesis would suggest presence of structural break. (b) Non-separable loss with shape parameter, $p = 2$.

Table 1.9: Tests of forecast breakdown for WASDE forecasts, 1988/89-2018/89

Forecast Horizon	Forecasting scheme					
	Fixed		Rolling		Recursive	
	test statistic	p-value	test statistic	p-value	test statistic	p-value
	Corn					
May	0.150	0.880	0.344	0.731	0.118	0.906
June	1.029	0.303	0.453	0.651	0.787	0.431
July	0.910	0.363	-0.315	0.752	0.494	0.621
August	0.424	0.672	0.141	0.888	0.229	0.819
September	0.438	0.662	0.257	0.797	0.336	0.737
October	0.332	0.740	0.273	0.785	0.310	0.757
November	0.505	0.613	-0.002	0.998	0.131	0.896
December	0.073	0.942	-0.043	0.966	-0.002	0.998
	Soybean					
May	0.571	0.568	0.979	0.328	0.299	0.765
June	0.943	0.346	0.608	0.543	0.452	0.651
July	0.388	0.698	0.355	0.723	0.207	0.836
August	0.141	0.888	-0.102	0.919	-0.004	0.996
Sepember	0.353	0.724	0.206	0.837	0.208	0.836
October	-0.431	0.667	-0.691	0.489	-0.939	0.348
November	-1.964	0.049	-2.184	0.029	-1.882	0.060
December	-2.825	0.005	-4.027	0.000	-3.118	0.002
	Wheat					
May	-0.334	0.738	0.085	0.932	-0.439	0.661
June	-0.171	0.864	-0.182	0.855	-0.446	0.656
July	-0.079	0.937	-0.094	0.925	-0.283	0.778
August	-0.235	0.814	-0.181	0.857	-0.385	0.700
September	0.279	0.780	0.354	0.723	0.115	0.909

Note: (a) T-test statistics and corresponding p-values of structural break test using surprise loss are reported. Rejection of the null hypothesis would suggest presence of structural break. (b) Non-separable loss with shape parameter, $p = 2$.

Chapter 2: The Accuracy and Informativeness of Agricultural Baselines

2.1 Introduction

Long-term market projections play a vital role in both policy and investment decisions. Federal government statistical agencies, such as the United States Department of Agriculture's (USDA) Economic Research Service (ERS), are tasked with "collecting, producing, and disseminating data that the public, businesses, and governments use to make informed decisions"(Office of Management and Budget, 2020). To satisfy this mandate, ERS leads a team from 10 USDA agencies to produce annual projections of key measures of agricultural market conditions for the next decade. These projections facilitate comparisons of policy alternatives by providing a conditional "baseline" scenario based on specific macroeconomic, weather, policy, and trade assumptions. In addition, the Food and Agricultural Policy Research Institute (FAPRI) produces similar ten-year projections of key agricultural variables. USDA baseline projections are typically released in February, with the FAPRI baseline projections following in March. Over the years, the baseline projections have been used for a variety of purposes, including estimating farm program costs and preparing the President's budget. Despite their growing role in shaping agricultural policy, the baselines have not been rigorously evaluated. In this study, we evaluate the accuracy and informativeness of the USDA and FAPRI baselines using novel econometric techniques.

Our study focuses on two important series of projections from USDA and FAPRI: 1) projected bottom-line net cash income and its components, and 2) projected harvested acres, farm price, and yield of three major commodities (corn, soybeans, and wheat). The projections are examined in three steps. First, we examine the accuracy of both USDA and FAPRI projections using standard measures of accuracy, such as mean absolute percent error (MAPE) and root mean square percent error (RMSPE). As part of our preliminary analysis, we also investigate the degree to which each projection exhibits systematic bias, following Holden and Peel (1990b). Previous studies have identified a systematic downward bias in USDA's initial forecasts of bottom-line net cash income, crop receipts, livestock receipts, and cash expenses (Bora, Katchova, & Kuethe, 2021; Isengildina-Massa, Karali, Kuethe, & Katchova, 2021). Since many USDA forecasts are used as an input for the beginning conditions of the USDA baseline models, baseline projections may show a similar tendency to under-predict.

Second, we examine the extent to which the value of information for each series of projections diminishes across the projection horizon. Both USDA and FAPRI baselines provide projections for ten years into the future, and we test the null hypothesis that the projections become uninformative beyond a given horizon using the encompassing approach developed by Breitung and Knüppel (2021). Our tests of predictive content use the unconditional mean as the uninformative (naïve) benchmark, and compare the mean square error of the projections to the unconditional variance of the target variable in a regression framework. The informativeness tests may be particularly useful for policymakers with an interest in long-run policy concerns, such as climate change, which exceed the current ten-year horizon. If the current projections are uninformative beyond a few years, it will be difficult to provide accurate projections for longer horizons.

Finally, we formally test whether USDA or FAPRI provide more accurate baseline projections. Traditional forecast evaluation tests examine predictions at a single horizon (e.g., Diebold & Mariano, 1995; Harvey, Leybourne, & Newbold, 1997). Since the full ten-year path of baseline projections are used in policy analysis, these tests may provide inaccurate evaluation of relative accuracy, as one set of projections may perform better than the other at some horizons and worse at the remaining horizons. As a result, we evaluate the relative predictive accuracy of USDA and FAPRI using a novel testing procedure developed by Quaedvlieg (2021) that includes information across all horizons jointly. We test for superior predictive ability using two forms of the Quaedvlieg (2021) test. The first specification tests whether one set of projections perform better than the other across all projection horizons (uniform predictive ability). The second specification relaxes the assumption of uniform predictive ability by testing for differences in accuracy using a weighted average of loss differentials across horizons (average predictive ability). Thus, the second specification allows one baseline to have superior predictive ability over the other, even if it performs worse in some horizons. We further perform regression-based tests to examine whether the FAPRI projections encompass the USDA projections and vice versa (Harvey, Leybourne, & Newbold, 1998).

Our analysis yields a number of significant findings. First, the accuracy measures suggest that projection errors increase across the horizon for most variables, with the notable exception of crop yield projection. Second, our analysis identifies a number of systematic biases. For example, soybean harvested acres are consistently under-predicted while wheat harvested acres are consistently over-predicted at all horizons. In addition, net cash income, crop receipts, livestock receipts, and cash expenses are biased downward, consistent with previously reported bias in ERS's farm income forecasts for the one-year horizon (Bora et al., 2021; Isengildina-Massa et al., 2021), but the magnitude of bias increases with the projection horizon. Third, the tests of predictive content show that, for most variables, the projections stay informative up to 4-5 years and diminish thereafter. Finally, the multi-horizon comparison tests suggest that neither USDA nor FAPRI projections outperform one another across the entire projection horizon (uniform predictive ability), except for farm-related income, where FAPRI performs better than USDA, and corn price and soybean yield, where USDA perform better than FAPRI. However, the FAPRI projections perform better at shorter horizons, which may be a result of the later release and the potential to include updated information. These findings may have important implications for the models and processes used to produce the baseline projections by both USDA and FAPRI, as well as for projection users.

The remainder of the study is organized as follows. The next section provides a detailed description of the agricultural baseline projections produced by USDA and FAPRI, followed by a summary of our data. Subsequent sections describe our empirical approach and findings. The final section provides concluding remarks.

2.2 Agricultural Baseline Projections

A number of government agencies, international organizations, and private firms produce long-run projections of key economic variables to help formulate policy and to support long-term planning. For example, within the agricultural sector, the Organisation for Economic Co-operation and Development (OECD) produces a ten-year global outlook report in collaboration with the Food and Agricultural Organization (FAO) which contains projections of agricultural indicators, such as market conditions and consumption (OECD & Food and Agriculture Organization of the United Nations, 2020). In addition, the Congressional Budget Office (CBO) produces long-run cost projections for several mandatory Federal farm programs, such as price loss coverage (PLC), agricultural risk coverage (ARC), crop insurance, disaster assistance, and conservation programs. Each of these reports provide projections on some indicators of U.S. agricultural market conditions, such as prices, acreage, and yields of key commodities. In this study, however, we focus on the baseline projections produced by USDA and FAPRI, which offer the most comprehensive coverage of US agricultural indicators over a ten-year horizon. The USDA and FAPRI baselines provide projections for key indicators of agricultural market conditions, including commodity prices and production, global agricultural trade, and farm income.

USDA baseline projections are produced by the Interagency Agricultural Projections Committee, comprised of experts from 10 USDA agencies and offices. USDA emphasizes that the baseline projections are "not intended to be a forecast of what the future will be" (USDA Office of the Chief Economist, 2020, pp. 1). Instead, the USDA baseline offers a "conditional, long-run scenario about what would be expected to happen under a continuation of current farm legislation and other specific assumptions" (USDA Office of the Chief Economist, 2020, pp. iii). The specific assumptions include normal weather and the absence of domestic or external shocks affecting global agricultural supply and demand. In addition, the macroeconomic conditions, productivity growth rates, and trade policies are assumed to persist throughout the projection period. USDA's baseline projections reflect a composite of model results and judgment-based analysis (USDA ERS, 2020). The projections are designed to provide "a neutral reference scenario that can serve as a point of departure for a discussion of alternative farm sector outcomes that could result under different domestic or international conditions" (USDA Office of the Chief Economist, 2020, pp. 1). Hjort, Boussios, Seeley, and Hansen (2018) provides a detailed description of the USDA baseline model and various processes followed during the preparation of the baseline report. ERS begins the baseline projection process in August and September of the preceding year by developing domestic and international macroeconomic assumptions. Over the next few months, the committee prepares detailed core domestic analysis for program commodities, projections for livestock and other non-program commodities, and commodity projections for foreign countries. ERS economists then prepare the sector-wide projections for farm income and agricultural trade in January before publication of the baseline report in February.

FAPRI also produces 10-year baseline projections for the U.S. agricultural sector every year. Over the years, the FAPRI baseline procedures have evolved to include five main steps, as outlined in Meyers, Westhoff, Fabiosa, and Hayes (2010). First, FAPRI personnel update baseline models, data, and assumptions to include the November

World Agricultural Supply and Demand Estimates (WASDE) and the latest macroeconomic projections. Second, FAPRI analysts deliberate and produce preliminary baseline projections in late November. Third, the initial projections are subject to peer review from analysts from government and international agencies, agribusinesses, and other universities. Fourth, in mid-January, FAPRI analysts revise the preliminary baseline projections based on comments received during the peer review and update the WASDE and macroeconomic projections. Fifth, the baseline projections are finalized, and a briefing is provided to the U.S. Congress, after which the FAPRI baseline is released to the public. Meyers and Westhoff (2010) stress that the "FAPRI approach" of producing the baseline projections focuses on developing good models, while underlining their use by skilled analysts.

Agricultural baselines produced by USDA and FAPRI are widely used in farm policy debates, particularly as they relate to farm bills and other legislation affecting the agricultural sector. Both agencies play an advisory role in providing long-term budgetary estimates to policymakers and program administrators. It is important to note that one set of baseline projections examined are produced by the USDA, a department of the executive branch of the U.S. Federal government, while the other set is produced by a research institute housed at a Land Grant university. FAPRI was established by the U.S. Congress, part of the legislative branch. Thus, our work is complimentary to previous studies that evaluate forecasts produced by agencies from different branches of the Federal government (for a recent review, see Ericsson & Martinez, 2019).

As previously stated, U.S. agricultural baseline projections have not been rigorously evaluated, despite their role in shaping agricultural policy. There are a few recent exceptions. Irwin and Good (2015) question the use of USDA baseline projections in Farm Bill program choice decisions by demonstrating that corn, soybeans, and wheat price projections tend toward a steady state, leading to high projection errors. Westhoff (2015) extends the analysis in Irwin and Good (2015) to the FAPRI baselines and finds that the projection errors are similar to the USDA baselines across commodities. Irwin and Good (2015) and Westhoff (2015) also compare the commodity price baselines with season-average prices derived from futures markets. In addition, Boussios, Skorbiansky, and MacLachlan (2021) show that USDA baseline projections consistently under-estimate corn harvested area and over-estimate wheat harvested area. Finally, Kuethe, Bora, and Katchova (2022) compare the current year projections of US net cash income and its components to ERS's forecasts released in the same month. The study suggests that USDA baseline projections outperform ERS forecasts for government payments and farm-related income. While Kuethe et al. (2022) underline the potential of USDA baseline projections for short-run predictions, the study examines only current year projections, ignoring all other horizons.

2.3 Data and Descriptive Analysis

2.3.1 Data

We examine a set of agricultural baseline projections from both the USDA and FAPRI from 1997 to 2020. Both organizations publish their projections in a similar format. The baseline projections include the most recent USDA estimate at the time of publication, provisional USDA estimates for the previous year, and projections for the year of the current release and the next nine years. For example, the February 2020 USDA baseline report contains realized estimates for 2018, provisional estimates for 2019, and projections for 2020–2029. For some aggregate indicators, such as farm income, the baselines report calendar year values, while for commodities, they report marketing year values.

In this study, we examine two main series of projections in the baseline reports. First, we examine the projections of bottom-line net cash income and its components, which include crop receipts, livestock receipts, direct government payments, farm-related cash income, and cash expenses. Net cash income is a sector-wide measure of cash earnings generated by farms that can be used to meet a wide range of obligations, including debt payments (McGath et al., 2009b). It is defined as gross cash income less cash expenses. Gross cash income includes crop and livestock cash receipts, direct government payments, and farm-related income. Direct government payments are limited to federal government funds paid directly to farmers to support farm incomes, conserve natural resources, or compensate for natural disasters (McGath et al., 2009b). Farm-related income includes machine hire and custom work, forest products, and other income from farm output and sales. Net cash income is calculated from its components using a bottom-up approach as per the accounting equation:

$$\begin{aligned}\text{Net cash income} =& (\text{Crop receipts} + \text{Livestock receipts} + \text{Cash farm-related income} \\ & + \text{Direct government payments}) - \text{Cash expenses}.\end{aligned} \tag{2.1}$$

Second, we analyze the projections of harvested acres, farm price, and yield for three commodities: corn, soybeans, and wheat. Together, these three field crops constituted about 70% of the principal crops area planted in the US in 2020 (USDA NASS, 2021).[7] The projections are averages for the marketing years, which differ

[7] Crops included in area planted are corn, sorghum, oats, barley, rye, winter wheat, Durum wheat, other spring wheat, rice, soybeans, peanuts, sunflower, cotton, dry edible beans, chickpeas, potatoes, sugarbeets, canola, and proso millet.

by crop. The marketing year for corn begins on September 1 and comprises four quarters. For example, the marketing year 2020/21 for corn and soybeans starts on September 1, 2020, and ends on August 31, 2021. The 2020/21 marketing year for wheat begins on June 1, 2020, and ends on May 31, 2021. It is important to note that, as a result, the estimates for the current year are still provisional, as the marketing years for the various crops have yet to conclude. Similarly, the farm income estimates for the current year will be finalized in August.

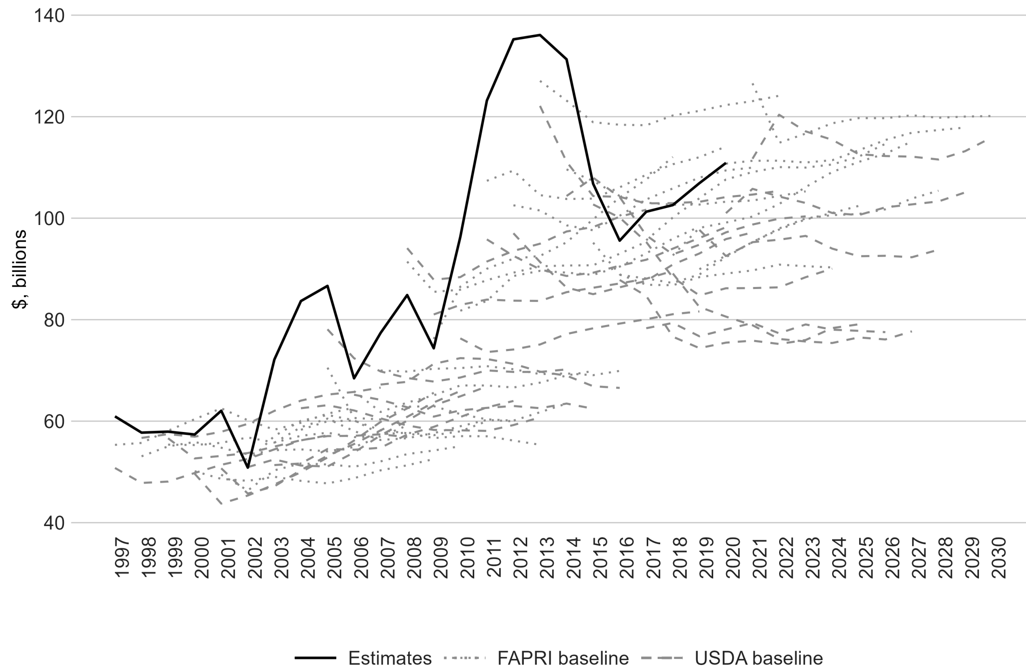
We compile our dataset from multiple online sources. The Albert R. Mann Library at Cornell University maintains an electronic archive of USDA baseline projections since 1997 (USDA ERS, 2021b). The majority of FAPRI baseline reports were obtained from the FAPRI website (FAPRI, 2021). For some early years, the baseline reports are available from the Iowa State University Digital Repository (Iowa State University, 2021).[8] The realized estimates for farm income indicators are taken from ERS's website (USDA ERS, 2021a). As mentioned in the previous discussion, the baseline reports also publish the realized values for two years before the release year. However, the realized estimates reported in the baseline report are subject to periodic adjustment as new information becomes available from multiple USDA agencies, such as the Census of Agriculture conducted once every five years. Therefore, instead of choosing the USDA or FAPRI release of realized estimates in their baseline reports, we use the most up-to-date information available at the ERS's website. Similarly, realized values for harvested acres, farm price, and yield of corn, soybeans, and wheat are obtained from the NASS Quickstats application programming interface (API) (USDA National Agricultural Statistics Service, 2021).

[8]FAPRI baseline projections are available for a few additional years before 1997, but for comparison with USDA, we limit our analysis to all years in which both sets of projections are available.

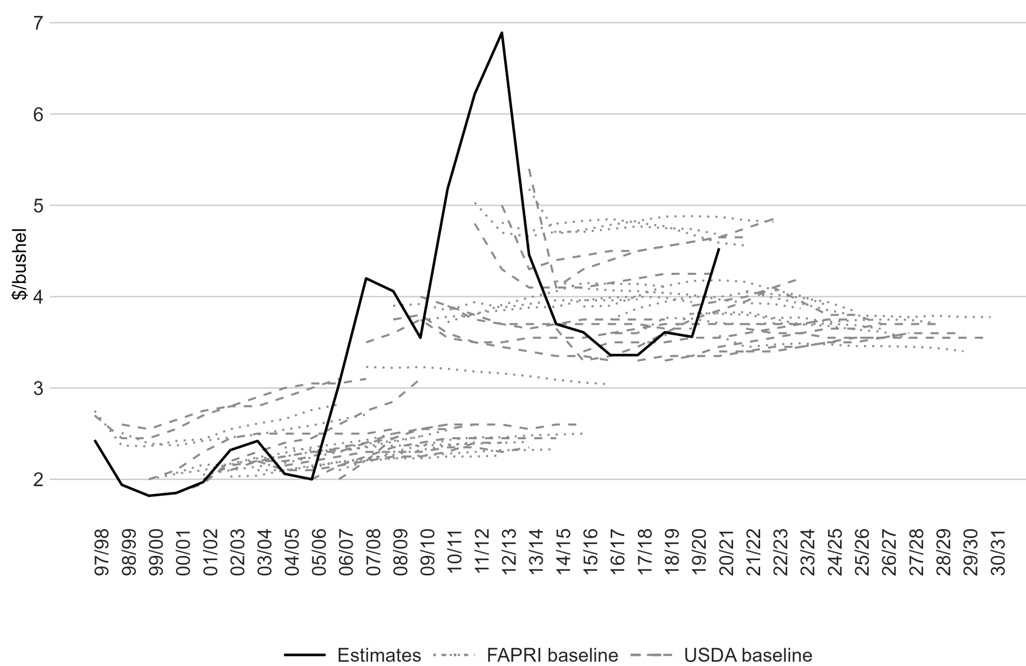
For each reference year (calendar or marketing year), we define Y_t as the realized value for year t for farm income and harvested acres, farm price, and yield for corn, soybeans, and wheat. We use the log transformations of the realized values: $y_t = \ln(Y_t)$ to eliminate the impact of changing forecast levels, following Isengildina-Massa et al. (2021). A projection made in year t for future year $t + h$ (at horizon h) by organization $i = \{USDA, FAPRI\}$ is denoted $\hat{Y}^i_{t+h|t}$. Again, we express the projection in natural logarithms of the variables for our analysis: $\hat{y}^i_{t+h|t} = \ln(\hat{Y}^i_{t+h|t})$. The projection horizon h can take values between $h = 0$ for the projection made during the reference year t and $h = 9$ for projections made for year $t + 9$. Again, for example, the 2020 baseline includes projections for 2020 ($h = 0$) to 2029 ($h = 9$).[9]

It is important to note that our dataset spans the baseline projections between 1997 and 2020, yet the evaluation period T differs for each projection horizon. The evaluation period for 0 years ahead horizon projections ($h = 0$) starts in 1997, and runs through 2020, resulting in a sample size of $T = 24$ observations. We lose one year from our sample size T for each year increase in the projection horizon h. For example, for $h = 1$, the length of the evaluation period is $T = 23$, as 1-year-ahead projections were not produced for the year $t = 1997$. Similarly, the sample size reduces to $T = 15$ observations for 9-years-ahead projections ($h = 9$), as $h = 9$ projections are available for the years 2006 to 2020. Figure 2.1 plots the baseline projections of net cash income and average farm prices of corn for the USDA and FAPRI reports between 1997 and 2021. As can be seen in the figure, the baseline projections are usually smoothed, particularly over longer horizons, and often fail to capture market shocks.

[9] In the forecast evaluation literature, projections made for $h = 0$ are sometimes referred to as *nowcasts*.



(a) Net cash income realized values and baseline projections, 1997-2021



(b) Corn price realized values and baseline projections, 1997-2021

Figure 2.1: Net cash income and corn price realized values and baseline projections between 1997 and 2021

2.3.2 Accuracy and Bias

Accuracy measures the difference between realized and predicted values. For each variable, the percent prediction 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 t is the reference year and $i = \{USDA, FAPRI\}$. We use two common measures of the relative accuracy of USDA and FAPRI projections: mean absolute percent error (MAPE) and root mean squared percent error (RMSPE) defined as,

$$\text{MAPE}^i_h = \frac{1}{T}\sum_t |e^i_{t+h|t}| \tag{2.2}$$

and

$$\text{RMSPE}^i_h = \sqrt{\frac{1}{T}\sum_t (e^i_{t+h|t})^2}. \tag{2.2'}$$

As MAPE is less susceptible to outliers, it is unaffected by the occasional large prediction errors. RMSPE, on the other hand, measures the square root average of squared errors and gives more weight to large prediction errors. Smaller MAPE or RMSPE values suggest more accurate projections.

In addition, we examine the degree to which the projections consistently differ from their realized values (bias) using the regression-based test of Holden and Peel (1990b). For each series of projections, we test for bias at each horizon $h = \{0, 1, \ldots 9\}$:

$$e^i_{t+h|t} = \alpha^i_h + \varepsilon^i_{t+h}. \tag{2.3}$$

where α^i_h is an unknown constant to be estimated and ε^i_{t+h} is white noise regression residual. The projections are unbiased if they do not consistently differ from realized values, or alternatively, their percent prediction error has an expected value of zero. We evaluate the null hypothesis that the projections are unbiased by testing the regression constraint $H_0 : \alpha^i_h = 0$. A positive and significant coefficient $\hat{\alpha}^i_h$ would suggest that the USDA or FAPRI projections consistently under-predict realized values. Similarly, a negative and significant coefficient $\hat{\alpha}_h^i$ implies that the projections systematically overestimate the realized values. For both USDA and FAPRI projections, we estimate equation (2.3) separately for each projection horizon h using ordinary least squares (OLS) with heteroskedasticity and autocorrelation consistent (HAC) standard errors (Newey & West, 1987).

Figures 2.2 and 2.3 plot the mean absolute percent error (MAPE, solid line) and root mean square percent error (RMSPE, dotted line) of the projections for field crop production and prices (figure 2.2) and net cash income and its components (figure 2.3) from 1997 through 2020. The vertical axis represents the MAPE and RMSPE, and the horizontal axis represents the projection horizon h, from 0 to 9.

As shown in figure 2.2, both MAPE and RMSPE increase with the projection horizon for harvested acres and farm price of corn, soybeans, and wheat for both USDA and FAPRI projections. This pattern, however, does not hold for crop yield projections. Corn yield projections exhibit smaller and more stable MAPE and RMSPE, and MAPE and RMSPE for wheat yields decreases across the projection horizon h. The stable or decreasing percent errors may be the result of small deviations in crop yields from long-term upward trends. Further, figure 2.2 suggests limited differences between USDA and FAPRI commodity price and production projections.

Figure 2.3 shows that projection errors for net cash income and its components increase with the horizon h. In addition, projection errors for net cash income, crop receipts, and livestock receipts are lower for the FAPRI baseline at shorter horizons, while USDA baseline projection errors are lower at longer horizons. For farm-related income, the FAPRI projection has lower errors for all horizons.

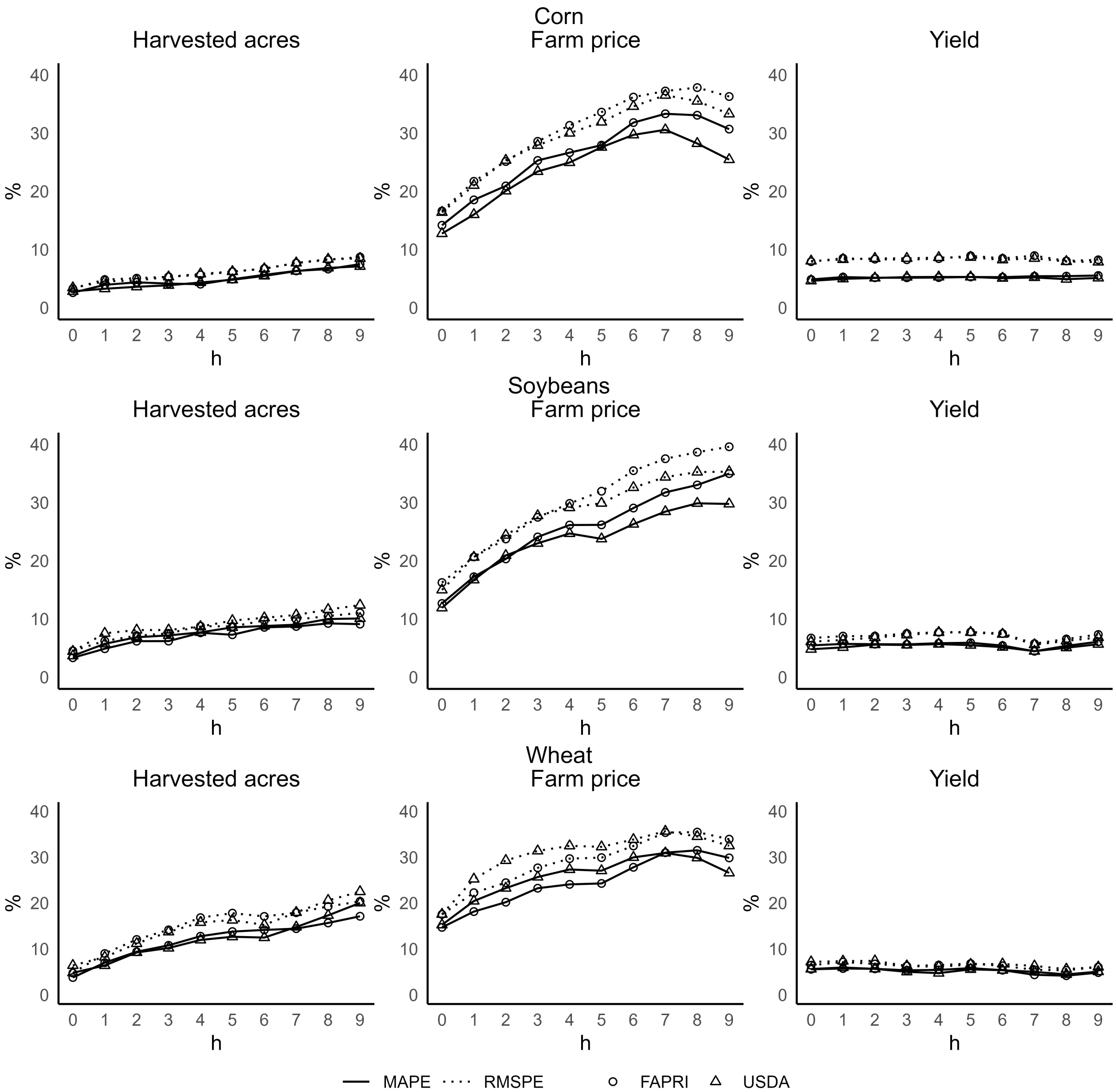


Figure 2.2: Mean absolute percent error (MAPE) and root mean square percent error (RMSPE) for baseline projections of corn, soybeans and wheat by projection horizon h, 1997–2020

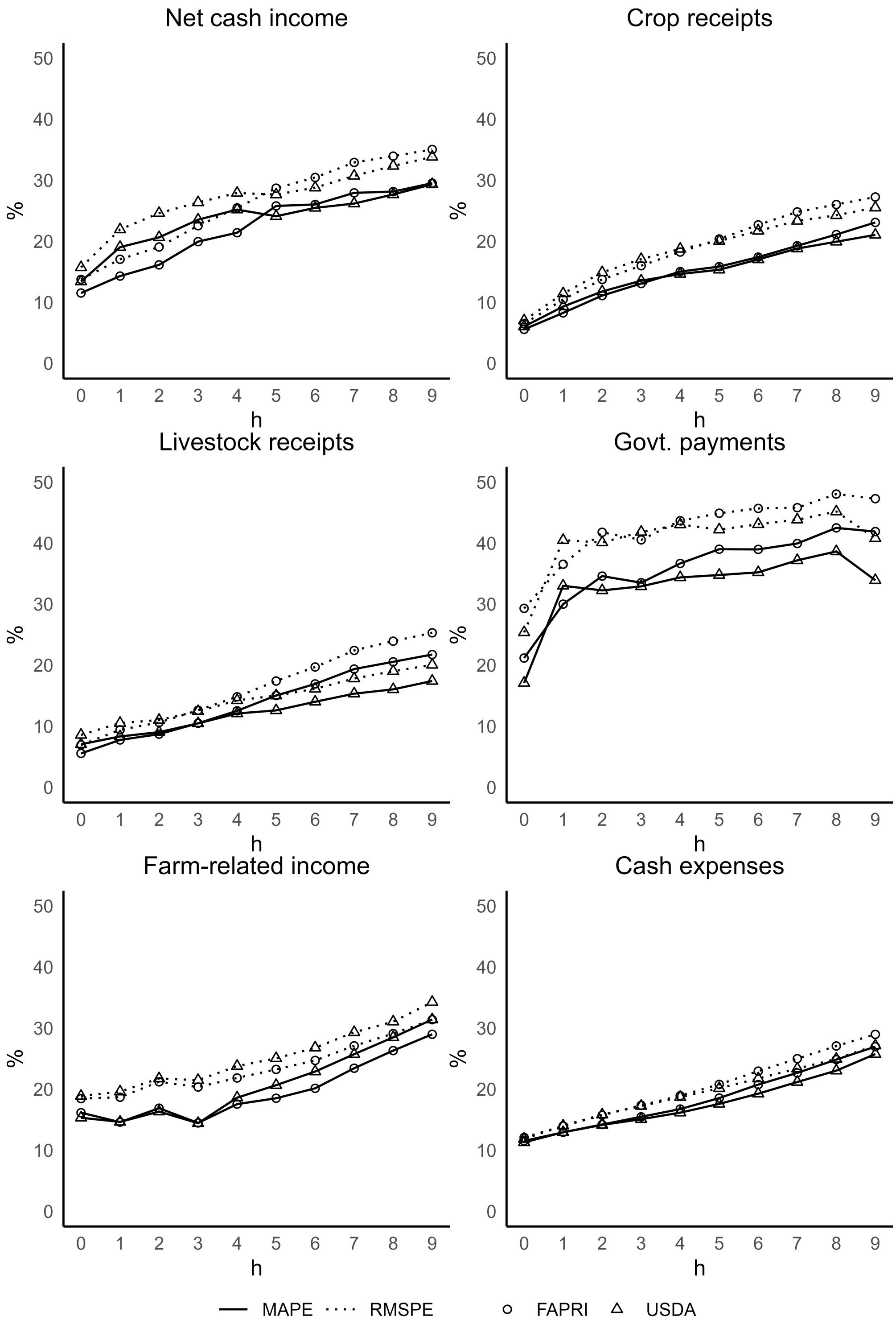


Figure 2.3: Mean absolute percent error (MAPE) and root mean square percent error (RMSPE) for baseline projections of net cash income and its components by projection horizon h, 1997–2020

The tests of bias for both commodity and net cash income projections show a similar pattern as reported in previous studies of USDA forecasts. In tables 2.1 and 2.2, we report the estimates of bias $\hat{\alpha}_h^i$ for projections $i = \{USDA, FAPRI\}$ at horizon h from equation (2.3) along with HAC standard errors. As reported in Boussios et al. (2021), the USDA baselines consistently under-predict soybean harvested acres and over-predict wheat harvested acres. The magnitude of bias increases with the projection horizon h. Corn harvested acres do not show such bias. Farm prices of the three commodities do not show significant bias for shorter horizons, but they tend to be under-predicted for horizons larger than four years. Crop yield predictions do not show significant bias for any of the three commodities. Both FAPRI and USDA projections of net cash income, crop receipts, livestock receipts, and cash expenses are biased downward at a 5% significance level, and the magnitude of bias increases with the horizon. This finding is consistent with previous findings of downward bias in USDA net cash income forecasts, which can be compared with projections at horizons $h = \{0, 1\}$ (Isengildina-Massa et al., 2021; Kuethe et al., 2022). As the short-term, one year USDA forecasts are an input for short-term baseline projections, it is not surprising that baselines are also biased downward, and that the bias carries forward to longer horizons. USDA projections of government payments show downward bias at longer horizons, while FAPRI projections of government payments do not show bias. Farm-related income projections are biased downward at longer horizons for both FAPRI and USDA projections.

Table 2.1: Estimates of bias in USDA baseline projections, 1997–2020

	Projection horizon									
	h=0	h=1	h=2	h=3	h=4	h=5	h=6	h=7	h=8	h=9
Corn										
Harvested acres	−0.009	−0.001	0.008	0.011	0.015	0.021	0.028	0.035	0.043	0.054
	(0.007)	(0.009)	(0.012)	(0.014)	(0.018)	(0.021)	(0.025)	(0.027)	(0.030)	(0.031)
Farm price	0.049	0.080	0.095	0.113	0.136*	0.160**	0.181	0.207*	0.262**	0.322**
	(0.039)	(0.069)	(0.095)	(0.084)	(0.072)	(0.069)	(0.105)	(0.098)	(0.095)	(0.129)
Yield	−0.010	−0.008	−0.007	−0.006	−0.006	−0.005	0.000	0.002	−0.003	−0.002
	(0.018)	(0.020)	(0.021)	(0.022)	(0.023)	(0.023)	(0.024)	(0.024)	(0.022)	(0.023)
Soybeans										
Harvested acres	0.008	0.028*	0.040**	0.055***	0.064***	0.072***	0.078***	0.085**	0.090**	0.098**
	(0.009)	(0.016)	(0.019)	(0.018)	(0.019)	(0.020)	(0.025)	(0.030)	(0.032)	(0.034)
Farm price	0.104***	0.133*	0.147	0.164	0.184***	0.212***	0.236**	0.251***	0.302***	0.360***
	(0.036)	(0.067)	(0.092)	(0.098)	(0.060)	(0.061)	(0.097)	(0.071)	(0.081)	(0.082)
Yield	−0.002	−0.002	−0.002	0.002	0.005	0.005	0.009	0.019	0.018	0.016
	(0.014)	(0.016)	(0.020)	(0.021)	(0.022)	(0.024)	(0.024)	(0.022)	(0.026)	(0.032)
Wheat										
Harvested acres	−0.040***	−0.046**	−0.056*	−0.071*	−0.090**	−0.102**	−0.112***	−0.134***	−0.155***	−0.179***
	(0.008)	(0.016)	(0.028)	(0.034)	(0.038)	(0.036)	(0.031)	(0.032)	(0.030)	(0.027)
Farm price	0.059	0.102	0.126	0.146*	0.166*	0.182**	0.189**	0.205*	0.240**	0.278*
	(0.047)	(0.088)	(0.101)	(0.080)	(0.094)	(0.069)	(0.084)	(0.116)	(0.112)	(0.137)
Yield	0.019	0.018	0.016	0.015	0.015	0.016	0.027	0.025	0.026*	0.028*
	(0.015)	(0.017)	(0.015)	(0.014)	(0.013)	(0.015)	(0.017)	(0.015)	(0.014)	(0.015)
Farm income										
Net cash income	0.132***	0.194***	0.238***	0.267***	0.284***	0.288***	0.312***	0.330***	0.347***	0.376***
	(0.027)	(0.041)	(0.045)	(0.038)	(0.055)	(0.065)	(0.072)	(0.088)	(0.099)	(0.096)
Crop receipts	0.036*	0.060	0.081	0.097***	0.113***	0.134***	0.152***	0.172***	0.198***	0.249***
	(0.020)	(0.043)	(0.055)	(0.033)	(0.015)	(0.012)	(0.026)	(0.042)	(0.036)	(0.066)
Livestock receipts	0.022	0.042*	0.059**	0.072**	0.084*	0.096*	0.121**	0.147**	0.163**	0.189***
	(0.018)	(0.023)	(0.028)	(0.033)	(0.044)	(0.049)	(0.052)	(0.053)	(0.057)	(0.058)
Govt. payments	0.164	0.297*	0.361*	0.445**	0.474**	0.451**	0.469**	0.472*	0.491**	0.436*
	(0.096)	(0.168)	(0.188)	(0.196)	(0.210)	(0.209)	(0.222)	(0.228)	(0.214)	(0.209)
Farm-related income	0.048	0.092	0.124	0.151*	0.193**	0.235**	0.274***	0.314***	0.351***	0.397***
	(0.061)	(0.074)	(0.090)	(0.087)	(0.089)	(0.084)	(0.076)	(0.063)	(0.060)	(0.064)
Cash expenses	0.121***	0.140***	0.156***	0.168***	0.183***	0.201***	0.222***	0.246***	0.270***	0.305***
	(0.011)	(0.021)	(0.030)	(0.035)	(0.040)	(0.020)	(0.034)	(0.032)	(0.038)	(0.042)

Notes: The bias term $\hat{\alpha}_h^{USDA}$ is estimated from the equation (2.3). ***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively. Standard errors (in parentheses) are heteroskedasticity and autocorrelation consistent (HAC)(Newey & West, 1987). The sample sizes of regressions for h=0,1,2,...,9 are T=24, 23,..., 15 respectively. For farm income variables, sample size for h=9 is 14 as the 1997 USDA baseline didn't publish projections for the year 2006.

Table 2.2: Estimates of bias in FAPRI baseline projections, 1997–2020

	Projection horizon									
	h=0	h=1	h=2	h=3	h=4	h=5	h=6	h=7	h=8	h=9
Corn										
Harvested acres	−0.005	−0.006	−0.003	0.004	0.007	0.014	0.022	0.031	0.039	0.047
	(0.005)	(0.010)	(0.013)	(0.015)	(0.020)	(0.021)	(0.025)	(0.028)	(0.032)	(0.033)
Farm price	0.026	0.047	0.071	0.094	0.124	0.156	0.189	0.223*	0.284**	0.350**
	(0.045)	(0.076)	(0.101)	(0.078)	(0.087)	(0.109)	(0.120)	(0.105)	(0.112)	(0.140)
Yield	−0.003	−0.001	−0.002	−0.002	−0.002	−0.003	0.001	0.001	−0.004	−0.003
	(0.018)	(0.020)	(0.022)	(0.022)	(0.023)	(0.024)	(0.025)	(0.026)	(0.024)	(0.025)
Soybeans										
Harvested acres	0.003	0.020	0.034**	0.042**	0.054**	0.062***	0.072***	0.079***	0.084***	0.089***
	(0.009)	(0.013)	(0.016)	(0.018)	(0.019)	(0.020)	(0.020)	(0.023)	(0.025)	(0.028)
Farm price	0.077*	0.091	0.105	0.137	0.168**	0.211**	0.249*	0.273**	0.329***	0.391***
	(0.039)	(0.078)	(0.097)	(0.105)	(0.076)	(0.094)	(0.127)	(0.117)	(0.109)	(0.128)
Yield	0.009	0.009	0.009	0.012	0.013	0.013	0.014	0.024	0.023	0.022
	(0.016)	(0.019)	(0.022)	(0.023)	(0.023)	(0.023)	(0.024)	(0.022)	(0.027)	(0.034)
Wheat										
Harvested acres	−0.026***	−0.048**	−0.067**	−0.084**	−0.103**	−0.111**	−0.116***	−0.130***	−0.141***	−0.154***
	(0.008)	(0.017)	(0.029)	(0.036)	(0.036)	(0.042)	(0.038)	(0.032)	(0.036)	(0.034)
Farm price	0.038	0.064	0.086	0.102	0.128	0.155	0.178	0.205	0.248	0.291*
	(0.059)	(0.086)	(0.092)	(0.083)	(0.108)	(0.106)	(0.110)	(0.145)	(0.150)	(0.143)
Yield	0.026*	0.023	0.021	0.019	0.018	0.020	0.031**	0.030**	0.031**	0.035**
	(0.013)	(0.015)	(0.014)	(0.013)	(0.014)	(0.016)	(0.014)	(0.012)	(0.011)	(0.014)
Farm income										
Net cash income	0.116***	0.147***	0.164***	0.190***	0.228***	0.261***	0.299***	0.326**	0.354***	0.372***
	(0.023)	(0.032)	(0.041)	(0.055)	(0.068)	(0.087)	(0.094)	(0.122)	(0.100)	(0.086)
Crop receipts	0.030*	0.048	0.062	0.078**	0.096***	0.121***	0.147***	0.170**	0.203***	0.239***
	(0.017)	(0.040)	(0.053)	(0.035)	(0.026)	(0.036)	(0.049)	(0.066)	(0.061)	(0.072)
Livestock receipts	0.030**	0.044**	0.059*	0.080**	0.107**	0.136**	0.175**	0.206**	0.230***	0.252***
	(0.014)	(0.020)	(0.028)	(0.037)	(0.047)	(0.058)	(0.063)	(0.071)	(0.072)	(0.071)
Govt. payments	0.156	0.226	0.272	0.292	0.303	0.287	0.290	0.287	0.303	0.267
	(0.094)	(0.157)	(0.202)	(0.178)	(0.239)	(0.242)	(0.249)	(0.246)	(0.218)	(0.241)
Farm-related income	0.032	0.056	0.087	0.113	0.156*	0.201**	0.240***	0.282***	0.321***	0.357***
	(0.061)	(0.061)	(0.085)	(0.089)	(0.088)	(0.082)	(0.077)	(0.068)	(0.062)	(0.058)
Cash expenses	0.123***	0.140***	0.156***	0.172***	0.189***	0.212***	0.240***	0.266***	0.296***	0.325***
	(0.011)	(0.018)	(0.026)	(0.034)	(0.038)	(0.038)	(0.037)	(0.042)	(0.049)	(0.050)

Notes: The bias term $\hat{\alpha}_h^{FAPRI}$ is estimated from the equation (2.3). ***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively. Standard errors (in parentheses) are heteroskedasticity and autocorrelation consistent (HAC)(Newey & West, 1987). The sample sizes of regressions for h=0,1,2,...,9 are T=24, 23,..., 15 respectively.

2.4 Methods

The analysis of accuracy and bias in the previous section suggests that the projections are less accurate at longer horizons. We conduct tests for the predictive accuracy of the projections at different horizons and determine the maximum informative projection horizon for each variable (Breitung & Knüppel, 2021). We then compare the USDA and FAPRI baseline models using multi-horizon tests developed by Quaedvlieg (2021).

2.4.1 Informativeness

The accuracy and bias measures consider the projections at each horizon independently. The baseline projections, however, are multi-horizon forecasts or *path forecasts*, as in Jordà and Marcellino (2010). An important evaluation criterion for path forecasts is the horizon up to which the projections provide meaningful information. Galbraith (2003) calls the maximum informative horizon of a path forecast the *content horizon.* A number of previous studies develop empirical tests to estimate the content horizon of path forecasts relative to an uninformative or naïve forecast (Galbraith & Tkacz, 2007; Isiklar & Lahiri, 2007).

A popular measure used for quantifying information content is the Theil's U statistic (Theil, 1958). Theil's U is a scaled version of the root mean square error (RMSE) that has the advantage of not being affected by the variance of the actual process. It is defined:

$$U_h^i(\hat{y}_{na\ddot{i}ve}) = \sqrt{\frac{\sum_{t=1}^{T}(y_{t+h} - \hat{y}_{t+h|t}^i)^2}{\sum_{t=1}^{T}(y_{t+h} - \hat{y}_{na\ddot{i}ve})^2}} \tag{2.4}$$

A common choice for the naïve projection, $\hat{y}_{na\ddot{i}ve}$, is a no-change projection using the previous year's estimate. Following Isiklar and Lahiri (2007), we use the previous 5-year's average as the naïve projection, and calculate $U_h^i(\hat{y}_{na\ddot{i}ve})$ for our selected variables for each horizon $h = \{0, 1, \ldots 9\}$, and agency $i = \{USDA, FAPRI\}$. If Theil's U is less than one, then the baseline is a better predictor than the naïve projection. Conversely, when the naïve benchmark is a better predictor than the agency baseline, Theil's U is larger than one.

The choice of the naïve projection $\hat{y}_{na\ddot{i}ve}$ greatly influences the Theil's U statistic. As a result, we also estimate the informativeness or content horizon for the agricultural baselines using a method recently proposed by Breitung and Knüppel (2021), which does not require a naïve forecast for comparison. Instead, the Breitung and Knüppel test directly compares the mean-squared forecast error to the unconditional variance of the forecasted variable. The Breitung and Knüppel testing framework is based on a limited set of assumptions. The test assumes that the realized values y_t are generated by a stationary and ergodic stochastic process. We further assume that the realized values y_t are generated by a linear process with constant variance, although this assumption may be relaxed in some conditions.

The Breitung and Knüppel test for the maximum informative prediction horizon compares the mean-squared prediction error of the projections to the variance of the realized values over the evaluation sample. Under quadratic loss, the optimal projection equals the conditional mean of the projection $\mu_{h,t}^i = E(\hat{y}_{t+h}^i | I_t)$, given the information set I_t available at reference year t. In particular, we test the following hypothesis:

$$H_0 : E(y_{t+h} - \hat{y}_{t+h|t})^2 \geq E(y_{t+h} - \mu)^2, \text{ for } h > h^* \tag{2.5}$$

$$H_1 : E(y_{t+h} - \hat{y}_{t+h|t})^2 < E(y_{t+h} - \mu)^2 \tag{2.6}$$

where, $\mu = E(y_t)$ is the unconditional mean of the realized values. The null hypothesis states that there exists a maximum projection horizon h^* beyond which the realized values y_t would be unpredictable with respect to the information set I_t. We term the null hypothesis as *no information* hypothesis, against the alternative hypothesis, which states that the projection remains informative as the mean-squared prediction error is lower than the variance of the realized values around their unconditional mean.

Another test of predictive content can be formulated based on the conditional mean of the projection being constant within the evaluation sample, or the *constant mean* hypothesis:

$$H_0 : E(\hat{y}^i_{t+h}|I_t) = \mu_{h,t} = \mu, \text{ for } h > h^* \tag{2.7}$$

$$H_1 : E(\hat{y}^i_{t+h}|I_t) \neq \mu_{h,t} = \mu. \tag{2.8}$$

This is a more relaxed criterion compared to the *no information* hypothesis as it requires the projection to be uncorrelated with the realized value for it to be uninformative. If the projection $\hat{y}^i_{t+h|t}$ is identical to the conditional mean $\mu_{h,t}$ of the target variable, then the *no information* hypothesis is equivalent to the *constant mean* hypothesis (Breitung & Knüppel, 2021).

Breitung and Knüppel (2021) suggest considering three scenarios based on how the projections are generated. The first scenario refers to projections generated from the expectations of individuals, and the expectation is identical to some conditional mean. The second scenario involves projections generated from survey expectations which are also contaminated by noise (e.g., macro-economic forecasts of Consensus Economics). The third scenario refers to projections generated from models. The baseline projections we consider here are unique in the sense that they are generated based on models, as well as expert opinions or expectations of individuals. Therefore, we consider the second and third scenarios. In both scenarios, the *no information* hypothesis and the *constant mean* hypothesis can be formulated in terms of testing coefficients in a Mincer-Zarnowitz regression (Mincer & Zarnowitz, 1969).

Breitung and Knüppel (2021) show that if the baseline projection is generated by a conditional mean of the projection and noise (η_t), $\hat{y}^i_{t+h|t} = \mu_{h,t} + \eta^i_t$, the *no information* hypothesis is equivalent to testing the null hypothesis $\beta^i_h \leq 0.5$ in the regression:

$$y^i_{t+h} = \beta^i_{0,h} + \beta^i_h \hat{y}^i_{t+h|t} + \nu^i_{t+h}. \tag{2.9}$$

Breitung and Knüppel (2021) further show that the *constant mean* hypothesis is equivalent to testing the null hypothesis $\beta^i_h \leq 0$ in the same regression, implying that the baseline projection is uncorrelated to the realized values. The tests of the parameters β^i_h can be performed using a HAC t-statistic constructed as:

$$\tau_a = \frac{1}{\hat{\omega}_a \sqrt{T}} \sum_t a_t \tag{2.10}$$

$$a_t = \left[y_{t+h} - \overline{y_{t+h}} - 0.5(\hat{y}_{t+h|t} - \overline{\hat{y}_{t+h}}) \right] (\hat{y}_{t+h|t} - \overline{\hat{y}_{t+h}}) \text{ for } H_0 : \beta_h = 0.5 \tag{2.11}$$

$$a_t = (y_{t+h} - \overline{y_{t+h}})(\hat{y}_{t+h|t} - \overline{\hat{y}_{t+h}}) \text{ for } H_0 : \beta_h = 0 \tag{2.12}$$

where $\hat{\omega_a}^2$ is a consistent estimator of the long-run variance of a_t. The Lagrange Multiplier statistic has an asymptotic standard normal distribution.

While constructing the HAC t-statistic, we use the in-sample mean of the baseline projections. While alternative versions of the test use a recursive mean in place of the in-sample mean, they require more information prior to the evaluation period, which is not available in our case. To determine the maximum informative projection horizon h^*, we begin by testing the null hypothesis for the $h = 0$ horizon projection. If the null hypothesis is rejected, we test the $h = 1$ horizon projection, and so on. We stop when the null hypothesis is no longer rejected. The maximum informative projection horizon h^* is the penultimate horizon before the null hypothesis is not rejected.

An advantage of the tests proposed by Breitung and Knüppel (2021) is that they do not require a naïve benchmark, as they directly compare the mean-squared prediction error to the unconditional variance of the realized values. Another advantage is that when we apply the tests with in-sample mean, additional information prior to the evaluation period is not required, therefore these tests are suitable for our limited observation period. The baseline projections share properties of both survey forecasts and model-based forecasts, as they are a combination of model prediction and expert opinions. On the other hand, a limitation of these tests is that they can be sensitive to the transformations of the variables. In addition, the maximum information projection horizon is a conservative estimate and is subject to the process used to produce the projections. The maximum informative projection horizon could be longer if the projection process did not fully incorporate available information. Another limitation is that the tests at longer projection horizons may have less power due to smaller sample size.

2.4.2 Comparing USDA and FAPRI Baseline Projections

Multi-horizon Comparison

The final step in our evaluation compares the relative accuracy of the baseline projections produced by USDA and FAPRI. First, we follow the forecast comparison test procedure developed by Harvey et al. (1997) for each projection at each horizon. The Harvey et al. (1997) test is a modified version of the test procedure introduced by Diebold and Mariano (1995) which incorporates a modified student t distribution and bias correction to improve small sample properties of the tests. The comparison of the USDA and FAPRI projections is based on the mean loss differential between them, $\boldsymbol{\mu} = \lim_{T\to\infty} \frac{1}{T}\sum_t E(\boldsymbol{d}_t)$.

The modified Diebold-Mariano tests for single horizons compare the USDA and FAPRI projections by calculating a standard t-test:

$$t_{DM}^{h} = \frac{\sqrt{T}\bar{d}_h}{\hat{\omega_h}} \tag{2.13}$$

where $\bar{d}_h = \frac{1}{T}\sum d_{t,h}$, and $\hat{\omega_h^2}$ is a HAC estimate of the variance of $d_{t,h}$. We first test the null hypothesis that the mean loss differential at horizon h is less than or equal to zero ($H_0 : \mu_h \leq 0$). A failure to reject the null hypothesis of $\mu_h \leq 0$ suggests that the FAPRI projections do not perform better than USDA, and a rejection of the null would indicate that the FAPRI projections perform better than USDA. We then test the null hypothesis that the mean loss differential at horizon h is greater than or equal to zero ($H_0 : \mu_h \geq 0$). For this test, a failure to reject the null hypothesis of $\mu_h \geq 0$ would indicate that the USDA projections do not perform better than FAPRI, and a rejection of the null would indicate that the USDA projections perform better than FAPRI.

The modified Diebold-Mariano test compares the USDA and FAPRI projections at each horizon. As a result, the test may yield contradictory results for multi-horizon projections, as one set of projections may provide more accurate projections at some horizons but not at others. This shortcoming may limit the use by policymakers who are interested in the relative accuracy of the entire path forecast from horizons 0 through 9. As a result, we also examine the relative accuracy along the entire projection path.

A number of recent studies propose methods to compare the relative accuracy of path forecasts (Capistrán, 2006; Martinez, 2020; Patton & Timmermann, 2012). In our analysis, we use the tests of multi-horizon superior predictive ability proposed by Quaedvlieg (2021) which jointly consider all horizons along the entire projection path. Following Giacomini and White (2006), the procedure developed by Quaedvlieg (2021) tests for finite-sample multi-horizon predictive ability using estimated values of parameters. To conduct multi-horizon comparison tests, we start by using a vectorized version of the previous notations, denoting the USDA and FAPRI projections $i \in \{USDA, FAPRI\}$ as, $\hat{\boldsymbol{y}}_t^i = [\hat{y}_{t|t-0}^i, \hat{y}_{t|t-1}^i, \ldots, \hat{y}_{t|t-9}^i]$, where $\hat{y}_{t|t-h}^i$ is the projection of y_t based on the information set at time $t-h$. We are interested in comparing the USDA and FAPRI projections in terms of their loss differentials, following the approach in Diebold and Mariano (1995). We assume a general loss function $\boldsymbol{L}_t^i = L(\boldsymbol{y_t}, \hat{\boldsymbol{y}}_t^i)$ which maps the prediction errors into a 10-dimensional vector since there are 10 projection horizons. For our analysis, we use mean squared error (MSE) and mean absolute error (MAE) loss function, however, these can be generalized to allow multivariate loss function. We calculate the loss differential for year t between the USDA and

FAPRI projections as a 10-dimensional vector:

$$\boldsymbol{d}_t = \boldsymbol{L}_t^{USDA} - \boldsymbol{L}_t^{FAPRI}. \tag{2.14}$$

Quaedvlieg (2021) provides two alternative definitions of multi-horizon predictive ability. First, a path forecast is said to have *uniform* superior predictive ability (uSPA) if it has smaller loss at each horizon when compared to the alternative path forecast. Uniform SPA, however, is a very strict criterion which may not be realistic in practice. As a result, Quaedvlieg (2021) develops the concept of *average* superior predictive ability (aSPA) for a path forecast with larger loss at some horizons that is compensated by superior performance at other horizons when compared to the alternative path forecast. Thus, average SPA relaxes the stringent requirements of uniform SPA. Quaedvlieg construct bootstrap test statistics for both uniform and average SPA, which reduce to the standard DM tests at a single horizon.

The uniform SPA test is based on the minimum loss differential:

$$\mu^{uSPA} = \min_h \mu_h. \tag{2.15}$$

The uniform SPA test is given by the null hypothesis $H_0 : \mu^{uSPA} \leq 0$ against the alternative hypothesis $H_a : \mu^{uSPA} > 0$. Rejecting the null hypothesis will suggest that the FAPRI projection has uniform superior predictive ability over the USDA projection. In other words, the minimum loss differential between the USDA and FAPRI projection across horizons h should be significantly greater than zero if the FAPRI projection is to be uniformly superior to the USDA projection. To test for uSPA of the USDA projection over FAPRI, we use the same equation (2.15) for minimum loss differential but reverse the two projections in the loss differentials equation (i.e. $\boldsymbol{d}_t = \boldsymbol{L}_t^{FAPRI} - \boldsymbol{L}_t^{USDA}$). In this case, rejecting the null hypothesis will suggest that the USDA projection has uniform superior predictive ability over the FAPRI projection.

The average SPA test, by contrast, is based on a weighted average of losses across all horizons or whether, for example, the FAPRI baseline projection is on average superior to the USDA baseline projection across all horizons. The average SPA test is based on the minimum loss differential:

$$\mu^{aSPA} = \boldsymbol{w}'\boldsymbol{\mu} = \sum_{h} w_h \mu_h. \tag{2.16}$$

The average SPA allows losses at different horizons to compensate for one another. For example, the FAPRI projection may perform worse at some horizons but still be superior compared to the USDA projection, on average. We test the null hypothesis $H_0 : \mu^{aSPA} \leq 0$ (FAPRI projection does not have aSPA) against the alternative $H_a : \mu^{aSPA} > 0$ (FAPRI projection has aSPA). We also test the null hypothesis $H_0 : \mu^{aSPA} \geq 0$ (USDA projection does not have aSPA) against the alternative $H_a : \mu^{aSPA} > 0$ (USDA projection has aSPA).

The choice of weights (w_h) is flexible but is chosen *a priori*. To make sure our findings are robust to this choice, we examine alternative weighting procedures. We first use equal weights for each horizon h but also consider weighing the loss differentials by the variance of the loss differential at the horizon that is being compared divided by the sum of variances across all horizons. The test statistic for the multi-horizon comparison tests are given by:

$$t_{uSPA} = \min_{h} \frac{\sqrt{T}\bar{d}_h}{\hat{\omega_h}} \tag{2.17}$$

and,

$$t_{aSPA} = \frac{\sqrt{T}\bar{d}_h}{\hat{\zeta}_h}, \tag{2.17'}$$

respectively. For the uSPA tests, we calculate two t-statistics: one testing whether the FAPRI projection has uSPA over the USDA and the other testing whether the USDA projection has uSPA over FAPRI, as the minimum loss differentials are different for these two hypotheses. However, for the aSPA tests, we need to calculate only one t-statistic and conduct one-tailed tests in both directions to test aSPA of the FAPRI projection over USDA and vice versa.

We obtain estimates of variances $\hat{\omega_h^2}$ for uSPA from the diagonal elements of the covariance matrix of loss differential $\boldsymbol{d}$ calculated using an HAC-type estimator (Newey & West, 1987). Similarly, we get the estimates of variance $\hat{\zeta_h^2}$ for aSPA as the diagonal elements of the weighted covariance matrix of $\boldsymbol{d}$. The test-statistic for the uniform SPA is the minimum of Diebold-Mariano test statistic for all horizons. The average SPA test is simply a Diebold-Mariano test on average loss differential (Quaedvlieg, 2021). The critical values and p-values for the uSPA and aSPA tests are obtained using a moving block bootstrap (MBB) technique. By computing either of the test statistics on many MBB re-samples, we approximate the distribution of the original statistics under the null hypothesis. The critical values at α significance level are obtained by calculating the α percentile of the bootstrap distribution.

Encompassing Tests

As previously stated, USDA baseline projections are typically released in February and FAPRI baseline projections in March. As a result, FAPRI analysts have the advantage of using more recent information to prepare their projections. The updated information set of the FAPRI analysts includes information from reviewer comments, the January WASDE and associated reports, and the February USDA farm income estimates. In comparison, the USDA baseline projections are based on the October

WASDE (USDA Office of the Chief Economist, 2020). Therefore, one might expect the FAPRI baseline projections to contain new information beyond the USDA baseline projections. On the other hand, there is a bi-directional flow of information between USDA and FAPRI analysts through official meetings, review sessions, and informal conversations, which may lead to herding in the projections produced by both agencies. FAPRI usually finalizes its projections by the time USDA releases its report, and the USDA report does not act as a significant input to FAPRI's forecasting process. We test whether the information content of USDA or FAPRI baseline projections dominates the other using the encompassing test developed by Harvey et al. (1998).

When two competing sets of projections are available for the same variable, a relevant question to ask is whether one set of projections *encompasses* another, that is, the informational content of the preferred projection dominates the other. Harvey et al. (1998) frame this question as a problem of forming a combined projection from the weighted average of the individual ones and estimating the optimal weights assigned to each projection. In this framework, a projection would be preferred if its optimal weight is unity in the weighted average, and the combined projection consists entirely of the preferred projection. Harvey et al. (1998) develop a regression-based test to estimate the optimal weights for the combined projection. For our study, the regression is expressed as:

$$e^{USDA}_{t+h|t} = \alpha_h + \lambda_h(e^{USDA}_{t+h|t} - e^{FAPRI}_{t+h|t}) + \varepsilon_{t+h|t}. \tag{2.18}$$

where $e^{USDA}_{t+h|t}$ is the prediction error at horizon h of USDA baselines, and $e^{FAPRI}_{t+h|t}$ is the prediction error at horizon h of FAPRI baselines projections. The coefficients α_h and

λ_h at horizon h are estimated by OLS regression, and $\varepsilon_{t+h|t}$ is a white noise regression error.

The coefficient λ_h in the regression equation (2.18) determines the optimal weights assigned to the USDA and FAPRI projections to form a combined projection that would have a smaller expected squared error than either of the two projections. The combined projection is formed by assigning weights $(1 - \lambda_h)$ and λ_h to the USDA projections and FAPRI projections, respectively. We test the null hypothesis that USDA baselines encompass the FAPRI projections using a two-tailed t-test of the restriction $\lambda_h = 0$. If we fail to reject $\lambda = 0$, it implies that USDA is preferred to FAPRI (i.e., the combined projection consists entirely of the USDA baseline). Alternatively, we test the hypothesis that the combined projection consists entirely of the FAPRI baseline by using a two-tailed t-test of the restriction $\lambda_h = 1$. A failure to reject $\lambda_h = 1$ would suggest that the FAPRI baseline is preferred. If we reject both $\lambda_h = 0$ and $\lambda_h = 1$, a combined projection is formed by weighting FAPRI baseline by $\hat{\lambda}_h$ and USDA baseline by $(1 - \hat{\lambda}_h)$. In this case, both baselines contain unique information to contribute to the combined projection. Finally, if we fail to reject both $\lambda_h = 0$ and $\lambda_h = 1$, the optimal composite projection can be either the USDA or the FAPRI baseline, as the projections are very similar. We perform encompassing tests for our selected variables for each horizon separately.

2.5 Results

The following section presents the primary findings of our analysis. First, we measure the accuracy of USDA and FAPRI baseline projections for major field crops, as well as U.S. net cash farm income and its components, and test for bias. Second, we estimate the content horizon of each set of projections. Finally, we empirically test the degree to which USDA or FAPRI has superior predictive ability.

2.5.1 Informativeness

The Theil's U statistic for the commodities and net cash income components are plotted in figures 2.4 and 2.5 for USDA and FAPRI baselines. As previously discussed, Theil's U compares the predictive accuracy of FAPRI or USDA baseline projections relative to a naïve prior based on a 5-year moving average, following Isiklar and Lahiri (2007). The USDA or FAPRI baseline projection is a better predictor than the naïve prior if Theil's U is less than 1, which is represented by the horizontal dashed line in figures 2.4 and 2.5. Both USDA and FAPRI projections are better predictors of corn harvested acres across all horizons, relative to the naïve prior. However, for harvested acres of soybeans and wheat, the naïve prior is preferred at longer horizons. The predictive accuracy of both FAPRI and USDA baseline projections relative to the naïve prior diminish at longer horizons for all farm price projections. Interestingly, yield projections perform better for both agencies at larger horizons relative to the naïve projection. For net cash income, FAPRI baselines are preferred to the naïve prior for at shorter horizons, yet the naïve prior is preferred to USDA baseline projections beyond the reference year projections. For both crop and livestock receipts, USDA and FAPRI baseline projections are preferred to the naïve prior at all horizons. The projections of government payment, on the other hand, fail to beat the naïve beyond the current year for both agencies. Overall, Theil's U statistics suggest that the baselines beat the naïve projection for most variables across horizons, underlining that baselines contain information. We investigate the informativeness of the baselines further with our empirical tests of predictive content.

Our estimates of the content horizon of each projection series, following Breitung and Knüppel (2021), are presented in tables A.1 and A.2. As previously discussed, the empirical test of Breitung and Knüppel (2021) is based on the traditional Mincer-Zarnowitz regression (equation (2.9)). The two hypotheses tested are $H_0 : \beta_h^i \leq 0.5$ for *no information* and $H_0 : \beta_h^i \leq 0$ for *constant mean*.

As shown in tables A.1 and A.2, the estimates of $\hat{\beta}_h^i$ are closer to unity for shorter horizons, but decrease for longer horizons. For example, for the USDA projections of corn harvested acres, the estimates of $\hat{\beta}_h^{USDA}$ decrease from 0.98 for the next year projection (horizon $h = 1$) to 0.07 for the ten years ahead projection ($h = 9$), which suggests a reduction in the predictive content of the USDA projections at longer horizons (table A.1). Similarly, for the FAPRI projections of corn harvested acres, the estimates of $\hat{\beta}_h^{FAPRI}$ decrease from 0.996 for horizon $h = 1$ to -0.044 for $h = 9$ (table A.2). The statistical significance of the coefficients tested with a one-tail test show that the projections for corn harvested acres become uninformative after $h = 5$ and then constant mean after $h = 7$.

We further plot the p-values for the *no information* and *constant mean* tests for predictive content against the projection horizon h for the commodities and net cash income components in figures 2.6 and 2.7. The horizontal dashed line stands for significance at a 5% level. These figures mirror and confirm the results in tables A.1 and A.2. In general, the results show that yield is better predicted than harvested acres, which is better predicted than farm price in terms of becoming uninformative and constant mean at longer horizons.

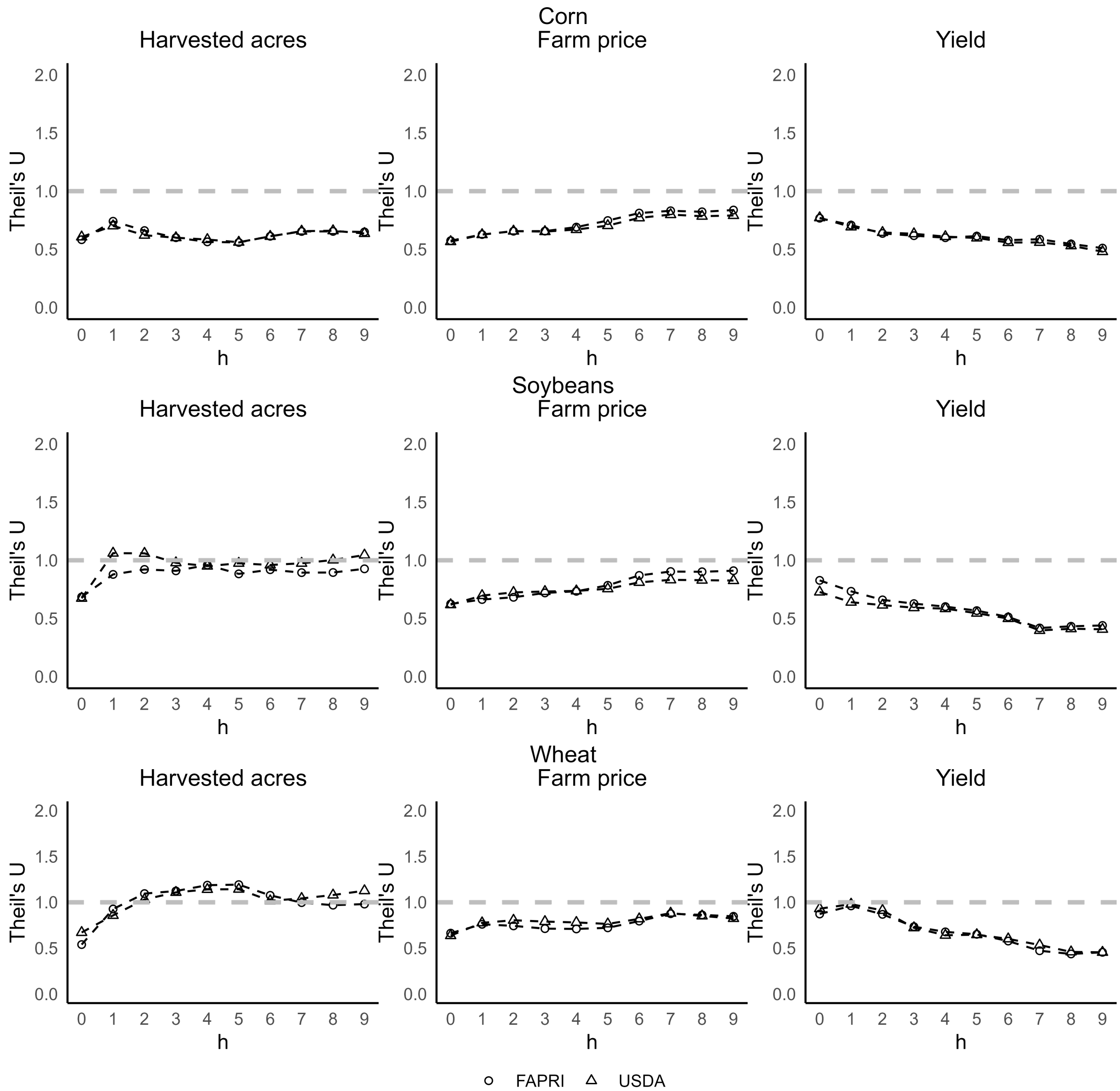


Figure 2.4: Theil's U for USDA and FAPRI baseline projections of corn, soybeans and wheat by projection horizon h, 1997–2020

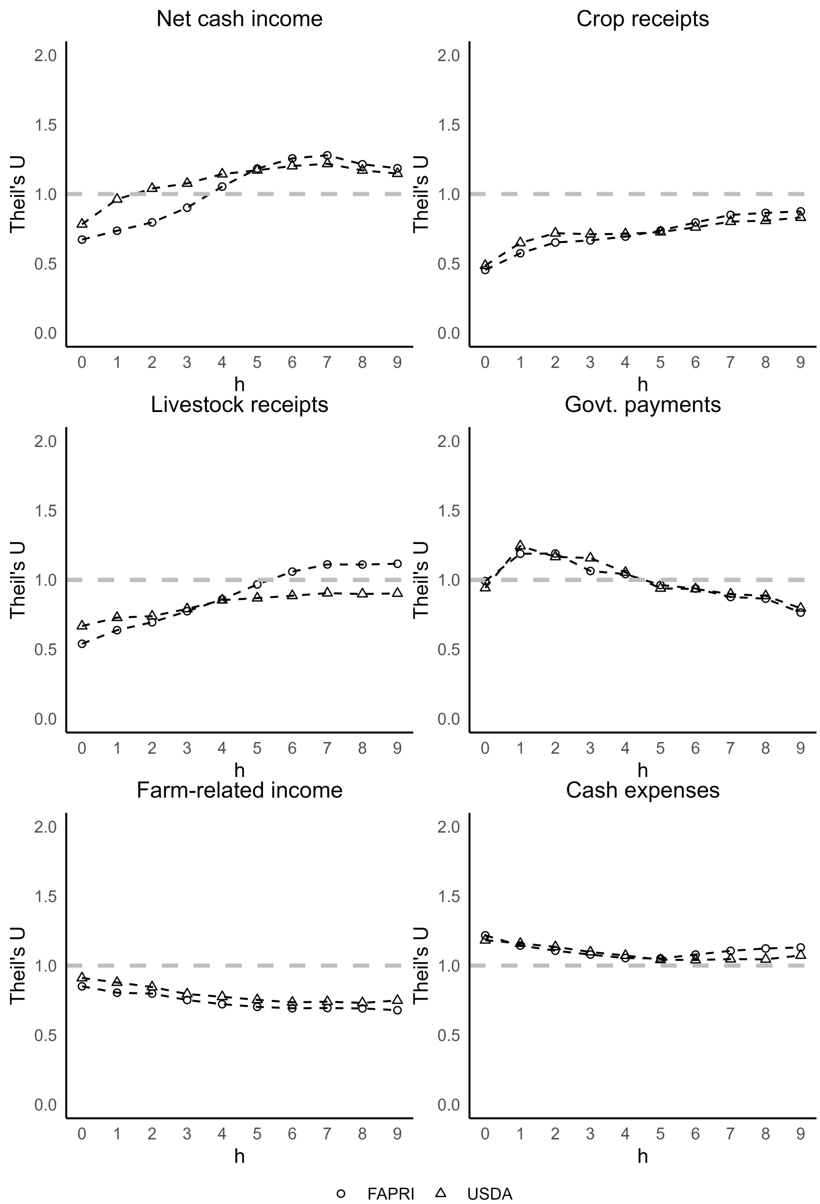


Figure 2.5: Theil's U for USDA and FAPRI baseline projections of net cash income and its components by projection horizon h, 1997–2020

Finally, we calculate the maximum informative projection horizons h^* for both tests at a 5% significance level in table 2.3. The maximum informative projection horizon is calculated as the penultimate horizon, after which the null hypothesis is not rejected for the first time. For example, using the *no information* hypothesis test, $h^* = 5$ for corn harvested acres projections by both USDA and FAPRI as *no information* test is significant at 5% level until $h = 5$. Similarly, using the *constant mean* hypothesis test, $h^* = 7$ for corn harvested acres projections by both USDA and FAPRI as no information test is significant at 5% level until $h = 7$. Because $\hat{\beta}_h^i$ are generally decreasing with the horizon h and the *no information* hypothesis tests whether the coefficient estimate is less than 0.5 versus the *constant mean* hypothesis that tests whether the coefficient estimate is less than 0, the results imply that the *no information* hypothesis is not rejected at shorter horizons than the *constant mean* hypothesis. In other words, projections for the shortest horizons are both informative and do not have constant mean, and for medium horizons, the projections become uninformative. For the longest horizons, the projections are also constant mean.

For most variables, the informative content of the projections starts diminishing after 4-5 years from the current year, using the more conservative *no information* test results. These results vary greatly across variables. Both USDA and FAPRI are able to predict yield per acre for the longest horizons of 9 years ahead, with reduced predictive ability for harvested acres of about 5-7 years ahead and the lowest predictive for farm price of only 2-4 years ahead. These results are not surprising because predicting yield around a long-term trend has proven to be easier than predicting farm prices, which are more volatile. The bottom-line net cash income also remain informative 4-6 years into the future, while some individual components such as crop receipts and cash expenses generally remain informative for shorter horizons of about 2 years. Government payments are notably difficult to predict even in the current year and are not informative after the current year, consistent with previous studies (Bora et al., 2021; Isengildina-Massa et al., 2021). The findings, however, do not suggest that the projections cannot be improved beyond the reported maximum horizon, as our test results are subject to the projection process. Our results only suggest that the projections may stay informative for a longer period using improved models.

There may be several explanations why a variable might not stay informative beyond a few years. It may be that the variable under examination is difficult to predict. For example, it is not surprising that the government payments do not stay informative beyond the current year, as policy decisions are often unpredictable. The opposite is true for crop yield projections, where even a linear trend model may predict future yield with low percent errors. Our findings of a short content horizon would suggest that the projection may be improved by using better projection models, more rigorous review processes, and robust information sets. However, errors in the baseline projections may come from two distinct sources. First, the assumptions about macroeconomic conditions, weather, trade policies while producing the baselines may not be realized in the future. Second, even if correct assumptions were made, the models used in the projections may be inadequate or inaccurate. Our tests of predictive content do not pinpoint whether a short content horizon may result from incorrect assumptions or incorrect models and analysis, and would merely suggest that future revisions of the baselines should try to improve both the assumptions and the modeling process. The same limitation applies to other tests used in this study.

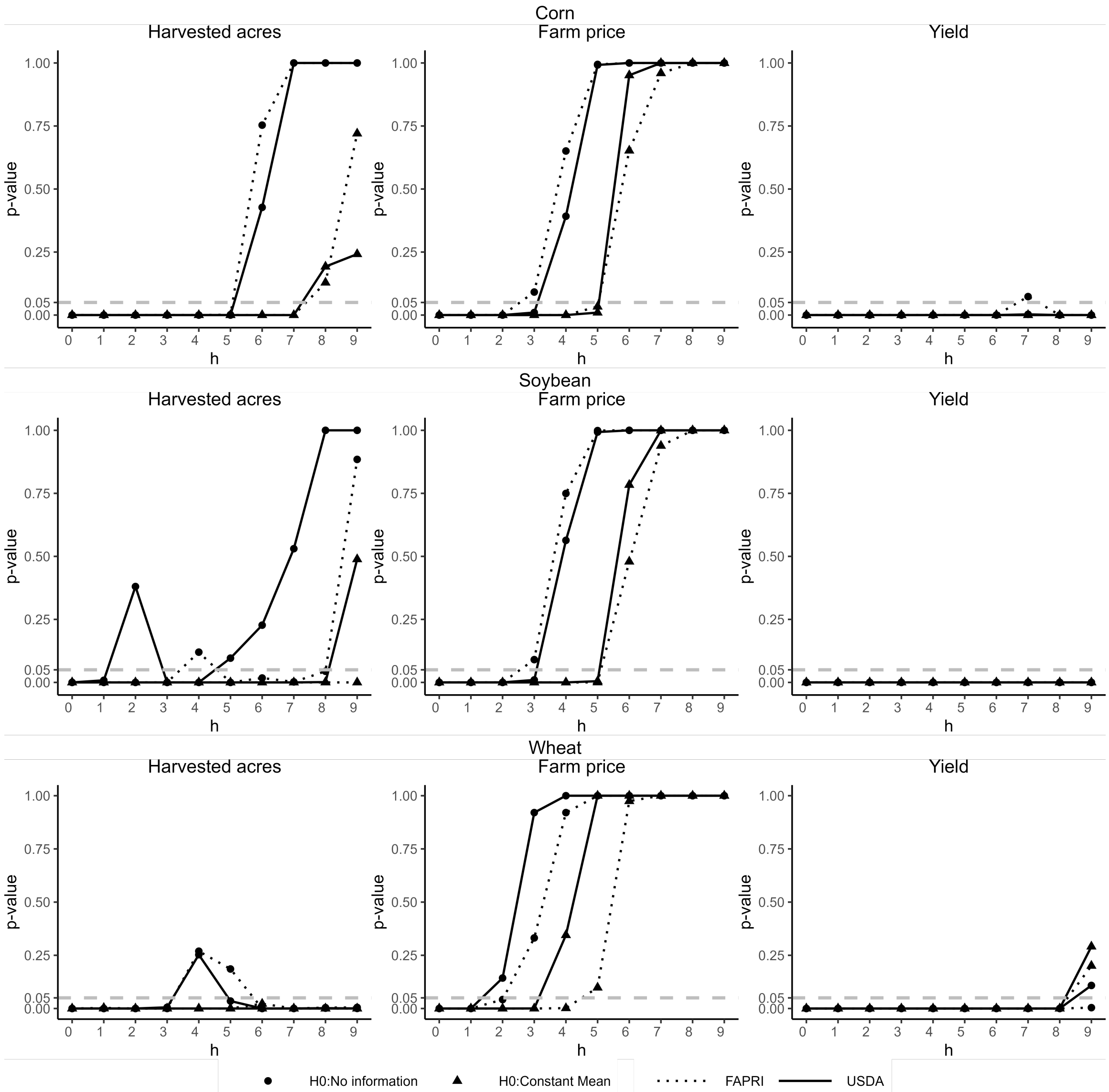


Figure 2.6: P-values for the tests of predictive content of the USDA and FAPRI commodity projections by horizon, 1997–2020

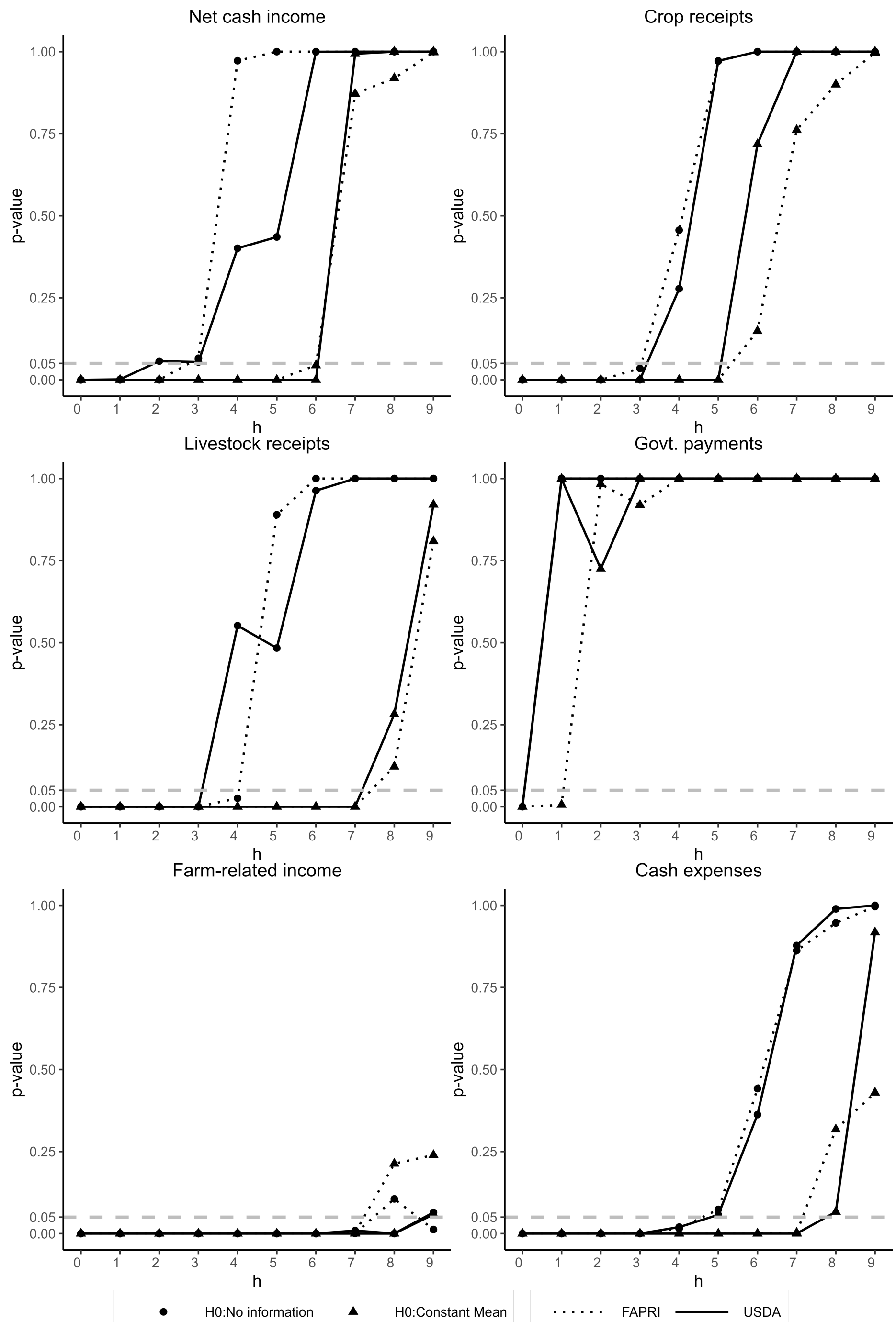


Figure 2.7: P-values for the tests of predictive content of the USDA and FAPRI farm income components projections by horizon, 1997–2020

Table 2.3:
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

2.5.2 Comparing USDA and FAPRI Baseline Projections

We first compare the FAPRI and USDA baselines using the modified Diebold-Mariano (MDM) test of Harvey et al. (1997) using a root mean square error loss function (table A.3). For this MDM test, we compare USDA and FAPRI projections at each horizon separately using the test statistic from equation (2.13). We then perform multi-horizon uniform SPA test using the test statistic from equation (2.17) to test whether the FAPRI projections perform better than the USDA projections or whether the USDA projections perform better than FAPRI (table A.4). Then, we conduct two versions of the average SPA test using the test statistic from the equation (2.17'). The first average SPA test assigns equal weights to each horizon while calculating loss differentials (table A.5). Table A.6 presents the results of the average SPA test using weights based on variances of loss differentials of the horizons. The multi-horizon tests of uniform SPA and average SPA are performed for all horizons up to h. Thus, at the last horizon $h = 9$, we run the full version of the multi-horizon comparison test by including all horizons. Figures 2.8 and 2.9 plot the p-values of the MDM test and the multi-horizon comparison tests.

The p-values of all four multi-horizon comparison tests for the commodities projections in figure 2.8 suggest that the FAPRI projections do not outperform the USDA projections for most variables, as we cannot reject the null hypothesis. Notable exceptions are that the FAPRI projections perform better than USDA for soybean harvested acres and wheat price. The multi-horizon comparison test results shown in figure 2.9 suggest that the FAPRI projections perform better in shorter horizons ($h \leq 4$) for net cash income and crop receipts, while FAPRI consistently predicts better than USDA farm-related income for all horizons $h \leq 9$. One reason the FAPRI projections may perform better at shorter horizons is that they use the most recent forecasts available in November as inputs to their projections, while the USDA uses forecasts available in October. Also, USDA releases their projections a couple of weeks earlier than FAPRI, so FAPRI may contain additional information, especially expert opinions. Since expert opinions mostly influence shorter horizons of the projections, the FAPRI projections are better for some variables. Additionally, the three multi-horizon comparison tests (uSPA, aSPA equal weights, and aSPA variance weights) yield similar results, and the findings are consistent with the results of the single-horizon MDM test.

The results of multi-horizon comparison tests in tables A.4, A.5, and A.6 provide additional insights to the single-horizon MDM tests presented in table A.3. The MDM test results show that the FAPRI projections perform better in shorter horizons for net cash income, crop receipts, and wheat price. For farm-related income, the FAPRI projection performs better than USDA across the entire projection horizon. The USDA projection performs better at longer horizons for corn price and yield, soybean price, crop receipts, livestock receipts, and cash expenses. The multi-horizon tests, on the other hand, aggregate the loss differential across multiple horizons. We start our multi-horizon tests with the projection for the current year ($h = 0$) and progressively include additional horizons until we cover the entire projection horizon ($h \leq 9$). This allows us to observe how the addition of more horizons affects the results. For shorter horizons, the multi-horizon tests yield similar results to the MDM test. However, as we keep adding horizons, in a multi-horizon framework, the results differ from the single-horizon tests. For example, the tests of uSPA in table A.4 show that, over the projection path ($h \leq 9$), the FAPRI projection performs better for farm-related

Table 2.4: Encompassing Tests

	Projection Horizon									
item	h=0	h=1	h=2	h=3	h=4	h=5	h=6	h=7	h=8	h=9
Corn										
Harvested acres	0.599	0^{++}	0.061^{+++}	0.319	0.631	-0.013	-0.149	0.271	0.188	-1.091
	(0.393)	(0.365)	(0.280)	(0.405)	(0.538)	(0.840)	(0.723)	(1.267)	(0.750)	(1.333)
Farm Price	0.132^{++}	0.009^{+}	0.353	0.189	-0.247	-0.985	-1.18	-0.729	-1.043	-0.939^{+}
	(0.372)	(0.550)	(0.838)	(1.000)	(1.165)	(1.600)	(1.826)	(1.565)	(1.194)	(1.022)
Yield	-0.185	-2.067	1.201	2.058^{*}	1.244	-1.546	-1.996	-2.831	-0.53	-1.628^{++}
	(1.462)	(2.932)	(1.839)	(1.113)	(1.462)	(1.978)	(1.803)	(2.391)	(1.565)	(0.934)
Soybean										
Harvested acres	0.298	1.761^{***}	1.539^{***}	0.369^{++}	-0.597^{++}	1.018	0.792	1.574^{***}	2.396^{+++***}	2.204^{+***}
	(0.650)	(0.618)	(0.496)	(0.293)	(0.568)	(0.715)	(0.701)	(0.394)	(0.371)	(0.678)
Farm Price	-0.888^{+++}	-0.884^{++}	-0.002	-0.287	-0.479	-1.094^{+}	-1.578^{+}	-1.529^{+++**}	-1.615^{+++***}	-2^{+++***}
	(0.519)	(0.717)	(0.840)	(1.415)	(1.150)	(1.029)	(1.220)	(0.681)	(0.418)	(0.458)
Yield	-2.578^{+++**}	-2.215^{+++*}	-0.768^{++}	-0.286	0.417	-0.041	0.026	0.073	-0.323	-1.199^{++}
	(1.021)	(1.140)	(0.714)	(1.337)	(0.855)	(1.387)	(0.933)	(1.575)	(1.356)	(0.927)
Wheat										
Harvested acres	0.745^{***}	-0.081^{+}	0.489	0.747	0.484	0.071^{++}	-0.326^{+++}	0.216^{+}	0.371	0.085^{+++}
	(0.222)	(0.612)	(0.549)	(0.744)	(0.533)	(0.359)	(0.435)	(0.443)	(0.407)	(0.231)
Farm Price	0.301	0.896	1.509	1.748^{**}	1.847^{**}	1.486^{**}	1.444	1.006	0.657	0.47
	(0.486)	(0.798)	(0.883)	(0.695)	(0.787)	(0.693)	(0.858)	(1.038)	(1.202)	(0.776)
Yield	1.93^{**}	1.493^{*}	2.089^{*}	0.733	-0.124^{+}	0.48	1.463^{*}	2.012^{***}	1.207^{**}	0.907
	(0.725)	(0.780)	(1.086)	(0.691)	(0.562)	(0.681)	(0.805)	(0.623)	(0.409)	(0.618)
Farm Income										
Expenses	0.374	0.691	1.252	1.261	1.141	0.733	0.482	-0.139^{++}	-0.41^{+++**}	-0.956^{+++*}
	(0.597)	(0.629)	(0.881)	(0.791)	(0.777)	(0.698)	(0.690)	(0.400)	(0.160)	(0.526)
Crop Receipts	1.134	1.421	0.88	0.56	0.186	-0.296	-0.631	-0.761	-0.87	-1.491^{+++**}
	(0.700)	(1.099)	(1.233)	(1.250)	(1.140)	(1.076)	(1.151)	(1.176)	(1.132)	(0.673)
Farm-related Income	1.511	1.813	0.781	0.486	1.125	0.619	0.392	0.944	0.678	1.673^{***}
	(0.967)	(1.301)	(1.218)	(0.914)	(1.057)	(0.917)	(0.884)	(0.639)	(0.737)	(0.449)
Government Payments	0.156^{+++}	0.697^{**}	0.153^{++}	0.363	-0.283^{++}	-0.739^{++}	-0.626^{++}	-0.392^{++}	-0.445^{++}	-0.238^{+++}
	(0.198)	(0.309)	(0.397)	(0.487)	(0.603)	(0.610)	(0.633)	(0.590)	(0.553)	(0.380)
Livestock Receipts	1.469^{***}	1.34^{**}	0.671	0.672	0.942	0.197	-0.152^{++}	-0.531^{+++}	-0.571^{+++}	-0.958^{+++***}
	(0.437)	(0.595)	(0.657)	(0.667)	(0.674)	(0.636)	(0.487)	(0.431)	(0.391)	(0.284)
Net Cash Income	0.858^{***}	1.196^{***}	1.061^{***}	0.484^{+++**}	-0.026	-0.747^{+++*}	-0.67^{+++*}	-0.503^{+++}	-0.161^{++}	-0.351^{++}
	(0.286)	(0.275)	(0.268)	(0.202)	(0.595)	(0.406)	(0.382)	(0.370)	(0.485)	(0.475)

Notes: *, **, and *** denote statistical significance at 10%, 5%, and 1% respectively for testing the null hypothesis $H_0 : \lambda = 0$. Likewise, +, ++, and +++ denote statistical significance at 10%, 5%, and 1% respectively for testing the null hypothesis $H_0 : \lambda = 1$.

income, whereas the USDA projection performs better for corn price and soybean yield at 5% significance level. The tests of aSPA in table A.5 and A.6 yield similar conclusions. Interestingly, the full-horizon ($h \leq 9$) multi-horizon comparison tests do not suggest that either projection performs better than the other for net cash income, crop receipts, and livestock receipts. The single-horizon tests in table A.3 show that the FAPRI projections perform better in shorter horizons and the USDA projections perform better in longer horizons. As the multi-horizon tests consider performance over the entire projection horizon, they conclude that neither the USDA nor the FAPRI projection is superior to the other projection.

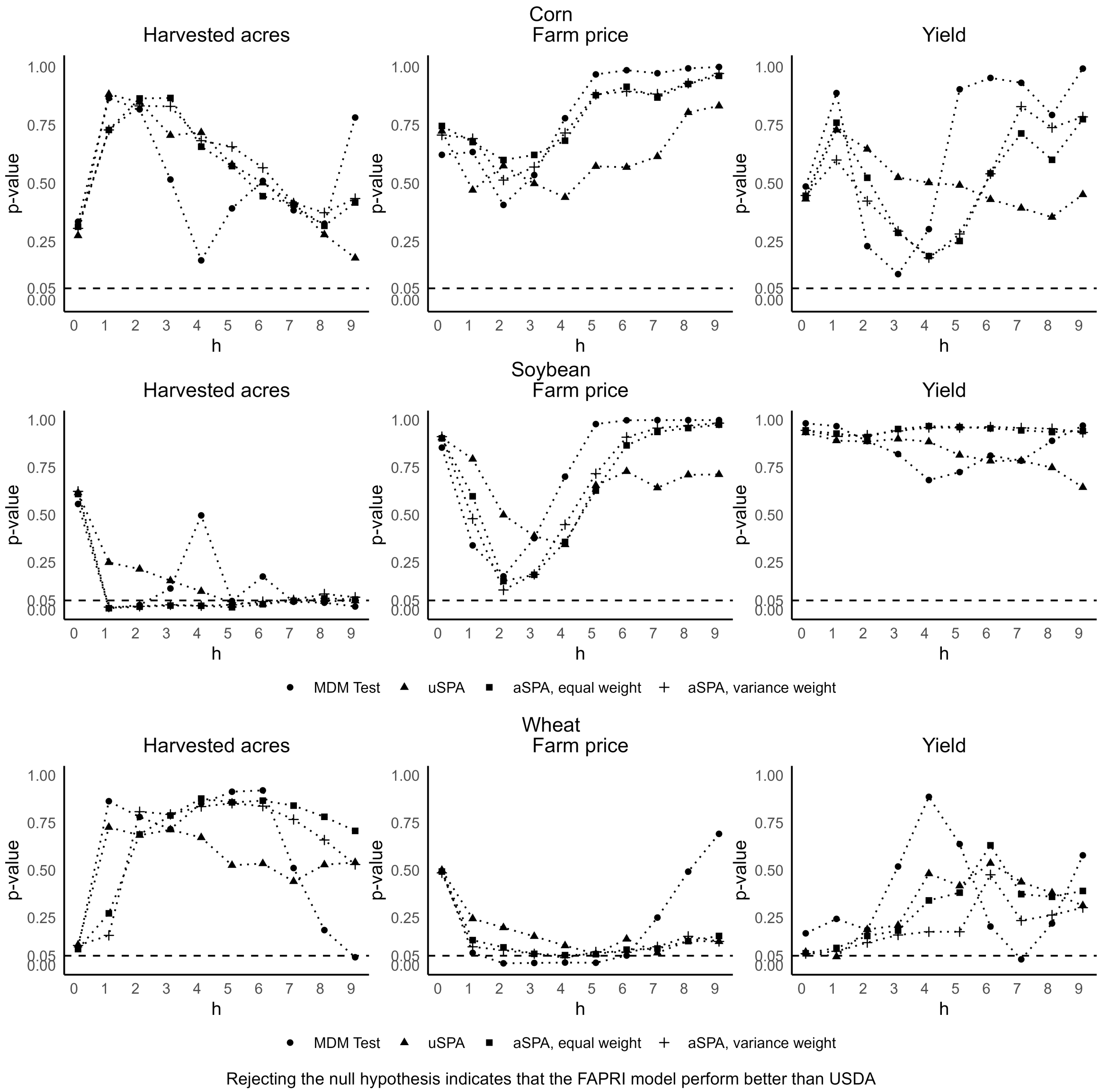


Figure 2.8: Multi-horizon comparison tests of USDA and FAPRI commodity projections by horizon, 1997–2020

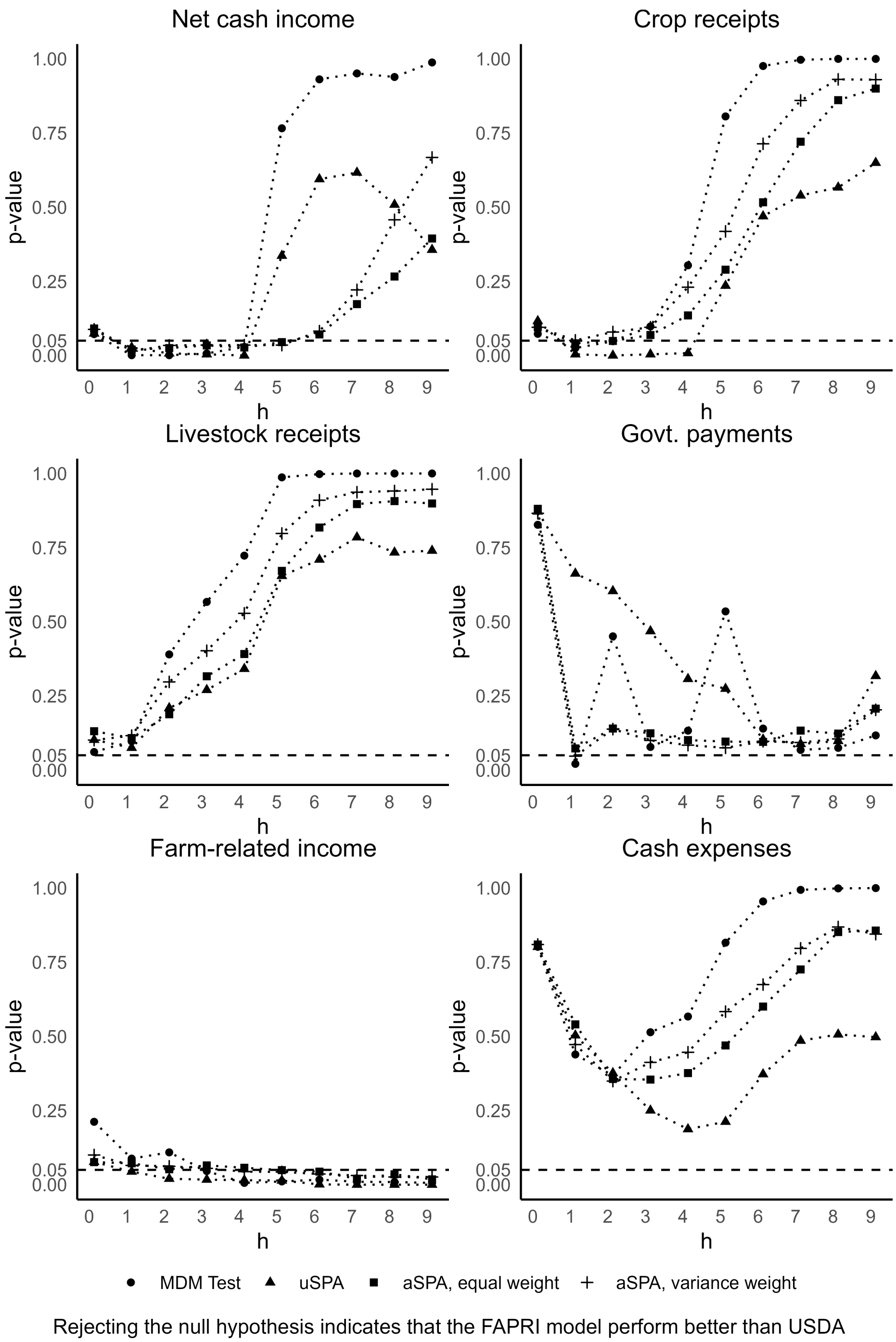


Figure 2.9: Multi-horizon comparison tests of USDA and FAPRI net cash income projections by horizon, 1997–2020

The estimates of optimal weight $\hat{\lambda}$ of our encompassing tests in equation (2.18) is presented in table 2.4. For corn, either the USDA and FAPRI baseline projections can generally be substituted for one another. For soybean prices, USDA projections are preferred in the short term, while a composite projection can be created by taking the weighted average of both projections at larger horizons ($h = 7$ to 9). The net cash income projections of the FAPRI baseline are preferred in the shorter horizons ($h =$ 1 to 3), while USDA net cash income projections are preferred in the larger horizons ($h = 7$ to 9). This finding is consistent with our multi-horizon comparison tests. The government payments of the USDA baseline encompass the FAPRI projections over the length of the horizons. The composite projections created using the encompassing weights are more accurate than either of the two projections (Kuethe et al., 2022).

2.6 Conclusion

Both USDA and FAPRI baseline projections play an important role in shaping agricultural policy in the U.S. The baseline projections provide a conditional scenario against which alternative policies can be evaluated. In recent years, policymakers, agricultural businesses, and program administrators have used these projections extensively in their policy and investment decisions. Given the importance of the baseline projections in determining the long-term outlook of the farm economy, this study examines the accuracy and informativeness of both sets of baseline projections using a number of forecast evaluation techniques.

Our measures of prediction error show that the projections become less accurate as the projection horizon increases, with crop yields being a notable exception. Our tests of bias suggest that the baselines show similar bias as USDA's short-term forecasts documented in the existing literature (Bora et al., 2021; Isengildina-Massa et al., 2021), and the magnitude of the bias increases as the projection horizon increases. This finding is not surprising given the fact that inputs for many baseline models come from USDA forecasts, such as WASDE and farm income forecasts. Our tests of predictive content show that the information content of most of the projected variables starts to diminish after 4-5 years from the current year, with farm price projections becoming uninformative only after 2-3 years and yield remaining informative for the entire projection horizon. The findings suggest that the projections may be improved using better models and processes. The single-horizon tests comparing the two projections suggest that the FAPRI projections perform better at shorter horizons for net cash income and crop receipts, potentially due to the updated information available to the FAPRI projection process, which follows the USDA report by a few weeks. On the other hand, the USDA projection performs better at longer horizons for corn price and yield, soybean price, crop receipts, livestock receipts, and cash expenses. However, our multi-horizon comparison tests suggest that neither USDA nor FAPRI baselines outperform one another for most projected variables if we consider the full projection path. A notable exception is the FAPRI projection for farm-related income, which has uniform superior predictive ability over the USDA projection. Similarly, the USDA projection for corn price and soybean yield has uniform superior predictive ability over the FAPRI projection. For the rest of the variables, neither projection performs better than the other.

The findings of this study also underline the importance of stochastic analysis while producing the baselines. While the figures published in the baseline reports are point estimates, both the USDA and FAPRI perform additional stochastic analysis to project distributions for different future scenarios. One can expect the point estimates in the baseline reports to differ from actual values, as many of the analysts' assumptions may not realize. However, the stochastic analysis should account for such changed scenarios, and actual values should ideally lie within the projected distribution. The agencies have not always published the stochastic projections, or the stochastic projections have not received the same attention from users. The agencies may consider releasing stochastic projections in addition to their point projections to allow users to adapt the projections to different scenarios.

One limitation of our study is that some of our findings may be influenced by the projections made in the previous decade(s) as opposed to more recent projections. The baseline models and processes for both agencies have evolved and, hopefully, improved over time. The baseline projection process at both agencies has also been subject to changes in personnel and information technology infrastructure. The newer reports may have already addressed some issues related to bias or informativeness found in this study. Given our small sample size, we cannot undertake sub-sample analysis to see if our estimates of bias and informativeness remain steady over time.

Our findings provide valuable insights which may help improve the models and processes used to produce the projections by each organization. Our tests of informativeness might be especially useful for the desire to provide agricultural sector projections at longer horizons to examine issues related to technology adoption or climate change. The balance between empirical models and the judgment of a panel of experts employed by the baseline may also prove beneficial to other short-term USDA forecasts, including those of commodity production and trade. Furthermore, our findings provide important information to various market participants who use these projections.

To our knowledge, this is the first study to look into the accuracy and usefulness of agricultural baselines. There are various directions in which agricultural baselines research could go in the future. Using a more comprehensive information set is one way to enhance the projections. For example, distant futures contract prices may be useful in projecting commodity prices, as suggested by Irwin and Good (2015) and extending the approach of Hoffman, Etienne, Irwin, Colino, and Toasa (2015b) beyond one year. Future revisions of the baseline projections may also benefit from examining the factors that may have contributed to systematic deviations from observed values in the past, such as failures to anticipate the ethanol boom, the growth in Chinese soybean demand, and Russia's emergence as a major wheat exporter. Another option is to improve the methodology, potentially using recent advances in machine learning.

Chapter 3: Multi-step Commodity Forecasts using Deep Learning

3.1 Introduction

The availability of long-term information about commodity markets plays a vital role in policy and investment decisions by market participants. The forecasts of season-average farm prices of major field crops such as corn, soybeans, and wheat are widely used to inform decisions by farmers, agricultural businesses, and the government. Similarly, the forecasts of harvested area and yield provide information about the production of the commodities for the marketing year and help anticipate the ending stocks. The USDA's World Agricultural Supply and Demand Estimates (WASDE) provide forecasts about commodities for the current marketing year. However, market participants may require information about market trends beyond the current marketing year to inform their decisions. For example, forecasts for the next few years can facilitate comparisons of policy alternatives by government agencies. Similarly, long-term forecasts can help estimate the outlays of various farm program costs under the federal budget. The Farm Bill programs are typically implemented in five-year cycles, and having information for the next five years will help immensely in planning the budget. Similarly, long-term prices and crop yield forecasts may help farmers inform their long-term decisions about planting, crop choice, and land use. For example, the decision to enroll farmland in federal programs like conservation reserve programs (CRP) may be informed by crop prices and yield forecasts for multiple years into the future. The importance of reliable long-term forecasts became evident when the pandemic hit the economy, and policymakers required information deep into the future to plan the recovery process.

The USDA's baseline projections, published every year in February, are one of the principal sources of long-term information about the US farm sector. The baselines are produced by a team from ten USDA agencies, including the Economic Research Service (ERS), and contain annual projections of key measures of agricultural market conditions for the next decade. These projections facilitate comparisons of policy alternatives by providing a conditional "baseline" scenario based on specific macroeconomic, weather, policy, and trade assumptions. Over the years, the baseline projections have been used for a variety of purposes, including estimating farm program costs and preparing the President's budget. In addition to USDA, the Food and Agricultural Policy Research Institute (FAPRI), University of Missouri, produces similar ten-year projections of key agricultural variables. The baseline projections are produced through a mixture of the output of quantitative models and expert opinions. Previous studies show that many variables in the USDA baseline projections are biased and that the predictive content of the baselines diminishes after a few years (Bora, Katchova, & Kuethe, 2022). As the evaluation of the baselines has shown its limited predictive content, an investigation of alternative methods to improve the long-term projections becomes essential.

This study aims to forecast the harvested area, yield, and farm price of three major field crops in the US for the next five years using deep learning models. Our investigation is performed in three steps. First, we formulate a supervised learning problem for the forecasting process and develop a test harness to compare the performance of various methods based on a train-test split of the sample. The last ten years were used as a test sample using a walk-forward validation approach. Second, we benchmark the performance of traditional methods such as a naïve no-change forecast, exponential smoothing, and USDA baseline reports. Finally, we implement a suite of deep learning models to predict the commodity market indicators, with particular emphasis on long short-term memory (LSTM) recurrent neural networks (RNN), convolutional neural networks (CNN), and their hybrids. We train the deep learning models using a large number of input features reflecting macroeconomic indicators, demographic trends, weather variability, global trade, demand, and supply of key commodities.

Our study contributes to the literature in several ways. We use state-of-the-art deep learning methods to improve the long-term forecasts of commodity market indicators. While deep learning methods have shown great promise in forecasting in other fields (Borovykh, Bohte, & Oosterlee, 2018; Huang et al., 2020; Kim & Won, 2018; Lara-Benítez, Carranza-García, Luna-Romera, & Riquelme, 2020; Wan, Mei, Wang, Liu, & Yang, 2019; Wang, Shen, Mao, Chen, & Zou, 2019), their use in predicting long-term agricultural statistics such as the USDA baselines has been limited. This study aims to bridge this gap. Our results suggest that deep learning networks may perform better than the official USDA baselines at longer forecast horizons. In particular, when the USDA baselines perform well, deep learning models match the accuracy, but if the USDA baselines do not perform well, deep learning models perform better. These findings may have important implications for future revisions of the USDA baseline models and processes. Deep learning models with improved accuracy may complement the existing models for the baselines. The current baseline process is time-consuming, and the adoption of deep learning techniques as an alternative method may help in reducing this timeline. The existing process of producing the baseline reports involves many agencies, which work on specific components of the report and create inputs for the composite model. Deep learning methods have the potential to make the baseline projection process more straightforward, faster, and more accurate.

The remainder of this article is organized as follows. The next section describes the various datasets used in this study. The third section describes the methodology, followed by results and discussion. The final section contains concluding remarks.

3.2 Data

Our dataset of the target variables consists of historical values of harvested area, yield, and farm price of corn, soybeans, and wheat in the US since 1961. Together, these three field crops constitute a significant share of the area under cultivation in the US. The values are averages for the marketing years, which differ by crop. The marketing year for corn and soybeans begins on September 1 and comprises four quarters. For example, the marketing year 2021/22 for corn and soybeans starts on September 1, 2021, and ends on August 31, 2022. The 2021/22 marketing year for wheat begins on June 1, 2021, and ends on May 31, 2022. All this information was obtained using the NASS Quickstats API (USDA National Agricultural Statistics Service, 2021).

Figure 3.1 shows the plots of harvested area, yield, and farm price of the three crops for the period 1961–2021. The figures suggest that many of these indicators are highly correlated, and they may be related to each other or other macroeconomic, weather, or trade indicators. For example, the loss of wheat harvested area over the years is accompanied by a contemporaneous increase in soybeans harvested area.

[**FIGURE 3.1 ABOUT HERE**]

An archive of the USDA agricultural baseline projections since 1997 is available at The Albert R. Mann Library at Cornell University (USDA ERS, 2021b). The baseline reports typically include estimates of the previous year(s) and projections for the next ten years. For example, the February 2022 USDA report contains realized estimates for 2020, provisional estimates for 2021, and projections for 2022–2031 (USDA Office of Chief Economist, 2022). The exact information set available to the committee producing the projections in the early years is difficult to retrieve due to lack of information on the variables that were in the information set of the committee, and the revisions made to the realized values over time. As the organizations involved with the projections process go through personnel and information technology infrastructure changes over the years, the exact information used to produce the baselines is challenging to ascertain. The projections and estimates are often revised long after they are first published. For example, there is no way of accessing the exact data used as the information set of the committee when the baseline projections were prepared for 1997. We can assume that the committee made the best use of the information they had. To mimic the forecasting process of the committee, we try to provide many macroeconomic, population, trade, and weather information as input features to train our deep learning models. The committee may have had a different

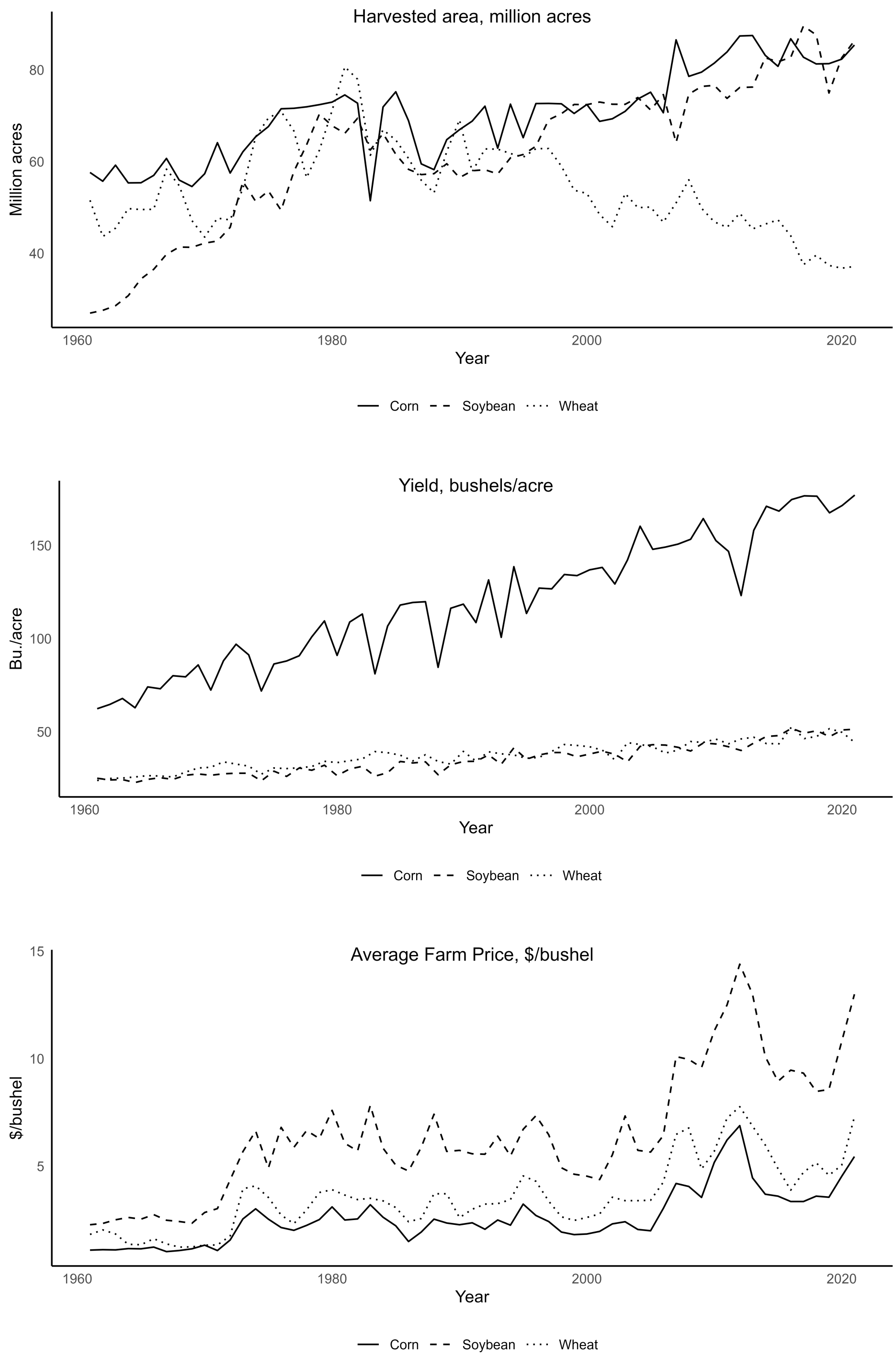


Figure 3.1: Historical harvested area, yield, and farm price of corn, soybeans, and wheat, 1961-2021

set of variables and/or different values for these variables that were later revised to what is available today. Our goal is to use deep learning methods to produce the forecasts using a similar information set, and examine whether these forecasts have a superior performance over the USDA baselines.

We use data from several sources as input features to train the deep learning models. First, we use lagged values of commodity indicators to forecast their future values. Apart from the lagged commodity market indicators, we include several macroeconomic, population, trade, and weather variables for the World and the US as input features to our models. These include growth rates for gross domestic product (GDP) and population. For the US economy, we also include inflation, unemployment, labor market participation, and interest rates. We also include features that represent changes in weather in the World and the US over time. To account for temperature changes all over the World, we include global annual average temperature anomalies, measured as deviation from 20^{th} century average. The macroeconomic data are taken from the World Bank Open Data Catalog. Similarly, we include US annual average temperature, maximum temperature, minimum temperature, precipitation, and heating and cooling degree days. All weather information was obtained from National Oceanic and Atmospheric Administration (NOAA) (NOAA National Centers for Environmental information, 2022). Finally, we add commodity balance sheet variables representing domestic use, imports, exports, and ending stocks of corn, soybeans, and wheat as input features. The commodity balance sheet information is extracted from the Production, Supply, and Distribution (PSD) Database published by USDA Foreign Agricultural Service (USDA Foreign Agricultural Service, 2021).

3.3 Methodology

In this section, we define our prediction problem and proceed to develop a test harness for comparing the performance of the methods used in this study. We then describe different traditional and deep learning methods used in this study.

3.3.1 The Prediction Problem

We denote the realized or actual values of commodity indicators of harvested acres, yield, and farm price for corn, soybeans, and wheat in year t by $\boldsymbol{y}_t$. At year t, the forecaster makes a forecast $\hat{\boldsymbol{y}}_{t+h|t}$ for horizon $h \in \{0, 1, \ldots, H-1\}$ for H future years, including year t using lagged values of the commodity indicators and a set of other covariates such as macroeconomic, population, and weather variables. Although the baselines are for $H = 10$ years, we limit our attention to forecasts of up to five years due to the small length of the time period, i.e. $h = \{0, \ldots, 4\}$. Similarly, we assume that up to five years of lagged values of input features are used to produce the forecasts.

We first transform the prediction problem into a supervised learning problem where a set of input features $\boldsymbol{X}$ are mapped to an output variable $\boldsymbol{y}$. For year t, our input $\boldsymbol{X}_t$ consists of vectors of all input features up to lag five, and $\boldsymbol{y}_t$ consists of vectors of the next five years of values of the target variables (harvested acres, yield, and farm price of corn, soybeans, and wheat). From our dataset for the time period 1961-2021, we construct $\{\boldsymbol{X_t}, \boldsymbol{y_t}\}$ pairs for 52 years between 1966-2017. This yields a three-dimensional array of input features $\boldsymbol{X}$ with dimensions $(52, 5, n\_features)$, where *n_features* is the total number of input features. This is important since the deep learning models used in this study accept three-dimensional input. We use a total of 44 input features in this study, however, this number can be augmented by including additional features.

3.3.2 Developing a Test Harness

A test harness ensures that all deep learning methods used in this study are evaluated using a consistent approach for comparability. The important components of our test harness are the train-test split validation strategy and the evaluation criteria.

Train-test split

Our dataset contains commodity market variables of harvested area, yield, and farm price for corn, soybeans, and wheat between 1961 and 2021. Since we use up to five-year lagged features in our deep learning algorithms to produce five-year ahead forecasts, this results in a complete dataset of features (X) and output (y) between 1966 and 2017, a total of 52 years. We use the last ten years of the data as our test sample between 2008–2017, representing close to 20% of the entire sample. As preferred in time-series applications, we use a walk-forward validation strategy, allowing updated information to train the model as we progress through the years in the test sample. We use an expanding training window approach, which means the training sample increases as we walk through the test sample. For example, we train a model using 42 samples between 1966-2007 to produce forecasts for 2008. We then add the sample for 2008 back to the training sample to produce forecasts for 2009 and so on. This validation strategy closely follows how the USDA produces the baseline reports as forecasters make use of new information as it becomes available. Another choice is to use a sliding window, where the oldest training sample is dropped as we add a new sample, keeping the length of the training sample constant. However, we prefer an expanding window as we would like to make use of all information available, and our sample size is small.

Evaluation Criteria

We will use two widely adopted error metrics to measure the performance of the proposed methods: root mean squared error (RMSE) and mean absolute percent error (MAPE). The RMSE is calculated at the level of the variables, while the MAPE is calculated relative to the actual level of the variables according to the following formulas:

$$RMSE_h = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(y_{t+h} - \hat{y}_{t+h|t})^2} \tag{3.1}$$

$$MAPE_h = \frac{100}{T}\sum_{t=1}^{T}\left|\frac{y_{t+h} - \hat{y}_{t+h|t}}{y_{t+h}}\right| \tag{3.2}$$

where y_{t+h} are the realized values, $\hat{y}_{t+h|t}$ are forecasts of the target variable at horizon h, and T is the sample size of the test or the training sample. For calculating in-sample forecast errors, we use the sample size $T = 42$ for the training sample, while for out-of-sample errors, we use the test sample $T = 10$.

3.3.3 Benchmarking with Traditional Methods

Naïve Benchmark

We first develop a benchmark model to improve upon using deep learning methods. A natural choice is to use a naïve no-change forecast, where we consider the most recent year's value as the forecast for the next five years. This is a fairly naïve benchmark that would result in high forecast errors. Any econometric or deep learning method is expected to perform better than this naïve benchmark, as the methods are supposed to add some skill to forecasting.

Simple Exponential Smoothing (ETS)

We also use the simple exponential smoothing (ETS) method, which is useful for forecasting when the time series have no clear trend or seasonal pattern. The ETS forecast is a weighted average of past observations, where the weights decay exponentially for older observations. The ETS method can be expressed in terms of the following equations (Hyndman & Athanasopoulos, 2021),

$$\hat{y}_{t+h|t} = \ell_t \tag{3.3}$$

$$\ell_t = \alpha y_t + (1 - \alpha)\ell_{t-1} \tag{3.4}$$

where ℓ_t is the level of the variable at time t. The smoothing parameter α represents the rate at which the weight placed on past observations decreases.

Exponential Smoothing (ETS) with Trend

We then use an extension of the simple exponential smoothing method, which allows a trend (Holt, 2004). Some of our data series, such as crop yield, shows a clear time trend, and farm price may also be trending upward over the years. The ETS method with a trend can be expressed as (Hyndman & Athanasopoulos, 2021),

$$\hat{y}_{t+h|t} = \ell_t + hb_t \tag{3.5}$$

$$\ell_t = \alpha y_t + (1 - \alpha)(\ell_{t-1} + b_{t-1}) \tag{3.6}$$

$$b_t = \beta(\ell_t - \ell_{t-1}) + (1 - \beta)b_{t-1} \tag{3.7}$$

where β is an additional smoothing parameter for the trend. We use the implementations of ETS and ETS with trend methods in Python *statsmodels* library to produce the forecasts (Seabold & Perktold, 2010).

USDA Baseline Report

Our final choice for comparison is the projections produced by the USDA in their baseline report. These projections are produced using a mixture of economic/econometric models, survey information, and expert opinions. We calculate the error metrics for baseline projections up to five years for the test period 2008-2017 for comparison with the other methods used in our study. As mentioned earlier, the exact information set used to produce these projections is challenging to ascertain. Therefore, the comparison with deep learning methods using the current training set may not be entirely justifiable.

3.3.4 Deep Learning Methods

The methods discussed in the previous section are traditional time-series forecast models. However, in recent years, deep neural networks have become popular in forecasting time series (Schmidhuber, 2015). Neural networks are a collection of algorithms used in pattern recognition. Deep learning refers to a subset of neural networks which consists of more than three layers.

The most basic deep learning networks are feed-forward neural networks (FNN) that do not allow recursive feedback, such as the Multi-layer Perceptron (MLP). The computational architecture of FNNs consists of three layers: an input layer, the hidden layer(s), and an output layer. Since two consecutive layers have only direct forward connections, FFNs ignore the temporal nature of the data and treat each input independently. Therefore, they are of limited use in dealing with our data which are inherently temporal, sequential data. We consider two main families of deep learning methods that account for temporal dependence in sequences, namely recurrent neural networks (RNN) and convolutional neural networks (CNN). We also explore hybrid deep learning models, which have seen increased popularity in recent years.

Recurrent Neural Networks

Recurrent neural networks (RNN) are popular in time series prediction applications. An RNN allows recursive feedback, and each RNN unit can take the current and previous input simultaneously. They are widely used for prediction in different fields, including stock price forecasting (Kim & Won, 2018), wind speed forecasting (Huang et al., 2020), and solar radiation forecasting (Wang et al., 2019). Moreover, RNNs have done remarkably well at forecasting competitions, such as the recent M4 forecasting competition (Makridakis, Spiliotis, & Assimakopoulos, 2018). In a recent study, (Medvedev & Wang, 2022) used RNNs to predict the volatility of the S&P 500 Index (SPX) for pricing options, with good success. However, we are not aware of any studies applying RNNs to forecast long-term information about agricultural markets.

Elman (1990) proposed an early RNN which generalizes feedforward neural networks by using recurrent links in order to provide networks with dynamic memory. This type of network is more suitable for handling ordered data sequences like financial time series. While the Elman's RNN model is simple, training these models is difficult due to inefficient gradient propagation. In particular, the problem of vanishing and exploding gradients makes it challenging to learn long-term dependencies. Due to vanishing gradients, it may take a long time to train the model, while the exploding gradients may cause the model's weights to oscillate (Lara-Benítez, Carranza-García, & Riquelme, 2021).

Long Short-Term Memory (LSTM) networks were proposed to address the vanishing and exploding gradients problems faced by standard RNNs (Hochreiter & Schmidhuber, 1997). LSTMs can model long-term temporal dependencies without compromising short-term patterns. LSTM networks have a similar structure to the Elman's RNN, but differ in the composition of the hidden layer, known as the LSTM memory cell. Each LSTM cell has three gates: a multiplicative input that controls memory units, a multiplicative output that protects other cells from noise, and a forget gate. Gated Recurrence Units (GRU) are simplified versions of LSTMs that replace the forget and input gates with a single update gate to reduce trainable parameters. An RNN can also have stacked recurrent layers to form a deep RNN.

Convolutional Neural Networks

Convolutional neural networks (CNN) are mainly used in classification applications such as speech recognition, object recognition, and natural language processing (NLP). However, with some adjustments, they can be used for time-series predictions as well. A CNN uses the convolutional operation to extract meaningful features from raw data and create feature maps (Lara-Benítez et al., 2021). A CNN consists of convolution layers, pooling layers, and fully connected layers. The pooling layers lower the spatial dimension of the feature maps, while the fully connected layers combine the local features to form global features. As CNNs have a smaller number of trainable parameters, the learning process is more time-efficient than RNNs (Borovykh et al., 2018). In addition, different convolutional layers can be stacked together to allow the transformation of raw data (Chen, Kang, Chen, & Wang, 2020).

Hybrid models are a recent trend in time series forecasting using deep learning. For example, depending upon the application, LSTMs can be used with RNNs or CNNs. Also, deep learning models can be used with traditional econometric methods to achieve superior results. The winning entry of the M4 forecasting competition in 2018 used a hybrid ETS-LSTM model (Smyl, 2020). While the exponential smoothing component captured seasonality, the LSTM focused on non-linear trends and cross-learning from related series.

In this study, we use three deep learning architectures to forecast the commodity market indicators.

(a) **Vanilla LSTM**: The first architecture that we use is a simple LSTM model with one LSTM layer.

(b) **Encoder-decoder LSTM (ED-LSTM)**: The second architecture that we use is an encoder-decoder LSTM with two layers. The first layer reads the input sequence and encodes it into a fixed-length vector, and the second layer decodes the fixed-length vector and outputs the predicted sequence.

(c) **CNN-LSTM**: The last architecture that we use consists of an LSTM preceded by a convolution layer at the input.

We train the deep learning networks using *keras* (Chollet et al., 2015) and *TensorFlow* (Abadi et al., 2015) libraries in Python. We chose the hyperparameters of the models using trial and error. We train each model for 1000 training epochs with a batch size of 16. Due to the stochastic nature of the deep learning models, we consider the average of 100 models. As a standard practice, we normalize the input features using a min-max scaler so that all feature values are in the range $[0, 1]$. We introduce a 20% dropout regularization layer after each LSTM layer in our models. We compile the models using the Adam optimizer (Kingma & Ba, 2014) and a Huber loss function, which is less susceptible to outliers (Huber, 1964).

3.4 Results and Discussions

We present the forecast accuracy metrics for the harvested area, yield, and farm price of the three commodities for the models described above in tables 3.1, 3.2, and 3.3. The naïve benchmark is a low bar, and any model that yields smaller errors than this naïve benchmark will be considered skillful. USDA baselines have smaller RMSE and MAPE than those of the naïve benchmark for harvested area, yield, and farm price for all three crops across all horizons. Any candidate algorithm to improve the baselines would need to have a couple of desirable properties. At a minimum, it must perform better than the naïve benchmark. Second, it should improve the performance of the USDA baselines, at least for some horizons. In particular, smaller forecast errors at longer horizons would be a good contribution, as USDA baselines tend to be less informative at longer horizons (Bora et al., 2022). Figures 3.2, 3.3, and 3.4 show the comparison of the forecast errors of all methods for harvested area, yield, and farm price, respectively, of the three commodities.

Table 3.1: Forecast accuracy for corn, soybeans, and wheat harvested area

		Corn		Soybeans		Wheat	
Horizon	Method	RMSE	MAPE(%)	RMSE	MAPE(%)	RMSE	MAPE(%)
h=0							
	Naive	4.03	4.06	4.65	4.13	3.86	7.31
	USDA	1.84	1.70	3.40	3.79	2.02	3.69
	ETS	3.33	3.61	4.60	4.17	3.86	7.31
	ETS Trend	3.16	3.40	3.93	3.58	3.85	7.31
	LSTM	3.91	3.71	3.20	3.15	4.14	7.78
	ED-LSTM	4.04	3.97	4.92	4.89	3.60	6.49
	CNN-LSTM	3.31	3.43	5.36	5.50	3.92	7.14
h=1							
	Naive	4.84	5.51	5.68	5.49	4.85	8.74
	USDA	2.77	2.61	5.60	5.86	2.55	4.34
	ETS	4.18	4.22	5.65	5.63	4.85	8.74
	ETS Trend	3.99	4.17	4.31	4.39	4.83	8.61
	LSTM	1.82	1.73	2.37	2.32	2.27	4.12
	ED-LSTM	3.06	2.99	3.80	3.87	3.46	6.37
	CNN-LSTM	3.62	3.61	5.47	5.76	4.09	7.73
h=2							
	Naive	4.76	4.61	6.71	6.84	5.58	10.52
	USDA	3.03	2.87	7.20	7.93	4.01	8.39
	ETS	4.63	4.58	6.66	6.76	5.58	10.52
	ETS Trend	4.50	4.76	5.45	5.97	5.49	10.20
	LSTM	1.37	1.29	2.00	2.07	1.89	3.37
	ED-LSTM	3.12	2.91	4.62	4.76	3.05	5.56
	CNN-LSTM	3.48	3.41	5.49	5.78	4.01	7.90
h=3							
	Naive	4.55	4.31	6.97	6.99	6.19	13.58
	USDA	3.21	2.95	7.10	7.62	4.08	8.44
	ETS	4.99	4.84	7.00	7.18	6.19	13.58
	ETS Trend	4.95	5.31	5.42	5.53	6.00	13.11
	LSTM	1.07	1.03	2.47	2.29	1.51	2.93
	ED-LSTM	2.92	2.72	4.86	4.96	2.53	4.73
	CNN-LSTM	3.64	3.35	5.51	5.73	3.86	7.88
h=4							
	Naive	4.11	3.86	8.20	8.58	7.28	15.57
	USDA	2.90	2.73	7.55	7.81	4.72	10.05
	ETS	5.14	4.72	8.14	8.63	7.28	15.57
	ETS Trend	5.10	5.44	5.80	5.67	7.07	14.94
	LSTM	9.77	7.29	7.14	6.45	10.20	15.53
	ED-LSTM	3.80	3.54	6.12	6.29	3.00	5.85
	CNN-LSTM	3.94	3.54	6.40	6.74	3.92	8.03

Table 3.2: Forecast accuracy for corn, soybeans, and wheat yield

		Corn		Soybeans		Wheat	
Horizon	Method	RMSE	MAPE(%)	RMSE	MAPE(%)	RMSE	MAPE(%)
h=0							
	Naive	13.87	6.43	2.60	4.85	3.94	6.68
	USDA	14.69	6.88	2.55	4.72	2.82	4.70
	ETS	15.52	8.80	3.20	5.90	3.55	6.08
	ETS Trend	13.73	6.96	3.16	5.69	2.60	4.41
	LSTM	11.44	6.33	2.65	4.84	2.63	4.26
	ED-LSTM	14.88	7.15	3.03	5.75	3.67	6.18
	CNN-LSTM	16.40	8.58	3.51	6.26	3.56	5.77
h=1							
	Naive	18.27	8.20	3.50	6.77	4.15	7.27
	USDA	14.79	6.99	2.65	4.77	2.67	4.20
	ETS	18.54	10.24	4.02	7.13	3.41	5.72
	ETS Trend	14.31	7.44	3.48	6.24	2.49	3.99
	LSTM	6.59	3.39	2.01	3.28	1.78	2.67
	ED-LSTM	13.71	6.64	3.00	5.16	3.14	4.25
	CNN-LSTM	17.32	8.93	3.91	6.91	3.57	5.31
h=2							
	Naive	20.62	9.83	4.32	7.80	3.46	6.16
	USDA	14.66	6.77	2.82	5.07	2.83	4.65
	ETS	19.22	9.85	4.67	8.05	3.37	5.85
	ETS Trend	13.75	6.80	3.66	6.21	2.77	4.67
	LSTM	4.12	2.07	1.08	1.82	1.38	2.24
	ED-LSTM	13.64	6.53	2.58	4.30	3.37	4.97
	CNN-LSTM	14.50	7.29	3.23	5.49	3.84	5.69
h=3							
	Naive	20.60	9.38	4.41	5.72	4.29	7.29
	USDA	14.20	6.50	3.00	5.29	2.88	4.47
	ETS	19.53	9.88	5.12	8.34	3.98	6.41
	ETS Trend	13.38	6.48	3.75	6.48	2.72	4.31
	LSTM	4.47	2.25	1.26	2.06	1.39	2.05
	ED-LSTM	11.15	5.59	2.37	4.10	2.81	4.39
	CNN-LSTM	14.55	7.38	3.31	5.72	3.77	5.45
h=4							
	Naive	20.62	8.76	4.93	7.66	5.36	9.23
	USDA	13.42	5.39	3.18	5.66	3.28	5.48
	ETS	20.39	11.14	5.76	10.55	4.78	8.12
	ETS Trend	12.81	5.73	4.01	7.32	3.16	5.52
	LSTM	23.70	11.18	4.86	8.87	8.04	11.20
	ED-LSTM	14.86	7.28	3.32	5.84	5.04	8.37
	CNN-LSTM	14.81	7.69	3.54	6.23	4.28	6.74

Table 3.3: Forecast accuracy for corn, soybeans, and wheat farm price

		Corn		Soybeans		Wheat	
Horizon	Method	RMSE	MAPE(%)	RMSE	MAPE(%)	RMSE	MAPE(%)
h=0							
	Naive	1.06	18.54	1.69	12.57	1.20	19.49
	USDA	0.92	13.25	1.55	10.34	0.93	13.58
	ETS	1.05	16.07	1.55	11.72	1.08	17.99
	ETS Trend	1.07	16.86	1.62	12.32	1.10	18.25
	LSTM	0.95	18.66	2.04	15.52	1.34	20.35
	ED-LSTM	1.05	16.21	1.48	10.89	1.00	16.32
	CNN-LSTM	1.35	24.41	2.00	15.94	1.24	19.81
h=1							
	Naive	1.63	29.91	2.72	21.29	1.89	29.46
	USDA	1.22	15.20	2.07	13.43	1.11	13.89
	ETS	1.61	28.14	2.44	18.52	1.64	26.92
	ETS Trend	1.66	29.64	2.53	18.93	1.71	28.92
	LSTM	0.44	8.33	0.93	7.35	0.63	9.46
	ED-LSTM	0.91	13.34	1.29	9.65	1.00	15.86
	CNN-LSTM	1.24	21.17	1.99	15.23	1.23	20.18
h=2							
	Naive	1.83	32.57	3.05	24.98	1.90	28.34
	USDA	1.40	17.47	2.44	16.02	1.29	14.43
	ETS	1.89	32.57	2.96	23.76	1.77	27.80
	ETS Trend	1.97	35.04	3.10	26.10	1.89	29.23
	LSTM	0.35	6.17	0.65	5.05	0.42	6.16
	ED-LSTM	0.96	13.28	1.26	8.91	0.84	12.20
	CNN-LSTM	1.11	16.76	1.72	12.92	1.04	15.35
h=3							
	Naive	1.91	38.23	3.32	28.26	1.76	29.28
	USDA	1.46	20.56	2.60	16.03	1.52	19.82
	ETS	1.92	37.38	3.15	26.19	1.82	29.64
	ETS Trend	2.04	40.98	3.38	29.47	1.98	31.44
	LSTM	0.29	4.99	0.55	4.17	0.29	4.43
	ED-LSTM	0.72	10.86	1.20	8.45	0.74	9.96
	CNN-LSTM	1.00	15.01	1.60	11.10	0.99	13.57
h=4							
	Naive	2.17	40.37	3.83	32.06	2.14	33.24
	USDA	1.40	20.35	2.71	18.55	1.66	23.17
	ETS	1.91	37.19	3.42	29.14	2.02	31.77
	ETS Trend	2.08	40.93	3.73	32.59	2.23	36.00
	LSTM	1.43	28.92	2.43	18.72	1.72	26.37
	ED-LSTM	0.96	15.65	1.70	11.73	1.06	15.43
	CNN-LSTM	0.98	15.76	1.86	12.91	1.09	14.45

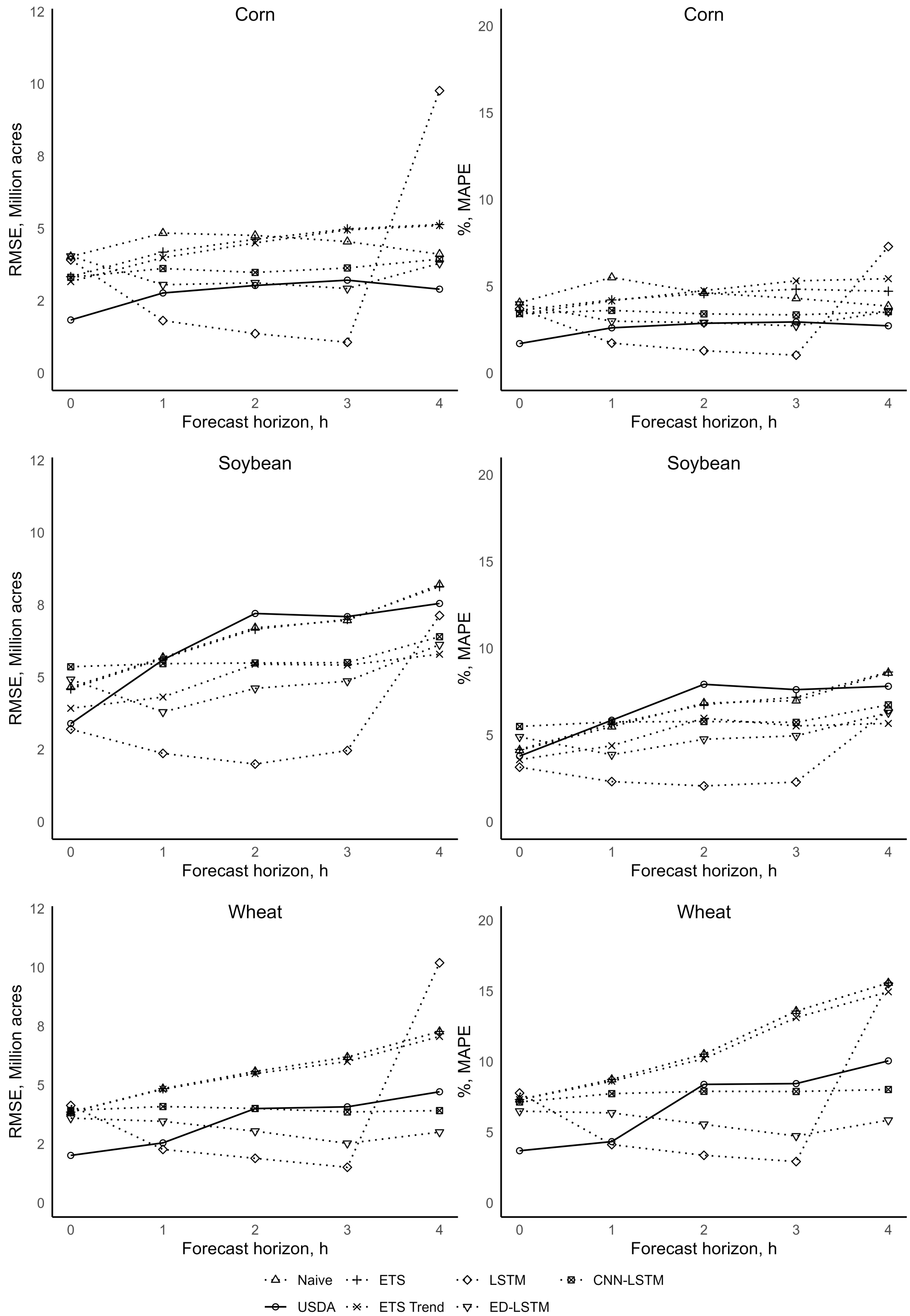


Figure 3.2: Root mean square errors (RMSE) and Mean Absolute Percent Errors (MAPE) for harvested area of corn, soybeans, and wheat

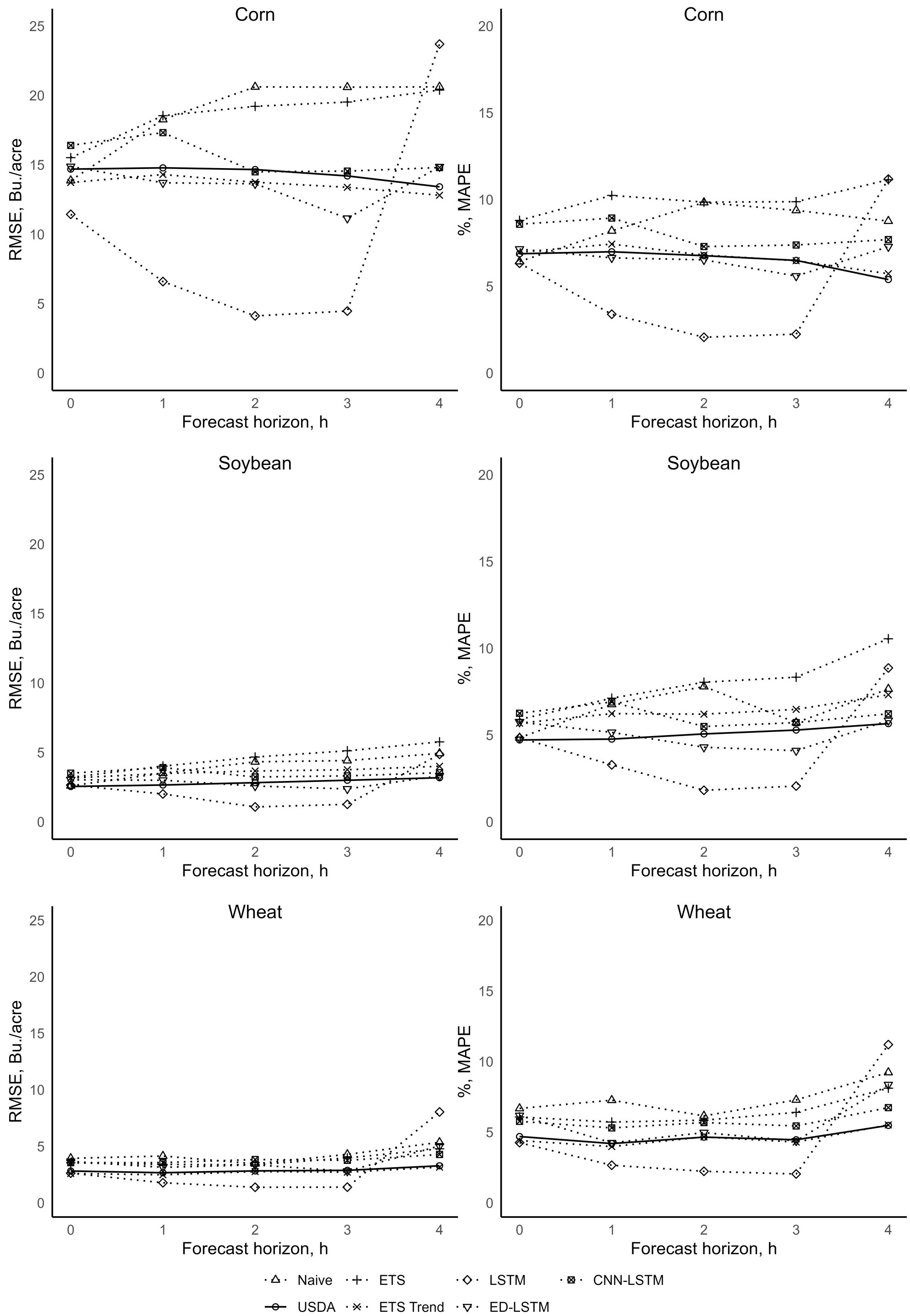


Figure 3.3: Root mean square errors (RMSE) and Mean Absolute Percent Errors (MAPE) for yields of corn, soybeans, and wheat

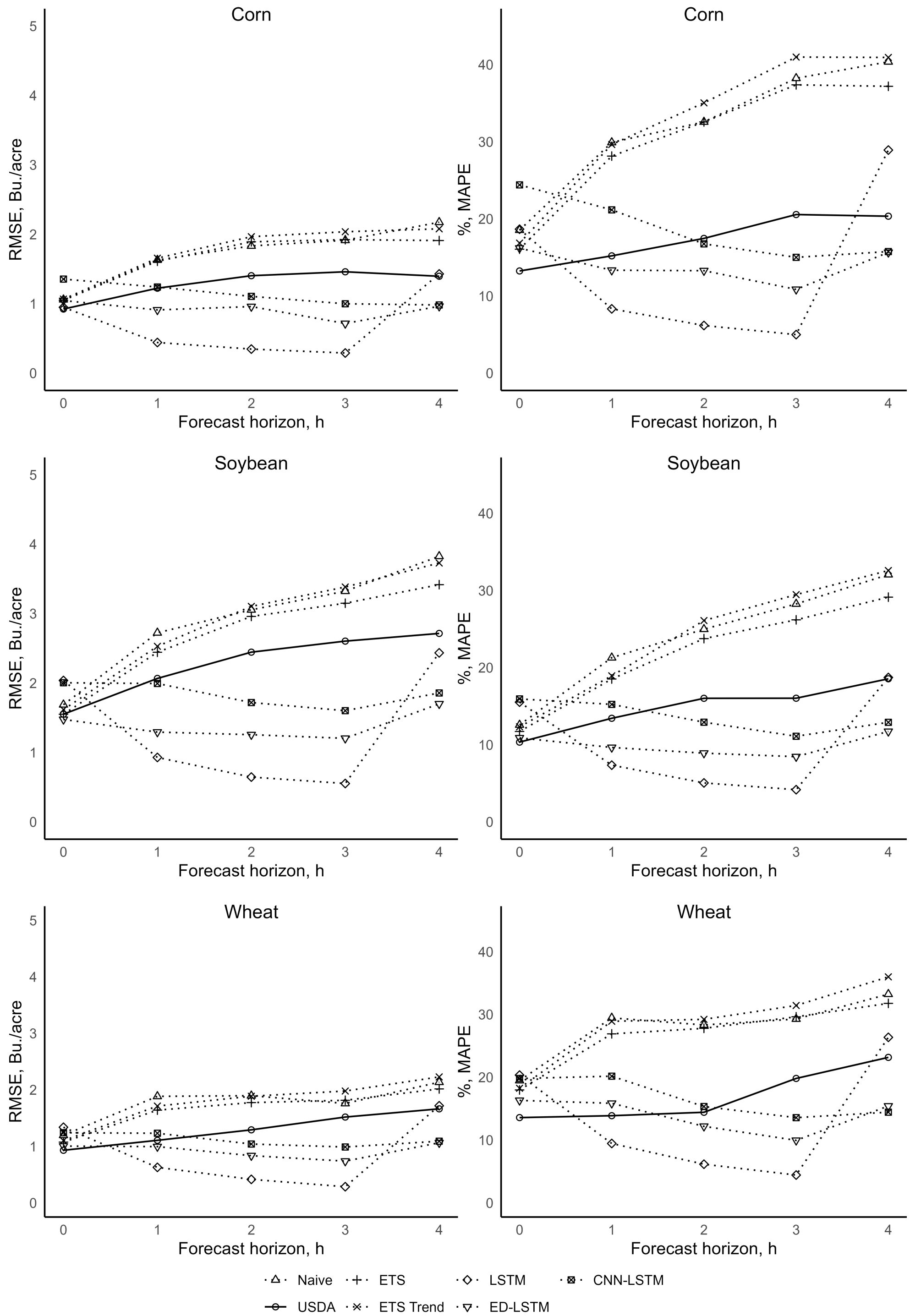


Figure 3.4: Root mean square errors (RMSE) and Mean Absolute Percent Errors (MAPE) for farm prices of corn, soybeans, and wheat

As expected, the RMSEs and MAPEs of the naïve benchmark are very high for most indicators across all horizons. The ETS methods, with or without trend, do not result in a considerable improvement in accuracy and have errors that are comparable to the naïve benchmark for forecasts of harvested area and farm price. For crop yield forecasts, the ETS with trend model performs well. The USDA baseline and the deep neural networks generally show superior skill compared to the naïve benchmarks. We focus the rest of our discussion on the performance of the USDA baseline and the three deep learning models.

The USDA predicts more accurately the harvested area of crops for the current year compared to the other methods (table 3.1 and figure 3.2). At $h = 0$, the MAPEs of the USDA baselines for corn, soybean, and wheat harvested area are 1.70%, 3.79%, and 3.69%, respectively. The MAPE of USDA baselines of corn harvested area remains low at longer horizons, with ED-LSTM and CNN-LSTM forecasts matching its performance closely for $h = \{1, 2, 3, 4\}$. The vanilla LSTM shows better accuracy than all other models for $h = \{1, 2, 3\}$, but its MAPEs are quite high for $h = \{0, 4\}$. The USDA baselines do not perform well in predicting harvested area of corn and soybeans for longer horizons, with a large increase in MAPE between $h = 0$ and $h = 4$ for both crops. For $h = \{2, 3, 4\}$, both ED-LSTM and CNN-LSTM forecasts have higher accuracy for soybean and wheat harvested area, with the ED-LSTM model performing slightly better.

The USDA projections of crop yields are fairly accurate across horizons, with MAPE around 5% (table 3.2 and figure 3.3). As observed in figure 3.1, crop yield has a strong time trend for all crops, making it easier to predict if the trend is correctly identified. The ETS trend model closely matches in performance with the

USDA model for all three crops, suggesting USDA might be using a similar model that includes trend to predict crop yield. The LSTM model has a lower MAPE than that of the USDA baselines for horizons $h = \{1, 2, 3\}$, but its accuracy drops sharply at $h = 4$. The MAPE for the yield forecasts from the CNN-LSTM and ED-LSTM models are of similar magnitude as the USDA baselines, at around 4–7%.

The deep learning models show noticeable improvement in accuracy while predicting farm price, which is an indicator that the USDA baselines struggle to predict accurately at longer horizons (table 3.3 and figure 3.4). At $h = 0$, the MAPE forecast errors of the USDA baselines are the lowest among all models. However, the MAPEs of the USDA baselines increase for longer horizons. Between $h = 0$ and $h = 4$, the MAPE of USDA corn price baselines increases from 13.25% to 20.35%. For the same horizons, the MAPE of the soybean and wheat price baselines increases from 10.34% to 18.55% and from 13.58% to 23.17%, respectively. The price forecasts from the vanilla LSTM model show very low MAPEs at horizons $h = \{1, 2, 3\}$, but its performance decreases drastically at $h = 4$, making it somewhat unreliable. Both CNN-LSTM and ED-LSTM models, however, perform much better at longer horizons compared to the USDA baseline in predicting farm prices of all three commodities for all forecast horizons except $h = 0$. For example, at $h = 4$, MAPEs of corn, soybean and wheat price forecasts of ED-LSTM model are 15.65%, 11.7%, and 15.43%, respectively. The MAPEs of the same forecasts of the CNN-LSTM model are 15.76%, 12.91% and 14.45%, respectively. Given that farmers frequently choose between various crops when planting, being able to reliably predict long-term commodity prices has implications for estimating outlays for federal programs.

The USDA baselines generally perform better than all other methods for the current-year forecasts ($h = 0$). For example, the current-year USDA baselines for harvested areas of corn, soybeans, and wheat have MAPEs of 1.7%, 3.79%, and 3.69%, respectively, which are among the lowest of all models. The current-year USDA crop yield baselines have low MAPEs as well, though the deep learning methods have comparable performance at $h = 0$. Similarly, the MAPEs of the USDA baselines are the lowest for the current year forecasts of farm prices of the three crops (13.25% for corn price, 10.34% for soybean price, and 13.58% for wheat price). These findings show that for indicators like yield, for which the USDA baselines are relatively accurate, the deep learning methods do not show much advantage in their performance. However, for indicators like farm prices and, to some extent, harvested areas that are more difficult to predict and have high errors, the deep learning methods can be used as an alternative method to improve the performance of the USDA baselines. This is not surprising since the USDA enjoys rich market and survey information and expert judgments for making predictions for the current year. On the other hand, all other methods rely solely on past patterns for predicting the values for the current year. The value of survey information and expert judgments diminishes as we move into the longer horizons. At longer horizons, the CNN-LSTM and ED-LSTM models, and sometimes the LSTM model, perform better than the USDA baselines or at least match them.

Among the three deep learning methods, the ED-LSTM and CNN-LSTM models show the most accurate performance across indicators over the horizons. The vanilla LSTM model performs well up to $h = 3$, but its accuracy deteriorates at $h = 4$, even worse than the naïve benchmark in some cases. However, the vanilla LSTM model is straightforward, with only one LSTM layer. The slightly more complex CNN-LSTM and ED-LSTM with additional layers consistently perform better than the other methods, especially at longer horizons. The extra layers make the deep neural networks more suited to learning complex non-linear relationships between several indicators. Between the two, the accuracy of the ED-LSTM model is marginally better than the CNN-LSTM model, although the difference is minimal.

Our study provides a working example to demonstrate that deep learning methods may produce more accurate multi-step commodity forecasts. One way to improve the predictions may be to add more input features to the problem, such as variables for additional crops. As in many high-dimensional small sample size applications of deep learning (Shen, Er, & Yin, 2022; Vabalas, Gowen, Poliakoff, & Casson, 2019), incorporating additional features may help overcome challenges posed by limited training samples and facilitate better forecast performance. Such high-dimensional networks might need a more complex architecture than the ones used in this study. The three deep learning models used in this study are still relatively simple compared to what a production-ready model with more input features and additional target variables to cover the entire baseline report would entail.

3.5 Conclusions

In this study, we developed three deep learning models for predicting harvested area, yield, and farm price of three major field crops for five years into the future and compared their performance against a naïve benchmark, exponential smoothing with and without trend, and USDA baselines. Except for ETS with a trend model for crop yields, the exponential smoothing methods do not significantly improve forecast accuracy over the naïve benchmark. The USDA baselines perform well in forecasting crop yield but struggle in forecasting harvested area and farm price, especially at longer horizons. The deep learning models show better accuracy than the USDA baselines in forecasting at longer horizons, most notably in predicting farm prices, where the USDA baselines show poor accuracy. The results suggest that deep learning methods can, at the very least, match the accuracy of USDA baselines for most indicators while offering significant improvement in accuracy for indicators that the USDA baselines do not predict well.

Deep learning methods have shown great promise in forecasting in other fields, but their use in predicting long-term agricultural statistics such as the USDA baselines have been limited. This study aims to bridge this gap. Efficient deep learning methods can have important implications for USDA baseline models and processes. The current baseline process is time-consuming, as it takes more than eight months to produce the baseline report. The use of deep learning methods as an alternative method may help in producing baseline forecasts that may have comparable accuracy but on a shorter timeline. In addition, several different sub-committees work on specific parts of the baseline report and produce inputs for the composite model. Deep learning methods, on the other hand, have the potential to make the process more straightforward and hence improve transparency. Therefore, the deep learning methods can act as a complement to the existing baseline projections models.

One of the limitations of this study is that the training sample is relatively small in the number of years. Deep neural networks often perform better when the training sample is large, and at smaller samples they may lead to overfitting. While our time period is limited, the number of features can be made much larger than in the current study. The USDA baseline report publishes hundreds of indicators, representing a high-dimensional prediction problem where the sample size is much smaller than the number of features. Future research may incorporate more input features and produce forecasts for additional target variables. However, it may require careful feature selection and dimensionality reduction strategies to overcome the challenge of high-dimensionality. Another limitation about the small sample is that we were able to produce forecasts for only five years. Yet another limitation of the deep learning methods is their "black-box" nature, making them difficult to explain when compared to economic modeling. However, advances in explainable deep learning methods may be able to address this issue in the future.

Bibliography

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., ... Zheng, X. (2015). *TensorFlow: Large-scale machine learning on heterogeneous systems.* Retrieved from `https://www.tensorflow.org/` (Software available from tensorflow.org)

Adjemian, M. K. (2012). Quantifying the wasde announcement effect. *American Journal of Agricultural Economics*, *94*(1), 238–256.

Adjemian, M. K., & Irwin, S. H. (2018). Usda announcement effects in real-time. *American Journal of Agricultural Economics*, *100*(4), 1151–1171.

Ahn, Y. B., & Tsuchiya, Y. (2019). Asymmetric loss and the rationality of inflation forecasts: evidence from south korea. *Pacific Economic Review*, *24*(4), 588–605.

Aretz, K., Bartram, S. M., & Pope, P. F. (2011). Asymmetric loss functions and the rationality of expected stock returns. *International Journal of Forecasting*, *27*(2), 413–437.

Auffhammer, M. (2007). The rationality of eia forecasts under symmetric and asymmetric loss. *Resource and Energy Economics*, *29*(2), 102–121.

Baghestani, H. (2013). Evaluating federal reserve predictions of growth in consumer spending. *Applied Economics*, *45*(13), 1637–1646.

Bailey, D., & Brorsen, B. W. (1998). Trends in the accuracy of usda production forecasts for beef and pork. *Journal of Agricultural and Resource Economics*, 515–525.

Batchelor, R. (2007). Bias in macroeconomic forecasts. *International Journal of Forecasting*, *23*(2), 189–203.

Batchelor, R., & Dua, P. (1990). Product differentiation in the economic forecasting industry. *International Journal of Forecasting*, *6*(3), 311–316.

Batchelor, R., & Peel, D. A. (1998). Rationality testing under asymmetric loss. *Economics Letters*, *61*(1), 49–54.

Beaudry, P., & Willems, T. (2018). *On the macroeconomic consequences of over-optimism* (Tech. Rep.). National Bureau of Economic Research.

Bora, S. S., Katchova, A. L., & Kuethe, T. H. (2021). The rationality of usda forecasts under multivariate asymmetric loss. *American Journal of Agricultural Economics*, *103*(3), 1006-1033. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/ajae.12142 doi: https://doi.org/10.1111/ajae.12142

Bora, S. S., Katchova, A. L., & Kuethe, T. H. (2022). *The accuracy and informativeness of agricultural baselines.* https://aede.osu.edu/sites/aede/files/publication_files/AgBaselinesInformativeness.pdf. (Working paper)

Borovykh, A., Bohte, S., & Oosterlee, C. W. (2018). Dilated convolutional neural networks for time series forecasting. *Journal of Computational Finance*. Retrieved from https://doi.org/10.21314/jcf.2019.358 doi: 10.21314/jcf.2019.358

Boussios, D., Skorbiansky, S. R., & MacLachlan, M. (2021, February). *Evaluating U.S. Department of Agriculture's Long-Term Forecasts for U.S. Harvested Area* (Economic Research Report No. 285). Economic Research Service, USDA.

Breitung, J., & Knüppel, M. (2021). How far can we forecast? statistical tests of the predictive content. *Journal of Applied Econometrics*, *36*(4), 369–392. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1002/jae.2817 doi: https://doi.org/10.1002/jae.2817

Capistrán, C. (2008). Bias in federal reserve inflation forecasts: Is the federal reserve irrational or just cautious? *Journal of Monetary Economics*, *55*(8), 1415–1427.

Capistrán, C. (2006). On comparing multi-horizon forecasts. *Economics Letters*, *93*(2), 176-181. Retrieved from https://www.sciencedirect.com/science/article/pii/S0165176506001480 doi: https://doi.org/10.1016/j.econlet.2006.04.010

Caunedo, J., Dicecio, R., Komunjer, I., & Owyang, M. T. (2020). Asymmetry, complementarities, and state dependence in federal reserve forecasts. *Journal of Money, Credit and Banking*, *52*(1), 205–228.

Caunedo, J., Dicecio, R., Komunjer, I., & Owyang, T., Michael. (2018). Asymmetry, complementarities, and state dependence in federal reserve forecasts. *Journal of Money, Credit and Banking*. Retrieved from https://onlinelibrary.wiley

.com/doi/abs/10.1111/jmcb.12590 doi: 10.1111/jmcb.12590

Chen, Y., Kang, Y., Chen, Y., & Wang, Z. (2020). Probabilistic forecasting with temporal convolutional neural network. *Neurocomputing*, *399*, 491–501. Retrieved from https://www.sciencedirect.com/science/article/pii/S0925231220303441 doi: https://doi.org/10.1016/j.neucom.2020.03.011

Chollet, F., et al. (2015). *Keras.* GitHub. Retrieved from https://github.com/fchollet/keras

Christodoulakis, G. A. (2020). Estimating the term structure of commodity market preferences. *European Journal of Operational Research*, *282*(3), 1146–1163.

Christodoulakis, G. A., & Mamatzakis, E. C. (2008). An assessment of the eu growth forecasts under asymmetric preferences. *Journal of Forecasting*, *27*(6), 483–492.

Christoffersen, P. F., & Diebold, F. X. (1996). Further results on forecasting and model selection under asymmetric loss. *Journal of applied econometrics*, *11*(5), 561–571.

Christoffersen, P. F., & Diebold, F. X. (1997). Optimal prediction under asymmetric loss. *Econometric theory*, *13*(6), 808–817.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, *74*(366a), 427-431. Retrieved from https://doi.org/10.1080/01621459.1979.10482531 doi: 10.1080/01621459.1979.10482531

Diebold, F. X., & Lopez, J. A. (1996). Forecast evaluation and combination. *Handbook of Statistics*, *14*, 241–268.

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

Dorfman, J. H., & Karali, B. (2015). A nonparametric search for information effects from usda reports. *Journal of Agricultural and Resource Economics*, 124–143.

Dubman, R., McElroy, R., & Dodson, C. (1993). *Forecasting farm income: Documenting usda's forecast model* (Technical Bulletin No. 1825). Washington DC: US Department of Agriculture, Economic Research Service.

Elliott, G., & Timmermann, A. (2008). Economic forecasting. *Journal of Economic Literature*, *46*(1), 3–56. Retrieved from http://www.jstor.org/stable/

27646946

Elliott, G., Timmermann, A., & Komunjer, I. (2005, 10). Estimation and Testing of Forecast Rationality under Flexible Loss. *The Review of Economic Studies*, *72*(4), 1107-1125. Retrieved from https://doi.org/10.1111/0034-6527.00363 doi: 10.1111/0034-6527.00363

Ellison, M., & Sargent, T. J. (2012). A defense of the fomc. *International Economic Review*, *53*(4), 1047–1065.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*(2), 179–211. Retrieved from https://www.sciencedirect.com/science/article/pii/036402139090002E doi: https://doi.org/10.1016/0364-0213(90)90002-E

Ericsson, N. R., & Martinez, A. B. (2019). Evaluating government budget forecasts. In D. Williams & T. Calabrese (Eds.), *The palgrave handbook of government budget forecasting* (pp. 37–69). Palgrave Macmillan.

Estrin, S., & Holmes, P. (1990). Indicative planning in developed economies. *Journal of Comparative Economics*, *14*(4), 531–554.

FAPRI. (2021). *Fapri-mu baseline outlook.* Food & Agricultural Policy Research Institute, University of Missouri. Retrieved from https://www.fapri.missouri.edu/publications/outlook/ (Accessed on April 1, 2021)

Fortenbery, T. R., & Sumner, D. (1993). The effects of usda reports in futures and options markets. *Journal of Futures Markets*, *13*(2), 157–173.

Francis, J., Hanna, J. D., & Philbrick, D. R. (1997). Management communications with securities analysts. *Journal of Accounting and Economics*, *24*(3), 363–394.

Francis, J., & Philbrick, D. (1993). Analysts' decisions as products of a multi-task environment. *Journal of Accounting Research*, *31*(2), 216–230.

Frankel, J. (2011). Over-optimism in forecasts by official budget agencies and its implications. *Oxford Review of Economic Policy*, *27*(4), 536–562.

Fritsche, U., Pierdzioch, C., Rülke, J.-C., & Stadtmann, G. (2015). Forecasting the brazilian real and the mexican peso: Asymmetric loss, forecast rationality, and forecaster herding. *International Journal of Forecasting*, *31*(1), 130–139.

Galbraith, J. W. (2003). Content horizons for univariate time-series forecasts. *International Journal of Forecasting*, *19*(1), 43–55. Retrieved from https://www.sciencedirect.com/science/article/pii/S0169207001001248 doi:

https://doi.org/10.1016/S0169-2070(01)00124-8

Galbraith, J. W., & Tkacz, G. (2007). Forecast content and content horizons for some important macroeconomic time series. *Canadian Journal of Economics/Revue canadienne d'économique*, *40*(3), 935-953. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2966.2007.00437.x doi: https://doi.org/10.1111/j.1365-2966.2007.00437.x

Giacomini, R., & Rossi, B. (2009, 04). Detecting and Predicting Forecast Breakdowns. *The Review of Economic Studies*, *76*(2), 669-705. Retrieved from https://doi.org/10.1111/j.1467-937X.2009.00545.x doi: 10.1111/j.1467-937X.2009.00545.x

Giacomini, R., & White, H. (2006). Tests of Conditional Predictive Ability. *Econometrica*, *74*(6), 1545–1578. Retrieved 2021-05-19, from http://www.jstor.org/stable/4123083

Giovannelli, A., & Pericoli, F. M. (2020). Are gdp forecasts optimal? evidence on european countries. *International Journal of Forecasting*.

Graham, J. R. (1999). Herding among investment newsletters: Theory and evidence. *The Journal of Finance*, *54*(1), 237–268.

Granger, C. W. J. (1969). Prediction with a generalized cost of error function. *Journal of the Operational Research Society*, *20*(2), 199–207.

Granger, C. W. J. (1999). Outline of forecast theory using generalized cost functions. *Spanish Economic Review*, *1*(2), 161–173.

Granger, C. W. J., & Pesaran, M. H. (2000). A decision theoretic approach to forecast evaluation. In W. Chan, W. Li, & H. Tong (Eds.), *Statistics and finance: An interface* (pp. 261–278). Imperial College Press.

Harvey, D. I., Leybourne, S., & Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of forecasting*, *13*(2), 281–291.

Harvey, D. I., Leybourne, S. J., & Newbold, P. (1998). Tests for forecast encompassing. *Journal of Business & Economic Statistics*, *16*(2), 254–259.

Higgins, M. L., & Mishra, S. (2014). State dependent asymmetric loss and the consensus forecast of real us gdp growth. *Economic Modelling*, *38*, 627–632.

Hjort, K., Boussios, D., Seeley, R., & Hansen, J. (2018, November). *The ERS Country-Commodity Linked System: Documenting Its International Country* *and Regional Agricultural Baseline Models* (Technical Bulletin No. 1951). Economic Research Service, USDA.

Hochreiter, S., & Schmidhuber, J. (1997, November). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780. Retrieved 2021-07-11, from https://doi.org/10.1162/neco.1997.9.8.1735 doi: 10.1162/neco.1997.9.8.1735

Hoffman, L. A., Etienne, X. L., Irwin, S. H., Colino, E. V., & Toasa, J. I. (2015a). Forecast performance of wasde price projections for us corn. *Agricultural economics*, *46*(S1), 157–171.

Hoffman, L. A., Etienne, X. L., Irwin, S. H., Colino, E. V., & Toasa, J. I. (2015b). Forecast performance of wasde price projections for u.s. corn. *Agricultural Economics*, *46*(S1), 157-171. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/agec.12204 doi: https://doi.org/10.1111/agec.12204

Holden, K., & Peel, D. A. (1990a). On testing for unbiasedness and efficiency of forecasts. *The Manchester School*, *58*(2), 120–127.

Holden, K., & Peel, D. A. (1990b). On testing for unbiasedness and efficiency of forecasts. *The Manchester School of Economic & Social Studies*, *58*(2), 120–127.

Holt, C. C. (2004). Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting*, *20*(1), 5-10. Retrieved from https://www.sciencedirect.com/science/article/pii/S0169207003001134 doi: https://doi.org/10.1016/j.ijforecast.2003.09.015

Huang, Y., Li, J., Hou, W., Zhang, B., Zhang, Y., Li, Y., & Sun, L. (2020). Improved clustering and deep learning based short-term wind energy forecasting in large-scale wind farms. *Journal of Renewable and Sustainable Energy*, *12*(6), 066101. Retrieved from https://doi.org/10.1063/5.0016226 doi: 10.1063/5.0016226

Huber, P. J. (1964). Robust Estimation of a Location Parameter. *The Annals of Mathematical Statistics*, *35*(1), 73 – 101. Retrieved from https://doi.org/10.1214/aoms/1177703732 doi: 10.1214/aoms/1177703732

Huffstutter, P., & Polansek, T. (2019). Farmer's threat prompts U.S. Agriculture Department to pull staff from crop tour. *Reuters*. Retrieved 2020-05-10, from https://www.reuters.com/article/us-usa-grains-tour-threat/farmers-threat-prompts-usda-to-pull-staff-from-crop-tour-idUSKCN1VB242

Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice (3rd ed).* https://otexts.com/fpp3/. (Accessed on 07.18.2022)

Iowa State University. (2021). *Fapri staff reports.* Iowa State University Digital Repository. Retrieved from https://lib.dr.iastate.edu/fapri_staffreports/ (Accessed on April 1, 2021)

Irwin, S., & Good, D. (2015, February). Long-term corn, soybeans, and wheat price forecasts and the farm bill program choice. *farmdoc daily*, *(5)*(20). Retrieved from https://farmdocdaily.illinois.edu/2015/02/long-term-forecasts-and-farm-bill-progr\protect\@normalcr\relaxam-choice.html

Isengildina, O., Irwin, S. H., & Good, D. L. (2004). Evaluation of usda interval forecasts of corn and soybean prices. *American Journal of Agricultural Economics*, *86*(4), 990 - 1004. Retrieved from https://proxy.lib.ohio-state.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=ecn&AN=0754042&site=ehost-live

Isengildina, O., Irwin, S. H., & Good, D. L. (2006). Are revisions to usda crop production forecasts smoothed? *American Journal of Agricultural Economics*, *88*(4), 1091–1104.

Isengildina-Massa, O., Cao, X., Karali, B., Irwin, S. H., Adjemian, M., & Johansson, R. C. (2020). When does usda information have the most impact on crop and livestock markets? *Journal of Commodity Markets*, 100137.

Isengildina-Massa, O., Irwin, S. H., Good, D. L., & Gomez, J. K. (2008). The impact of situation and outlook information in corn and soybean futures markets: Evidence from wasde reports. *Journal of Agricultural and Applied Economics*, *40*(1), 89–103.

Isengildina-Massa, O., Karali, B., & Irwin, S. H. (2013). When do the usda forecasters make mistakes? *Applied Economics*, *45*(36), 5086-5103. Retrieved from https://doi.org/10.1080/00036846.2013.818213 doi: 10.1080/00036846.2013.818213

Isengildina-Massa, O., Karali, B., Kuethe, T. H., & Katchova, A. L. (2020). Joint evaluation of the system of usda's farm income forecasts. *Applied Economic Perspectives and Policy*. doi: 10.1002/aepp.13064

Isengildina-Massa, O., Karali, B., Kuethe, T. H., & Katchova, A. L. (2021). Joint evaluation of the system of usda's farm income forecasts. *Applied Economic Perspectives and Policy*.

Isengildina-Massa, O., MacDonald, S., & Xie, R. (2012). A comprehensive evaluation of usda cotton forecasts. *Journal of Agricultural and Resource Economics*, 98–113.

Isengildina-Massa, O., & Sharp, J. L. (2012). Evaluation of usda interval forecasts revisited: Asymmetry and accuracy of corn, soybean, and wheat prices. *Agribusiness*, *28*(3), 310–323.

Isiklar, G., & Lahiri, K. (2007). How far ahead can we forecast? Evidence from cross-country surveys. *International Journal of Forecasting*, *23*(2), 167–187. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0169207007000040` doi: https://doi.org/10.1016/j.ijforecast.2007.01.003

Jonung, L., & Larch, M. (2006). Improving fiscal policy in the eu: the case for independent forecasts. *Economic Policy*, *21*(47), 492–534.

Jordà, O., & Marcellino, M. (2010). Path forecast evaluation. *Journal of Applied Econometrics*, *25*(4), 635-662. Retrieved from `https://onlinelibrary.wiley.com/doi/abs/10.1002/jae.1166` doi: https://doi.org/10.1002/jae.1166

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological review*, *80*(4), 237.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–292.

Karali, B., Isengildina-Massa, O., Irwin, S. H., Adjemian, M. K., & Johansson, R. (2019). Are USDA reports still news to changing crop markets? *Food Policy*, *84*, 66–76.

Keane, M. P., & Runkle, D. E. (1990). Testing the rationality of price forecasts: New evidence from panel data. *American Economic Review*, *80*(4), 714–735.

Keane, M. P., & Runkle, D. E. (1998). Are financial analysts' forecasts of corporate profits rational? *Journal of Political Economy*, *106*(4), 768–805.

Kim, H. Y., & Won, C. H. (2018). Forecasting the volatility of stock price index: A hybrid model integrating lstm with multiple garch-type models. *Expert Syst. Appl.*, *103*, 25-37.

Kingma, D. P., & Ba, J. (2014). *Adam: A method for stochastic optimization.* arXiv. Retrieved from `https://arxiv.org/abs/1412.6980` doi: 10.48550/ARXIV.1412.6980

Komunjer, I., & Owyang, M. T. (2012). Multivariate forecast evaluation and rationality testing. *The Review of Economics and Statistics*, *94*(4), 1066-1080. Retrieved from `https://doi.org/10.1162/REST{_}a{_}00215` doi: 10.1162/REST_a_00215

Krol, R. (2013). Evaluating state revenue forecasting under a flexible loss function. *International Journal of Forecasting*, *29*(2), 282–289.

Krüger, J. J., & LeCrone, J. (2019). Forecast evaluation under asymmetric loss: A monte carlo analysis of the ekt method. *Oxford Bulletin of Economics and Statistics*, *81*(2), 437–455.

Kuethe, T. H., Bora, S. S., & Katchova, A. L. (2022, 5). Improving ers's net cash income forecasts using usda baseline projections. *Journal of Agricultural and Resource Economics*, *47*, 246-261.

Kuethe, T. H., Hubbs, T., & Sanders, D. R. (2018). Evaluating the usda's net farm income forecast. *Journal of Agricultural and Resource Economics*, *43*(3), 276505.

Lara-Benítez, P., Carranza-García, M., Luna-Romera, J. M., & Riquelme, J. C. (2020). Temporal convolutional networks applied to energy-related time series forecasting. *Applied Sciences*, *10*(7). Retrieved from `https://www.mdpi.com/2076-3417/10/7/2322` doi: 10.3390/app10072322

Lara-Benítez, P., Carranza-García, M., & Riquelme, J. C. (2021). An experimental review on deep learning architectures for time series forecasting. *International Journal of Neural Systems*, *31*(03), 2130001. Retrieved from `https://doi.org/10.1142/S0129065721300011` (PMID: 33588711) doi: 10.1142/S0129065721300011

Laster, D., Bennett, P., & Geoum, I. S. (1999). Rational bias in macroeconomic forecasts. *The Quarterly Journal of Economics*, *114*(1), 293–318.

Lewis, K. E., & Manfredo, M. R. (2012). An evaluation of the usda sugar production and consumption forecasts. *Journal of Agribusiness*, *30*(345-2017-1168), 155–172.

Lim, T. (2001). Rationality and analysts' forecast bias. *The journal of Finance*, *56*(1), 369–385.

Lucier, G., Chesley, A., & Ahearn, M. C. (1986). *Farm income data: A historical perspective* (Statistical Bulletin No. 740). Washington DC: US Department of Agriculture, Economic Research Service.

Maddox, W. T., & Bohil, C. J. (1998). Base-rate and payoff effects in multidimensional perceptual categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(6), 1459.

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). The m4 competition: Results, findings, conclusion and way forward. *International Journal of Forecasting*, *34*, 802-808.

Mamatzakis, E., & Koutsomanoli-Filippaki, A. (2014). Testing the rationality of doe's energy price forecasts under asymmetric loss preferences. *Energy Policy*, *68*, 567–575.

Mamatzakis, E., & Tsionas, M. G. (2015). How are market preferences shaped? the case of sovereign debt of stressed euro-area countries. *Journal of Banking & Finance*, *61*, 106 - 116. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0378426615002150` doi: https://doi.org/10.1016/j.jbankfin.2015.08.004

Martinez, A. (2020, January). *Testing for Differences in Path Forecast Accuracy* (Working Paper). Office of Macroeconomic Analysis, U.S. Department of the Treasury.

McGath, C., McElroy, R., Strickland, R., Traub, L., Covey, T., Short, S., … Vogel, S. (2009a). *Forecasting farm income: Documenting usda's forecast model* (Technical Bulletin No. 1924). Washington DC: US Department of Agriculture, Economic Research Service.

McGath, C., McElroy, R., Strickland, R., Traub, L., Covey, T., Short, S., … Vogel, S. (2009b). *Forecasting farm income: Documenting usda's forecast model* (Technical Bulletin No. 1924). Washington DC: US Department of Agriculture, Economic Research Service.

McKenzie, A. M. (2008). Pre-harvest price expectations for corn: The information content of usda reports and new crop futures. *American Journal of Agricultural Economics*, *90*(2), 351–366.

Medvedev, N., & Wang, Z. (2022). Multistep forecast of the implied volatility surface using deep learning. *Journal of Futures Markets*, *42*(4), 645-667. Retrieved from `https://onlinelibrary.wiley.com/doi/abs/10.1002/fut.22302` doi: https://doi.org/10.1002/fut.22302

Merola, R., & Pérez, J. J. (2013). Fiscal forecast errors: governments versus independent agencies? *European Journal of Political Economy*, *32*, 285–299.

Meyers, W. H., & Westhoff, P. (2010). The FAPRI approach: A few key principles. *Nova Science Publishers, Inc.*, *6*(1), 2–19.

Meyers, W. H., Westhoff, P., Fabiosa, J. F., & Hayes, D. J. (2010). The FAPRI Global Modeling System and Outlook Process. *Nova Science Publishers, Inc.*, *6*(1), 2–19.

Mincer, J. A., & Zarnowitz, V. (1969). The evaluation of economic forecasts. In J. Mincer (Ed.), *Economic forecasts and expectations: Analysis of forecasting behavior and performance* (pp. 3–46). NBER.

Muth, J. F. (1960). Optimal properties of exponentially weighted forecasts. *Journal of the american statistical association*, *55*(290), 299–306.

Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation. *Econometrica*, *55*(3), 703–708.

NOAA National Centers for Environmental information. (2022, July). *Climate at a glance: National time series.* `https://www.ncei.noaa.gov/cag/`. (retrieved on July 18, 2022)

OECD, & Food and Agriculture Organization of the United Nations. (2020). *OECD-FAO Agricultural Outlook 2020-2029*. OECD. Retrieved 2021-03-27, from `https://www.oecd-ilibrary.org/agriculture-and-food/oecd-fao-agricultural-outlook-\protect\@normalcr\relax2020-2029_1112c23b-en` doi: 10.1787/1112c23b-en

Office of Management and Budget. (2020). Statistical programs of the United States government. *Annual Report*.

Patton, A. J., & Timmermann, A. (2007). Properties of optimal forecasts under asymmetric loss and nonlinearity. *Journal of Econometrics*, *140*(2), 884–918.

Patton, A. J., & Timmermann, A. (2012, January). Forecast Rationality Tests Based on Multi-Horizon Bounds. *Journal of Business & Economic Statistics*, *30*(1), 1–17. Retrieved 2021-05-19, from `https://www.tandfonline.com/doi/full/10.1080/07350015.2012.634337` doi: 10.1080/07350015.2012.634337

Pierdzioch, C., Reid, M. B., & Gupta, R. (2016). Forecasting the south african inflation rate: On asymmetric loss and forecast rationality. *Economic Systems*, *40*(1), 82–92.

Pierdzioch, C., Rülke, J.-C., & Stadtmann, G. (2013). Forecasting us housing starts under asymmetric loss. *Applied Financial Economics*, *23*(6), 505–513.

Pierdzioch, C., Rülke, J.-C., & Stadtmann, G. (2015). Central banks' inflation forecasts under asymmetric loss: Evidence from four latin-american countries. *Economics Letters*, *129*, 66–70.

Quaedvlieg, R. (2021). Multi-horizon forecast comparison. *Journal of Business & Economic Statistics*, *39*(1), 40-53. Retrieved from `https://doi.org/10.1080/07350015.2019.1620074` doi: 10.1080/07350015.2019.1620074

Said, S. E., & Dickey, D. A. (1984). Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, *71*(3), 599–607. Retrieved from `http://www.jstor.org/stable/2336570`

Sanders, D. R., & Manfredo, M. R. (2003). Usda livestock price forecasts: A comprehensive evaluation. *Journal of Agricultural and Resource Economics*(1835-2016-148820), 19. Retrieved from `http://ageconsearch.umn.edu/record/31101`

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, *61*, 85–117. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0893608014002135` doi: https://doi.org/10.1016/j.neunet.2014.09.003

Seabold, S., & Perktold, J. (2010). statsmodels: Econometric and statistical modeling with python. In *9th python in science conference.*

Shen, L., Er, M. J., & Yin, Q. (2022). Classification for high-dimension low-sample size data. *Pattern Recognition*, *130*, 108828. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0031320322003090` doi: https://doi.org/10.1016/j.patcog.2022.108828

Smyl, S. (2020). A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting. *International Journal of Forecasting*, *36*(1), 75–85. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0169207019301153` doi: https://doi.org/10.1016/j.ijforecast.2019.03.017

Sumner, D. A., & Mueller, R. A. (1989). Are harvest forecasts news? usda announcements and futures market reactions. *American Journal of Agricultural Economics*, *71*(1), 1–8.

Theil, H. (1958). *Economic Forecasts and Policy.* North-Holland Publishing Company.

Tsuchiya, Y. (2016a). Assessing macroeconomic forecasts for Japan under an asymmetric loss function. *International Journal of Forecasting*, *32*(2), 233–242.

Tsuchiya, Y. (2016b). Asymmetric loss and rationality of Chinese renminbi forecasts: An implication for the trade between China and the US. *Journal of International Financial Markets, Institutions and Money*, *44*, 116–127.

Ulan, M., Dewald, W. G., & Bullard, J. B. (1995). Us official forecasts of g-7 economies, 1976-90. *Federal Reserve Bank of St. Louis Review*, *77*(2), 39.

Ulu, Y. (2013). Multivariate test for forecast rationality under asymmetric loss functions: Recent evidence from mms survey of inflation–output forecasts. *Economics Letters*, *119*(2), 168–171.

USDA ERS. (2020). *Usda ers, agricultural baseline, questions & answers.* USDA Economic Research Service. Retrieved from https://www.ers.usda.gov/topics/farm-economy/agricultural-baseline/questions-answers/ (Accessed on April 1, 2021)

USDA ERS. (2021a). *Data files: U.s. and state-level farm income and wealth statistics.* Economic Research Service, USDA. Retrieved from https://www.ers.usda.gov/data-products/farm-income-and-wealth-statistics\/data-files-us-and-state-level-farm-income-and-wealth-statistics/ (Accessed on April 1, 2021)

USDA ERS. (2021b). *Usda agricultural projections.* USDA Economic, Statistics and Market Information System. Retrieved from https://usda.library.cornell.edu/concern/publications/qn59q396v?locale=en (Accessed on April 1, 2021)

USDA Foreign Agricultural Service. (2021). *Production, supply and distribution database.* Retrieved from https://data.nal.usda.gov/dataset/production-supply-and-distribution-database (Foreign Agricultural Service, Department of Agriculture)

USDA NASS. (2021, June). *Acreage.* National Agricultural Statistics Service (NASS), Agricultural Statistics Board, United States Department of Agriculture (USDA). Retrieved from https://downloads.usda.library.cornell.edu/usda-esmis/files/j098zb09z/00000x092/kw52k657g/acrg0621.pdf

USDA National Agricultural Statistics Service. (2021). *Nass - quick stats: Ag data commons.* USDA National Agricultural Statistics Service. Retrieved from https://data.nal.usda.gov/dataset/nass-quick-stats (Accessed on April 1, 2021)

USDA Office of Chief Economist. (2022). *Usda agricultural projections to 2031.*

prepared by the interagency agricultural projections committee long-term projections report oce-2022-1 (United States Department of Agriculture).

USDA Office of the Chief Economist. (2020). *Usda agricultural projections to 2029* (Pepared by the Interagency Agricultural Projections Committee No. Long-term Projections Report OCE-2020-1). United States Department of Agriculture.

Vabalas, A., Gowen, E., Poliakoff, E., & Casson, A. J. (2019, 11). Machine learning algorithm validation with a limited sample size. *PLOS ONE*, *14*(11), 1-20. Retrieved from `https://doi.org/10.1371/journal.pone.0224365` doi: 10.1371/journal.pone.0224365

Varian, H. R. (1975). A bayesian approach to real estate assessment. In S. E. Fienber & A. Zellner (Eds.), *Studies in bayesian econometric and statistics in honor of leonard j. savage* (pp. 195–208). Amsterdam: North Holland.

Vogel, F., & Bange, G. (1999). *Understanding usda crop forecasts* (Technical Bulletin No. 1554). Washington DC: US Department of Agriculture, National Agricultural Statistics Service and Office of the Chief Economist, World Agricultural Outlook Board.

Wan, R., Mei, S., Wang, J., Liu, M., & Yang, F. (2019). Multivariate temporal convolutional network: A deep neural networks approach for multivariate time series forecasting. *Electronics*, *8*(8). Retrieved from `https://www.mdpi.com/2079-9292/8/8/876` doi: 10.3390/electronics8080876

Wang, Y., Shen, Y., Mao, S., Chen, X., & Zou, H. (2019). Lasso and lstm integrated temporal model for short-term solar intensity forecasting. *IEEE Internet of Things Journal*, *6*, 2933-2944.

Weber, E. U. (1994). From subjective probabilities to decision weights: The effect of asymmetric loss functions on the evaluation of uncertain outcomes and events. *Psychological Bulletin*, *115*(2), 228.

West, K. D., Edison, H. J., & Cho, D. C. (1993). A utility-based comparison of some models of exchange rate volatility. *Journal of International Economics*, *35*(1-2), 23–45.

Westhoff, P. (2015, February). Price projections and farm bill program choices: Adding FAPRI-MU projections to the mix. *farmdoc daily*, *(5)*(26). Retrieved from `https://farmdocdaily.illinois.edu/2015/02/price-projections-and-farm-bill-program-choices.html`

Xiao, J., Hart, C. E., & Lence, S. H. (2017). Usda forecasts of crop ending stocks: How well have they performed? *Applied Economic Perspectives and Policy*, *39*(2), 220–241.

Zellner, A. (1986a). Bayesian estimation and prediction using asymmetric loss functions. *Journal of the American Statistical Association*, *81*(394), 446–451.

Zellner, A. (1986b). Biased predictors, rationality and the evaluation of forecasts. *Economics Letters*, *21*(1), 45–48.

Appendix A: Additional Figures and Tables

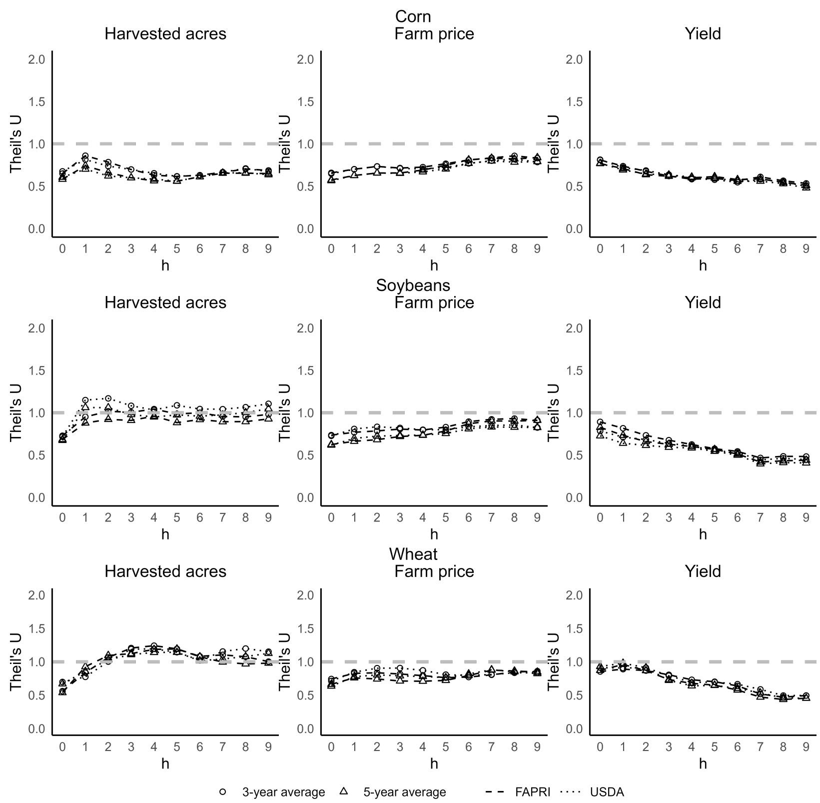


Figure A.1: Theil's U for USDA and FAPRI baseline projections of corn, soybeans and wheat by projection horizon h, 1997–2020

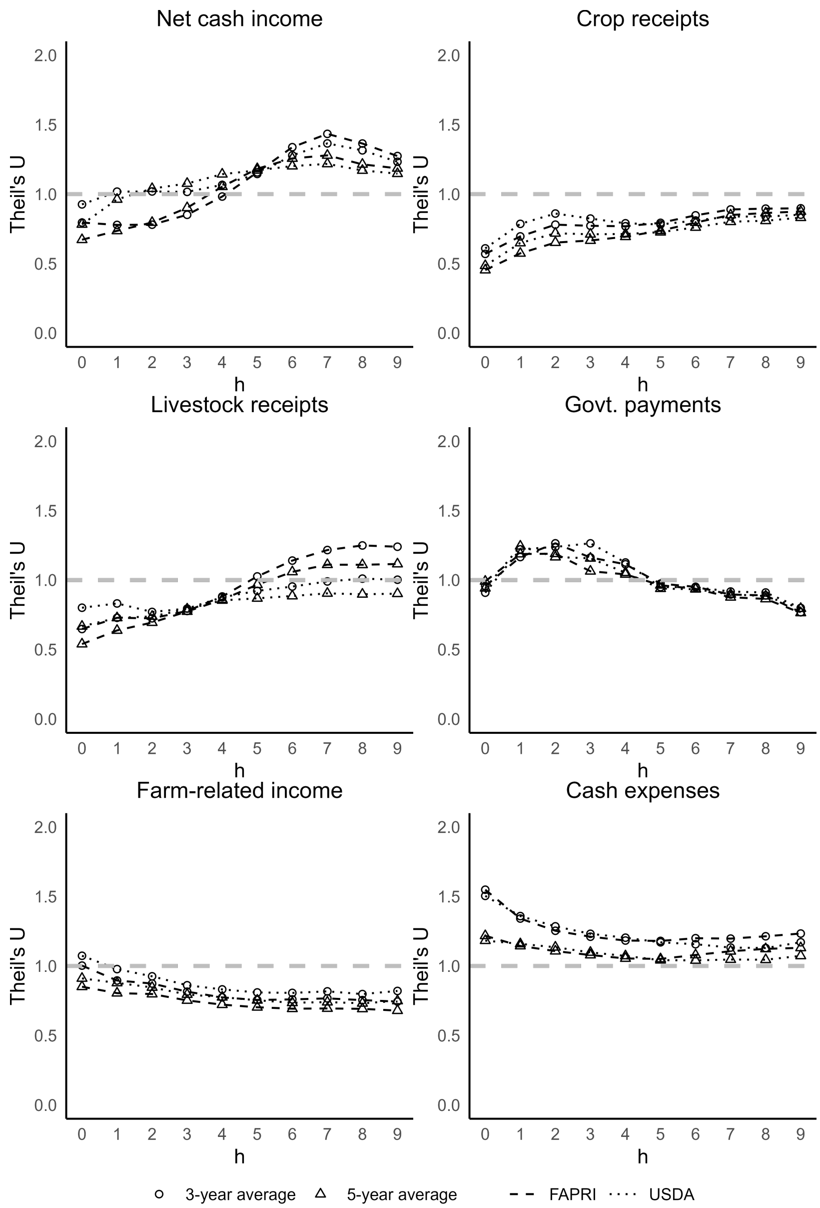


Figure A.2: Theil's U for USDA and FAPRI baseline projections of net cash income and its components by projection horizon h, 1997–2020

Table A.1: Estimates of the parameter $\hat{\beta}_h^{USDA}$ in the Mincer-Zarnowitz (MZ) regression for USDA projections

	h=0	h=1	h=2	h=3	h=4	h=5	h=6	h=7	h=8	h=9
Corn										
Harvested acres	0.978^{+++}_{**}	0.96^{+++}_{**}	0.994^{+++}_{**}	0.975^{+++}_{**}	0.906^{+++}_{**}	0.727^{+++}_{**}	0.509^{**}	0.193^{**}	0.043	0.065
	(0.079)	(0.139)	(0.174)	(0.215)	(0.308)	(0.293)	(0.227)	(0.150)	(0.181)	(0.221)
Farm price	1.058^{+++}_{**}	1.12^{+++}_{**}	1.003^{+++}_{**}	0.827^{++}_{**}	0.532^{**}	0.198^{**}	-0.15	-0.369	-0.525	-0.567
	(0.127)	(0.213)	(0.254)	(0.238)	(0.258)	(0.252)	(0.199)	(0.271)	(0.398)	(0.370)
Yield	0.816^{+++}_{**}	0.769^{+++}_{**}	0.762^{+++}_{**}	0.749^{+++}_{**}	0.739^{+++}_{**}	0.727^{+++}_{**}	0.659^{+++}_{**}	0.621^{+++}_{**}	0.748^{+++}_{**}	0.798^{+++}_{**}
	(0.131)	(0.148)	(0.158)	(0.158)	(0.159)	(0.158)	(0.132)	(0.160)	(0.140)	(0.193)
Soybean										
Harvested acres	0.809^{+++}_{**}	0.55^{+++}_{**}	0.507^{**}	0.65^{+++}_{**}	0.681^{+++}_{**}	0.568^{+}_{**}	0.547^{**}	0.496^{**}	0.231^{**}	0.004
	(0.119)	(0.195)	(0.142)	(0.177)	(0.181)	(0.213)	(0.282)	(0.236)	(0.284)	(0.418)
Farm price	0.983^{+++}_{**}	0.911^{+++}_{**}	0.804^{+++}_{**}	0.665^{+++}_{**}	0.485^{**}	0.225^{**}	-0.069	-0.249	-0.393	-0.435
	(0.100)	(0.153)	(0.126)	(0.150)	(0.168)	(0.203)	(0.194)	(0.092)	(0.239)	(0.267)
Yield	1.16^{+++}_{**}	1.261^{+++}_{**}	1.405^{+++}_{**}	1.433^{+++}_{**}	1.543^{+++}_{**}	1.797^{+++}_{**}	2.073^{+++}_{**}	1.871^{+++}_{**}	1.911^{+++}_{**}	2.027^{+++}_{**}
	(0.157)	(0.185)	(0.232)	(0.250)	(0.280)	(0.333)	(0.394)	(0.317)	(0.346)	(0.232)
Wheat										
Harvested acres	0.991^{+++}_{**}	0.917^{+++}_{**}	0.693^{+++}_{**}	0.597^{+++}_{**}	0.538^{**}	0.586^{++}_{**}	0.72^{+++}_{**}	0.67^{+++}_{**}	0.651^{+++}_{**}	0.698^{+++}_{**}
	(0.051)	(0.131)	(0.216)	(0.249)	(0.250)	(0.225)	(0.207)	(0.214)	(0.194)	(0.174)
Farm price	0.968^{+++}_{**}	0.863^{+++}_{**}	0.606^{**}	0.338^{**}	0.037	-0.166	-0.421	-0.639	-0.76	-0.712
	(0.098)	(0.170)	(0.036)	(0.048)	(0.040)	(0.039)	(0.115)	(0.192)	(0.217)	(0.180)
Yield	0.836^{+++}_{**}	0.827^{+++}_{**}	0.9^{+++}_{**}	1.142^{+++}_{**}	1.273^{+++}_{**}	1.189^{+++}_{**}	0.808^{+++}_{**}	0.988^{+++}_{**}	1.244^{+++}_{**}	1.216
	(0.199)	(0.289)	(0.234)	(0.298)	(0.161)	(0.300)	(0.307)	(0.289)	(0.280)	(0.263)
Farm income										
Net cash income	0.899^{+++}_{**}	0.647^{+++}_{**}	0.587^{+}_{**}	0.615^{+}_{**}	0.521^{**}	0.51^{**}	0.254^{**}	-0.204	-0.503	-0.403
	(0.150)	(0.188)	(0.211)	(0.283)	(0.284)	(0.203)	(0.253)	(0.293)	(0.243)	(0.189)
Crop receipts	1.03^{+++}_{**}	1.055^{+++}_{**}	0.97^{+++}_{**}	0.803^{+++}_{**}	0.572^{**}	0.295^{**}	-0.046	-0.372	-0.59	-0.67
	(0.093)	(0.160)	(0.126)	(0.086)	(0.018)	(0.006)	(0.008)	(0.016)	(0.103)	(0.196)
Livestock receipts	0.863^{+++}_{**}	0.78^{+++}_{**}	0.773^{+++}_{**}	0.659^{+++}_{**}	0.495^{**}	0.502^{**}	0.42^{**}	0.199^{**}	0.037	-0.196
	(0.126)	(0.197)	(0.186)	(0.208)	(0.201)	(0.282)	(0.313)	(0.249)	(0.274)	(0.444)
Govt. payments	0.84^{+++}_{**}	-0.135	-0.037	-0.398	-0.668	-0.607	-0.67	-0.69	-0.751	-0.484
	(0.189)	(0.316)	(0.229)	(0.230)	(0.203)	(0.251)	(0.272)	(0.286)	(0.315)	(0.319)
Farm-related income	0.853^{+++}_{**}	0.892^{+++}_{**}	0.876^{+++}_{**}	0.988^{+++}_{**}	0.987^{+++}_{**}	1.13^{+++}_{**}	1.129^{+++}_{**}	1.078^{+++}_{**}	1.163^{+++}_{**}	0.558^{+}
	(0.112)	(0.169)	(0.208)	(0.278)	(0.257)	(0.273)	(0.272)	(0.275)	(0.257)	(0.186)
Cash expenses	0.986^{+++}_{**}	0.98^{+++}_{**}	0.9^{+++}_{**}	0.815^{+++}_{**}	0.713^{++}_{**}	0.658^{+}_{**}	0.53^{**}	0.398^{**}	0.198	-0.188
	(0.078)	(0.114)	(0.140)	(0.172)	(0.210)	(0.153)	(0.169)	(0.129)	(0.098)	(0.144)

Notes: *, **, and *** denote statistical significance at 10%, 5%, and 1% respectively for testing the null hypotheses $H_0 : \beta_h^{USDA} \leq 0$. Likewise, +, ++, and +++ denote statistical significance at 10%, 5%, and 1% respectively for testing the null hypothesis $H_0 : \beta_h^{USDA} \leq 0.5$.

Table A.2: Estimates of the parameter $\hat{\beta}_h^{FAPRI}$ in the Mincer-Zarnowitz (MZ) regression for FAPRI projections

	h=0	h=1	h=2	h=3	h=4	h=5	h=6	h=7	h=8	h=9
Corn										
Harvested acres	0.995^{+++}_{**}	0.94^{+++}_{**}	0.925^{+++}_{**}	0.876^{+++}_{**}	0.833^{+++}_{**}	0.655^{+++}_{**}	0.469_{**}	0.207_{**}	0.059	-0.047
	(0.075)	(0.147)	(0.193)	(0.210)	(0.264)	(0.263)	(0.207)	(0.162)	(0.198)	(0.186)
Farm price	1.071^{+++}_{**}	0.962^{+++}_{**}	0.824^{+++}_{**}	0.678^{+}_{**}	0.456_{**}	0.187_{*}	-0.047	-0.173	-0.332	-0.358
	(0.128)	(0.192)	(0.190)	(0.237)	(0.261)	(0.277)	(0.251)	(0.216)	(0.227)	(0.263)
Yield	0.826^{+++}_{**}	0.76^{+++}_{**}	0.775^{+++}_{**}	0.767^{+++}_{**}	0.744^{+++}_{**}	0.696^{+++}_{**}	0.618^{+++}_{**}	0.566^{+}_{**}	0.698^{+++}_{**}	0.704^{+++}_{**}
	(0.138)	(0.167)	(0.163)	(0.150)	(0.154)	(0.171)	(0.141)	(0.152)	(0.140)	(0.182)
Soybean										
Harvested acres	0.767^{+++}_{**}	0.683^{+++}_{**}	0.613^{+++}_{**}	0.63^{+++}_{**}	0.523_{**}	0.67^{+++}_{**}	0.616^{++}_{**}	0.721^{+++}_{**}	0.626^{++}_{**}	0.423_{**}
	(0.120)	(0.205)	(0.180)	(0.148)	(0.148)	(0.147)	(0.245)	(0.294)	(0.211)	(0.288)
Farm price	0.92^{+++}_{**}	0.838^{+++}_{**}	0.746^{+++}_{**}	0.599^{+}_{**}	0.438_{**}	0.228_{**}	0.004	-0.117	-0.209	-0.257
	(0.098)	(0.151)	(0.128)	(0.162)	(0.156)	(0.176)	(0.185)	(0.085)	(0.097)	(0.158)
Yield	1.183^{+++}_{**}	1.283^{+++}_{**}	1.5^{+++}_{**}	1.56^{+++}_{**}	1.693^{+++}_{**}	1.83^{+++}_{**}	1.926^{+++}_{**}	1.656^{+++}_{**}	1.516^{+++}_{**}	1.442^{+++}_{**}
	(0.174)	(0.211)	(0.244)	(0.206)	(0.279)	(0.345)	(0.350)	(0.216)	(0.228)	(0.273)
Wheat										
Harvested acres	0.982^{+++}_{**}	0.903^{+++}_{**}	0.706^{+++}_{**}	0.653^{+++}_{**}	0.55_{**}	0.55_{**}	0.672^{+++}_{*}	0.696^{+++}_{**}	0.696^{+++}_{**}	0.714^{+++}_{**}
	(0.048)	(0.129)	(0.219)	(0.256)	(0.299)	(0.275)	(0.245)	(0.245)	(0.244)	(0.249)
Farm price	1.056^{+++}_{**}	0.926^{+++}_{**}	0.769^{++}_{**}	0.551_{**}	0.323_{**}	0.1	-0.141	-0.363	-0.528	-0.533
	(0.132)	(0.233)	(0.209)	(0.216)	(0.227)	(0.186)	(0.185)	(0.209)	(0.229)	(0.187)
Yield	0.967^{+++}_{**}	0.908^{+++}_{**}	1.028^{+++}_{**}	1.17^{+++}_{**}	1.192^{+++}_{**}	1.155^{+++}_{**}	0.888^{+++}_{**}	1.128^{+++}_{**}	1.258^{+++}_{**}	1.234^{+++}
	(0.213)	(0.252)	(0.249)	(0.258)	(0.224)	(0.307)	(0.226)	(0.234)	(0.148)	(0.243)
Farm income										
Net cash income	0.919^{+++}_{**}	0.832^{+++}_{**}	0.781^{+++}_{**}	0.582^{+}_{**}	0.388_{**}	0.235_{**}	0.096_{*}	-0.089	-0.141	-0.235
	(0.110)	(0.146)	(0.214)	(0.240)	(0.174)	(0.140)	(0.205)	(0.207)	(0.238)	(0.112)
Crop receipts	1.038^{+++}_{**}	0.978^{+++}_{**}	0.849^{+++}_{**}	0.695^{++}_{**}	0.513_{**}	0.305_{**}	0.085	-0.091	-0.222	-0.327
	(0.084)	(0.124)	(0.119)	(0.114)	(0.059)	(0.051)	(0.052)	(0.013)	(0.026)	(0.056)
Livestock receipts	0.897^{+++}_{**}	0.832^{+++}_{**}	0.789^{+++}_{**}	0.671^{+++}_{**}	0.554^{++}_{**}	0.461_{**}	0.349_{**}	0.169_{**}	0.061	-0.068
	(0.105)	(0.162)	(0.214)	(0.224)	(0.205)	(0.173)	(0.174)	(0.175)	(0.192)	(0.233)
Govt. payments	0.675^{+++}_{**}	0.123_{**}	-0.16	-0.104	-0.461	-0.612	-0.56	-0.448	-0.462	-0.296
	(0.099)	(0.269)	(0.275)	(0.291)	(0.207)	(0.229)	(0.203)	(0.189)	(0.233)	(0.177)
Farm-related income	0.944^{+++}_{**}	0.884^{+++}_{**}	0.824^{+++}_{**}	0.893^{+++}_{**}	0.925^{+++}_{**}	1.054^{+++}_{**}	1.034^{+++}_{**}	1.056^{+++}_{**}	1.109	0.87^{++}
	(0.167)	(0.170)	(0.189)	(0.247)	(0.253)	(0.237)	(0.241)	(0.230)	(0.276)	(0.210)
Cash expenses	1.006^{+++}_{**}	0.949^{+++}_{**}	0.872^{+++}_{**}	0.784^{+++}_{**}	0.694^{++}_{**}	0.627^{+}_{**}	0.517_{**}	0.364_{**}	0.208	0.063
	(0.079)	(0.131)	(0.136)	(0.133)	(0.161)	(0.174)	(0.180)	(0.164)	(0.163)	(0.135)

Notes: *, **, and *** denote statistical significance at 10%, 5%, and 1% respectively for testing the null hypotheses $H_0 : \beta_h^{FAPRI} \leq 0$. Likewise, +, ++, and +++ denote statistical significance at 10%, 5%, and 1% respectively for testing the null hypothesis $H_0 : \beta_h^{FAPRI} \leq 0.5$.

Table A.3: Modified Diebold-Mariano (MDM) test comparing FAPRI and USDA projections

	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 5$	$h = 6$	$h = 7$	$h = 8$	$h = 9$
Corn										
Harvested acres	0.429	-1.135	-0.926	-0.043	0.976	0.275	-0.029	0.296	0.455	-0.796
Farm price	-0.318	-0.35	0.234	-0.092	-0.787	-1.948^{++}	-2.338^{++}	-2.028^{++}	-2.736^{+++}	-4.005^{+++}
Yield	0.034	-1.247	0.748	1.257	0.521	-1.342^{+}	-1.746^{++}	-1.541^{+}	-0.834	-2.639^{+++}
Soybean										
Harvested acres	-0.148	2.526***	2.043**	1.245	0.005	1.75**	0.949	1.799**	1.852**	2.233**
Farm price	-1.085	0.418	0.947	0.315	-0.537	-2.157^{++}	-3.715^{+++}	-3.956^{+++}	-4.684^{+++}	-5.946^{+++}
Yield	-2.254^{++}	-1.942^{++}	-1.254	-0.939	-0.486	-0.609	-0.905	-0.808	-1.269	-1.989^{++}
Wheat										
Harvested acres	1.377*	-1.126	-0.789	-0.583	-1.07	-1.411^{+}	-1.457^{+}	-0.03	0.915	1.805**
Farm price	0.008	1.566*	2.511***	2.413**	2.349**	2.389**	1.7*	0.683	0.017	-0.509
Yield	0.985	0.704	0.906	-0.05	-1.245	-0.359	0.841	1.959**	0.784	-0.203
Farm income										
Net cash income	1.514*	3.442***	3.516***	2.652***	1.998**	-0.739	-1.54^{+}	-1.722^{++}	-1.606^{+}	-2.432^{++}
Crop receipts	1.505*	2.064**	1.733**	1.333*	0.52	-0.88	-2.09^{++}	-3.064^{+++}	-3.796^{+++}	-6.265^{+++}
Livestock receipts	1.605*	1.338*	0.281	-0.17	-0.6	-2.393^{++}	-3.157^{+++}	-4.209^{+++}	-4.19^{+++}	-4.563^{+++}
Govt. payments	-0.961	2.149**	0.124	1.465*	1.141	-0.089	1.108	1.543*	1.488*	1.224
Farm-related income	0.812	1.393*	1.267	1.753**	2.707***	2.447**	2.243**	2.512***	2.534***	2.704***
Cash expenses	-0.865	0.156	0.342	-0.035	-0.17	-0.92	-1.764^{++}	-2.707^{+++}	-3.466^{+++}	-3.886^{+++}

Notes: The estimates are DM-statistic of Modified DM test comparing FAPRI and USDA projections. (***), (**) and (*) suggests that FAPRI projection performed better than USDA at 1%, 5% and 10% significance levels respectively. Similarly, (+++), (++) and (+) suggests that the USDA projection performed better than FAPRI at 1%, 5% and 10% significance levels respectively. Tests were conducted for each horizon separately. Root mean square error was used to calculate loss differentials.

Table A.4: Tests of uniform superior predictive ability (uSPA) comparing FAPRI and USDA projections

	$h=0$	$h\leq 1$	$h\leq 2$	$h\leq 3$	$h\leq 4$	$h\leq 5$	$h\leq 6$	$h\leq 7$	$h\leq 8$	$h\leq 9$
Corn										
Harvested acres	0.533	-1.676	-1.987	-2.008	-1.944	-1.626	-1.495	-1.417	-1.207	-1.18
	-0.533	-0.522	-0.286	-0.21	-1.149	-1.931	-1.775	-1.52	-1.44	-1.423
Farm price	-0.538	-0.448	-0.898	-0.895	-0.886	-1.499	-1.716	-1.972	-3.378	-4.21
	0.538	0.319	-0.183	-0.079	0.199^{+}	0.45^{+++}	0.436^{+++}	0.245^{++}	-0.127^{+}	-0.212^{++}
Yield	0.034	-1.534	-1.457	-1.238	-1.046	-1.266	-1.323	-1.359	-1.289	-1.726
	-0.034	-0.035	-0.667	-1.07	-1.183	-1.464	-1.242	-1.34	-1.777	-1.659
Soybean										
Harvested acres	-0.165	-0.165	-0.276	-0.29	-0.326^{*}	-0.285^{**}	-0.393^{**}	-0.5^{*}	-0.521^{**}	-0.576^{**}
	0.165	-3.114	-2.924	-2.541	-2.694	-2.807	-2.98	-3.073	-3.063	-3.292
Farm price	-1.165	-1.117	-0.608	-0.407	-0.433	-1.79	-2.666	-2.609	-3.09	-3.582
	1.165^{+}	-0.416	-1.324	-1.274	-1.202	-1.397	-0.975	-1.547	-1.494	-1.447
Yield	-1.672	-1.649	-1.724	-2.284	-2.544	-2.57	-2.558	-2.191	-2.223	-2.146
	1.672^{+}	1.534^{++}	0.912^{+}	0.824^{++}	0.391^{++}	0.44^{++}	0.646^{+++}	0.728^{+++}	0.728^{+++}	0.696^{+++}
Wheat										
Harvested acres	1.381	-1.058	-1.116	-1.477	-1.417	-1.228	-1.316	-1.204	-1.406	-1.707
	-1.381	-1.392	-1.466	-1.438	-1.451	-1.393	-0.892	-0.875	-0.782	-1.361
Farm price	0.011	0.011	-0.069	-0.095	-0.065	-0.061^{*}	-0.414	-0.31^{*}	-0.755	-0.965
	-0.011	-1.341	-1.789	-1.965	-1.868	-1.982	-1.873	-1.71	-1.609	-1.645
Yield	1.281^{*}	0.713^{**}	0.031	-0.195	-1.107	-1.283	-1.738	-1.205	-1.313	-1.376
	-1.281	-1.251	-0.898	-0.988	-1.112	-1.071	-0.819	-1.469	-1.189	-1.124
Farm income										
Net cash income	1.568^{*}	1.358^{**}	1.439^{***}	1.471^{***}	1.457^{***}	-0.735	-1.801	-1.914	-1.851	-1.459
	-1.568	-2.959	-2.804	-2.869	-2.854	-2.576	-2.539	-2.397	-2.13	-2.681
Crop receipts	1.31	1.165^{***}	1.046^{***}	0.91^{***}	0.353^{***}	-0.605	-1.45	-1.904	-2.291	-3.3
	-1.31	-1.531	-1.53	-1.297	-1.269	-1.425	-1.438	-1.699	-1.733	-1.715
Livestock receipts	1.44	1.092^{*}	0.23	-0.127	-0.435	-1.52	-2.089	-2.552	-2.493	-2.529
	-1.44	-1.445	-1.47	-1.515	-1.511	-1.463	-1.095	-1.019	-1.918	-2.597
Govt. payments	-1.018	-1.021	-1.112	-1.018	-0.808	-0.767	-0.237	-0.193^{*}	-0.191^{*}	-0.861
	1.018	-1.873	-1.887	-1.911	-2.041	-1.97	-1.798	-1.522	-1.707	-1.662
Farm-related income	0.873^{*}	0.904^{**}	0.912^{**}	0.86^{**}	0.921^{**}	0.98^{**}	0.989^{***}	0.951^{***}	0.956^{***}	1.019^{***}
	-0.873	-1.042	-1.152	-1.44	-2.605	-2.831	-3.103	-4.094	-5.036	-4.474
Cash expenses	-0.79	-0.502	-0.5	-0.322	-0.392	-0.597	-1.182	-1.671	-2.121	-2.232
	0.79	-0.136	-0.617	-0.76	-0.964	-0.827	-0.803	-0.821	-0.898	-0.717

Notes: The estimates are t-statistics for test of uSPA. The first estimate in each cell refers to t-statistic for the null hypothesis $H_{0,uSPA}$: $\mu^{uSPA} \leq 0$. (***), (**) and (*) indicate the rejection of the null hypothesis, suggesting that the FAPRI projection has uSPA over the USDA projection at 1%, 5% and 10% significance levels respectively. The second estimate is the t-statistic when we switched the projections in the null hypothesis. Similarly, (+++), (++) and (+) indicate that the USDA projection has uSPA over the FAPRI projection at 1%, 5% and 10% significance levels respectively. Horizon $\leq h$ means the tests are performed using projections up to horizon h. A square loss function was used to calculate loss differentials.

Table A.5: Tests of average superior predictive ability (aSPA) between FAPRI and USDA projections, equal weights

	$h=0$	$h\leq 1$	$h\leq 2$	$h\leq 3$	$h\leq 4$	$h\leq 5$	$h\leq 6$	$h\leq 7$	$h\leq 8$	$h\leq 9$
Corn										
Harvested acres	0.533	-0.636	-1.353	-1.162	-0.516	-0.195	0.064	0.127	0.316	0.132
Farm price	-0.538	-0.428	-0.202	-0.267	-0.629	-1.411	-1.96^{+}	-2.148	-3.047^{+}	-4.318^{++}
Yield	0.034	-0.716	-0.019	0.469	0.787	0.597	-0.142	-0.565	-0.214	-0.722
Soybean										
Harvested acres	-0.165	3.99***	3.173**	3.036**	3.036**	3.611**	2.459**	2.248**	2.374*	2.669*
Farm price	-1.165^{+}	-0.241	0.817	0.69	0.252	-0.453	-1.486	-2.166^{+}	-3.002^{++}	-4.067^{++}
Yield	-1.672^{+}	-1.632^{+}	-1.513^{+}	-1.862^{++}	-2.17^{++}	-2.363^{++}	-2.506^{++}	-2.33^{+}	-2.336^{+}	-2.35^{+}
Wheat										
Harvested acres	1.381*	0.486	-0.416	-0.603	-0.869	-0.901	-1.015	-0.859	-0.648	-0.482
Farm price	0.011	1.185	1.638*	1.79*	1.767*	1.775*	1.832*	1.807*	1.564	1.293
Yield	1.281*	1.331*	0.793	0.624	0.368	0.206	-0.422	0.162	0.308	0.234
Farm income										
Net cash income	1.568*	2.706**	2.934**	2.88**	2.997**	2.53**	2.028*	1.264	0.744	0.393
Crop receipts	1.31*	2.01**	1.639**	1.312*	0.937	0.45	-0.151	-0.73	-1.322	-2.093
Livestock receipts	1.44	1.401	0.898	0.551	0.25	-0.418	-1.133	-1.764	-2.082^{+}	-2.175
Govt. payments	-1.018	1.113*	0.848	1.298	1.423	1.408*	1.495*	1.426	1.363	0.823
Farm-related income	0.873*	1.008*	1.094*	1.271*	1.532*	1.791**	2.166**	2.607**	2.817**	3.056**
Cash expenses	-0.79	-0.158	0.278	0.273	0.212	-0.001	-0.352	-0.772	-1.21	-1.398

Notes: The estimates are t-statistics for the null hypothesis $H_0 : \mu^{aSPA} \leq 0$. (***), (**) and (*) indicate the rejection of the null hypothesis, suggesting that the FAPRI projection has aSPA over the USDA projection at 1%, 5% and 10% significance levels respectively. Similarly, (+++), (++) and (+) indicate that the USDA projection has aSPA over the FAPRI projection at 1%, 5% and 10% significance levels respectively. Horizon $\leq h$ means the tests are performed using projections up to horizon h. Each horizon was given equal weights. A square loss function was used to calculate loss differentials.

Table A.6: Tests of average superior predictive ability (aSPA) between FAPRI and USDA projections, variance weights

	$h=0$	$h\leq 1$	$h\leq 2$	$h\leq 3$	$h\leq 4$	$h\leq 5$	$h\leq 6$	$h\leq 7$	$h\leq 8$	$h\leq 9$
Corn										
Harvested acres	0.533	-0.571	-1.177	-1.043	-0.611	-0.467	-0.131	0.115	0.28	0.082
Farm price	-0.538	-0.387	-0.024	-0.11	-0.565	-1.486	-2.012	-2.266	-3.119^{+}	-4.046^{++}
Yield	0.034	-0.326	0.1	0.455	0.697	0.449	-0.121	-0.92	-0.614	-0.767
Soybean										
Harvested acres	-0.165	3.355**	2.896**	2.856**	2.959**	2.928**	1.712**	1.708*	1.919*	2.046*
Farm price	-1.165^{+}	-0.02	1.186	0.76	0.046	-0.797	-1.931^{+}	-2.585^{++}	-3.25^{++}	-4.113^{++}
Yield	-1.672^{+}	-1.616^{+}	-1.544^{+}	-1.854^{+}	-2.16^{++}	-2.351^{++}	-2.619^{++}	-2.21^{++}	-2.211^{++}	-2.213^{+}
Wheat										
Harvested acres	1.381	0.825	-0.657	-0.58	-0.824	-0.871	-0.985	-0.725	-0.263	0.007
Farm price	0.011	1.3*	1.721*	1.864*	1.836**	1.75*	1.805*	1.805*	1.429	1.134
Yield	1.281*	1.312*	0.823	0.739	0.661	0.639	-0.053	0.597	0.599	0.553
Farm income										
Net cash income	1.568*	2.887**	2.822**	2.601**	2.687**	2.412**	1.871*	0.825	0.146	-0.503
Crop receipts	1.31*	1.663*	1.351*	1.088*	0.682	0.123	-0.67	-1.314	-1.914^{+}	-2.518^{+}
Livestock receipts	1.44	1.361	0.599	0.235	-0.089	-0.961	-1.783^{+}	-2.272^{+}	-2.414^{+}	-2.421^{+}
Govt. payments	-1.018	1.617**	0.951	1.317	1.341*	1.388*	1.483*	1.471*	1.406	0.855
Farm-related income	0.873	0.985*	1.066*	1.154*	1.323**	1.516**	1.83**	2.224**	2.377**	2.629**
Cash expenses	-0.79	0.008	0.288	0.168	0.041	-0.238	-0.701	-1.116	-1.574	-1.704

Notes: The estimates are t-statistics for the null hypothesis $H_0 : \mu^{aSPA} \leq 0$. (***), (**) and (*) indicate the rejection of the null hypothesis, suggesting that the FAPRI projection has aSPA over the USDA projection at 1%, 5% and 10% significance levels respectively. Similarly, (+++), (++) and (+) indicate that the USDA projection has aSPA over the FAPRI projection at 1%, 5% and 10% significance levels respectively. Horizon $\leq h$ means the tests are performed using projections up to horizon h. Each horizon was given weights equal to the ratio of the variance of loss differential for the horizon to the sum of variances of loss differentials across all horizons. A square loss function was used to calculate loss differentials.

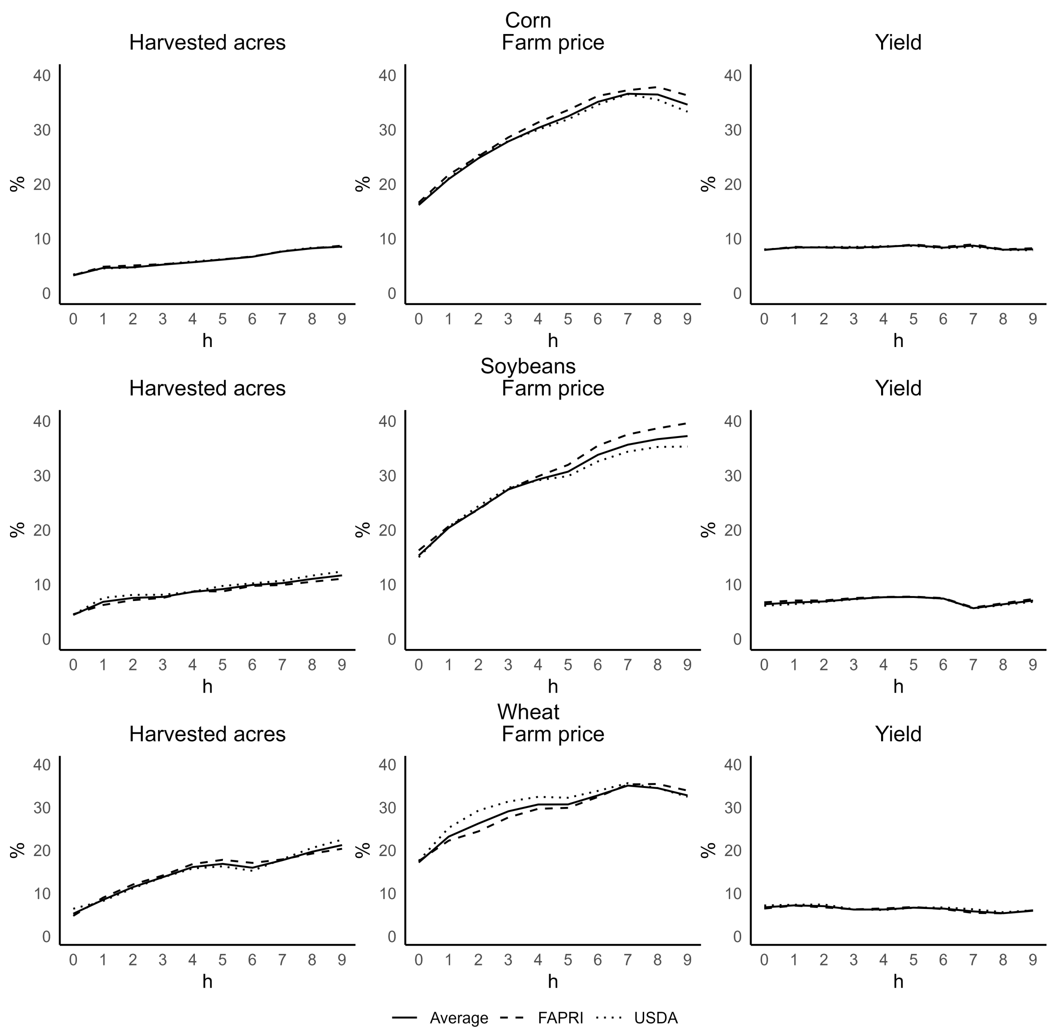


Figure A.3: RMSPE of a simple average of USDA and FAPRI baseline projections of corn, soybeans and wheat by projection horizon h, 1997–2020

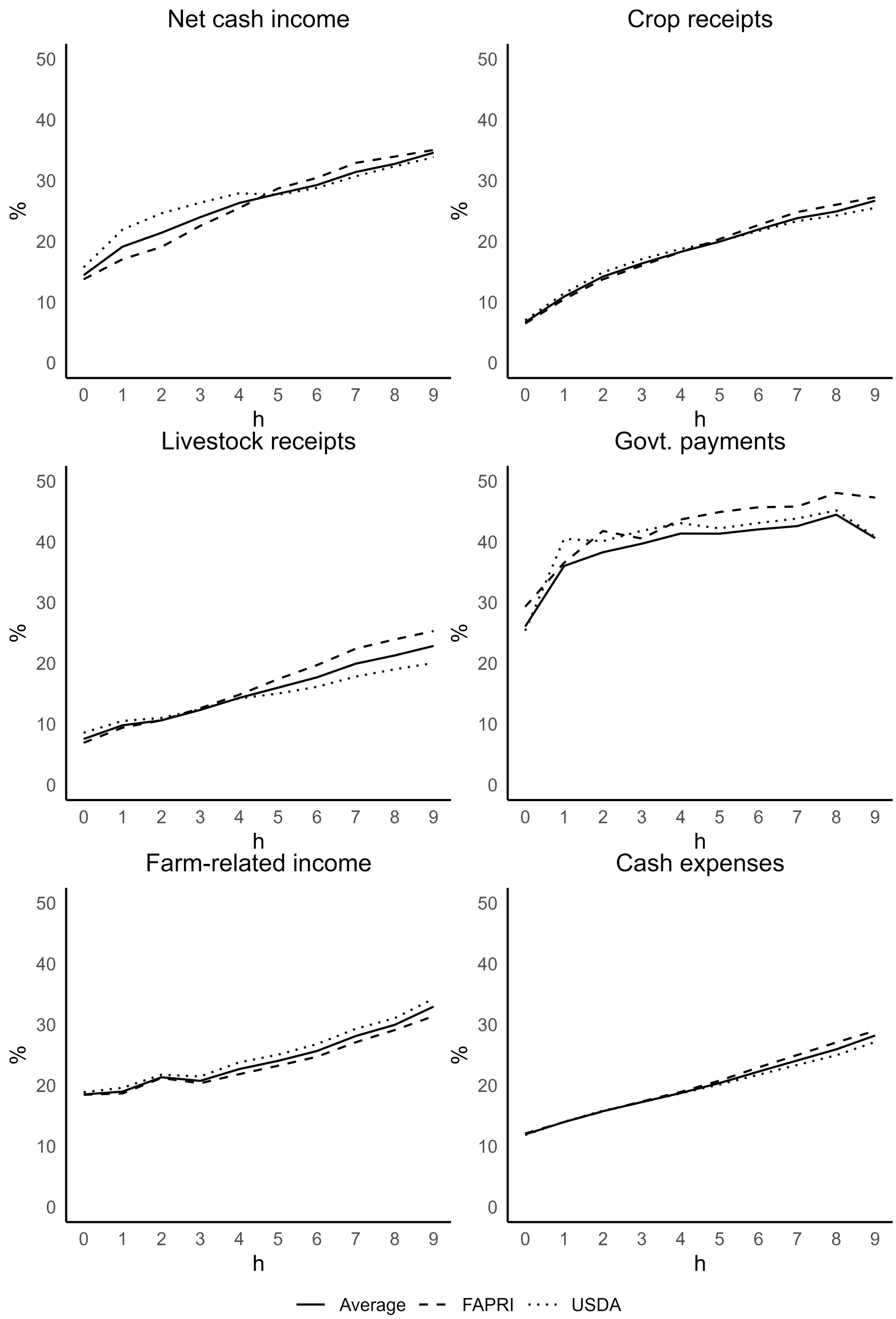


Figure A.4: RMSPE of a simple average of USDA and FAPRI baseline projections of net cash income components by projection horizon h, 1997–2020