Selected Poster/Paper prepared for presentation at the Agricultural & Applied Economics Association's 2017 AAEA Annual Meeting, Chicago, Illinois, July 30-August 1, 2017

A Model of Overconfidence in Subjective Probabilities

Clayton P. Michaud, Thomas W. Sproul and Calum G. Turvey

Abstract

We develop a model of probabilistic forecasts to examine evidence of systematic overconfidence among Chinese wheat and corn farmers (Turvey et al. 2013). We depart from the familiar probability weights approach of cumulative prospect theory (Kahneman and Tversky 1992) by explicitly modeling forecasts as shifting and scaling the historical distribution, with parameters that may incorporate a reference point. We find that forecasts are anchored to historical positive experience, and that forecasts are systematically overconfident. However, like Sproul and Michaud (2017), we find evidence of multi-modal heterogeneity such that population average parameter estimates do not represent individuals well. We estimate a finite Gaussian mixture model using expectation-maximization (Dempster, Laird and Rubin 1977) and find strong evidence for three basic forecast types: optimistic (about two thirds), realistic or neutral (about 20%), and pessimistic (the remainder). We find further clustering within the optimistic group, without about half highly optimistic and about half mildly so. The group-wise means and mixture weights are robust to inclusion of additional elicited data, and also to inclusion of a shape parameter in the forecast model. We also find evidence that forecast classifications do not perfectly map to a classification of the person making it. Instead, recent severe loss experience appears to be the strongest single predictor of pessimism, though we do find statistically significant gender differences. Our results have important implications regarding the need for crop insurance subsidies to induce participation. Future research should focus on discovering potential ways to reduce the discrepancy between subjective and objective risk.

Keywords: Behavioral Economics, Crop Insurance, Crop Yields, Expectation Maximization, Farm Risk Management, Forecasts, Gaussian Mixture Model, Optimism Bias, Overconfidence, Subjective Probability

JEL Codes: C38, D81, Q12

Acknowledgements

Michaud (corresponding author: claytonmichaud@my.uri.edu) is a graduate student and Sproul is Associate Professor, both in the Department of Environmental & Natural Resource Economics at the University of Rhode Island. Turvey is W.I. Myers Professor of Agriculture Finance in the Charles H. Dyson School of Applied Economics and Management at Cornell University. Dr. Sproul acknowledges financial support from USDA National Institute of Food and Agriculture Hatch Projects: No. RI00H-108, Accession No. 229284, and No. RI0017-NC1177, Accession No. 1011736, and USDA Economic Research Service Cooperative Research Agreement No. 58-6000-5-0091. Dr. Turvey acknowledges financial support from the W.I. Myers endowment funds, Cornell University, China National Nature Science Fund grant 70873096, and Humanities and Social Science Fund of China Education Department grant 07JA790027.

"Perhaps the most robust finding in the psychology of judgment is that people are overconfident" -De Bondt and Thaler (1995, pg. 389)

Evidence of overconfidence is rampant in the behavioral literature (see Sandroni and Squintani (2007), Slovic (2016), De Bondt and Thaler (1995), and Weinstein (1980) for overviews), and it is now generally accepted in the insurance literature that (consumers') risk misperceptions play a significant role in contributing to (demand) frictions in insurance markets, specifically the historic finding that large subsidies are often required to induce insurance participation (Barseghyan et al. (2013); Chiappori et al. (2006); Cutler and Zeckhauser (2004); Sandroni and Squintani (2007, 2013); Slovic (2016); Spinnewijn (2013, 2015, 2017); Turvey et al. (2013)). With the federal government currently subsidizing crop insurance premiums to the tune of 6 billion dollars a year (USDA ERS), a more complete understanding of the disconnect between perceived (subjective) risk, which governs demand, and actual (objective) risk, which governs supply, is crucial to designing more (cost) effective tools and policies.

This research will rigorously examine the relationship between the perceived future yield distribution and the distribution of historic yield outcomes at the farm(-crop) level (specifically with regard to location and scale) using self-reported expert elicitations of historic and expected yields from rural Chinese farmers. To that end, we have four specific objectives:

1. Develop a model of subjective forecasting based on historic outcomes.
2. Estimate the parameters of this model at the individual forecast level (not just per farm, but also per crop per farm)
3. Examine the population-level distribution of these parameters and identify potential clustering of forecasts into 'types', i.e. optimistic, realistic, pessimistic.

4. Examine potential factors influencing the parameters of our model.

We begin by developing a simple 2-parameter model for estimating subjective future yield forecasts based on transforming the location and scale of a historic (objective) yield distribution. We go on to use linear regression to compare the goodness-of-fit of potential functional forms (absolute vs. relative location change), as well as various potential location specifications for our model.

We estimate the parameters of our candidate model at the forecast level using self-reported elicitations regarding minimum, maximum, and most likely expected yields in the upcoming season, in combination with self-reported minimum, maximum, and most likely historic yield outcomes obtained from rural Chinese farmers. We then use the Expectation-Maximization (EM) algorithm to fit a finite Gaussian mixture model (GMM) over different numbers of components for our forecast level parameter estimates in order to identify potential heterogeneity/clustering in the distribution of these parameters at the population level.

We go on to use regression analysis to estimate potential factors influencing these parameters, such a demographics, farm tenure, and yield outcomes history (e.g. recency of historic minimum and maximum yields, including cross-crop outcomes).

Our data is sourced from Turvey et al. (2013) who used a survey method to directly elicit risk perceptions from rural farmers across three counties (25 villages) in Shaanxi Province in October and November 2010. Our final data set contains 879 observations and is comprised of 442 corn forecasts and 437 wheat forecasts.

When modeling future yield expectations as a function of historic outcomes, we find that location changes are best modeled via a relative, rather than an absolute, change in location. We also find that defining the location parameter in terms of an upper-bound (in our case, maximum historic yield) provided the best fit given our data and model. We find that, on average, the maximum yield value is relatively stable across the subjective and objective yield distributions, with farmers generally underestimating the extent of their downside risk. This is an important finding given that crop insurance is explicitly designed to help mitigate downside risk and can help explain why historically farmer's willingness-to-pay for crop insurance has historically fallen far below its actuarial cost.

We find strong evidence of distinct 'types' or forecasts, namely 'optimistic', 'realistic', and 'pessimistic' with 65% of forecasts classified as optimistic, 20% as 'realistic', and 15% classified as 'pessimists'. On average, optimistic forecasts had half the range (standard deviation) of their objective yield distribution, while pessimistic forecasts had twice the range of their objective distribution. Expanding our analysis to consider more classes, we find that the optimistic group cleanly separates into mildly and strongly optimistic groups of roughly equal size. In addition, we find that adding a third parameter to account for changes in shape did not produce significantly different results, and the mixture values provided little additional information in terms of heterogeneity. We conclude that transformations of the third moment are marginal enough to justify the more parsimonious, 2-parameter transformation model for future applied work.

Our regression analysis of factors influencing optimism provides us with three important results. Firstly, we see significant differences in the relative range of future forecasts by gender. Specifically, women appear to be much more optimistic with regard to yield risk than their male counterparts. On average, male forecasters expect future yields to have 80% the scale of their historic yield distribution, while female forecasters expect future yields to come from a distribution with roughly only 71% the scale of their historic distribution.

The second, and perhaps most important, finding is that a farmer's degree of optimism with regard to forecasting a specific crop is dramatically influenced by the recency of historic losses (i.e. years since a historic minimum yield), with each additional year since a historic loss predicting a 1% decrease in the scale of their forecasted yield distribution. In contrast, we do not find that historic gains have a significant impact on a farmer's forecast. In addition, we find no evidence of a cross-crop effect, i.e. wheat forecasts do not seem effected by historic losses for corn, and vice versa. These results are an important first step in better understanding the ways in which farmers might subscribe to the fallacy of the law of small numbers (LSN) when making yield forecasts, suggesting that farmers overweigh the statistical importance of recent 'losses' when estimating their perceived risk.

Lastly, we find that the length of a farmer's tenure (i.e. number of years farming) does not play a significant role in predicting his or her degree of optimism. This is an important finding because it tells us that the phenomenon we observe (that farmer's expect their future yields to come from a more favorable distribution than their historic yields) cannot be explained by technological advances over time causing more recent yields to outperform less recent ones. This supports our hypothesis that farmers are indeed being irrational in their expectation of more favorable future yields.

This paper is organized as follows: Section II provides a background of the relevant literature pertaining to overconfident beliefs and behavior. In addition, we discuss work looking at how such an effect influences demand for insurance. We conclude by discussing preliminary evidence of overconfidence with regard to future yield expectations. Section III introduces our model. Section IV presents and discusses our data set both in terms of how it was collected, as well as general summary statistics. Section V revisits our model and fits our data to it. This section includes the selection of the final form of our forecasting model. Section VI begins by motivation and outlining the Expectation Maximization procedure for fitting a (finite) Gaussian Mixture Model (GMM) to our data. The latter half of this section presents and discusses the results of our mixture model. Section VII examines potential influences of optimism bias, including demographic variables and yield outcome history. Section VIII concludes our paper with a discussion of our findings, potential areas of future research, and how our findings might be used to improve policy design.

Background

Overconfident Beliefs

There is a vast body of widely replicated experimental studies showing that individuals are systematically overly optimistic in the face of risk- by and large, individuals believe that they are more likely than average to experience positive future events, while being less likely to experience negative events (see Sandroni and Squintani (2007), Slovic (2016), De Bondt and Thaler (1995), and Weinstein (1980) for overviews).

Hoch (1985) found that MBA students overestimate the number of job offers they will receive and the magnitude of their salary. In a similar study, 258 college students estimated how their own chances of experiencing 42 events differed from the chances of their classmates. Overall, students rated their own chances as above average for positive events, and below average for negative events (Weinstein, 1980).

Statman and Tyebjee (1985) find widespread evidence of strong optimism bias when comparing forecasted and actual construction costs in a range of industries including energy, military hardware, pharmaceutical, chemical and other development projects.

Overconfidence is also not strictly an economic bias. For example, there is evidence that people underestimate their health risks in relation to medical exams (Kreuter and Strecher, 1995; Robb et al., 2004). Likewise, evidence of overconfidence has been widely reported among automobile drivers. Svenson (1981) finds that 82% of a sample of students placed themselves among the top 30% safest drivers. Similar results are found in Groeger and Grande (1996) who compare drivers' self-assessments skills with those assessed by an instructor. There is also strong specific evidence that overconfidence does not vanish with learning nor with experience. For example, Dalziel and Job (1997) find that even professional drivers underestimate their risk of automobile accident.

With regard to how overconfidence might be distributed at the population-level, in a study measuring optimism-pessimism from beliefs about future events, Wenglert and Rosén (2000) find that 72% of their subjects were classified as being overly optimistic about their personal future. Cooper, Woo, and Dunkelberg (1988) find that 68% of entrepreneurs think that their startup is more likely to succeed than comparable enterprises and that furthermore, only 5% believe that their odds of success are worse than their competitors, while a third of entrepreneurs believe their success is all but guaranteed (similar findings are reported in a survey of French entrepreneurs by Landier and Thesmar (2009). These results are consistent with this paper's finding that 67% of farmer's are overly optimistic with regard to future yield expectations, as well as the finding that there are both moderate and extreme optimists.

Overconfident Behavior

Additionally, there is evidence that individuals are not only optimistic with regard to beliefs about personal risk, but more importantly that they act upon these biased beliefs as well. For instance, there is evidence that overconfidence can lead to poor financial planning and economic decision-making. In an experiment about business entry Camerer and Lovallo (1999) find that much of the decision to start a new business is due to an innate appreciation of personal skills relative to compatriots. Entrepreneurs entering into a competitive market (over entry) believe that they have positive profits while other entrants will have negative profits. Optimists neglect to consider that the reference group views their opinions in the same way and will act accordingly. Another important finding of this study is that subjects are more likely to act with overconfidence when they were betting on their own skills rather than a random device.

Malmendier and Tate (2006, 2008) provide convincing evidence that overconfident CEOs are unambiguously more likely to overpay, to undertake value-destroying acquisitions, and to make acquisitions when their firm has abundant internal resources. Compared with their unbiased counterparts, optimistic CEOs are 65% more likely to complete mergers, are more likely to overpay for those target companies, and are more likely to undertake value-destroying mergers. A survey by Benartzi (2001) found that individuals are optimistic and/or overconfident about future returns to their employers stock relative to other firms. Only 16.4% percent of respondents believed that their company's stock was riskier than the stock market as a whole. This finding is used to help explain why employees often unwisely concentrate their portfolios with company stock.

There is also experimental evidence that parties involved in legal disputes are overly reluctant to settle out of court because they are overly optimistic about the outcome of their case (see Babcock, Loewenstein, and Issacharoff (1997)).

Overconfidence is not just bad for consumer's wallets. Just as economic biases lead to sub-optimal economic decision making, overconfidence has become recognized as both a major barrier preventing health behavior (Hoorens, 1994), and a major determinant of traffic safety (Hatakka et al., 2002; Bartl, 2000).

Lastly, perhaps the most famous treatise on overconfidence is Robert Shiller's best-selling book *Irrational Exuberance* (2000). Named for Alan Greenspan's imfamous caution against "unduly escalated asset values" (Greenspan, 1996), Shiller's work details the "wishful thinking… that blinds us to the truth of our situation". He discusses how overconfidence effects not only investment behavior, but can have serious macro-level effects, arguing that such beliefs and behavior have contributed greatly to the financial bubbles that have wreaked havoc on not just the U.S., but the global economy.

Overconfidence and the Demand for Insurance

A number of studies have looked at the impact of optimism bias and demand for insurance products. Cebulla (1999) conducted surveys on the perception of the risk of becoming unemployed and the willingness to purchase unemployment insurance concluding that risk underestimation reduced the willingness to buy insurance. As part of a study on secondary life-insurance markets, Bhattacharya et al. (2004) find evidence that patients who underestimate their risk of death are unwilling to hold insurance coverage.

Barseghyan et al. (2013) estimate a structural model of insurance deductible choices that incorporates standard risk aversion and probability distortions and find that probability distortions play an important role in explaining deductible choices.

Specific to the field of agricultural economics, Shaik et al. (2008) compare the cross-elasticity for crop revenue and crop yield demand using a subjective probability framework and find a clear tendency for farms with great perceived risk (either in yield or price) to prefer revenue insurance over yield insurance. Similarly, they find that the elasticity of demand for revenue insurance to be (relatively) greater then the elasticity of yield insurance.

Not surprisingly, farming risk has been shown to be positively correlated with the decision to purchase insurance. (Horowitz and Lichtenberg, 1993; Enjolras and Sentis, 2011; Wang et al., 2016).

Expanding the above finding to include subjective or perceived risk, both Sherrick et al. (2004) and Egelkraut et al. (2006) identify perceived yield risk as an important factor influencing farmers' crop insurance decisions. They find that producers who viewed their yields as better than average were less likely to buy crop insurance, while producers who perceived their yield variability to be larger than average are more likely to insurance.

None of these studies, however, compared subjective yield risk to objective yield risk at the farm level, nor how such a potential discrepancy might influences the necessary level of subsidies to induce participation.

Evidence of Overconfidence in Agriculture

Pease (1992) is the first study we are aware of to statistically compare subjective and historical crop yield probability distributions. His results indicated that in many individuals, there exist large differences in the moments of the subjective and historical distribution. Pease compared historical yields and subjective expectations for 1987 expected crop yields for 90 Western Kentucky grain farmers. Using this relatively small sample, he found that subjective forecasts tended to underestimate forecasting derived from corn yields, while conversely overestimating soybean forecasts. However, Pease also found that recent regional experiences seemed to play a more important role than the crop itself, suggesting that within-subject variation in forecasting bias results from crop-specific historic (specifically recent) outcomes. This result is consistent with our regression results from section VII.

In addition, the finding that corn farmers tended to be pessimistic with regard to future yield expectations (for the 1987 harvest) is not necessarily inconsistent with our overall findings due to the fact that drought conditions caused extremely low corn yields in 1980 and 1983. Given our result that the most significant factor mitigating a farmer's degree of optimism is the presence of recent historical losses, the finding that farmer's were more optimistic about soybean yields (which had suffered fewer recent losses) relative to corn yields actually supports our findings.

Egelkraut et al. (2006) asked farmers to asses their individual yields and yield variability relative to those of a typical farmer in their county. The majority of participants perceived their yields to be higher (46.1%) or similar (42.3%) relative to other farms in their county and only 11.6% believed they experienced lower yields. Likewise, most farmers viewed their yields as less variable (41.5) or similar (38.45) and only 20.1% reported that they experienced greater variability than was typical in their county.

In a similar study, Sherrick (2002) finds that farmer's subjective probability beliefs about important weather variables are systematically miscalibrated and demonstrate that significant errors in farmers' risk assessments and insurance valuation arise as a consequence of farmers' systematically inaccurate probability beliefs.

Turvey et al. (2013), the original source of our data, is the first paper we are aware of to explicitly argue that the need for subsidizing crop insurance premiums is a consequence of the dissonance between subjective and historical risks. They also conclude that the perception of downside risk is largely dictated by perceived mean and standard deviation. We expand upon these findings by explicitly modeling future forecasts as an affine transformation of the location and scale of a farmer's historic yield distribution. We then use our estimates of these transformation parameters to examine how optimism is distributed at the population level, identifying distinct 'types' of forecasts (specifically extremely optimistic, moderately optimistic, realistic, pessimistic). We go on to examine exogenous factors that may play a role in this forecasting process, such as demographic information and outcomes history.

A Location-Scale Model of Forecasts from Historical Data

In contrast to much of the established prospect theory literature, we explicitly model the forecast as being a transformation of historical experience. Specifically, we model the forecast distribution as an affine, "shift and scale" transformation of the historical distribution of the form:

$$Y = \alpha + \beta X \ . \qquad (1)$$

In the above equation, X and Y are random variables representing the historical distribution and forecast distribution, respectively, α is a shift parameter, and β is a scale parameter. We select this specification because it is *a priori* reasonable and consistent with basic statistical intuition, and because it applies broadly across the location-scale family of probability distributions. It's simplicity does come with the downside that higher moments of the distribution are not considered (e.g., a shape parameter), though we will show in what follows that allowing a skewness adjustment from historical to forecast does not meaningfully change our results.

The benefits of this simple specification appear to dramatically outweigh its downside: i) the model requires only the weak assumption that the historical and forecast distributions come from the same family, ii) it does not impose any directional bias on the nature of forecasts, and iii) it includes as a special case a normalization step, in which the historical X might be adjusted to have zero mean and unit variance before shifting and scaling. Past research has tended to find that forecasts deviate systematically from average, or typical, experience, and that they usually do so in an optimistic fashion for the majority of the population. Lack of directional bias is important in a model precisely because what is optimistic may change sign depending on the application: an optimistic earning forecast might include an upward shift, whereas an optimistic forecast of commute time might include a downward shift.

The generality of our model with respect to normalizing X is demonstrated as follows. First, assume that X has known mean and standard deviation, μ and σ, so the normalized forecast would be constructed with parameters a and b according to:

$$Y = a + b \cdot \left(\frac{X - \mu}{\sigma} \right) = \left(a - \frac{b\mu}{\sigma} \right) + \frac{b}{\sigma} X. \tag{2}$$

This is a special case of our model obtained by setting α equal to the parenthetical term and by setting $\beta = b/\sigma$. In fact, our model substantially generalizes this approach by the simple fact that μ and σ need not be known in order to make progress.

Rather, estimating the model's parameters requires only a minimum of two pairs of matching points from the support of each distribution. The pairs can be specific points in the support of the distribution, as well as the mean (even if it is not in the support, as for a categorical random variable), or they might be elicited according to specific percentiles (min, max, median, IQR, or endpoints of a confidence interval). The two modeling parameters, α and β, can be solved exactly from two such pairs of data points, or they can be estimated by regression in cases where more data are available and the system of equations is treated as over-identified with attendant errors. The specification in Equation 1 lends itself naturally to a regression approach.

The basic model can be extended in a behavioral economics setting to include a reference point. In this setting, the shift parameter is applied to a reference point, r, which might be a point in the support of X or a summary statistic derived from the historical distribution. This transformation takes the form:

$$Y = \alpha r + \beta X \,, \tag{3}$$

where X is not adjusted for r by the same reasoning as in the normalization discussion above. In practice, r might be the mean or median, or some other salient point that serves to anchor the forecast. In our application, we apply goodness-of-fit testing to determine that including a reference point outperforms a raw shift parameter, and to select a model in which forecasts are anchored to the reported historical maximum value.

IV. Data

Our data set was originally published in Turvey et al. (2013), in an effort to identify whether crop insurance could be rated and introduced in a setting with limited yield history available. The authors used a survey method to elicit information from corn and wheat farmers across three counties (25 villages) in Shaanxi Province, China, in October and November, 2010. Surveys were administered by 20 Chinese graduate students of the Northwest Agriculture and Forestry University (supervised by faculty researchers), who visited 780 households and collected 731 questionnaires. The completion rate was 93.72%. The validity rates are 78.11% and 98.93% for corn and wheat respectively.

The survey had 9 sections with 117 total questions, which were primarily devoted to understanding financial well-being of, and financial risk-taking by, the farmers. These questions included many potential sources of debt, corresponding interest rates, attitudes toward borrowing, different types of assets, percentage of off-farm income, etc. Basic demographic information was also collected, including gender, education and years of farming experience for the head of household. About 55% of respondents were male, with an average age of 48.72 years, and at least high school completion. On average, respondents had farmed for about 27 years but this ranged from first year farmers to about 60 years. Income averaged 23,796 Yuan per year (about $3,500 at the time) from all sources with the highest being 248,000 Yuan ($37,200). Summary statistics for the demographic information are included in table 1.

[Table 1 about here]

A small section of the survey was dedication to eliciting forecasts of crop yield distributions in the coming year (the 2011 crop year). Each farmer was asked for the minimum and maximum yield that might be achieved in the coming year, as well as the most likely yield outcome. Farmers were asked to make forecasts for both corn and wheat, so the data set contains a forecast for each crop in cases where the farmer was a grower of both. All yield questions were asked in Chinese units of Jin (1.102 lbs.) per Mu (about 0.165 acres, or 0.067 hectares), where one Jin/Mu is equivalent to 0.119 bushels per acre for corn, or 0.112 bushels per acre for wheat. The two key survey questions for eliciting the crop yield distributions are presented in figure 1.

To avoid inducing the farmers to anchor their forecasts to historic experience, the questions about historic experience were administered only *after* the forward-looking forecast information was collected. In this manner, the forecasts are intended to capture farmers' true risk assessments, which still would be based on historical experience, but would not be explicitly anchored to their reported historical experience in the same survey. The historical yield information collected from each farmer (for each crop, if applicable) included the lowest ever and highest ever yields in his/her memory, the years when those yields occurred, and the average yield in their experience. We eliminate incomplete questionnaires, farmers who do not grow corn or wheat, and also farmers with a degenerate distribution (minimum equal to maximum) reported for either their historical experience or their future forecast. Our final tally is 388 data records for corn and 374 data records for wheat, coming from 332 farmers growing both crops, 57 growing only corn, and 45 growing only wheat.

[Figure 1 about here]

The nature of the forecast and historical information elicited was determined by the authors' choice of the well-known Beta-PERT expert elicitation procedure to estimate the forecast and historical distributions (Malcolm et al., 1959; Bewley et al., 2010). The Beta-PERT procedure, originally developed for estimating project completion times, operates by asking experts for the upper bound, lower bound and most likely (modal) outcome. These estimates are then incorporated as parameters of a Beta distribution, resulting in something close to a Normal distribution when the minimum and maximum are symmetric about the modal value, but allowing for skewness otherwise. For more information on the advantages of the PERT method in estimating agricultural yields, see Turvey et al. (2013).

We selected this data set because the elicitation of future and historical minima and maxima are consistent with the data requirements of our model, and because many farmers gave multiple forecasts. The presence of multiple forecasts gives us the opportunity to test whether classifications of forecasts map to classifications of the people making them; in other words, are optimistic forecasts made only by optimistic people and vice versa, or do people make forecasts in different modes according to domain-specific information, such as historical experience?

There are also a number of potential downsides of the data collection method that we try to rule out empirically: i) the elicitation of future mode versus historical mean, ii) the lack of incentive compatibility, and iii) the potential for technological change affecting yields over time. First, the elicitation of the future distribution includes the mode but not the mean, whereas the elicitation of the historical distribution includes only the mean. We are confident this was not due to translation error since the survey was prepared in English, translated into Chinese, and then back-translated into English by two independent bilinguals. All students were trained prior to the survey, and were debriefed twice daily by attending faculty while in the field. Rather, the modal value was chosen for the forecast because it is specified in the Beta-PERT methodology, and it is likely the mean was chosen for the historical data because of relevance. With yields averaging 800-1,000 Jin/Mu for many farmers, the historical values might all have been unique (farmers were asked to round values to the nearest 10 Jin/Mu), rendering a request for the historical mode ambiguous. In the remainder of this article, we test inclusion of the mode/mean pair as a data point for regression fitting of our basic two parameter model, and find that it does not materially change the results. We also test a 3-parameter model that includes an adjustment for skewness net of the shift and scale parameters. While we find that subjects report, on average, a forecast mean about 5% higher than their historic mode (after accounting for the shift and scale adjustments), we find no meaningful change in goodness of fit or in the number of mixture components selected.

A key criticism of the elicitation method for these data is the lack of incentive compatibility. Namely, farmers were not incentivized to report truthfully and may have been ambivalent over multiple different values they might report. However, to the extent that farmers might have recognized they were being asked in order to estimate crop insurance premiums, they did not respond systematically to that incentive. A reasonable approach for a policy maker considering a new crop insurance program would be to use historical data to price the insurance (in the absence of trends due to technological change, which we will address momentarily), and to use farmers' forecasts to estimate insurance demand. As noted above, Turvey et al. (2013) were the first to explicitly note that forecast disparities from historical experience implied subsidies were needed, and their article was published after the survey data collection took place (it's their data). If farmers were systematically trying to influence crop insurance premiums in their favor (i.e., to make premiums cheaper), then they would have underreported their historic risks, and if they were trying to increase the likelihood that insurance would be offered, then they would have over-reported forward-looking risks. Instead, in our data set, the majority of farmers reported higher yields and lower risk in their forecasts than their own historical experience would imply. It is unlikely that these features of the data are attributable to perceived incentives, since in combination they are likely to lead to both a lower probability of crop insurance being offered, as well as higher premiums for them insurance when it is made available. We also point out that the Beta-PERT method has not been historically applied in an incentive compatible setting; the farmers are generally perceived as having expert knowledge in this domain, and standard practice expects them to reveal that knowledge truthfully. For these reasons, once trends from technological change can be ruled out, we expect that the lack of incentive compatibility might result in more *noisy* elicitations but not *biased* elicitations.

When modeling crop yields in the United States, it is common knowledge that technological improvements have played a role in dramatically increasing yield per acre over time. In the case of corn, national average yields per acre have increased by more than 7x in the last 100 years, in a manner that distinctly resembles a linear trend line. Simulated yield distributions often feature a detrended component (e.g., as in Cooper 2010), and crop insurance policies are available with a "trend adjustment" for similar reasons (RMA 2011). However, crop yields in China have not responded nearly as aggressively to technological change over time as their American counterparts. A further mitigating factor in our data set is the relatively high proportion of economically disadvantaged smallholder farmers, many of which do not report on-farm income as their primary income source. It is unlikely that these farmers have easy access to the latest seed, capital equipment and planting technologies that might be available to more commercialized operations.

Given this ambiguity, we are fortunate that an explicit statistical test is available: we can use the estimated parameters from each forecast, $\hat{\alpha}$ and $\hat{\beta}$, as dependent variables, and regress them (separately) on years farming and other covariates. This test is valid, in part, because we have already observed that farmers are optimistic on average, in the sense of predicting future yields that are higher and less variable. If the apparent optimism is simply due to yield trends arising from technological change, then we would expect that either $\hat{\alpha}$ or $\hat{\beta}$, or both, would vary systematically with the number of years spent farming (all farmers were currently engaged in farming at the time of the survey). In other words, more experience farming means both more time for extreme events to occur, and also more time for average yields to differ materially from the present. The details of these tests and their associated robustness checks are found in table 9 of our section on Factors Influencing Forecasts: we find no evidence that a farmer's estimated optimism is a function of time. Thus, we conclude that technological change is not a material factor affecting the measurement of optimism in our data set.

Model Selection

To identify whether a raw shift and scale model (Equation 1) or one incorporating a reference point (Equation 3) is more appropriate, we test regression fits for each and compare them using the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). In particular, we test the model from Equation 1 alongside several variants of the model from Equation 3, considering that the reference point might be any of the reported values from the historic distribution (the minimum, maximum, or mean), as well as the midpoint between the historic minimum and maximum. The regression results are presented in table 2. Both goodness of fit criteria unambiguously selected Model 2, indicating that the best fit is achieved using the reported historical maximum as a reference point. This is, in and of itself, a finding of optimism in the sense that forecasts are apparently anchored to the upside of historical experience.

[Table 2 about here]

As discussed above, there were three data points elicited from each farmer for each of their historical and forecast distribution. They were asked for the historical minimum, maximum and mean yields, and they were asked for minimum, maximum and modal yields in the upcoming year. The results in table 2 use only the minimum and maximum, reported for both future and historic, since they are clearly representative of the same points in the distribution. In fact, even if farmers were only reporting endpoints of a confidence interval (e.g., the 2.5^{th} and 97.5^{th} percentiles), the data are still valid for our approach. The same validity does not apply when comparing the historical mean against the forecast mode: we have no reasonable assurance that one will map to the other according to our basic shift and scale model. Nonetheless, we find some evidence that they are not far off, and that the subjects may not be differentiating between the two. To show this, and as a robustness check, we re-estimated the results in table 2 using all three data points per forecast in the regression. The new model results are shown in table 3: it is clear that the results are materially unchanged by the addition of mean/mode reporting.

[Table 3 about here]

Identifying Heterogeneity with a Gaussian Mixture Model

An examination of the empirical densities over the population of forecasts for estimates of α and β reveals substantial heterogeneity. Of particular interest, the empirical marginal densities appear to be multi-peaked and a joint plot indicates that the peaks may coincide. This is shown in figure 2, which is a bivariate hex-plot of the joint density, with marginal kernel densities along each edge. Accompanying summary statistics for forecast parameters by crop and by gender are presented in table 4. Together, these features give reason for concern that heterogeneity in the form of clustering may be present, which could be better represented by a mixture distribution over types or “classes” of forecasts. If so, simply reporting the population-mean values (implicitly as the center point of a single-peaked distribution) may result in estimates that do not represent any one peak in the distribution (e.g., as noted in Sproul and Michaud, 2017). To conduct an explicit test of whether the observed phenomenon is statistically meaningful, we compare goodness-of-fit statistics from a set of (Finite) Gaussian Mixture Models (GMMs) fitted with different numbers of components.

[Table 4 about here]

[Figure 2 about here]

Each GMM is fit using the well-known expectation-maximization (EM) algorithm of Dempster, Laird and Rubin (1977), a Bayesian procedure comprised of two steps. In the expectation step ("E-step"), the likelihood function is used to calculate so-called 'membership probabilities,' τ_{ic}, denoting the probability that each individual *i* belongs to each type (or 'class') *c*. The average of these membership probabilities becomes the updated mixture probability for each class, π_c. In the maximization step ("M-step"), the updated mixture probabilities are held fixed while the log likelihood is maximized by varying the parameters for each class, collectively referred to as the vector, θ_c. After the M-step, the algorithm repeats until suitable convergence is achieved.

Formally, let *t* denote an iteration of the algorithm. In the E-step, the updated membership probabilities for step *t + 1,* for each individual, *i*, and each class, *c,* are given according to:

$$\tau_{ic}^{t+1} = \frac{\pi_c^t \cdot f(x_i \mid \hat{\theta}_c^t)}{\sum_{j=1}^{C} \pi_j^t \cdot f(x_i \mid \hat{\theta}_j^t)} \ . \tag{4}$$

Here, x_i are the data for individual *i*, $\hat{\theta}_c^t$ are the most recent maximum likelihood estimates for the parameters describing type *c* (from the previous step), and *f* denotes the likelihood of the data given the estimated parameters. The updated mixture probabilities for each type (at the population level) are then simply the averages of the membership probabilities:

$$\pi_c^{t+1} = \frac{1}{I}\sum_{i=1}^{I} \tau_{ic}^{t+1} . \tag{5}$$

The M-step then maximizes the log-likelihood function, holding the mixture probabilities constant. The updated estimate, $\hat{\theta}^{t+1}$, solves

$$\max_{\theta} LL(\theta;\pi) = \sum_{i=1}^{I} \ln\left(\sum_{c=1}^{C} \pi_c^{t+1} \cdot f\left(x_i \mid \theta_c\right)\right). \tag{6}$$

Recall that the vector θ_c represents the parameters collectively describing the multivariate normal for each type, *c*. Since our forecasting model is a 2-parameter model, each vector θ_c contains two mean parameters, as well as two additional parameters to populate the covariance matrix, which we assume to be diagonal following Bruhin, Fehr-Duda and Epper (2010).

A known shortcoming of applying expectation-maximization to fit a mixture model is the researcher must specify *ex ante* the number of mixture components (types). The algorithm does have some capacity to "zero out" redundancies by giving near-zero weights (the mixture probabilities, π_c) to extraneous classes as it endogenously determines the classification of data points. In practice, however, specification of too many mixture components can lead to over-fitting, exhibited by ambiguous membership probabilities (e.g., τ_{ic} close to 0.5).

Information criteria such as the corrected Akaike information criterion (AICc), BIC, or comparable cross-validation approaches, are often used for model selection but may be inadequate in mixture model applications due to insufficient penalization of extra parameters (Celeux and Soromenho, 1996; Biernacki et al., 2000). The latter authors introduce the integrated completed likelihood (ICL) criterion, a modification of BIC with an additional penalty for entropy with the goal of achieving better out-of-sample classification. More entropy means more ambiguous assignments to the various classes, corresponding to τ_{ic} values that are far from both 0 and 1, an indication that the fitted mixture model is not effectively classifying the data into distinct types. In effect, ICL adjusts BIC for the ambiguity of classifications, and this ambiguity is conveniently measured by the posterior membership probabilities estimated in the course of the EM procedure. As an additional check on the robustness of classification, Sproul and Michaud (2017) also report percentages of individuals with at least one posterior membership probability greater than 0.90, 0.95 and 0.99, respectively.

Our estimation results for GMMs with up to 6 mixture components are presented below in table 5. We present results separately for corn forecasts (*N=442*), for wheat forecasts (*N=437*) and pooled (*N=879*), with boldface indicating model selection for each criterion. A number of features of the model selection process are apparent, including i) the selection by ICL of more parsimonious models with fewer components, ii) the ability of the EM algorithm to "zero out" extraneous outlier groups, iii) the "preference" of AICc and BIC for more components due to increases in the likelihood function, and iv) the consistent "preference" of ICL for models with high rates of unambiguous classifications, as evidenced by the percentage of individuals classified with τ_{ic} close to 1. Across data sets, each criterion selects $C > 1$ components, providing empirical justification of our mixture model approach to examine heterogeneity.

[Table 5 about here]

For each data set, ICL is minimized with a 4-component mixture, comprised of three main groups/classes and one small outlier group. The three larger classes are roughly the same size in each of the corn, wheat, and pooled data sets, and as will be seen below, they correspond to roughly identical model parameters as well. These classes correspond to optimistic forecasts (67%), neutral forecasts (20%) and pessimistic forecasts (11%), for which the interpretation of parameters into labels will be discussed momentarily. The final outlier group (2%) includes extremely pessimistic forecasts that dramatically differed from their peers. Because of the consistent small size of the outlier group, we will treat these results as representative of a 3-component mixture with outliers in what follows. The class breakdowns, indicated by the fitted mixture weights, are presented by crop and by number of components in table 6.

[Table 6 about here]

It can also be seen across the data sets that AICc and BIC are almost perfectly monotonically decreasing in the number of components. This is the very observation that led to the development of ICL by Biernacki et al. (2000), who, like others, observed that these criteria will very often select the maximum number of components in a given choice set. In fact, we tested up to 8 components and found that AICc and BIC would select the maximum number of components if it were 7 or 8, as well. Those extra tests are omitted because they are otherwise uninformative, and to save space.

In their 2010 article, Bruhin et al. indicate that a desirable feature of mixture model classification is the reliable "splitting" of classes into economically meaningful subgroups as the number of model components increases. In other words, it is an indicator of model quality if more model components, whether chosen by AICc, BIC, ICL or some other criterion, tend to subdivide the space of individuals being classified instead of adding new classes which adopt members out of multiple existing groups.

We observe this feature here. When increasing from a single component to a 2-component mixture, the pessimistic group is split off from the rest in each data set. For the corn forecasts, increasing from 2 to 3 components gives us our three main groupings of optimistic, pessimistic and neutral, whereas in wheat and in the pooled data, adding a third component results in splitting off the outliers (extremely pessimistic) from the pessimistic group. Adding a fourth component in corn splits off outliers from the pessimistic group, while in wheat and pooled we observe non-pessimists splitting into the optimistic and neutral groups. Across data sets, adding a fifth component results in a further split of the already-tiny outlier group, and adding a sixth component results in division of the optimistic group into mildly optimistic and extremely optimistic.

In each case described here, the parameter estimates for α and β support our characterizations of the groups in a manner indicative of consistent subdivisions when the number of components is increased. Of particular interest are the mixture models with four and six components respectively, which are best characterized as a 3-component mixture without outliers (due to the 4^{th} outlier group), and a 4-component mixture with outliers (due to the outlier group being subdivided into two smaller outlier groups). We now explore the parameters of these two models in detail. Contour plots visualizing our mixture results are shown in figure 3.

[Figure 3 about here]

[Table 7 about here]

Table 7 describes the 3-component mixture of optimistic, neutral (or "realistic") and pessimistic forecasts with the outlier group discarded. The optimistic group gets its label because it is characterized by (α, β) parameters of approximately (0.45, 0.54) across the data sets, indicating an upward shift of 45% of the reference point (the historic max) and reduction of range (risk) to 54% of historic. This group, comprising approximately 67% of forecasts, is therefore anticipating higher yields with less risk than historical experience dictates. The neutral, or realistic, group has mean (α, β) parameters of approximately (0.02, 1.00) across the data sets, indicating that approximately 20% of forecasts are almost perfectly in line with historical experience. Finally, the pessimistic group represents approximately 11% of forecasts with estimated parameters of (-0.43, 1.56) on average across the population, indicating expectations of lower and more variable yields than historical experience implies.

[Table 8 about here]

Table 8 describes the 4-component mixture of highly optimistic, mildly optimistic, realistic and pessimistic forecasts with two small outlier groups discarded (this was the fitted 6-component mixture from table 6). The key distinction between information presented in table 7 versus table 8 is the subdivision of the optimistic group into two groups, highly optimistic and mildly optimistic, for which the splits are relatively stable across corn, wheat and pooled forecasts. The highly optimistic group comprises about 35% of forecasts overall, with central (α, β) parameter estimates of (0.58, 0.40) indicating expected yields even higher and even less risky than those indicated by the mean estimates of the consolidated optimistic group from the 3-component model. The mildly optimistic group, on the other hand, exhibits parameter estimates of (0.31, 0.70) indicating a more neutral or realistic outlook but still clearly optimistic (30% reduction in risk over neutral). The mildly optimistic group also has consistent parameters across the data sets, and comprises 31% of forecasts overall.

Factors Influencing Forecasts

Having identified distinct types of forecasts, it is worthwhile to examine how these differing biases (or lack thereof) might arise. There are two key questions. First, are forecast types actually revealing types of the people making them? That is, must an optimistic forecast necessarily come from an optimistic person? Second, to the extent that forecast types might differ within individuals, what are the key sources of variation?

One reason for choosing our data set was the presence of multiple forecasts for many subjects interviewed, which gives the opportunity to test whether forecast classifications are generally consistent within individuals. A review of the available evidence makes it appear that they are not consistent, at least in one sense. If we take the view that the posterior membership probabilities, τ_{ic}, arising from the EM procedure represent *a posteriori* classification probabilities, then the forecast types cannot represent types of subjects. Simply put, nearly 20% of subjects growing both crops gave forecasts with distinct types, which is nearly impossible given the results in Table 5 (*Pooled Data, C=4*), where 95% of forecasts had $\tau_{ic} \geq 0.99$ and 99% had $\tau_{ic} \geq 0.95$. We are forced to conclude that variation in forecast type, while it may be influenced by individual factors, might also be affected by crop specific factors. To test for key sources of variation in forecast types, we conducted a number of regression tests depicted in table 9.

[Table 9 about here]

The Effect of Historic Yield Outcomes

Perhaps our most important finding pertains to the effect of historic outcomes on a farmer's level of optimism (in terms of the scale factor, β). The results show that for each year that a farmer goes without suffering a historic loss (i.e. their reported minimum historic yield), he or she becomes increasingly optimistic. For each year that passes since a historic loss, the average farmer reduces the relative scale (β) of their future yield risk by 1% (in the optimistic direction). Effects for the shift parameter (α) are in the opposite direction nominally, but also optimistic.

While these numbers may appear small, the effect can be dramatic. As shown in Table 1, the average time since a historic loss is 8.56 years (σ=8.48), with 3 years as the lower bound of the interquartile range (IQR), and 11 years as the upper bound. All else equal, the relative scale (compared to their own historic distribution) for the future yield forecast of a farmer at the top of the IQR for historic losses will be eight percentage points below that of a farmer who suffered a historic loss 3 years ago (eight percentage points is equivalent to a 13% deviation of β below the mean).

In contrast, future forecasts were *not* significantly affected by recent events in the gains domain (i.e. recent maximum historic yields). This suggests a finding similar to the perspective pervasive in prospect theory, that losses are more salient that gains. Explicitly testing this hypothesis, however, is beyond the scope of both our data and this paper.

We tested a number of history-related features of the data set, including cross-crop experience, but did not find statistically significant prediction from any of them, apart from recent severe losses in the crop being forecast. We also tested crop-level fixed effects and found no significant variation, implying that recent loss history is likely the dominant crop-specific factor in determining differential forecasts in our data set. At least in the case of cross-crop experiences, it is possible that effects are not identifiable due to correlation induced by weather. Naturally, strong positive correlation due to weather would result in collinearity in our regression tests; this is shown in figure 4, which plots the years since a historic loss in corn versus in wheat for those farmers growing both.

[Figure 4 about here]

Gender Differences in Forecasts

In addition to crop-specific effects arising from loss history, we also observe subject-specific effects in the form of statistically significant differences in parameter estimates by gender. Our regression results show that on average, women appear to be even more optimistic with regard to yield risk than their male counterparts. While male forecasters expect future yields to have 80% the scale of their historic yield distribution (i.e., $\hat{\beta} = 0.80$), female forecasters expect future yields to come from a distribution with roughly only 71% the scale of their historic distribution (i.e., $\hat{\beta} = 0.71$). This finding is consistent across both crops, despite the presence of within-subject heterogeneity in beliefs between crops. Our results regarding gender differences in forecast tend to stand somewhat in contrast with the literature on risk perceptions. Namely, the consensus is that on average, risk tends to be judged as lower by men than by women (see, for example, Gutteling and Wiegman, 1993; Stern et al., 1993; Flynn et al., 1994; Slovic et al., 1997; Finucane et al., 2000).

Flynn et al. (1994) and Finucane et al. (2000), in particular, find that both gender and race plays a critical role in making judgments about risk. In Flynn et al. (1994), 1,512 Americans were asked to rate their level of perceived risk for 25 (environmental) hazards. For all 25 hazards, white males' risk perception ratings were consistently much lower than the means of non-white males, and females, with non-white males and females generally being relatively similar in their perceptions of risk. In addition, this so-called `white male' effect appeared due to a group of roughly 30% of the white male sample who judged risks to be extremely low.

Finucane et al. (2000) expanded Flynn et al.'s analysis using data collected as part of a national telephone survey designed to test hypotheses about risk perceptions over a range of hazards. The survey contained questions about worldviews, trust, and a range of demographic variables. In addition to a 'white' dummy, they included three additional racial/ethnic groups: African-American, Hispanic, and Asian. They replicated the findings of Flynn (1994), and also found that risk perceptions varied considerably across African-American, Asian, and Hispanic males and females, implying that risk perceptions depend importantly on the characteristics of the individuals facing the risk. Their findings support the view that differences cannot be explained entirely from a biological perspective but rather that risk is a social construct that depends largely on the characteristics of the individuals facing the risk.

From the prospect theory literature, there has also been evidence of gender differences in risk preferences (as opposed to risk perception) and in probability weights.

For example, Fehr-Duda, de Gennaro and Schubert (2006) and Bruhin, Fehr-Duda and Epper (2010) both find that female subjects statistically differ in their application of probability weights from male subjects. Namely, the female subjects tended to show a substantially lower slope of the probability weighting function, indicating less "rational" probability weights in the sense of their elicited prospect values being relatively insensitive to marginal changes in probability of payment. Tanaka, Camerer and Nguyen (2010) do not find significant gender differences in risk or time preferences, but Liu (2013) does find that females were significantly more risk averse, in the sense of the curvature (exponent) of the prospect theory value function using the same methodology and a larger data set.

To summarize this discussion, our results indicate that women tend to forecast less risk than men, all else equal, but the literature on risk perception suggests that they generally perceive risks to be more severe. At the same time, research on prospect theory suggests that women are more averse to risk than men, but that they are insensitive to marginal changes in risk for probabilities that are not close to certain. Clearly, further research is needed to determine whether these associations are unified by some underlying mechanism or model.

Extending the Model to Include a Shape Parameter

While we have shown our basic shift-and-scale model to be fairly robust, we have not thus far tackled the question of whether the model might include a meaningful shape parameter to account for changes in higher moments. In particular, it might be the case the forecasts are not only classified according to optimistic, realistic or pessimistic shifting and scaling, but also that these classifications vary meaningfully with respect to shape changes in the transformation from the historical distribution to the forecast. To address this concern, we extend the model from a 2-parameter shift and scale form to a 3-parameter shift-scale-and-shape form. Specifically, we include a third parameter, γ, mapping the historic mean into the future mode, net of changes already induced by the shift and scale parameters, α and β. In terms of our regression specification, the shape parameter is interacted with a dummy, $\mathbf{1}_{mid}$, equal to one for rows in which central points of the historic and forecast distribution are matched:

$$Y = \alpha r + \beta X + \mathbf{1}_{mid} \cdot \gamma X \quad . \tag{7}$$

It is worth repeating that any results obtained with this model on our current data are exploratory in nature: our data set is based on an elicitation method that does not match mean-to-mean or mode-to-mode for the historical and forecast distributions, but only allows (potentially) matching historic mean to forecast mode. That said, our earlier results are remarkably robust to this extension of our model. Table 10 details the goodness of fit evaluation and the mixture weights in Panel A, and presents the mixture means in Panel B. The only meaningful difference between fitting our model with the shape parameter, versus without, is that there are no longer small outlier groups in the 3-component and 4-component models selected respectively by ICL and AICc/BIC. The group weights and their central parameter estimates remain virtually unchanged. In all cases, it can be clearly seen that the shape parameter is near zero, reaching its maximum of about a 3% increase (from mean to mode) in the pessimistic group.

[Table 10 about here]

Conclusion

In this article we develop a simple model of probabilistic forecasts to examine overconfidence. We use an explicit location-scale model of forecasts acting as transformations of historic experience, with parameters that may incorporate a reference point. We find that forecasts are anchored to historical positive experience, and that forecasts are systematically overconfident. However, like Sproul and Michaud (2017), we find evidence of multi-modal heterogeneity such that population average parameter estimates do not represent individuals well. We estimate a finite Gaussian mixture model using expectation-maximization (Dempster, Laird and Rubin 1977) and find strong evidence for three basic forecast types: optimistic (about two thirds), realistic or neutral (about 20%), and pessimistic (the remainder). We find further clustering within the optimistic group, without about half highly optimistic and about half mildly so. The group-wise means and mixture weights are robust to inclusion of additional elicited data, and also to inclusion of a shape parameter in the forecast model. We also find evidence that forecast classifications do not perfectly map to a classification of the person making it. Instead, recent severe loss experience appears to be the strongest single predictor of pessimism, though we do find statistically significant gender differences. Our results have important implications regarding the need for crop insurance subsidies to induce participation. Future research should focus on discovering potential ways to reduce the discrepancy between subjective and objective risk.

Works Cited

Babcock, Linda, George Loewenstein, and Samuel Issacharoff. "Creating Convergence: Debiasing Biased Litigants." *Law & Social Inquiry* 22.4 (1997): 913-925

Barseghyan, Levon, Et Al. "The Nature Of Risk Preferences: Evidence From Insurance Choices." *The American Economic Review* 103.6 (2013): 2499-2529.

Bartl, G., and G. Stummvoll. "Description Of Post Licensing Measures In Austria." Dan-Report. Results Of Eu-Project: Description and Analysis Of Post Licensing Measures For Novice Drivers. *Kuratorium Fur Verkehrssicherheit*, Wien (2000).

Benartzi, Shlomo, and Richard H. Thaler. "Naive Diversification Strategies In Defined Contribution Saving Plans." *American Economic Review* (2001): 79-98.

Bewley, J. M., M. D. Boehlje, A. W. Gray, H. Hogeveen, S. J. Kenyon, S. D. Eicher, and M. M. Schutz. "Assessing The Potential Value For An Automated Dairy Cattle Body Condition Scoring System Through Stochastic Simulation." *Agricultural Finance Review* 70.1 (2010): 126-150.

Bhattacharya, Jayanta, Et Al. "Disability Forecasts and Future Medicare Costs." *Frontiers In Health Policy Research* 7 (2004): 75-94.

Biernacki, Christophe, Gilles Celeux, and Gérard Govaert. "Assessing A Mixture Model For Clustering With The Integrated Completed Likelihood." *Ieee Transactions On Pattern Analysis and Machine Intelligence* 22.7 (2000): 719-725.

Bruhin, Adrian, Helga Fehr-Duda, and Thomas Epper. "Risk and Rationality: Uncovering Heterogeneity In Probability Distortion." *Econometrica* 78.4 (2010): 1375-1412.

Buckley, Neil J., Et Al. "The Transition From Good To Poor Health: An Econometric Study Of The Older Population." *Journal Of Health Economics* 23.5 (2004): 1013-1034.

Camerer, Colin, and Dan Lovallo. "Overconfidence and Excess Entry: An Experimental Approach." *The American Economic Review* 89.1 (1999): 306-318.

Cebulla, andreas. "A Geography Of Insurance Exclusion: Perceptions Of Unemployment Risk and Actuarial Risk Assessment." *Area* 31.2 (1999): 111-121.

Celeux, Gilles, and Gilda Soromenho. "An Entropy Criterion For Assessing The Number Of Clusters In A Mixture Model." *Journal Of Classification* 13.2 (1996): 195-212.

Chiappori, Pierre-André, Et Al. "Asymmetric Information In Insurance: General Testable Implications." *The Rand Journal Of Economics* 37.4 (2006): 783-798.

~~Claassen, Roger, and Richard E. Just. "Heterogeneity and Distributional Form Of Farm-Level Yields." *American Journal Of Agricultural Economics* 93.1 (2011): 144-160.~~

Cooper, Arnold C., Carolyn Y. Woo, and William C. Dunkelberg. "Entrepreneurs' Perceived Chances For Success." *Journal Of Business Venturing* 3.2 (1988): 97-108.

Cooper, Joseph C. "Average Crop Revenue Election: A Revenue-Based Alternative To Price-Based Commodity Payment Programs." American Journal Of Agricultural Economics (2010): Aaq048.

Cutler, David M., and Richard Zeckhauser. "Extending The Theory To Meet The Practice Of Insurance." *Brookings-Wharton Papers On Financial Services* 2004.1 (2004): 1-53.

Dalziel, James R., and Rf Soames Job. "Motor Vehicle Accidents, Fatigue and Optimism Bias In Taxi Drivers." *Accident Analysis & Prevention* 29.4 (1997): 489-494.

De Bondt, Werner Fm, and Richard H. Thaler. "Financial Decision-Making In Markets and Firms: A Behavioral Perspective." *Handbooks In Operations Research and Management Science* 9 (1995): 385-410.

Dempster, Arthur P., Nan M. Laird, and Donald B. Rubin. "Maximum Likelihood From Incomplete Data Via The Em Algorithm." *Journal Of The Royal Statistical Society. Series B (Methodological)* (1977): 1-38.

Egelkraut, Thorsten M., Philip Garcia, Joost ME Pennings, and Bruce J. Sherrick. "Producers' Yield and Yield Risk: Perceptions versus Reality and Crop Insurance." (2006).

Enjolras, Geoffroy, and Patrick Sentis. "Crop insurance policies and purchases in France." *Agricultural Economics* 42.4 (2011): 475-486.

Fehr-Duda, Helga, Manuele De Gennaro, and Renate Schubert. "Gender, financial risk, and probability weights." *Theory and decision* 60.2 (2006): 283-313.

Finucane, Melissa L., Paul Slovic, Chris K. Mertz, James Flynn, and Theresa A. Satterfield. "Gender, Race, and Perceived Risk: The 'White Male' Effect." *Health, Risk & Society* 2.2 (2000): 159-172.

Flynn, James, Paul Slovic, and Chris K. Mertz. "Gender, Race, and Perception Of Environmental Health Risks." *Risk Analysis* 14.6 (1994): 1101-1108

Greenspan, Alan. "The Challenge Of Central Banking In A Democratic Society." Annual Dinner and Francis Boyer Lecture Of The American Enterprise . Washington, D.C. 5 Dec. 1996. The Federal Reserve Board. Web. 23 May 2017.

Groeger, John A., and Gunn E. Grande. "Self-Preserving Assessments Of Skill?." *British Journal Of Psychology* 87.1 (1996): 61-79.

Güner, A. Burak, Ulrike Malmendier, and Geoffrey Tate. *The Impact Of Boards With Financial Expertise On Corporate Policies*. National Bureau Of Economic Research, 2006.

Gutteling, Jan M., and Oene Wiegman. "Gender-Specific Reactions To Environmental Hazards In The Netherlands." *Sex Roles* 28.7 (1993): 433-447

~~Harri, Ardian, Cumhur Erdem, Keith H. Coble, and Thomas O. Knight. "Crop Yield Distributions: A Reconciliation Of Previous Research and Statistical Tests For Normality." *Review Of Agricultural Economics* (2009): 163-182.~~

Hatakka, Mika, Et Al. "From Control Of The Vehicle To Personal Self-Control; Broadening The Perspectives To Driver Education." *Transportation Research Part F: Traffic Psychology and Behaviour* 5.3 (2002): 201-215.

Hoch, Stephen J. "Counterfactual Reasoning and Accuracy In Predicting Personal Events." *Journal Of Experimental Psychology: Learning, Memory, and Cognition* 11.4 (1985): 719.

Hoorens, Vera. "Unrealistic Optimism In Social Comparison Of Health and Safety Risks." *The Social Psychology Of Health and Safety* (1994): 153-174.

Horowitz, John K., and Erik Lichtenberg. "Insurance, moral hazard, and chemical use in agriculture." *American Journal of Agricultural Economics* 75.4 (1993): 926-935.

Kreuter, Matthew W., and Victor J. Strecher. "Changing Inaccurate Perceptions Of Health Risk: Results From A Randomized Trial." *Health Psychology* 14.1 (1995): 56.

Landier, Augustin, and David Thesmar. "Financial Contracting With Optimistic Entrepreneurs." *Review Of Financial Studies* 22.1 (2009): 117-150.

Liu, Elaine M. "Time to change what to sow: Risk preferences and technology adoption decisions of cotton farmers in China." *Review of Economics and Statistics* 95.4 (2013): 1386-1403.

Malcolm, D. G., J. H. Roseboom, C. E. Clark, and W. Fazar. "Application Of A Technique For Research and Development Program Evaluation." *Operations Research* (1959): 646-669.

Malmendier, Ulrike, and Geoffrey Tate. "Who Makes Acquisitions? Ceo Overconfidence and The Market's Reaction." *Journal Of Financial Economics* 89.1 (2008): 20-43.

Pease, James W. "A Comparison Of Subjective and Historical Crop Yield Probability Distributions." *Southern Journal Of Agricultural Economics* 24.02 (1992): 23-32.

RMA (United States Department of Agriculture Risk Management Agency). “Trend-adjusted Actual Production History (APH) Procedures.” (2011). Accessed online May 24, 2017: https://www.rma.usda.gov/bulletins/pm/2011/11-046b.pdf.

Sandroni, Alvaro, and Francesco Squintani. "Overconfidence and Asymmetric Information: The Case Of Insurance." *Journal Of Economic Behavior & Organization* 93 (2013): 149-165.

Sandroni, Alvaro, and Francesco Squintani. "Overconfidence, Insurance, and Paternalism." *The American Economic Review* 97.5 (2007): 1994-2004.

Shaik, Saleem, Keith Coble, Thomas Knight, Alan Baquet, and George F. Patrick. "Crop Revenue and Yield Insurance Demand: A Subjective Probability Approach." *Journal Of Agricultural and Applied Economics* 40, No. 3 (2008): 757-766.

Sherrick, Bruce J., Et Al. "Factors Influencing Farmers' Crop Insurance Decisions." *American Journal Of Agricultural Economics* 86.1 (2004): 103-114.

Shiller, Robert J. *Irrational Exuberance*. Princeton, N.J.:Princeton University Press, 2000.

Slovic, Paul. *The Perception Of Risk (Risk, Society and Policy)*. London: Earthscan Publications, 2000.

Slovic, Paul, Torbjörn Malmfors, C. K. Mertz, Nancy Neil, and Iain Fh Purchase. "Evaluating Chemical Risks: Results Of A Survey Of The British Toxicology Society." *Human & Experimental Toxicology* 16.6 (1997): 289-304.

Spinnewijn, Johannes. "Heterogeneity, Demand For Insurance, and Adverse Selection." *American Economic Journal: Economic Policy* 9.1 (2017): 308-343.

Spinnewijn, Johannes. "Insurance and Perceptions: How To Screen Optimists and Pessimists." *The Economic Journal* 123.569 (2013): 606-633.

Spinnewijn, Johannes. "Unemployed But Optimistic: Optimal Insurance Design With Biased Beliefs." *Journal Of The European Economic Association* 13.1 (2015): 130-167.

Sproul, Thomas, and Clayton Michaud "Heterogeneity In Loss Aversion: Evidence From Field Elicitations." *Agricultural Finance Review* 77.1 (2017): 196-216.

Statman, Meir, and Tyzoon T. Tyebjee. "Optimistic Capital Budgeting Forecasts: An Experiment." *Financial Management* (1985): 27-33.

Stern, Paul C., Thomas Dietz, and Linda Kalof. "Value Orientations, Gender, and Environmental Concern." Environment and Behavior 25.5 (1993): 322-34

Svenson, Ola. "Are We All Less Risky and More Skillful Than Our Fellow Drivers?." *Acta Psychologica* 47.2 (1981): 143-148.

Tanaka, Tomomi, Colin F. Camerer, and Quang Nguyen. "Risk and time preferences: linking experimental and household survey data from Vietnam." *American Economic Review* 100.1 (2010): 557-571.

Turvey, C. G., Gao, X., Nie, R., Wang, L., & Kong, R. "Subjective Risks, Objective Risks and The Crop Insurance Problem In Rural China." The Geneva Papers On Risk and Insurance Issues and Practice 38.3 (2013): 612-633.

Wang, Ming, Tao Ye and Peijun Shi. "Factors Affecting Farmers' Crop Insurance Participation in China." *Canadian Journal of Agricultural Economics.* 64 (2016): 479-492.

Weinstein, Neil D. "Unrealistic Optimism About Future Life Events." *Journal Of Personality and Social Psychology* 39.5 (1980): 806.

Wenglert, Leif, and Anne-Sofie Rosen. "Measuring Optimism–Pessimism From Beliefs About Future Events." *Personality and Individual Differences* 28.4 (2000): 717-728.

Figures

Figure 1: Turvey et al. (2013) survey questions

If you grow corn or wheat, identify the lowest yield you believe possible, the yield that you believe is most likely to be received, and the highest possible yield you believe possible (jin/mu) in ***the next crop year (2010/11)*** If you do not recall exacts, please answer to nearest within 10 jin/mu

Crop	Lowest possible yield (jin/mu)	Most likely yield (jin/mu)	Highest possible yield (jin/mu)
1 Corn			
2 Wheat			

If you grow corn and wheat, what is the lowest and highest yield (jin/mu) that you recall from your years in farming? If you do not recall exacts, please answer to nearest within 10 jin/mu.

Crop	Lowest historical yield (jin/mu)	Year it occurred	Highest historical yield (jin/mu)	Year it occurred	Average yield across year
1 Corn					
2 Wheat					

Figure 2: Hex plot with marginal densities

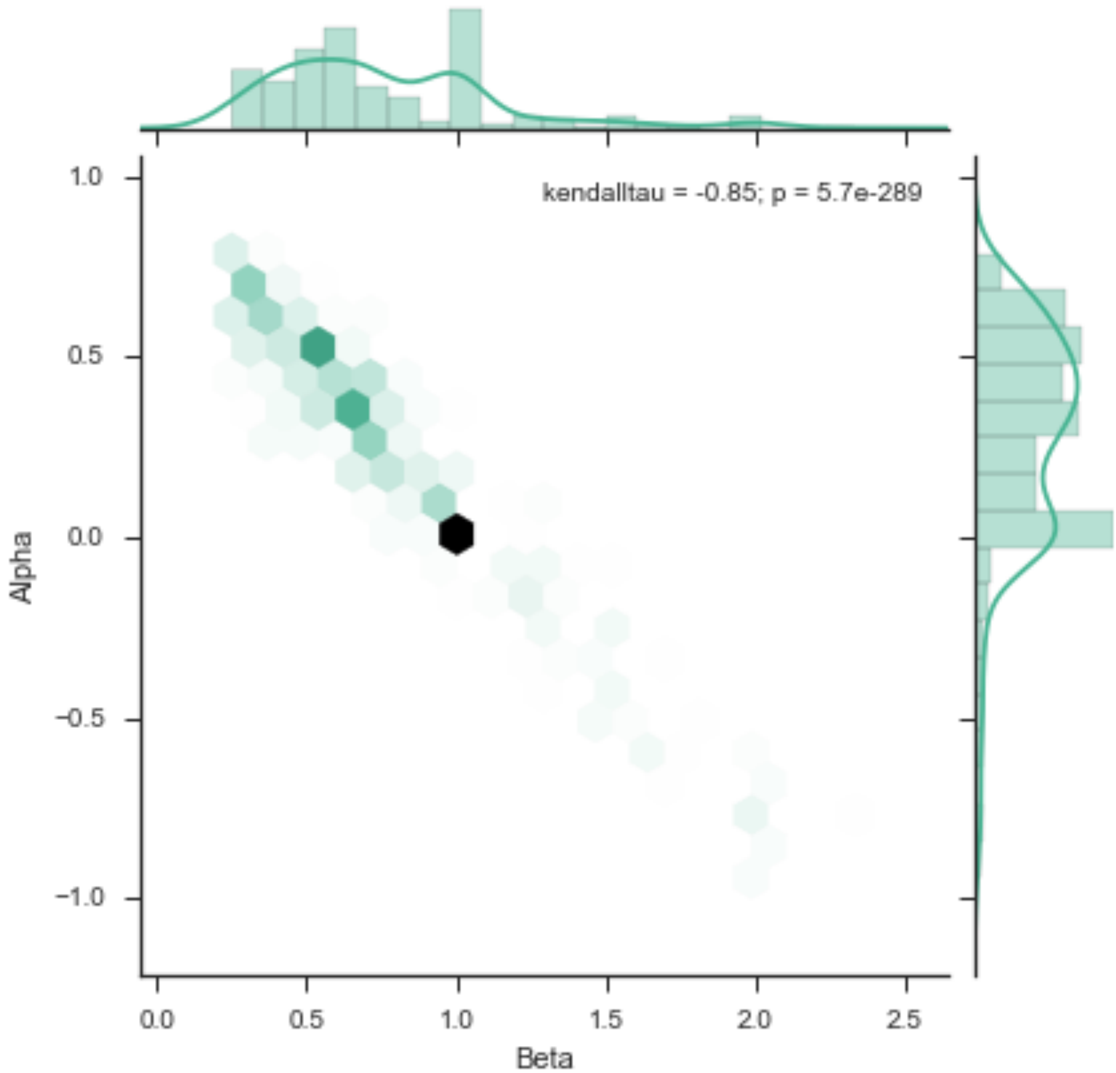
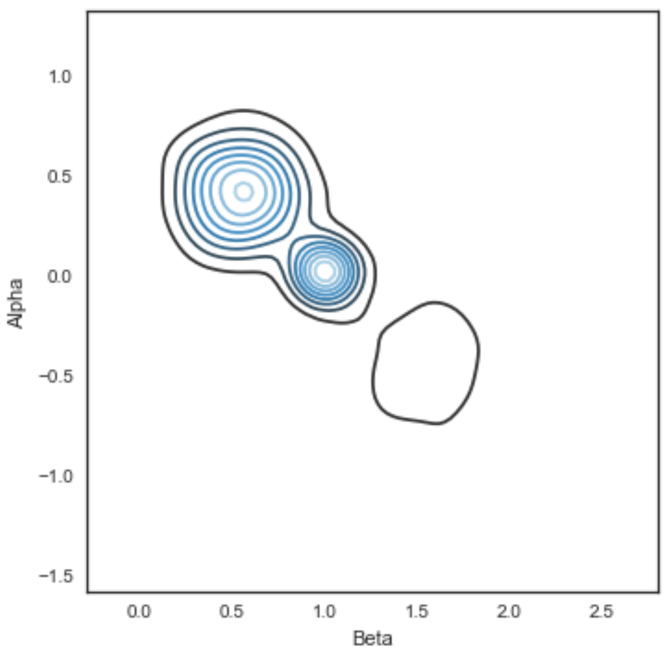


Figure 3: Contour plots for Fitted 3-Mixture (Panel A) and 4-Mixture (Panel B)

Panel A



Panel B

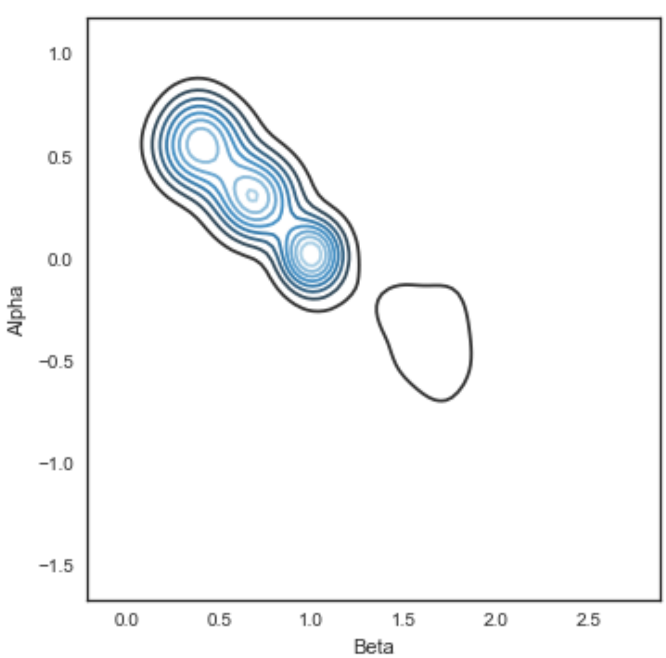
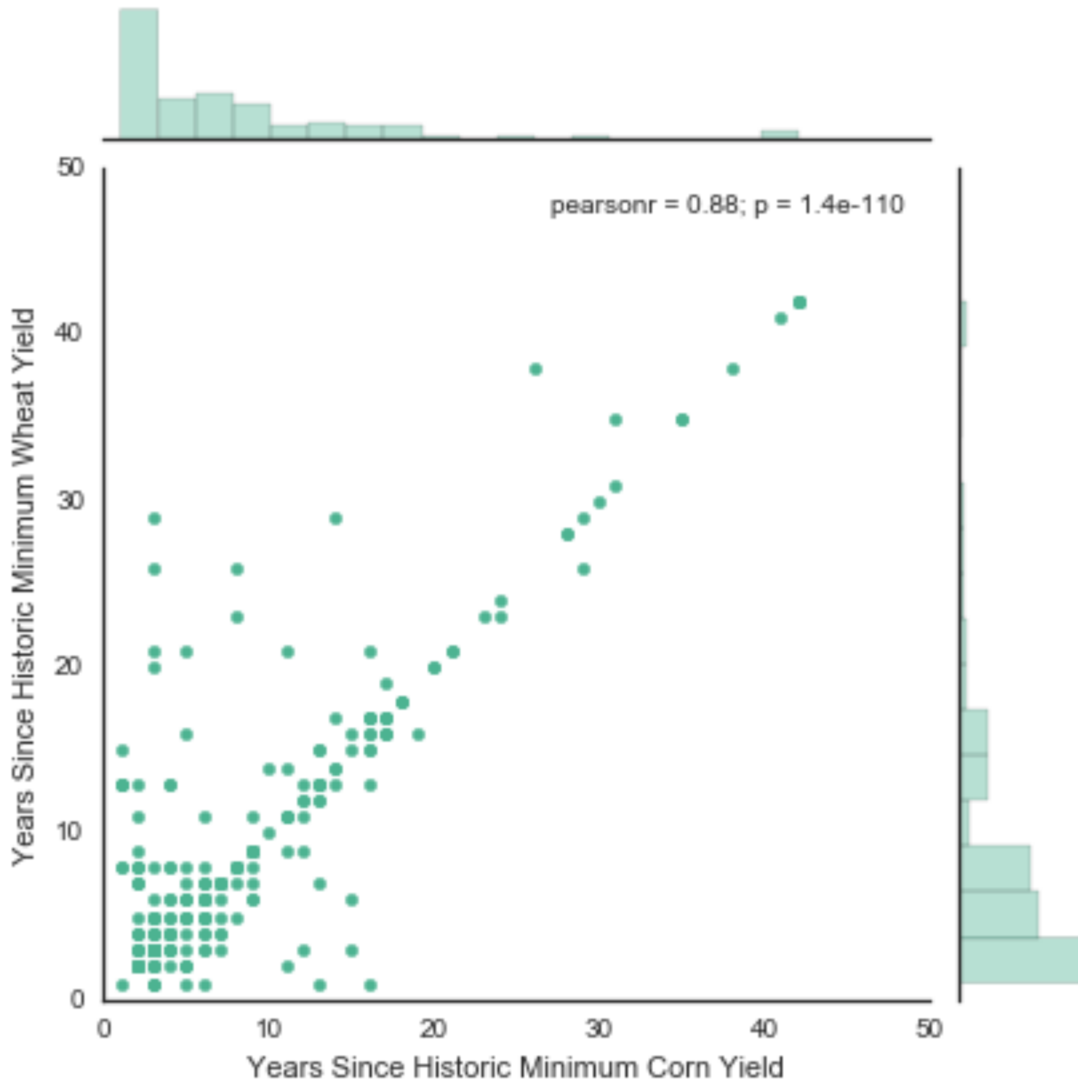


Figure 4: Years since worst yield, by crop, for farmers of both corn and wheat



Tables

Table 1: Summary Statistics (*N=762*)

	Mean	Std. Dev.	Min.	25%	50%	75%	Max.
Years since Min Yield	8.56	8.48	1	3	6	11	42
Years since Min Corn Yield (*N=388*)	8.21	8.41	1	3	5	10.25	42
Years since Min Wheat Yield (*N=374*)	8.92	8.55	1	3	6	12.75	42
Years since Max Yield	1.97	1.54	1	1	2	2	14
Years since Max Corn Yield (*N=388*)	1.90	1.53	1	1	1	2	14
Years since Max Wheat Yield (*N=374*)	2.03	1.54	1	1	2	2	14
Gender (1=Male)	0.55	0.50	0	0	1	1	1
Age	48.09	10.90	20	40	50	56	72
Education	4.34	1.77	0	3	5	5	8
Years Farming	27.16	12.93	0	18.0	30.0	37.0	65
Percent Farm Income	41.74	26.40	0	20.0	36.7	59.6	100
Household Income (Yuan/year)	22924	20101	1000	10000	20000	30000	248000

Table 2: Model Selection for Candidate Forecasting Models

	Model 1	Model 2	Model 3	Model 4
Scale	0.6673***	0.6144***	0.8421***	0.6309***
	(0.011)	(0.011)	(0.014)	(0.012)
Constant	315.38***			
	(8.321)			
r =				
Historic Max		0.3791***		
		(0.009)		
Historic Min			0.2943***	
			(0.018)	
Historic Mode				0.4405***
				(0.012)
Adj. R-sqaured	0.708	**0.986**	0.976	0.985
Log-Likelihood	-9794	**-9752**	-10188	-9819
AIC	19590	**19510**	20380	19640
BIC	19600	**19520**	20390	19650

Notes: *** Signifies statistical significance at the 99.9% level. N=812.

Table 3: Model Selection (2-params, 3 obs)

	Model 1	Model 2	Model 3	Model 4
Scale	0.6717***	0.6029***	0.8886***	0.6301***
	(0.01)	(0.01)	(0.012)	(0.012)
Constant	310.94***			
	(7.535)			
r =				
Historic Max		0.3877***		
		(0.009)		
Historic Min			0.233***	
			(0.015)	
Historic Mode				0.4394***
				(0.012)
Adj. R-sqaured	0.665	**0.987**	0.979	0.985
Log-Likelihood	-14597	**-14516**	-15134	-14694
AIC	29200	**29040**	30270	29390
BIC	29210	**29050**	30280	29400

Notes: *** Signifies statistical significance at the 99.9% level. N=812.

Table 4: Summary of α and β by Gender and by Crop

Corn Farmers (N=389)

	Males (*N=210*)		Females (*N=179*)	
	α	β	α	β
Mean	0.207	0.806	0.298	0.716
Std	0.336	0.401	0.294	0.334

Wheat Farmers (N=377)

	Males (*N=210*)		Females (*N=167*)	
	α	β	α	β
Mean	0.217	0.800	0.309	0.716
Std	0.363	0.397	0.290	0.322

Note: Of the male farmers, 183 grow both corn and wheat. Of the female farmers, 149 grow both.

Table 5**:** Model Selection Criteria by Number of Components, *C*

C	*N*	AICc	BIC	ICL	τ > .99	τ > .95	τ > .90
Corn							
1	442	1348.85	1365.19	682.59	1	1	1
2	442	570.12	606.85	431.47	0.78	0.95	0.96
3	442	-627.26	-570.17	92.82	0.88	0.96	0.98
4	442	-763.51	-686.11	**54.1**	**0.94**	**0.97**	**0.99**
5	442	-773.64	-675.96	60.24	0.93	0.97	0.98
6	442	**-987.73**	**-869.81**	173.88	0.7	0.76	0.9
Wheat							
1	437	1397.04	1413.34	706.67	1	1	1
2	437	618.85	655.48	475.9	0.77	0.96	0.97
3	437	521.05	577.97	472.19	0.94	0.97	0.97
4	437	-801.41	-724.22	**53.34**	0.96	0.98	1
5	437	-809.6	-712.2	68.22	**0.97**	**0.99**	**1**
6	437	**-1084.49**	**-966.91**	141	0.77	0.9	0.93
Pooled							
1	879	2739.49	2758.59	1379.29	1	1	1
2	879	1175.42	1218.38	889.89	0.77	0.95	0.96
3	879	961.83	1028.64	869.77	0.78	0.96	0.97
4	879	-1586.72	-1496.09	**58.34**	**0.95**	**0.99**	**0.99**
5	879	-2077.98	**-1963.54**	236.21	0.73	0.88	0.91
6	879	**-2095.77**	-1957.55	252.44	0.74	0.88	0.91

Table 6: Mixture Weights by Number of Components, C

C	Weights					
Corn						
1	1					
2	0.911	0.090				
3	0.670	0.131	0.199			
4	0.674	0.115	0.199	**0.011**		
5	0.673	0.122	0.199	0.005	0.002	
6	0.336	0.337	0.122	0.199	0.005	0.002
Wheat						
1	1					
2	0.890	0.110				
3	0.888	0.087	**0.025**			
4	0.659	0.103	0.213	0.026		
5	0.659	0.104	0.213	0.018	0.007	
6	0.370	0.289	0.104	0.213	0.018	0.007
Pooled						
1	1					
2	0.899	0.102				
3	0.887	0.095	**0.018**			
4	0.667	0.109	0.206	**0.018**		
5	0.352	0.314	0.109	0.206	0.018	
6	0.352	0.314	0.110	0.206	0.011	0.007

Table 7: 3-Component Mixture Model (with outliers omitted)

Corn Forecasts (N=442)

	Optimistic		Realistic		Pessimistic	
	α	β	α	β	α	β
Mean	0.439	0.542	0.027	1.000	-0.404	1.555
Std. Dev.	0.176	0.182	0.055	0.001	0.324	3.622
Weight	0.674		0.199		0.115	

Wheat Forecasts (N=437)

	Optimistic		Realistic		Pessimistic	
	α	β	α	β	α	β
Mean	0.460	0.538	0.022	1.000	-0.444	1.564
Std. Dev.	0.175	0.182	0.056	0.001	0.303	0.288
Weight	0.659		0.213		0.104	

Pooled Forecasts (N=879)

	Optimists		Realists		Pessimists	
	α	β	α	β	α	β
Mean	0.449	0.540	0.024	1.000	-0.428	1.564
Std. Dev.	0.176	0.182	0.055	0.001	0.318	0.331
Weight	0.667		0.206		0.109	

Table 8: 4-Component Mixture Model (with outliers omitted)

Corn Forecasts (N=442)

	Highly Optimistic		Mildly Optimistic		Realistic		Pessimistic	
	α	β	α	β	α	β	α	β
Mean	0.572	0.396	0.308	0.686	0.027	1.000	-0.443	1.590
Std. Dev.	0.122	0.112	0.109	0.104	0.055	0.001	0.402	0.427
Weight	0.336		0.337		0.199		0.122	

Wheat Forecasts (N=437)

	Highly Optimistic		Mildly Optimistic		Realistic		Pessimistic	
	α	β	α	β	α	β	α	β
Mean	0.584	0.406	0.303	0.707	0.022	1.000	-0.445	1.565
Std. Dev.	0.111	0.117	0.175	0.182	0.056	0.001	0.303	0.290
Weight	0.370		0.289		0.213		0.104	

Pooled Forecasts (N=879)

	Highly Optimistic		Mildly Optimistic		Realistic		Pessimistic	
	α	β	α	β	α	β	α	β
Mean	0.578	0.400	0.306	0.697	0.024	1.000	-0.425	1.562
Std. Dev.	0.117	0.114	0.105	0.095	0.055	0.001	0.318	0.332
Weight	0.352		0.314		0.206		0.110	

Table 9: Regression Results

	β			α		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Intercept	0.600***	0.603***	0.629***	0.405***	0.406***	0.392***
	(0.101)	(0.102)	(0.118)	(0.089)	(0.089)	(0.104)
Crop Dummy (Corn=1)		-0.006	-0.016		-0.002	0.004
		(0.026)	(0.027)		(0.023)	(0.023)
Sex (Male=1)	0.091***	0.091***	0.101***	-0.094***	-0.094***	-0.099***
	(0.029)	(0.029)	(0.029)	(0.025)	(0.025)	(0.026)
Age (Years)	0.002	0.002	0.002	-0.002	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Education	0.012	0.013	0.013	-0.012*	-0.012	-0.012
	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)
Years Farming	0.001	0.001	-0.004	-0.001	-0.001	0.002
	(0.002)	(0.002)	(0.004)	(0.001)	(0.001)	(0.003)
Years Farming^2			0.000			-0.000
			(0.000)			(0.000)
Years since Min Yield	-0.010***	-0.010***	-0.012***	0.010***	0.010***	0.010***
	(0.002)	(0.002)	(0.004)	(0.001)	(0.001)	(0.003)
Min Yield Last Year Dummy			0.035			-0.035
			(0.047)			(0.041)
Years since Max Yield	0.006	0.006	0.005	-0.006	-0.006	-0.001
	(0.009)	(0.009)	(0.014)	(0.007)	(0.008)	(0.012)
Max Yield Last Year Dummy			-0.003			0.003
			(0.045)			(0.039)
Min Yield Year > Max Yield Year			-0.069			0.011
			(0.066)			(0.058)
rs since Min Yield, Other Crop			0.002			-0.001
			(.003)			(0.003)
rs since Max Yield, Other Crop			0.013			-0.010
			(0.010)			(0.009)
Adj. R^2	0.07	0.06	0.06	0.08	0.08	0.08
N	762	762	759	762	762	759

Notes: Standard errors in parentheses. * p<.1, ** p<.05, ***p<.01

Table 10: Classification Results for the 3-parameter Model (Pooled Data)

Panel A: Model Selection

C	AICc	BIC	ICL	τ>.99	τ>.95	τ>.90	Weights			
1	-386.43	-358.24	-179.12	1	1	1	1			
2	-1369.95	-1308.91	-100.88	0.66	0.82	0.86	0.578	0.422		
3	-3599.37	-3505.49	**-1062.27**	**0.94**	**0.98**	**0.99**	0.670	0.114	0.217	
4	**-4087.35**	**-3960.64**	-909.62	0.74	0.89	0.92	0.336	0.335	0.112	0.217

Panel B: 3-Class Mixture Model with Shape Parameter

Pooled Forecasts (N=879)

	Optimistic		
	α	β	γ
Mean	0.44	0.55	-0.01
Std. Dev.	0.166	0.171	0.087
Weight		0.669	

	Realistic		
	α	β	γ
Mean	0.02	1.00	0.01
Std. Dev.	0.062	0.035	0.090
Weight		0.221	

	Pessimistic		
	α	β	γ
Mean	-0.37	1.51	0.03
Std. Dev.	0.290	0.327	0.156
Weight		0.111	