User Responses to Imperfect Forecasts: Findings from an Experiment with Kentucky Wheat Farmers

YOKO KUSUNOSE

Department of Agricultural Economics, University of Kentucky, Lexington, Kentucky

LALA MA

Department of Economics, University of Kentucky, Lexington, Kentucky

DAVID VAN SANFORD

Department of Plant and Soil Sciences, University of Kentucky, Lexington, Kentucky

(Manuscript received 10 December 2018, in final form 21 June 2019)

ABSTRACT

Weather and climate forecasts can help agricultural producers improve management choices in anticipation of uncertain growing conditions. Current literature conjectures that the extent to which forecasts are useful depends on their accuracy, that is, the probability with which a forecasted event, such as precipitation, is projected to occur. Too little accuracy can potentially render forecasts effectively useless, even if they convey some form of information. In this study, we collect farmer-based data through a questionnaire and a framed field experiment to test for the existence of an accuracy threshold for forecasts, below which forecasts do not induce any behavioral changes. We do this in the context of a very specific management choice—the timing and amount of nitrogen that Kentucky farmers apply to their wheat in early spring—in response to randomly generated 6-to-10-day forecasts of rainfall conditions. We find that forecasts provide economically significant value to decision-makers only when they depart dramatically from what is normally expected. These results have implications that extend beyond the nitrogen-application decision for winter wheat: if this type of behavior is widespread, at current accuracy levels, other types of forecasts may be of little value to decision-makers and therefore go unheeded.

1. Introduction

Weather and climate forecasts are viewed as holding great potential in aiding agricultural producers to increase profits and reduce weather-related losses. Literature is concerned with the estimation of this potential, typically using either the cost-loss model or the decision-analytic framework pioneered by Nelson and Winter (1964) and Hilton (1981). In the former, agents can decide whether to engage in costly measures of protection against the potential losses from an uncertain, adverse event (e.g., a drought). With the latter, agricultural managers receive forecasts, update their weather or climate expectations, and choose the utility-maximizing management choice(s) based on these updated expectations. In both approaches, the value of a forecast source is defined as the difference in the revenue

stream resulting from the manager heeding these forecasts compared to that which would result in their absence. Most such studies focus on seasonal climate forecasts, particularly those related to El Niño–Southern Oscillation [see Meza et al. (2008) for a recent review].

No forecast is perfect, meaning that even if the forecast itself is correct, the predicted event takes place with less than 100% certainty. Moreover, the estimated values of forecasts from the literature vary widely based on modeling assumptions about locales, crops, relevant management decisions, manager characteristics (e.g., objectives, risk preferences, and management constraints), forecast events, and forecast characteristics (e.g., lead time, format, and accuracy) (Meza et al. 2008). However, as long as the forecasts are correct and impart some information, the end user should never be worse off as a result of having consulted them. Despite this, the reported use of forecasts appears to be low, with a recent review finding user perceptions of low accuracy to be the most commonly cited

Corresponding author: Yoko Kusunose, yoko.kusunose@uky.edu

DOI: 10.1175/WCAS-D-18-0135.1

© 2019 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

barrier in the United States and Australia (Mase and Prokopy 2014).

The precise meaning of “accuracy” depends on the type of forecast. The analysis we present here uses probability forecasts, and we use the term accuracy to mean the probability with which a particular event is projected to occur. For example, for the forecast, “It will be an El Niño year with 60 percent probability,” the accuracy is 60%, assuming no bias or uncertainty in the forecast itself. The forecast is imperfect because the accuracy is not 100%. Another aspect of a forecast is “relative skill,” which measures the amount of *novel* information conveyed in a forecast. Using the example above, if the climatological likelihood of an El Niño event is 30%, then the 60-percent forecast is meaningful; however, if the climatological likelihood were 58%, the 60-percent forecast would present relatively little new information. For the purposes of this paper, if the accuracy of a forecast is high enough that it is able to provide some information beyond the baseline expectation (e.g., climatology), then we refer to that forecast as being (relatively) skilled.¹

Hu et al. (2006) address the relationship between forecast accuracy and use among U.S. row crop growers and conclude that farmer perceptions of accuracy and their valuation of these perceptions matter as much as accuracy itself. It is plausible, however, that managers find forecasts not worth consulting, even when they *correctly* perceive the accuracy of those forecasts. One obvious reason is because even a “correct” *ex ante* response can be the “wrong” choice *ex post*, at least some of the time. Another, less obvious, reason is that when forecasts have low accuracy, managers are less likely to take action in response to them; but when no action is taken, that forecast has zero value (Kusunose and Mahmood 2016).

Despite the accuracy improvements seen in recent decades, many types of forecasts—particularly those with longer lead times—are limited in their accuracy and may experience only marginal future improvement (O’Lenic et al. 2008; Bauer et al. 2015; Hoskins 2013; Westra and Sharma 2010). This could present an even greater reason for concern for forecast developers, given the notion from the field of economics that even a skilled forecast can be of zero value to a user (e.g., Radner and Stiglitz (1984)). Joseph Stiglitz, in his 2001 Nobel Prize lecture, famously stated, “. . . it never paid to buy just a little bit of information.” While this may be the case, the quantification of

“too little” or a more precise evaluation of this notion would be important among developers of forecast-based agricultural decision support tools. Indeed, Moeller et al. (2008), using a simulation model of a crop production system in Western Australia, show that low skill—which is related to low accuracy—can cause forecasts to have zero value. However, this and other simulation models effectively assume managers as making “rational” management choices, which may or may not be the case.

In this study, we use farmer-based data to test for the existence of an accuracy threshold for forecasts, below which forecasts do not induce any behavioral changes. We specifically examine 1) whether there exists an accuracy threshold below which managers appear not to act upon forecasts, 2) which types of events elicit changes in management, and 3) the relative importance of forecasts, compared to that of heuristics. We do this in the context of a very specific management choice—the timing and amount of nitrogen that Kentucky farmers apply to their wheat in early spring—that we presume farmers make in consultation with 6-to-10-day precipitation forecasts. We use a questionnaire to ask wheat producers about their forecast sources generally and typical spring nitrogen application practices, and then run a framed field experiment in which the same producers state their optimal nitrogen application rates (and the timing of the applications) in response to a series of hypothetical 6-to-10-day precipitation forecasts. These experimentally generated data are then analyzed to detect the existence of accuracy thresholds.

This study fits into the relatively small literature that examines user responses to weather forecast information using an experimental setting. This experimental literature is typically concerned with user responses to the format of forecasts (e.g., Roulston et al. 2006; Roulston and Kaplan 2009; Joslyn et al. 2009; Nadav-Greenberg et al. 2008). Our research question is closest to that of Morss et al. (2010), who use a nationwide U.S. survey to elicit the actions that respondents would take in response to forecasts of rain or low temperatures. For rain forecasts, the survey asked participants to choose whether to move a picnic indoors and whether to proactively release water from a reservoir. In the case of forecasts for low temperature, respondents chose whether to cover their garden plants and, in the second question, protect their orchard against frost. They, too, are interested in identifying accuracy thresholds at which point users will take actions in response to probability forecasts. As will be discussed in detail below, our field experiment and findings bring unique and valuable evidence to this question. For example, we present probability forecasts in a way that curtails any biases from uncertainty, (incorrect) *perceptions* of

¹ This is, of course, a simplification. More precisely, the relative skill of a probabilistic forecast can be measured using the Brier skill score (BSS), with BSS = 0 indicating no skill and BSS = 1 indicating perfect skill. What we call a skilled forecast is one that would have a BSS value that is greater than zero.

accuracy, and/or prior beliefs that users may have; all of our respondents are relatively knowledgeable and interested in precipitation forecasts; and we present these respondents with a highly salient scenario, which means that the stated answers of the respondents are more likely to represent their true actions accurately.

The organization of this paper is as follows. [Section 2](#) gives a brief background of winter wheat farming in Kentucky, provides the setup of the survey and the field experiment, and describes the regression model used to analyze the collected data. The results of the survey, experiment, and regression analysis are then presented in [section 3](#). Finally, [section 4](#) concludes with a discussion of the results and the limitations of our work.

2. Methods

This study uses a framed field experiment to examine farm management decisions relating to the timing and application rates of nitrogen for winter wheat. In the experiment, or “game,” we ask growers to state their optimal spring nitrogen application (or top dressing) decisions in response to a series of hypothetical precipitation forecasts of varying accuracy. The experiment is accompanied by a questionnaire that gathers additional information about the backgrounds of participants. This section begins by describing the context and background of our study, followed by a description of the survey and field experiment.

a. Area and context

In Kentucky, winter wheat is typically grown as part of a three-crop rotation (corn/wheat/double-crop soybeans). Winter wheat is harvested in June, providing an important source of cash in the summer months; it can aid in weed suppression, and provide valuable ground cover to prevent soil erosion. With wheat production, nitrogen is the nutrient requiring the most management, and the most critical application window is in late winter/early spring, just as the wheat crop begins to uptake nitrogen rapidly ([Lee et al. 2009](#)). Nitrogen is generally applied either through 1) two applications (in late February and in early March—this is called a “split” application) or 2) one application (typically in early March—called a “single” application). In either case, amounts are typically adjusted depending on soil and rainfall conditions ([Lee et al. 2009](#)).

Most relevant to our study, winter wheat is a rainfed crop, and the optimal amount of rainfall during the application period is crucial. Insufficient rainfall prevents applied nitrogen from reaching the root zone. However, too much rainfall causes crop yields to suffer because wheat is generally intolerant of poor drainage, nitrogen losses are higher, and farmers cannot access fields with

their application equipment ([Lee et al. 2009](#)). Thus both the timing and the amount of rainfall must be within a particular range for a manager to receive a return on his fertilizer application. Of course, managers cannot control the rainfall that they receive. They do, however, have some flexibility in terms of when and how much fertilizer is applied during this three-week window. Crucial to our research design, these decisions require some advance planning, as arrangements for machinery and even fertilizer purchases must be made ahead of time. Thus, producers are likely to rely on expectations of precipitation to decide how much (if any) fertilizer to apply.

The Ohio Valley² happens to be where the warmer, moisture-laden air from the Gulf of Mexico and continental polar air meet [R. Mahmood 2019, personal communication; M. Dixon 2019, personal communication (Dixon is an extension meteorologist for the University of Kentucky; Mahmood is an applied climatologist)]. Western Kentucky (at the western edge of the Ohio Valley) is where most of Kentucky’s wheat is grown. Here, the average monthly precipitation is 4.15 in., with the most rainfall being concentrated in April and May (which average 5.28 and 5.27 in., respectively) and the months with the least precipitation occurring in January (3.89 in.), July (3.96 in.), and November (3.95 in.).³ The precipitation averages for February and March, which comprise our window of interest, are 3.51 and 4.55 in., respectively. Early spring is a transition period, when the circulation patterns of winter give way to the patterns of summer. Indeed, the coefficient of variation (across years) is high in March (0.50) and then again in August, September, October, and November. In western Kentucky, precipitation is generally abundant year-round, but it is particularly difficult to predict at these times. The Climate Prediction Center’s (CPC) 6-to-10-day precipitation outlooks hardly ever show above- (or below-) average rainfall probabilities that exceed 50% at any time of the year (M. Dixon 2019, personal communication), much less in early spring.

b. Survey and experiment

The data for this study were collected in a one-hour session that was part of the 2015 Annual Winter Wheat

² This is the Ohio River valley drainage basin, which includes nearly all of Kentucky and West Virginia, as well as parts of Tennessee, Illinois, Ohio, and even Pennsylvania. It is considered a meaningful geographic unit in that it is a low area that is topographically homogeneous, relatively speaking.

³ These and all other monthly statistics in the remainder of this paragraph have been calculated by the authors using a 20-yr (1994–2014) monthly data series available on the Kentucky Climate Center website (<http://kyclimate.org/data.html>), which is maintained by Western Kentucky University.

Meeting in Hopkinsville, Kentucky, a conference hosted by the University of Kentucky and sponsored, in part, by the Kentucky Small Grain Growers Association. It is a day-long event in which researchers and extension specialists from various disciplinary fields and multiple Kentucky universities hold educational sessions. This conference is widely publicized and highly attended, with attendees comprising mostly producers and crop advisers, the latter being able to earn continuing education credits for attending. All attendees in the auditorium were invited to participate in the session. The session began with a 10-min presentation explaining the experiment, followed by 15 min to complete a questionnaire, and the experiment itself, which took 25 min. Participants responded to the questionnaire and experiment with paper and pens provided by the researchers. All responses were then manually digitized for analysis.

The questionnaire contained a total of 16 questions that assessed the background, experience, and management strategies of participants with respect to top-dressing winter wheat. For example, it inquired about the participants' sources of weather forecast information and their yield expectations. These were assessed through multiple choice questions such as "Which of these weather forecast products do you regularly consult? (Check all that apply.)" and open response questions like "On your farm, what is a 'very good' (good, typical, mediocre, bad) yield (farm average in bushels/acre)?" Expectations for rainfall were gauged by presenting participants with a picture of a rain gauge marked with inches and asking participants to shade in the ranges of rainfall that they consider to be "drought," "normal" rates, and "waterlogging" rates. The survey also asked participants to choose nitrogen application rates and timing under various weather conditions, assuming *those conditions were known with certainty*. Specifically, the questionnaire asked "Given each of these March conditions, what is your best nitrogen application strategy? (For each of the five conditions below, first check either 'single' or 'split,' then indicate the rate(s) that would apply under these conditions.)" The full questionnaire is included in the [appendix](#), with completion rates noted next to each question. These questions complement the results from the experiment (described next), as they provide more context about the participants and potentially the determinants of their sensitivity to forecast accuracy.

At the start of the experiment, we explained the rules of the "game" and discussed the hypothetical scenario to which farmers would be asked to repeatedly respond, or the framing. We asked farmers to imagine their wheat fields in late February, given their local soil type (which managers stated in the questionnaire), moderate-to-dry

soil conditions, medium-to-low yields on the preceding corn crop, a typical nitrogen price (also stated), and an "average" season so far in terms of all other growing conditions. Also as part of the framing, we asked farmers to reflect on their goals when managing their crops. The example goals we discussed in the explanatory session were achieving the highest yields possible, maximizing profit, earning enough to cover all cash costs, and clearing a particular per-acre profit margin.

The experiment comprised participants playing 12 rounds, or "seasons," within each of which players decide whether and how much nitrogen to apply at two junctures—late February and early March—based on precipitation forecasts for those two periods. Each round followed this sequence: 1) February forecast is announced and displayed. 2) Participant chooses February application rate. 3) February rainfall is revealed. The same three steps are repeated for March, after which participants self-evaluate their harvests. Farmers noted on a worksheet these harvest ratings, along with their application choices and the rainfall outcomes. We next provide additional details of this process.

Forecasts for both February and March mimicked the CPC's 6–10-day precipitation outlooks, where information was presented in a three-category format—"drier than normal," "wetter than normal," and "normal"—with the first two categories having a probability associated with it (e.g., "90% chance wetter than normal"). In addition to announcing these forecasts orally and visually, we displayed them on large (one meter in diameter) paper pie charts. On these pie charts, the three-category forecasts were broken down into the relative probabilities of five rainfall scenarios—very dry (VD), fairly dry (FD), normal (NO), fairly wet (FW), and very wet (VW)—with the goal of mitigating the likelihood that participants would interpret the three-category forecasts in different ways. We made sure that everybody could clearly see the pie chart before affixing it to the prize wheel, which we then turned. We are cognizant of the possibility that presenting each forecast using both formats might have caused participants to focus on different aspects of the forecasts.

It is important to note that we did not describe the five rainfall conditions using inches of rain; rather, we made clear that "very wet" connoted waterlogging conditions and that "very dry" connoted drought conditions. We did this under the assumption that the same conditions could be triggered at different rainfall rates depending on the soil types and even wheat varieties on participants' farms. Upon subsequent analysis of the questionnaire data, we would learn that the average respondent would consider a rainfall rate lower than 1.5 in. per month to be "very dry" and rainfall exceeding 6.7 in. per month to be "very wet."

TABLE 1. Forecasts used in experiment (source: author). This table presents the set of all forecast probability distributions from which we chose the specific forecasts for the experiment.

Forecast		% probability of				
		Very dry (VD)	Fairly dry (FD)	Normal (NO)	Fairly wet (FW)	Very wet (VW)
100% drier	(D100)	60	40	0	0	0
90% drier	(D90)	50	40	10	0	0
80% drier	(D80)	40	40	20	0	0
70% drier	(D70)	40	30	30	0	0
60% drier	(D60)	30	30	30	10	0
50% drier	(D50)	20	30	30	20	0
Normal	(NO)	10	20	40	20	10
50% wetter	(W50)	0	20	30	30	20
60% wetter	(W60)	0	10	30	30	30
70% wetter	(W70)	0	0	30	30	40
80% wetter	(W80)	0	0	20	40	40
90% wetter	(W90)	0	0	10	40	50
100% wetter	(W100)	0	0	0	40	60

The forecast probability distributions for the experiment were chosen from the set of distributions in Table 1. As an example, the three-category format of “90% chance wetter than normal” (the row labeled “90% wetter” in Table 1) has a probability distribution over rainfall categories of very dry, fairly dry, normal, fairly wet, and very wet that is respectively 0, 0, 10, 40, and 50. The February and March forecasts in each round are shown in Table 2.

While the distribution and sequence of *forecasts* (and pie charts) were determined beforehand, the realized outcomes (also shown in Table 2) themselves were random and determined by spinning the prize wheel. We followed Hayman et al. (2007) in using the pie charts and prize wheel to depict the forecasts. While their use may seem incidental to the game’s setup, we believe that they were integral in helping our participants understand 1) the nature of probability forecasts and 2) that the probability distribution itself—if not the outcome—is certain. With this format, we could reasonably assume that responses were minimally influenced by participant priors or uncertainty of the forecasts themselves. At both the February and March junctures, precipitation outcomes were revealed only after participants made their fertilizer decisions.

At the end of each season, participants self-evaluated their harvest on a five-point scale (by circling one: bad, mediocre, typical, good, or very good), based on the two rainfall outcomes received and the two application choices that they made. The self-evaluation component was included mainly for the purposes of improving the comprehension of the game, providing performance feedback, and maintaining participant interest. The player-specific definitions of these harvest categories were captured in the questionnaire.

The game began with three practice rounds, or seasons. Comprehension was checked after each of these practice rounds. Participants then played 12 “real” seasons. Because of the live format of the game, all participants saw the same sequence of forecasts and outcomes. To maintain the interest level for 12 seasons, every four seasons we asked players to volunteer their strategies behind their best and worst seasons. To those who did well and were intrepid enough to share their results with the group, we liberally handed out 1/87-scale model tractors; to those whose choices were, in hindsight, unlucky, we handed bags of peanuts. The participants sharing their outcomes reported which season was best (worst) and why, such as, “Round 6 was my best

TABLE 2. Experiment forecasts and outcomes, by round (source: author and experiment). This table presents the February and March forecasts shown to the participants, along with the outcomes indicated by the prize wheel. The blank lines indicate the breaks where participants shared their experiences with the group.

Round	February		March	
	Forecast	Outcome	Forecast	Outcome
1	D80	VD	NO	VD
2	W80	VW	W70	FW
3	D90	VD	D90	VD
4	W60	VW	D60	VD
5	W70	NO	W90	FW
6	NO	VW	D80	VD
7	D60	FD	W60	NO
8	D50	NO	D50	FW
9	W90	FW	W90	FW
10	D70	VD	D70	FD
11	W50	FW	W50	VW
12	NO	VD	W80	VW

season: I put down 50 pounds and then had normal rain, and then put 60 pounds and got fairly wet conditions.”

c. Regression model

We follow the survey and experiment by taking our collected data to a regression model in order to test for systematic behavioral responses to forecasts. Since we observe multiple choices per participant, observations in our regression model are at the participant-round level.⁴ We test for behavioral responses using a two-step approach. First, we estimate the following model of nitrogen application rate in pounds per acre ($Nfeb_{i,r}$) at the February juncture for participant i in round r :

$$Nfeb_{i,r} = \delta_0 + \delta_1 r + \delta_2 r^2 + \delta_3 \text{harvest}_{i,r-1} + u_{i,r}, \quad (1)$$

where r denotes the round in the game, $\text{harvest}_{i,r-1}$ is participant i 's subjective harvested amount in the last round (evaluated on a five-point scale), and $u_{i,r}$ represents any remaining factors that contribute to the February application rate that are not already explained by round number and previous yield performance. Second, we use the estimated coefficients from Eq. (1) to recover the residuals from the regression, $\hat{u}_{i,r}$, and calculate the *conditional* mean and variance of these residuals at various forecasted probabilities (10%, 20%, 30%, 40%, and 50%) of being “very wet” or “very dry.”

The inclusion of a quadratic term in r in the regression model controls the potential for nonlinear learning effects to occur between rounds. Any dynamic dependence in behavior would also be captured by $\text{harvest}_{i,r-1}$ if participants' nitrogen application decisions are driven by previous harvest outcomes. These controls ensure that the measured responses in nitrogen application to forecast probabilities are absent any learning or dependence between the rounds of the game. In addition, we augment the model to include a fixed effect for the respondent so that the regression analysis exploits only the within-person variation in behavior. This removes *time-invariant*, player-specific effects on the nitrogen application decision that are unobserved to us as researchers (e.g., years of experience growing wheat).

Assessment of the residuals in the second step then allows us to check whether individuals are systematically responding to forecasts. If participants indeed respond to forecasts in the same manner, then the *average* of the residuals from model 1 would reflect this. However, if participants are more responsive, but not necessarily in the same “direction,” then it would be more useful to

examine the *variance* of the responses (and therefore of the residuals) as the forecast probability for very wet (or dry) rainfall conditions increases. If, at some level of accuracy, the variance of residuals increases dramatically, this could suggest an accuracy threshold. With extreme weather conditions being the most threatening to winter wheat yields and thus the most salient to growers, we focus on whether the responses and residuals exhibit an accuracy threshold with respect to the forecasts for very wet (very dry) conditions in our analysis, the percentages of which are noted under the column titled “Very wet” (“Very dry”) in Table 1.

We repeat this procedure with our experimental outcomes at the March juncture, except we slightly modify the nitrogen application model in the first step to account for dependence on February application decisions ($Nfeb_{i,r}$) and rainfall ($Rfeb_r$):

$$\begin{aligned} Nmar_{i,r} = & \phi_0 + \phi_1 r + \phi_2 r^2 + \phi_3 \text{harvest}_{i,r-1} \\ & + \phi_4 Rfeb_r + \phi_5 Rfeb_r^2 + \phi_6 Nfeb_{i,r} \\ & + \phi_7 Nfeb_{i,r}^2 + e_{i,r}. \end{aligned} \quad (2)$$

We then estimate the mean and variance of the residuals from the March application model, $\hat{e}_{i,r}$, using the estimated coefficients in Eq. (2).

3. Results

a. Survey

There were approximately 150 participants present in the conference auditorium where the survey and experiment took place. A total of 69 participants began the questionnaire and 47 answered at least 90% of the questions; 52 people completed the game, where completion is defined as finishing all 12 rounds of the experiment. Table 3 presents the composition of respondents. The percentage of those reporting that they make the decisions as farmers—whether exclusively or in combination with farm consulting—is between one-third and 40%, depending on the activity (questionnaire or experiment). Consultants were just as prevalent as farmers, which was not surprising, given that consultants could earn continuing education credits for attending various sessions of the conference. The next most prevalent category is that of participants identifying as “other.” Participants identifying as “other” were asked to state their alternative capacity. These answers included researchers, fertilizer dealers/applicators, and a crop monitor. Finally, as Table 3 shows, a sizable portion of participants chose not to answer this question.

In general, the response rates for questions listed at the beginning of the questionnaire are higher than for

⁴ For example, if 5 participants completed all 12 rounds of the game, then there would be 60 observations.

TABLE 3. Composition of participants (source: survey). Column 1 shows the composition of participants who initiated the questionnaire, whereas column 2 describes those who completed the questionnaire. Column 3 describes those who participated in all 12 rounds of the game. N is the total number of participants in a column. The category Other includes researchers, industry representatives, crop monitors, and those involved in nonprofits. The most prevalent participant types were farmer and/or consultant. Not all participants who initiated the survey completed it, and of those samples the share that did not respond ranged between 4% and 13%.

Participants	Initiated survey		Completed survey		Completed game	
	Count ($N = 69$)	(%)	Count ($N = 47$)	(%)	Count ($N = 52$)	(%)
Farmer (exclusively)	19	(27.5)	15	(31.9)	14	(26.9)
Farmer + consultant	8	(11.6)	5	(10.6)	4	(7.7)
Consultant (exclusively)	18	(26.1)	17	(36.2)	15	(28.8)
Agent (exclusively)	6	(8.7)	3	(6.4)	5	(9.6)
Other	9	(13.0)	5	(10.6)	8	(15.3)
Not reported	9	(13.0)	2	(4.2)	6	(11.5)

those at the end. This could be because the questionnaire opened with questions that are easier to answer, and reserved the more cognitively taxing and/or sensitive questions for the end. We do not think that participants were time-constrained. Because very few respondents answered every single question in the questionnaire, summary statistics are taken over all participants who responded with an answer to the specific question. As such, different sets of statistics are based on slightly different subsets of the participants. (Recall that the response rates are noted next to each question, in the [appendix](#).)

Table 4 shows the rates at which participants responded in the affirmative to the eight listed sources of weather forecasts, as well as to “other” forecasts, which farmers then specified. The most consulted source is the local television news, which typically provides a five-day forecast. While participants may appear to consult shorter-term forecast products far more than longer-term products, the observed pattern could simply reflect a preference among forecast *sources*, rather than particular forecast *types*. We found other patterns in these responses, which are obscured by these summary statistics: Twenty-three percent of participants reported regularly relying on just one (listed) source. Invariably, these participants consulted seven-day daily forecast products. The median number of weather forecast sources is two; 38% consult exactly two of the listed sources. These two-source participants typically consult two different seven-day forecast products, although a handful in this category consult the 45-day Accuweather forecast and a seven-day forecast. Those consulting the month-ahead and three-month ahead forecasts are those who consult four and five different sources.

In Table 5, we summarize the “best” nitrogen application strategies stated (in response to five March conditions, if conditions were known with certainty). These responses are categorized into four general types: 1) those who always split their application but adjust the

total amount applied based on the condition; 2) those who switch their application timing (split vs. single) but apply the same total nitrogen; 3) those who switch their timing and adjust the total applied; and 4) those who always split their application and always apply the same amount of total nitrogen. Given that the goal of the experimental portion is to detect the effect of forecasts and their accuracy on nitrogen application strategies, it was important to know if and how farmers would modify

TABLE 4. Weather forecast products regularly consulted (source: survey). This table presents sources of weather forecast that respondents selected. Respondents were asked to list their source if their forecast tool was not listed in the questionnaire. These listed sources are categorized under Other in the table. Because these are respondent-specified sources, their percentages are likely to understate the actual use. Generally, the most consulted source was the local television news, although the median participant consulted two sources.

Weather product	Total = 69	(%)
Weather forecast on local television news	49	(71.0)
The Weather Channel, 7-day daily forecast	39	(56.5)
National Weather Service, 7-day daily forecast	30	(43.3)
Accuweather 45-day forecast	16	(23.2)
Climate Prediction Center, 6-to-8-day-ahead forecast	5	(7.2)
Climate Prediction Center, 8-to-14-day-ahead forecast	2	(2.9)
Climate Prediction Center one-month-ahead forecast	3	(4.4)
Climate Prediction Center three-month-ahead forecast	1	(1.4)
Other (specified by respondent):		
Intellicast 7-to-10-day forecast	5	(7.2)
DTN Weather, unspecified product	2	(2.9)
Weather Underground, unspecified product	2	(2.9)
Accuweather 7-day forecast	2	(2.9)
Smartphone app. (e.g., AeroWX, MyCast, Weather Bug)	3	(4.3)

TABLE 5. Optimal application rates, by producer type and rainfall (source: survey). The rainfall conditions are for March. The responses were elicited using the question, “Given each of these March conditions, what is your best nitrogen application strategy?” Each row represents a participant’s response, where we have categorized producers into four “types.” A cell with a bolded number indicates that the manager’s optimal response under a particular condition would be to do a single application (apply once, in early March). The nonbolded numbers indicate split applications (divide the total among two applications, in late February and early March); the number in any of these cells is the sum of the optimal February and March rates. Note that these nonbolded numbers obscure any adjustments made in terms of the proportion applied in February vs the proportion applied in March. Visually, it is clear that most managers would split their applications under most conditions. And if it were known with certainty that conditions would be “normal,” nearly all participants state that they would split their applications.

Total nitrogen applied (lb per acre)			Single	Split
Very dry	Fairly dry	Normal	Fairly wet	Very wet
Type 1: Switch methods, adjust total nitrogen				
80	90	120	100	100
90	90	90	80	90
100	100	100	100	110
100	100	110	110	—
110	100	100	110	—
150	150	150	147	150
Type 2: Always split, adjust total nitrogen				
35	35	25	35	35
75	80	100	120	125
80	90	80	90	100
90	100	100	100	105
100	100	100	100	110
100	105	120	105	100
120	110	110	110	120
100	110	115	140	140
110	110	110	115	115
100	100	100	120	120
100	110	110	120	120
120	110	110	125	110
Type 3: Switch methods, same total nitrogen				
90	90	90	90	90
150	150	150	150	150
150	150	150	150	150
Type 4: Always split, same total nitrogen				
100	100	100	100	100
100	100	100	100	100
110	110	110	110	110
120	120	120	120	120
120	120	120	120	120
120	120	120	120	120
120	120	120	120	120

their application methods in response to different conditions, if these conditions were known with certainty. If farmers do not change their application timing/rates in this context, then they would be unlikely to vary their behavior in response to variations in forecast accuracy in

the experiment. Of the 28 respondents who responded to this question and completed the experiment, seven participants stated that they would always split their application and apply the same total amount of fertilizer, regardless of the rainfall condition. However, *most* managers would adjust their application practices in response to forecasts if managers believed that they were perfect.⁵

In terms of yield expectations, the average respondent considered a yield of 91 bushels per acre to be very good. What was considered to be a very good yield ranged from 50 to 145 bushels per acre, with the standard deviation (s.d.) being 16. A fairly good yield was, on average, 79 bushels (per acre), with an s.d. of 15; a typical yield was 71 bushels (s.d.: 14); a mediocre yield was 61 bushels (s.d.: 12); and a bad yield was 50 bushels per acre (s.d.: 14). The large variation in yield definitions likely arises from the heterogeneity in seed varieties, soil types, the extent to which farming contributes to the managers’ livelihoods (i.e., whether it is the primary income source), and the fertilizer application practices themselves.

In terms of rainfall definitions, the average respondent considered a (monthly) rainfall rate of less than 1.5 in. (s.d.: 0.69) to be drought conditions and a rate greater than 6.7 in. per month (s.d.: 1.37) to be waterlogging conditions, as mentioned previously. Respondents considered the lower bound of “normal” to be 3.5 in. (s.d.: 0.82) and the upper bound 4.9 in. (s.d.: 0.91). The high variation in rainfall definitions is likely to arise from the heterogeneity in local microclimates, as well as the drainage properties of the different soil types found across the farms. While an analysis of the factors influencing yield and rainfall definitions would be interesting, we consider it beyond the scope of this paper. Last, participants reported requiring anywhere from zero days to 180 days of advance notice to plan or arrange for nitrogen applications, with 17% stating either 30 or 45 days, and 47% with responses between 5 and 10 days. It is for this last group that week-ahead forecasts are likely to be of greatest relevance. This confirmed that most managers needed lead time for their top-dressing operations, and that this adjustment time may create costs for farmers who do not act strategically with respect to nitrogen application.

b. Experiment

We begin by graphically examining the relationship between nitrogen application responses and the probabilities

⁵ Of the 47 total respondents to this question, 10 responded that they would not change their nitrogen application choices in response to the different conditions. We do not show all 47 responses in the interest of space.

FIG. 1. Application rates by forecast probability (source: experiment). (a) Summary of distribution of responses for very wet and very dry February forecasts by forecast probability, ranging from 10% to 50%. Nitrogen amounts in pounds per acre are displayed on the y axis, and probability of rainfall conditions are on the x axis. (b) As in (a), but for March forecasts. Boxplots denote the median value (middle line), the 75th and 25th percentile values (the top and bottom of the box, respectively), the so-called maximum and minimum values (the top and bottom whiskers, respectively), and outliers (the dots). Calling the 75th-percentile, or upper-quartile, value UQ, and the 25th-percentile, or lower quartile, value LQ, the “maximum” is calculated as $UQ + 1.5(UQ - LQ)$ and the “minimum” is calculated as $LQ - 1.5(UQ - LQ)$. The distribution of application rates is largely unchanged for probabilities of a very wet outcome of up to 40%, but then becomes much more dispersed once the probability of a very wet event reaches 50%. This pattern suggests the existence of an accuracy threshold for modifying application rates.

associated with the two extreme cases of “very wet” and “very dry.”⁶ Figure 1a summarizes the distribution of responses in the experiment for very wet and very dry February forecasts by forecast probability, ranging from 10% to 50%. These box-and-whiskers plots suggest the existence of an accuracy threshold for modifying February application rates, at least for very wet forecasts. As the right-hand half of Fig. 1a shows, the distribution of application rates is largely unchanged as the probability of a very wet outcome increases from 10% (corresponding to a “normal” forecast) to 40%. However, when the probability of a very wet event is 50%, the distribution of responses is wider, with more participants choosing to forgo February application altogether (i.e., opting for a single application). This pattern suggests that participants may indeed follow a rule-of-thumb

(static) strategy, deviating only once the forecast probability for very wet conditions is 50% or higher. The left-hand half of Fig. 1a shows a similar response with respect to very dry forecasts; the effect, however, is more subtle. We similarly visualize March responses (Fig. 1b), but keep in mind that these results are driven not only by forecasts, but also by February application rates and February rainfall outcomes.

While informative, these aggregate plots may mask important heterogeneity in behavioral responses. We thus check the response patterns of individual players. Figure 2 shows the February application responses to both very wet and very dry forecast probabilities for 16 randomly selected participants (approximately every fourth participant). From these individual examples, we can draw several conclusions: First, some participants employed perfectly static strategies—that is, they applied the same rate of fertilizer irrespective of the forecasts (e.g., players 114 and 170). Indeed, eight out of 52 participants did just that. Second, extreme conditions appear to cause some participants to skip the February application altogether (e.g., players 138, 245, and 132).

⁶ For example, per Table 1, the response to the forecast “100% drier” will be treated as the forecast “60% very dry,” since this is the probability associated with “very dry” on the pie chart depicting this forecast.

FIG. 2. February fertilizer application responses, a randomly selected sample (source: experiment). The February application responses to both very wet (solid dots) and very dry (hollow dots) forecast probabilities are shown for 16 randomly selected participants. Nitrogen amounts in pounds per acre are displayed on the y axis, and probability of rainfall conditions are on the x axis. These plots reveal that respondents are heterogeneous, but that some patterns also emerge: 1) some do not alter their strategies at all, 2) extreme events cause some players to skip the February application, and 3) some respondents appear to increase fertilizer in response to very wet forecasts.

Such a pattern is observed in 15 cases. Third, some participants show a pattern of systematically increasing fertilizer applications in response to forecasts of higher-probability “very wet” forecasts, applying the most in response to the wettest forecasts (e.g., players 117 and 130). Last, we observe that the responses are “noisy” even within individuals, possibly because they are “gambling” during the game (e.g., player 249) or simply learning and recalibrating their responses with each round.

c. Regression analysis

Regression results are shown in Table 6, and the means and variances of the residuals are shown in Table 7. The variation in these residuals can be thought of as the variation in the player responses once the effects due to learning, previous-round performance, and February outcomes (in the case of the March decisions) have been purged. In Table 6, we see that there is weak evidence of learning (as evidenced by the coefficients on

round and *round, squared*). There is some evidence that participants are increasing their nitrogen applications with each round, but the estimated coefficients for these variables are not statistically different from zero at conventional levels of significance. However, the performance in the previous round appears to affect the amount of nitrogen applied in February (5% level of statistical significance). Specifically, doing poorly in the previous round appears to cause participants to apply more fertilizer in February. We also see that the March application rate is strongly influenced by the February application rate (5% level of statistical significance). In fact, the coefficient is almost exactly unity, which is consistent with a model of participants having a benchmark *total* amount that they aim to apply between the two applications. Finally, the rainfall experienced in February has a statistically significant effect on the amount applied in March; the greater the rainfall in February, the more is applied in March.

TABLE 6. Regression results (source: experiment). This table presents the regression results from estimating Eq. (1) without and with respondent fixed effects (respectively, columns 1 and 1'), and Eq. (2) without and with respondent fixed effects (respectively, columns 2 and 2'). Standard errors are given in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Performance in the previous round is a strong predictor of the amount of nitrogen applied. Furthermore, the March application rate is strongly influenced by the February application rate. The 628 observations are generated by 52 respondents.

	February application		March application	
	(1)	(1')	(2)	(2')
Round	−0.083 (0.788)	−0.049 (0.666)	−0.352 (0.733)	0.014 (0.567)
Round, squared	0.037 (0.058)	0.033 (0.049)	0.033 (0.054)	0.008 (0.418)
Previous harvest (1 = bad; 5 = very good)	−1.595 (0.591)	** −2.195 (0.554)	** 0.973 (0.473)	* 0.069 (0.409)
Feb rainfall rate (in. per month)			2.214 (0.858)	** 1.473 (0.698)
Feb rainfall rate, squared			−0.192 (0.096)	* −1.077 (0.078)
Feb application (lb per acre)			−1.000 (0.096)	** −1.080 (0.093)
Feb application, squared (lb per acre)			0.007 (0.001)	** 0.008 (0.001)
Observations	628	628	628	628
Respondent fixed effect	No	Yes	No	Yes
Regression F statistic	4.46	5.89	51.08	18.83
Adjusted R squared	0.0163	0.3003	0.3586	0.6265

As for the residuals themselves, their variances and means are presented in Table 7, by forecast probability. These residuals are from estimating the specifications in Eqs. (1) and (2). This information is also depicted graphically as box-and-whiskers plots in Fig. 3a (February) and Fig. 3b (March).⁷ Visually, it is apparent that the variance in February application rates dramatically increases when the probability of very wet conditions reaches 50% (Fig. 3a, right). As evidenced by the 25th-percentile value of −42 pounds per acre, applying zero nitrogen suddenly becomes a common response. This effect is apparent in Table 7 as well. In Fig. 3a, we see that when the probability of very wet conditions is 50%, the mean of the residuals falls to −8.73 pounds per acre and their variance jumps to 23.78.

When forecasts are for near-normal conditions, the fertilizer application rates given by respondents appear to be largely uninfluenced by the forecasts. However, when forecasts predict a particular adverse event (e.g., very wet conditions) with “high-enough” probability, respondents change their fertilizer application rates. The case of February responses to very wet forecasts is our only example of possible threshold behavior, but it is

also compelling. Our data strongly suggest the existence of one particular accuracy threshold: between 40% and 50% probability, for very wet conditions.

4. Summary and discussion

We draw on data from a framed field experiment and a questionnaire to answer these questions: 1) Does there exist an accuracy threshold below which managers appear not to act upon forecasts? 2) Which types of events elicit changes in management? 3) What is the relative importance of forecasts, compared to that of heuristics? The answers to these questions shed light upon the precise role that accuracy plays in managers' decisions to consult forecasts. Our experimental results suggest the existence of an accuracy threshold for very wet conditions that lies between 40% and 50%. They also corroborate anecdotal evidence that—when it comes to spring nitrogen applications on wheat—managers are most concerned with very wet conditions, an intuitive finding given the sensitivity of wheat to waterlogging, the difficulty of moving machinery when conditions are wet, and the increased risk of fertilizer runoff with excessive rainfall. Forecasts for very dry conditions elicited no discernible responses, at least over the probabilities presented in our game. Last, both experimental and survey results show that a portion of managers is unlikely to alter their management strategies no matter how their expectations of precipitation change. We observed that

⁷Note that the lines on the box-and-whiskers plots represent percentiles of the distribution (25th, 50th, and 75th) rather than the mean, and so the lines in the graphs will not necessarily align with the mean and standard deviations presented in Table 7.

TABLE 7. Regression residuals, by forecast probability (source: experiment). This table presents the mean and variance of residuals (in units of lb per acre) by probability of a very wet (VW) and a very dry (VD) forecast for (top) the February application and (bottom) the March application. Once the probability (pr) of a very wet forecast reaches 50% [$\text{pr(VW)} = 0.5$], there is a drastic change in both the mean and variance of the amount of nitrogen applied in February, indicative of threshold behavior.

Residuals: February application					
	pr(VW) = 0.1	pr(VW) = 0.2	pr(VW) = 0.3	pr(VW) = 0.4	pr(VW) = 0.5
Mean	1.433	-0.568	0.946	2.811	-8.726
Variance	11.043	14.199	17.640	17.186	23.775
	pr(VD) = 0.1	pr(VD) = 0.2	pr(VD) = 0.3	pr(VD) = 0.4	pr(VD) = 0.5
Mean	1.433	3.083	-0.728	0.661	-1.645
Variance	11.043	12.078	14.998	13.929	14.739
Residuals: March application					
	pr(VW) = 0.1	pr(VW) = 0.2	pr(VW) = 0.3	pr(VW) = 0.4	pr(VW) = 0.5
Mean	-0.852	-2.261	2.242	1.594	-0.164
Variance	15.862	16.986	9.041	11.983	12.829
	pr(VD) = 0.1	pr(VD) = 0.2	pr(VD) = 0.3	pr(VD) = 0.4	pr(VD) = 0.5
Mean	-0.852	-1.667	-0.064	0.439	-1.063
Variance	15.861	9.698	9.370	8.456	10.979

some game participants simply applied the same amount of nitrogen in response to all forecasts. Were they being “lazy” game players? The questionnaire responses suggest not; nearly a quarter of respondents stated that they

would not change their application amounts, even in the presence of perfectly accurate forecasts for specific events. For these respondents, fertilizer application choices are simply independent of expectations of precipitation.

FIG. 3. Application rate residuals by forecast probability (source: experiment). (a) Box-and-whisker plots of residuals from estimating Eq. (1) for very wet and very dry February forecasts by forecast probability. Nitrogen amounts in pounds per acre are displayed on the y axis, and probability of rainfall conditions are on the x axis. (b) As in (a), but from estimating Eq. (2) (March application rates). Boxplots denote the median value (middle line), the 75th and 25th percentile values (the top and bottom of the box, respectively), the co-called maximum and minimum values (the top and bottom whiskers, respectively), and outliers (the dots). Calling the 75th-percentile, or upper-quartile, value UQ, and the 25th-percentile, or lower-quartile, value LQ, the “maximum” is calculated as $UQ + 1.5(UQ - LQ)$ and the “minimum” is calculated as $LQ - 1.5(UQ - LQ)$. As with Fig. 1, the variance in February application rates increases dramatically when the probability of very wet conditions reaches 50% even after controlling for learning effects through the regression, and is indicative of a behavioral response with respect to an accuracy “threshold.”

Please note: By turning in this form, you agree to participate in this session and to let us study your answers.

Sources of weather information

1. Which of these weather forecast products do you regularly consult? (Check all that apply.)

N=69 (100%)

- ☐ Weather forecast on local television news
- ☐ The Weather Channel, 7-day daily forecast
- ☐ Accuweather 45-day forecast
- ☐ National Weather Service, 7-day daily forecasts
- ☐ Climate Prediction Center, 6 to 8 day ahead forecast
- ☐ Climate Prediction Center, 8 to 14 day ahead forecast
- ☐ Climate Prediction Center, one-month ahead forecast
- ☐ Climate Prediction Center, three-month ahead forecast
- ☐ other—please specify: _____

Wheat production practices

2. In what form do you *usually* apply nitrogen? (Check one.)

N=66 (95%)

- ☐ ammonium nitrate
- ☐ urea
- ☐ UAN solution
- ☐ prill
- ☐ manure--please specify type: _____
- ☐ other—please specify type: _____

3. What is the average price that you pay for this fertilizer? \$ _____/(unit _____)

N=34 (49%)

4. What is your *main* soil type? _____

N=49 (71%)

5. What variety(ies) do you currently grow? _____

N=53 (76%)

FIG. A1. Questionnaire, page 1 of 4.

We conclude that accuracy played a substantial role in explaining participants' choices in the game. The implication is that unless week-ahead forecasts for "very wet" conditions can be made with probabilities exceeding 50% (something that is presently impossible), managers will not alter their spring nitrogen application choices in response to forecasts. The broader implication is that users of forecasts and decision support tools are sensitive to the accuracy of information.

However, accuracy alone will not explain whether a user pays attention to forecasts. These findings echo those of [Morss et al. \(2010\)](#), who also find a forecast

probability threshold for rainfall at which there is an upward jump in the number of respondents stating that they would take some protective action (i.e., move a planned picnic indoors). They conclude that most respondents appear to understand probability forecasts (at least for germane events such as rainfall), but that comprehension of probability forecasts is not necessarily predictive of their use. Consulting and digesting information such as forecasts likely takes both time and cognitive energy. Indeed, [McCown \(2012\)](#) observes that managers often rely on heuristics and strategies based on intuition, saving their deliberative consultation and

Past Experiences**6. Wheat yields on your farm***N=55 (79%)*

- On your farm, what is a 'very good' yield (farm average)? _____ bu/acre
- On your farm, what is a 'good' yield? _____ bu/acre
- On your farm, what is a 'typical' yield? _____ bu/acre
- On your farm, what is a 'mediocre' yield? _____ bu/acre
- On your farm, what is a 'bad' yield? _____ bu/acre

7. Spring precipitation*N=43 (62%)*

Here's a drawing of a rain gauge. Think about the **month of March**. Now think about **total monthly precipitation** (including snow) at that time of year. For reference, Lexington got 4.1 inches during all of last March.

*Note that 1 inch of snow is approximately .1 inch of precipitation.

- First, shade in the area on the gauge that you would consider 'normal' or typical levels.
- Now shade in the area that constitutes 'waterlogging' levels, *given your soil type*.
- Finally shade in the area that constitutes 'drought' levels.

FIG. A2. Questionnaire, page 2 of 4.

strategies (in their case, using the FARMSCAPE yield simulator) for out-of-the-ordinary situations.

Our study provides insights that can help public entities to direct their efforts when creating decision support tools. First, some outcomes are more important to predict than others. Second, managers likely rely on rules-of-thumb, rather than recalculating their choices with every piece of information, most of the time. Third, farmers are heterogeneous in their general attitudes toward forecasts, which is likely influenced by heterogeneity in the "stakes" of making the correct choices.

We discuss several limitations to our study that should be kept in mind. First, many concerns are rightfully raised with respect to the validity of the responses collected through experiments of any sort. In our case, participants might have treated the experiment as a mere game and may not have considered their choices as carefully as they would have in real life. This being a framed field experiment, however, the risk of unrealistic responses was reduced. The game was concrete and relevant, as it asked professional growers of wheat to respond to realistic and plausible hypothetical situations

8. How often do you see these March conditions on your farm? *N=59 (85%)*

Very wet (waterlogging): _____ out of ten years

Fairly Wet: _____ out of ten years

Normal: _____ out of ten years

Fairly dry: _____ out of ten years

Very dry (drought conditions): _____ out of ten years

9. Given each of these March conditions, what is your best nitrogen application strategy? *N=47 (68%)*

Assume the following:

- The preceding corn crop had low to average yields.
- Early spring soil conditions are moderately dry.
- High and low temperatures are predicted to be 'normal' for this time of year.

Very wet ☐ single: _____ lbs/acre **or** ☐ split: _____/_____ lbs/acre

Fairly wet ☐ single: _____ lbs/acre **or** ☐ split: _____/_____ lbs/acre

Normal ☐ single: _____ lbs/acre **or** ☐ split: _____/_____ lbs/acre

Fairly dry ☐ single: _____ lbs/acre **or** ☐ split: _____/_____ lbs/acre

Very dry ☐ single: _____ lbs/acre **or** ☐ split: _____/_____ lbs/acre

10. How far in advance do you need to plan or arrange for your nitrogen applications? *N=53 (81%)*

_____ days

FIG. A3. Questionnaire, page 3 of 4.

within the familiar framework of the spring top-dressing decision. Thus it satisfied the first two criteria listed by [Harrison and List \(2004\)](#) of *knowledgeable players* and *salient scenarios*. Moreover, we believe that self-evaluation of yields after each of the 12 rounds also kept participants engaged and cognizant of the consequences of their decisions during the game; this feedback mechanism was built into the game in order to bolster Harrison and List's third criterion of *consequentiality*.

Second, aside from the self-reported participant types in [Table 3](#), it is difficult to know the extent to which our participants represent the "typical" agricultural manager.

The questionnaire did not ask questions about operator age, gross farm income, and education level. While these questions would have allowed a comparison of sample characteristics to state-level summary statistics for the farmers, these questions might have also been perceived as being sensitive and hence increased nonresponse. Despite the uncertainty surrounding the representativeness of our participants, we present our results because they are interesting in their own right; to the best of our knowledge, these types of questions have never been posed to any group of agricultural managers. Future work that examines the forecast responses of

Farm and operator

11. In what capacity do you make these decisions? *N=60 (89%)*
(Check all that apply.)

- ☐ farmer
- ☐ consultant
- ☐ ag. agent
- ☐ other—please specify: _____

12. In what county or counties is your wheat operation? _____ County
N=55 (79%)

13. How many acres was your total wheat operation last year? _____ acres
N=45 (65%)

14. Do you consider farming your *primary* income source? ☐ Yes ☐ No
N=56 (81%)

15. How often do you access the internet (using a computer or a smartphone)?
(Check one.) *N=62 (89%)*

- ☐ rarely
- ☐ a few times a month
- ☐ a few times a week
- ☐ at least once a day

16. Have you purchased any of these crop insurance products for your *wheat crop* in the past 5 years? If so, check all that apply and indicate your coverage level. *N=58 (84%)*

(Check all that apply and indicate coverage level.)

☐ Yield protection	coverage level: _____ %
☐ Revenue protection	coverage level: _____ %
☐ Revenue protection w/ harvest price exclusion	coverage level: _____ %

FIG. A4. Questionnaire, page 4 of 4.

agricultural managers in other areas would be helpful to assess the extent to which our results are more generally applicable.

Third, although the questionnaire reveals great heterogeneity among forecast users, our limited sample size means that the study is unable to formally test the effects of this heterogeneity on forecast use and threshold behavior. Forecast usage likely varies depending on factors from soil type and acres managed to lead time and crop insurance enrollment. While we capture some of these variables in our questionnaire, we do not incorporate

these into the regression analysis, since not all game players completed the questionnaire, and doing so would have further reduced our sample size. Moreover, forecast usage could be based on factors that were not assessed in our questionnaire, such as differing risk preferences or information costs. Any or all of these sources of heterogeneity can contribute to the existence of and variation in accuracy thresholds and the decision to rely on heuristics instead of forecasts. We echo many others in the field by concluding that a better understanding of the various mechanisms that lead to threshold behavior would

certainly be informative to forecast designers, and additional research in this regard would be valuable.

Last, this study is unable to discern between 1) managers ignoring the forecast and simply relying on their heuristics and 2) managers choosing the same actions after having paid attention to and consulted the forecasts. While it is beyond the scope of this analysis, future work could estimate the value of “paying attention” to forecasts by combining a theoretical model with a crop simulation model. Use of a crop simulation model would enable us to make these values farm-specific. One could then compare stated management responses against these estimated gains from paying attention.

Public institutions have, for decades, been working on improving the accuracy of their forecast products and disseminating these as public goods, often lamenting the apparent indifference with which these are met. Is a turning point at hand? The fact that private entities—the Climate Corporation being an example—have recently joined the public institutions in providing weather and climate forecasts could be interpreted as a sign that some information products are indeed starting to surpass minimum-accuracy thresholds. This, combined with more strategic choices in generating and designing tools, could herald a new era in the development of forecast-based decision support tools.

Acknowledgments. The authors wish to thank the late Don Halcomb for his valuable feedback on the study design, as well as the organizers of the 2015 University of Kentucky Winter Wheat Conference for allowing us to carry out this study. We are deeply appreciative of Matt Dixon and Rezaul Mahmood, who shared their time and expertise in helping us to understand the meteorological literature and the local precipitation data. We are also grateful to Carl Dillon at the University of Kentucky, and the faculties of the departments of Food and Resource Economics and Agricultural and Biological Engineering at the University of Florida for input on earlier versions of this work, particularly Senthil Asseng. This work was funded by the Kentucky Agricultural Experiment Station and the National Institute of Food and Agriculture, U.S. Department of Agriculture, under Hatch Project 1006174.

APPENDIX

The Questionnaire

The complete questionnaire appears in [Figs. A1–A4](#).

REFERENCES

- Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525**, 47–55, <https://doi.org/10.1038/nature14956>.
- Harrison, G. W., and J. A. List, 2004: Field experiments. *J. Econ. Lit.*, **42**, 1009–1055, <https://doi.org/10.1257/0022051043004577>.
- Hayman, P., J. Crean, J. Mullen, and K. Parton, 2007: How do probabilistic seasonal climate forecasts compare with other innovations that Australian farmers are encouraged to adopt? *Aust. J. Agric. Res.*, **58**, 975–984, <https://doi.org/10.1071/AR06200>.
- Hilton, R. W., 1981: The determinants of information value: Synthesizing some general results. *Manage. Sci.*, **27**, 57–64, <https://doi.org/10.1287/mnsc.27.1.57>.
- Hoskins, B. J., 2013: The potential for skill across the range of the seamless weather–climate prediction problem: A stimulus for our science. *Quart. J. Roy. Meteor. Soc.*, **139**, 573–584, <https://doi.org/10.1002/qj.1991>.
- Hu, Q., and Coauthors, 2006: Understanding farmers’ forecast use from their beliefs, values, social norms, and perceived obstacles. *J. Appl. Meteor. Climatol.*, **45**, 1190–1201, <https://doi.org/10.1175/JAM2414.1>.
- Joslyn, S. L., L. Nadav-Greenberg, M. U. Taing, and R. M. Nichols, 2009: The effects of wording on the understanding and use of uncertainty information in a threshold forecasting decision. *Appl. Cognitive Psychol.*, **23**, 55–72.
- Kusunose, Y., and R. Mahmood, 2016: Imperfect forecasts and decision making in agriculture. *Agric. Syst.*, **146**, 103–110, <https://doi.org/10.1016/j.agsy.2016.04.006>.
- Lee, C., J. Herbek, and R. Trimble, 2009: A comprehensive guide to wheat management in Kentucky. Tech. Rep. ID125, University of Kentucky, Cooperative Extension Service, 72 pp., <http://www2.ca.uky.edu/agcomm/pubs/id/id125/id125.pdf>.
- Mase, A. S., and L. S. Prokopy, 2014: Unrealized potential: A review of perceptions and use of weather and climate information in agricultural decision making. *Wea. Climate Soc.*, **6**, 47–61, <https://doi.org/10.1175/WCAS-D-12-00062.1>.
- McCown, R., 2012: A cognitive systems framework to inform delivery of analytic support for farmers’ intuitive management under seasonal climatic variability. *Agric. Syst.*, **105**, 7–20, <https://doi.org/10.1016/j.agsy.2011.08.005>.
- Meza, F. J., J. W. Hansen, and D. Osgood, 2008: Economic value of seasonal climate forecasts for agriculture: Review of ex-ante assessments and recommendations for future research. *J. Appl. Meteor. Climatol.*, **47**, 1269–1286, <https://doi.org/10.1175/2007JAMC1540.1>.
- Moeller, C., I. Smith, S. Asseng, F. Ludwig, and N. Telcik, 2008: The potential value of seasonal forecasts of rainfall categories—Case studies from the wheatbelt in Western Australia’s Mediterranean region. *Agric. For. Meteorol.*, **148**, 606–618, <https://doi.org/10.1016/j.agrformet.2007.11.004>.
- Morris, R. E., J. K. Lazo, and J. L. Demuth, 2010: Examining the use of weather forecasts in decision scenarios: Results from a US survey with implications for uncertainty communication. *Meteor. Appl.*, **17**, 149–162, <https://doi.org/10.1002/met.196>.
- Nadav-Greenberg, L., S. Joslyn, and M. U. Taing, 2008: The effect of weather forecast uncertainty visualization on decision-making. *J. Cogn. Eng. Decis. Mak.*, **2**, 24–47, <https://doi.org/10.1518/155534308X284354>.
- Nelson, R. R., and S. G. Winter Jr., 1964: A case study in the economics of information and coordination: The weather forecasting system. *Quart. J. Econ.*, **78**, 420–441, <https://doi.org/10.2307/1879475>.

- O'Lenic, E., D. Unger, M. Halpert, and K. Pelman, 2008: Developments in operational long-range climate prediction at CPC. *Wea. Forecasting*, **23**, 496–515, <https://doi.org/10.1175/2007WAF2007042.1>.
- Radner, R., and J. Stiglitz, 1984: A nonconcavity in the value of information. *Bayesian Models of Economic Theory*, M. Boyer, and R. Kihlstrom, Eds., Elsevier, 33–52.
- Roulston, M. S., and T. R. Kaplan, 2009: A laboratory-based study of understanding of uncertainty in 5-day site-specific temperature forecasts. *Meteor. Appl.*, **16**, 237–244, <https://doi.org/10.1002/met.113>.
- , G. E. Bolton, A. N. Kleit, and A. L. Sears-Collins, 2006: A laboratory study of the benefits of including uncertainty information in weather forecasts. *Wea. Forecasting*, **21**, 116–122, <https://doi.org/10.1175/WAF887.1>.
- Westra, S., and A. Sharma, 2010: An upper limit to seasonal rainfall predictability? *J. Climate*, **23**, 3332–3351, <https://doi.org/10.1175/2010JCLI3212.1>.