

2018

Discrete Classification Problems in Agricultural and Behavioral Economics

Clayton P. Michaud
University of Rhode Island, claytonmichaud@uri.edu

Follow this and additional works at: https://digitalcommons.uri.edu/oa_diss

Terms of Use

All rights reserved under copyright.

Recommended Citation

Michaud, Clayton P., "Discrete Classification Problems in Agricultural and Behavioral Economics" (2018).
Open Access Dissertations. Paper 698.
https://digitalcommons.uri.edu/oa_diss/698

This Dissertation is brought to you by the University of Rhode Island. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons-group@uri.edu. For permission to reuse copyrighted content, contact the author directly.

DISCRETE CLASSIFICATION PROBLEMS IN AGRICULTURAL AND
BEHAVIORAL ECONOMICS

BY

CLAYTON P. MICHAUD

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

IN

ENVIRONMENTAL AND NATURAL RESOURCE ECONOMICS

UNIVERSITY OF RHODE ISLAND

2018

DOCTOR OF PHILOSOPHY DISSERTATION

OF

CLAYTON P. MICHAUD

APPROVED:

Dissertation Committee:

Major Professor Thomas Sproul

Todd Guilfoos

Gavino Puggioni

Nasser H. Zawia
DEAN OF GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND

2018

Abstract

The focus of this dissertation is on discrete classification problems in agricultural and behavioral economics. In my first two manuscripts, I take up the issue of producer misperceptions of yield risk relative to their objective, a well-established phenomenon in which farmers tend to be overly optimistic in their perceptions of yield risk, forecasting yields with higher mean and lower variance than historical outcomes would suggest.

Manuscript 1 focuses on estimating both how such misperceptions are distributed across individual forecasts, as well as how such misperceptions might arise. Manuscript 2 goes on to look at how these misperceptions of yield risk affect farm-level crop insurance coverage level choices, simulating cross-coverage crop insurance demand across a broad set of scenarios. In my third chapter, I present a hierarchical Bayesian methodology for disaggregating, or downscaling, aggregated count data using an outside statistical sample. As an application, this chapter demonstrates how stakeholders can use readily available yet incomplete land use (e.g. agricultural) count data in combination with censored/aggregated census data provided at the county level to recover/estimate land use count data at the municipality (town, city, or any other sub-county region) level.

In Manuscript 1, we estimate the distribution of miscalibrated perceptions of yield risk, using the expectation maximization algorithm to perform a latent class analysis to uncover potential heterogeneity (clustering) in the parameters of our yield miscalibration model. Using self-reported yield forecasts and yield history from rural Chinese farmers, we estimate miscalibration parameters for each of our 879 forecast/history observations, using expectation maximization to fit these parameters to a Gaussian mixture model. We find that forecasts can best be described as coming from three distinct distributions or clusters that can best be described as ‘optimistic’, ‘unbiased’, and ‘pessimistic’. We find that roughly 67% of forecasts can be defined as optimistic with producers perceiving that, on

average, yields face only half (55%) of their true risk. 12% of forecasts can be defined as pessimistic with producers perceiving that, on average, yields face 50% more risk than that of their objective risk. The remaining 21% of forecasts can be classified as unbiased, with perceived yield risk being largely in line with objective yield risk. In addition, we find that our optimistic group separates cleanly into two distinct clusters of roughly equal size – one comprised of ‘mild optimists’, and another comprised of ‘extreme’ optimists.

We go on to examine the possible causes of these misperceptions of risk, finding that such misperceptions are not inherent to the producer, but rather result from crop-specific yield experience. Using regression methods, we find statistically significant evidence that recent historic losses increase the amount of producer’s level of perceived risk, while increases in the length of time since experiencing a historic loss decrease the level of perceived risk.

These results have important implications for crop insurance demand modeling. These findings also suggest that a targeted subsidy approach based on outcome history may be more cost-effective at inducing insurance participation than subsidies that are fixed across locations. Namely, Not only is it important to incorporate miscalibrated perceptions of risk in crop insurance demand models, it is also important to include heterogeneity with regard to those misperceptions.

Manuscript 2 takes up the question of how misperceptions of yield risk effect producers’ decisions regarding which crop insurance coverage level to participate in. We simulate cross-coverage level crop insurance demand for both yield and revenue insurance across four potential models of risk misperception and three potential models of decision-making - one based on expected utility and two based on cumulative prospect theory yet differentiated by whether decisions are framed within the broader context of farm risk-management, or whether decisions/outcomes are more narrowly framed - for a total of

twelve choice models. Optimal coverage level choices are simulated for both corn (based on data from York County, NE and considered to be ‘low risk’) and wheat (based on data from Sumner County, KS and considered ‘high risk’). We find that increases in optimism bias drive down the optimal choice of coverage level, eventually inducing producers not to participate in crop insurance at all. Conversely, pessimism causes producers to increase their coverage level to the point of maximum coverage. We also find that this effect is strongest in the case of yield insurance, although the effect is still significant for revenue insurance. Further sensitivity analyses suggest that these results are not highly sensitive to correlations between prices and yields.

The aim of Manuscript 3 is to help stakeholders obtain policy-critical micro-level statistical data in cases where such data may only exist at a higher level of aggregation than is desired (e.g. aggregated census data). In this manuscript, published in the December 2017 edition of *Agricultural and Resource Economics Review*, we develop a hierarchical Bayesian methodology for downscaling regional count data to the sub-region level through the incorporating of an outside statistical sample in the form of sub-regional lower bounds (e.g. sub-region_{*i*} has at least x_i farms, sub-region_{*j*} has at least x_j farms, etc...). Our methodology combines numerical simulations with exact calculations of combinatorial probabilities in order to determine which values of sub-regional counts are most likely to have resulted in the available statistical sample given the information contained in our two data sources. Although our method is designed to provide municipality count data based on county level data, as a proof of concept, we demonstrate our approach by estimating Rhode Island county level farm counts (which are known, but are not used in the estimation procedure) using state level farm count data provided by the Ag Census, along with a sample of Rhode Island farm locations collected by the University of Rhode Island. By estimating values that are known we are able to measure

the accuracy of our estimates. We are able to show that not only do our estimates outperform those obtained via maximum likelihood, but that they are robust to sampling variability across heterogeneous population sizes. We go on to expand our model to incorporate spatial considerations and demonstrate how the use of an informative prior based on relevant sub-region characteristics (land area, in our application) can further improve the estimates.

Acknowledgements

I would first like to thank my parents for their unwavering support. Next I would like to thank Tom Sproul, my thesis adviser, for all of his support and mentorship and for teaching me everything I know about research. Lastly, I would like to thank the rest of my dissertation/oral examination committee, Todd Guilfoos, Gavino Puggioni, David Bidwell, and Hiro Uchida for their invaluable time and support.

Preface

Formatting: This dissertation has been arranged according to the manuscript format.

List and Publication Status of Manuscripts Herein:

Manuscript 1

Heterogeneous Perceptions of Objective Yield Risk: Clusters and Causes

This manuscript is currently being revised for resubmission at the American Journal of Agricultural Economics.

Manuscript 2

The Effect of Optimism Bias on Cross-Coverage Level Crop Insurance Demand

This manuscript has not yet been submitted for publication at this time.

Manuscript 3

Bayesian Downscaling Methods for Aggregated Count Data

This manuscript was published in the December 2017 edition of Agricultural and Resource Economics Review

Table of Contents

Abstract.....	ii
Acknowledgements.....	vi
Preface.....	vii
Table of Contents.....	viii
List of Figures.....	x
List of Tables.....	xi
Manuscript 1	1
1.0 Abstract.....	2
1.1 Introduction.....	3
1.2 Background.....	6
1.3 A Location-Scale Model of Forecasts from Historical Data.....	12
1.4 Data.....	14
1.5 Model Selection.....	21
1.6 Identifying Heterogeneity with a Gaussian Mixture Model.....	23
1.7 Factors Influencing Forecasts.....	30
1.8 Conclusion.....	36
1.9 Works Cited.....	40
1.10 Figures.....	50
1.11 Tables.....	55
Manuscript 2	67
2.0 Abstract.....	68
2.1 Introduction.....	69

2.2 Modeling Optimism Bias.....	72
2.3 Modeling Farmer Crop Insurance Cover Level Decision-Making.....	75
2.4 Crop Insurance Premiums and Indemnities.....	80
2.5 Simulating Optimal Coverage Level Choices.....	80
2.6 Results.....	82
2.7 Conclusion.....	87
2.8 Works Cited.....	89
2.9 Figures.....	92
2.10 Tables.....	96
Manuscript 3	105
3.0 Abstract.....	106
3.1 Introduction.....	107
3.2 Bayesian Downscaling of Aggregated Count Data.....	111
3.3 Choosing a Prior.....	114
3.4 Sample Data.....	117
3.5 Methods.....	118
3.6 Results.....	119
3.7 Discussion.....	121
3.8 Conclusion.....	126
3.9 Works Cited.....	128
3.10 Figures.....	130
3.11 Tables.....	131

List of Figures

Manuscript 1

Figure 1.1: Turvey et al. (2013) Survey Questions.....	50
Figure 1.2: Elicited Historic vs. Future Yield Distributions.....	51
Figure 1.3: Hex Plot of Alpha and Beta with Marginal Densities.....	52
Figure 1.4: Contour Plots for Fitted 3-Mixture (Panel A) and 4-Mixture (Panel B)....	53
Figure 1.5: Scatter Plot: Years Since Historic Minimum Yield (Corn and Wheat).....	54

Manuscript 2

Figure 2.1: Example Comparisons of the Objective and Perceived/Forecasted Yield Distributions According to our Four Models of Optimism Bias.....	91
Figure 2.2: The Probability Weighting Function ($\gamma = 0.61, \delta = 0.69$).....	92
Figure 2.3: The Prospect Theory Value Function.....	93
Figure 2.4: Simulated Price-Yield and Yield Distribution Plots for Corn and Wheat..	94

Manuscript 3

Figure 3.1: Nested Data Structure.....	130
--	-----

List of Tables

Manuscript 1

Table 1.1: Household-Level Summary Statistics.....	55
Table 1.2: Testing for the Presence of Yield Trends.....	56
Table 1.3: Model Selection for Candidate Forecasting Models.....	57
Table 1.4: Model Selection (2-parameters, 3 observations).....	58
Table 1.5: Summary of α and β by Gender and by Crop.....	59
Table 1.6: Model Selection Criteria by Number of Components, C	60
Table 1.7: Mixture Weights by Number of Components, C	61
Table 1.8: 3-Component Mixture Model (with outliers omitted).....	62
Table 1.9: 4-Component Mixture Model (with outliers omitted).....	63
Table 1.10: Correlation Coefficients of Classes Across Forecasts.....	64
Table 1.11: Regression Results.....	65
Table 1.12: Classification Results for the 3-parameter Model (Pooled Data).....	66

Manuscript 2

Table 2.1: Parameters for $F(y, P)$ from Babcock (2015).....	95
Table 2.2A: Optimal Coverage Level Choice by Crop, Insurance Type, Decision Model: Optimism Bias Model 1.....	96
Table 2.2B: Optimal Coverage Level Choice by Crop, Insurance Type, Decision Model: Optimism Bias Model 2.....	97
Table 2.2C: Optimal Coverage Level Choice by Crop, Insurance Type, Decision Model: Optimism Bias Model 3.....	98
Table 2.2D: Optimal Coverage Level Choice by Crop, Insurance Type, Decision Model: Optimism Bias Model 4.....	99

Table 2.3A: Optimal Revenue Coverage by Crop, Decision Model and Price-yield Correlation: Optimism Bias Model 1.....	100
Table 2.3B: Optimal Revenue Coverage by Crop, Decision Model and Price-yield Correlation: Optimism Bias Model 2.....	101
Table 2.3C: Optimal Revenue Coverage by Crop, Decision Model and Price-yield Correlation: Optimism Bias Model 3.....	102
Table 2.3D: Optimal Revenue Coverage by Crop, Decision Model and Price-yield Correlation: Optimism Bias Model 4.....	103
Table 2.4: Optimal Coverage Level Choice by Crop, Insurance Type, Decision Model: Optimism Bias Model 1 Under Equal Proportionate Subsidies.....	104
Manuscript 3	
Table 3.1: Number of Eligible Combinations per County.....	131
Table 3.2: Farms, Municipalities, and Land Area (mi ²) per County.....	132
Table 3.3: Mean Normalized Root Mean Square Error by Estimate Type: Fixed Population Size (1,250) with Increasing Sample Size.....	133
Table 3.4: Mean Normalized Root Mean Square Error by Estimate Type: Fixed Sample Proportion (0.20) and Increasing Population Size.....	134
Table 3.5: Mean Normalized Root Mean Square Error by Estimate Type Using Simulated Sample: Fixed Population Size (1,250) and Increasing Sample Size.....	135
Table 3.6: Mean Normalized Root Mean Square Error by Estimate Type Using Simulated Sample: Fixed Sample Proportion (0.20) and Increasing Population Size.....	136

Manuscript 1

Heterogeneous Perceptions of Objective Yield Risk: Clusters and Causes

This manuscript is currently being revised for resubmission at the American Journal of Agricultural Economics, however this version has been revised to reflect comments from dissertation committee members.

Authors are listed and footnoted as follows:

Clayton P. Michaud, Thomas W. Sproul¹, Calum G. Turvey², Rong Kong²

¹ Dr. Sproul is Clayton Michaud's Thesis Adviser at the University of Rhode Island.

² Dr. Turvey and Dr. Kong were responsible for the initial data collection and appear as coauthors for that contribution.

Abstract

Understanding how farmers perceive risk is crucial in designing effective risk-management tools and policies. Federal crop insurance subsidies have grown dramatically over time to address low program participation, and it has been suggested that the widespread behavioral phenomenon of optimism bias may play a role. Nonetheless, no model currently exists to map optimistically biased forecasts into crop insurance demand. We develop a new behavioral model of perceived yield risk as shifting and scaling the objective distribution of yield risk. We fit the model to yield forecasts and yield histories of wheat and corn farmers in China, and find that forecasts are anchored to historical positive experience and they are optimistically biased, on average. To evaluate multi-modal heterogeneity in the forecasts, we estimate a Gaussian mixture model using expectation-maximization and find evidence for three basic forecast types: optimistic (roughly 67%), realistic or neutral (about 20%), and pessimistic (about 12%), with a small, extremely pessimistic outlier group. We find further clustering within the optimistic group, with 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. Moreover, we find no evidence that forecast classifications map to a classification of the individual farmer making them. Instead, recent severe loss experience in the crop of interest appears to be the strongest single predictor of pessimism, though we also find statistically significant gender differences.

1.1 Introduction

Since passage of the Federal Crop Insurance Act of 1980, policy makers and agricultural economists have struggled to understand the factors driving farmer adoption of crop insurance in the United States. Despite the obvious benefits of risk management, participation in crop insurance programs has historically been puzzlingly low unless heavily subsidized by the federal government (Coble and Barnett, 2013). Congress passed reform bills in 1994 and 2000 that dramatically increased the size and scope of the Federal Crop Insurance Program because of failures to attract adequate participation at sufficiently high coverage levels (Glauber and Collins, 2002), and program costs have now grown to almost 30 times 1980s levels: the federal government is projected to spend over 6 billion dollars on crop insurance subsidy programs in 2017 (USDA ERS).

Common hypotheses for the presence of “demand frictions” limiting crop insurance participation have included adverse selection, expectations of *ad hoc* disaster relief, and availability of alternative risk management tools (see, for example, Coble and Barnett, 2013; Coble et al., 1999; Just, Calvin and Quiggin, 1999; Smith and Goodwin, 1996; Skees and Reed, 1986), though these factors are observed in other insurance markets without the same severe effects. Some recent studies, however, posit that lower than expected demand for crop insurance results from a systematic miscalibration of subjective yield risks by farmers (Egelkraut et al. 2006; Turvey et al., 2013). There is empirical evidence that, on average, producers are optimistically biased with respect to yields and yield variability, expecting better than average yields and below average yield risk (Pease, 1992; Umarov and Sherrick,

2005; Egelkraut et al., 2006; Turvey et al., 2013), and a number of studies have shown that perceived yield risk affects crop insurance demand around the world, including farmers in the United States (Horowitz and Lichtenberg, 1993; Sherrick et al., 2004; Egelkraut et al., 2006; Shaik et al., 2008), in France (Enjolras and Sentis, 2011) and in China (Wang, Ye and Shi, 2016). While extensive work has been done attempting to model the distribution of crop yields, especially for crop insurance rating purposes (see Woodard and Verteramo-Chiu, 2017; Ker et al., 2015; and Woodard and Sherrick, 2011, for recent examples), modeling the transformation of perceived yields due to optimism bias remains an open question. Addressing this gap in the literature is a critical step in the development of behavioral models of crop insurance demand going forward.

In this article, we formulate and test a parsimonious and distribution-agnostic model in which subjective yield risks (herein, “forecasts”) are derived as shift-and-scale transformations of the historical distribution of outcomes. The model can be estimated by regression on a per-farmer and per-crop basis using simple elicitation data, and it allows for inclusion of a reference point anchoring the forecast to a summary statistic of the historical distribution. Using data from corn and wheat farmers in Shanxi Province, China, we find statistical evidence that farmers anchor to past positive yield experience and that they are optimistically biased on average. Critically, the distribution of model parameters appears to be multi-modal, meaning that there are heterogeneous clusters among forecasts which are not well-described by the population average (cf., Bruhin, Fehr-Duda and Epper, 2010; Sproul and Michaud, 2017).

Using a finite Gaussian mixture model, we find evidence for three basic forecast types: optimistic (about two thirds), realistic or neutral (about 20%), and pessimistic (most of the remainder), plus a small outlier group of extreme pessimists. On average, optimistic forecasts had half the risk (a scale parameter, proportional to standard deviation) of their objective historical yield distribution, while pessimistic forecasts had approximately twice the risk 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. Our results are robust to changing the number of data points per elicitation, and to adding a shape parameter to the model.

Since our data set contains many farmers who make more than one forecast, we are able to test whether forecast classifications are also classifications of the farmers themselves. We do not find statistical evidence in support of this hypothesis. To gain further insight, we conduct a regression analysis to identify factors influencing the degree of optimism bias in the parameters, and find that the recency of historic losses is the most statistically and economically significant. Farmers' forecasts exhibit a 1% decrease in scale (risk) for each additional year since their historic minimum yield. We find no similar effect for historic gains, and also find no evidence of a cross-crop effect within farmers, i.e. wheat forecasts do not seem affected by the same farmer experiencing corn losses, and vice versa. We also find statistically significant differences in optimism bias by gender: men forecast about 85% of their historical yield risk, while women forecast only 74% on average, after

controlling for other factors. Finally, we find that optimism bias appears to be persistent over time, as it does not vary with years of farming experience.

The remainder of this article is organized as follows: the next section provides a background of the literature pertaining to optimistically biased forecasting by farmers, and of the relevant behavioral economics literature more generally. The following two sections introduce our model and our data set. We then estimate the basic model, address the presence of reference points in forecasts, and estimate a mixture model to evaluate heterogeneity. Finally, we examine predictors of optimism bias, including demographic variables and yield history. The last section discusses implications for policy design and future research, and concludes.

1.2 Background

In the words of De Bondt and Thaler (1995, p. 389), “perhaps the most robust finding in the psychology of judgment is that people are overconfident.” 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 (cf. Slovic, 2016; Sandroni and Squintani, 2007). The finding is consistent across domains as diverse as post-college job prospects (Weinstein, 1980; Hoch, 1985), construction costs (Statman and Tyebjee, 1985), health risks (Kreuter and Strecher, 1995; Robb et al., 2004), entrepreneurship (Cooper et al., 1988; Camerer and Lovallo, 1999; Landier and Thesmar, 2009), and the now-canonical popular culture example, in which over 80% of drivers think they have

above average ability (Svenson, 1981; Groeger and Grande, 1996). A unifying feature of overconfident beliefs is their dependence on the perception of control or ability to exert skill over the outcome (Sandroni and Squintani, 2004). Though the terms ‘optimism bias’ and ‘overconfidence’ are often used interchangeably, there is a noteworthy distinction that is particularly relevant to this analysis. Overconfidence generally refers to the belief that one’s expected outcome will be more favorable than the expected outcome of one’s peers. Optimism bias, on the other hand, is a more general term that refers to the belief that one’s future outcomes will be more favorable than past outcomes, possibly including those experienced by others. Unfortunately this distinction is not always carefully applied in the literature, so to avoid constant context-switching and qualification of terms, we will refer to these phenomena herein under the umbrella term ‘optimism bias’.

The above examples are not simply cases of wishful thinking or otherwise erroneous reporting on surveys, nor are they confined to the behavioral laboratory: acting on misperceptions about risk has been observed as affecting health outcomes (Hoorens, 1994), legal settlements (Babcock et al., 1997), mergers and acquisitions (Malmendier and Tate, 2006; 2008), and insurance purchasing decisions (Bhattacharya et al., 2004). In a theory paper, Spinnewijn (2013) presents a model in which individuals are overly optimistic with regard to probabilities of discrete health outcomes as a function of effort, demonstrating how such a belief decreases an individual’s willingness to pay for health insurance. One feature of behavioral phenomena is that some can be “unlearned” through market or professional experience, like the endowment effect (List, 2003), while others persist, like myopic

loss aversion (Haigh and List, 2005). It appears that optimism bias falls in the latter category, and persists in the face of experience (Dalziel and Job, 1997), consistent with our results.

Despite the widespread evidence, optimistically biased beliefs do not apply to everyone in a given domain. Nearly every article on optimism bias has a subtext in which a surprisingly large percentage of subjects are overly optimistic and a surprisingly small (but non-zero) percentage are overly pessimistic. For example, Svenson (1981) finds that 82% of students placed themselves among the top 30% of drivers, and Wenglert and Rosén (2000) find that 72% of their subjects were classified as being overly optimistic about their personal future. Cooper et al. (1988) find that 68% of entrepreneurs believe 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 (similar findings are reported by Landier and Thesmar, 2009). In the investment domain, a survey by Benartzi (2001) found that only 16.4% percent of respondents believed that their company's stock was riskier than the stock market as a whole.

Despite the breadth of this literature, there are two key areas where it lacks depth that we hope to address here. The same challenges exist in the relevant agricultural economics literature, discussed below. First, most approaches rely on estimating optimism bias on average, either by regression or comparison of distribution moments, but none of them (to our knowledge) explicitly estimate how past experience is transformed into an optimistically biased forecast. Second, there is a consistent finding of widespread optimism but always at least some pessimism, and

little work has been done to estimate the incidence of these beliefs more broadly. A key contribution in this article is the classification of forecasts into optimistic, neutral and pessimistic types.

Within agricultural economics, Pease (1992) is the first study we are aware of to statistically compare farmers' subjective and historical crop yield probability distributions, evaluating historical yields and subjective expectations for the 1987 crop year for 98 Western Kentucky grain farmers. Pease finds that in many individuals, there exist large differences in the moments of the subjective and historical distributions. He also finds that corn forecasts were mildly pessimistic, on average, while soybean forecasts were optimistic. These differences were driven primarily by regional differences, with severe drought conditions affecting yields in both 1980 and 1983 for areas predominately planted in corn, while drought conditions affected yields only in 1983 for areas planted predominately in soybeans. The result that pessimistic forecasting bias is influenced by recent negative yield outcomes is consistent with our results as discussed above.

In a conference paper, Umarov and Sherrick (2005) test the “better than average” effect in a survey of 870 corn and soybean farmers in Illinois, Iowa and Indiana. They find strong evidence of optimism bias in terms of above average yields and below average stability (lack of variance) relative to their primary county, consistent across both crops. For each crop, approximately 60% of farmers reported mean yields different from their county average, and of these, 90% reported above average yields. Only 2% of large farms in their sample reported below average mean yields. With respect to variance, about 45% of farmers reported yield stability

different than county average, and of these, about 75% reported above average yield stability. Their results are telling in much of the literature we review here: optimism bias is a widespread phenomenon, but nearly every paper observing optimism bias also observes some pessimistic members of the population (and possibly some neutral ones as well).

In another conference paper, Egelkraut et al. (2006) replicate the results in Umarov and Sherrick on outside data while adding a probability elicitation task. Farmers were asked to compare their own yields and yield risks to other farmers in their county. Only 12% reported below-average yields, while 42% and 46% reported average or above-average yields, respectively. On the risk side, 20% reported above-average yield risk, while 38% and 42% reported average and below-average yield risk, respectively. In a first approach to (implicitly) model farmers' optimism bias in yields, the authors elicit a forecast distribution as a Weibull and compare it to a fitted Weibull distribution at the county level, using a Q-Q plot (quartile-quartile). Unfortunately, their data set lacks the farm-level yield histories necessary for the type of evaluation done by Pease, or herein.

Turvey et al. (2013), the source of our data set, elicit yield forecasts and yield histories for 570 corn farmers in China using the Beta-PERT expert elicitation method. Similar methods have been used elsewhere in agricultural economics to fit Beta distributions to yield forecasts (e.g., Shaik et al., 2008). Turvey et al. indicate two aims of their study: first, to propose a solution for estimating the historical yield distribution for insurance pricing in the absence of annual data, and second, to evaluate the degree to which that insurance would be demanded by farmers according

to their subjective forecasts. The authors find strong evidence of optimism bias, with 82% of households expecting average yields to exceed their historical experience, and 72% forecasting lower risk. These results are significantly correlated with farmers' self-reported "interest" in crop insurance, suggesting that optimism bias may be a key explanation for the need to subsidize crop insurance premiums.

In addition to their yields, there is consistent evidence that farmers are optimistically biased with respect to the prices they will receive. Eales et al. (1990) find that Illinois farmers and grain merchandisers had accurate forecasts of futures prices but were optimistically biased in consistently underestimating the futures price volatility, while Kenyon (2001) finds producer price expectations to be skewed toward higher prices and consistently underestimating the risk of large price changes during the season. More recently, in a conference paper, Riley and Anderson (2009) find the forecasts of Mississippi corn, cotton and soybean farmers to overestimate price and underestimate volatility relative to commodity futures and options markets. As mentioned above, the perception of skill or control over outcomes is a critical feature of optimism bias. While it is not reasonable to expect that farmers think they influence commodity prices, there may be a framing effect in which farmers recognize that they influence their income through skill (via yields), and this may carry over to optimism bias being exhibited in price forecasts when elicited separately. In a case without this framing issue, Sherrick (2002) finds farmers to be pessimistic on average with respect to weather forecasts.

To summarize, there is extensive evidence of optimism bias in the population at large and in farmers, as well. The agricultural economics literature is by-and-large

consistent with the mainstream behavioral economics literature, finding optimism bias to be widespread and persistent in the face of experience. There is not, to our knowledge, a mainstream model of optimism bias in terms of the way in which historical experience is explicitly transformed into an optimistically biased forecast. In the paper of Spinnewijn (2013) discussed above, there is a model of overconfidence about effort affecting a model of binary outcomes of success or failure, but this model does not generalize to continuous outcomes or non-binary distributions. There is some evidence in the above literature that optimism about average outcomes (e.g., mean yields) may materially differ from pessimism about risk (e.g., higher yield variance), but no model exists to bring these findings together under a cohesive structure. In what follows, we aim to satisfy these gaps in the literature and provide a structural model for future researchers to build upon.

1.3 A Location-Scale Model of Forecasts from Historical Data

We model the forecast distribution as an affine, “shift and scale” transformation of the historical distribution of the form:

$$Y = \alpha + \beta X . \tag{1.1}$$

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. The model’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 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:

$$Y = a + b \cdot \left(\frac{X - \mu}{\sigma} \right) = \left(a - \frac{b\mu}{\sigma} \right) + \frac{b}{\sigma} X. \quad (1.2)$$

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 1.1 lends itself naturally to a regression approach.

The basic model can be extended 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:

$$Y = \alpha r + \beta X, \tag{1.3}$$

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.

It is important to point out that our model is not designed to be an alternative to the probability weights used in cumulative prospect theory (Tversky and

Kahneman, 1992), despite the familiar use of the reference point terminology. Prospect theory probability weights, often referred to as decision weights, are simply another measure of risk preferences, and are not designed to model an individual's subjective beliefs about risk. In contrast, our forecasting model is designed to model beliefs about future risk as a function of a historic outcome distribution, *in which the support of the distribution is allowed to vary from history to forecast*. Probability weights transform the relative probabilities associated with each outcome, but do not contemplate the potential for changing support of the distribution. That being said, our forecasting model could easily be incorporated into a cumulative prospect theory framework by layering probability weights on top of elicited, subjective probabilities.

1.4 Data

The observations for corn farmers in our data set were 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 Shanxi 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 730 questionnaires. Turvey et al. report the results of 570 complete questionnaires from corn farmers, but their data collected from wheat farmers are published for the first time here. From these totals we removed all forecasts by farmers not actively growing that crop, as well as 19 nonsense forecasts with zero

risk. The resulting data set contains surveys from 483 households, 396 of which farmed both corn and wheat, 46 of which farmed only corn, and 41 of which farmed only wheat, giving us a total of 879 usable forecasts.

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 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 23,354 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 1.1.

[Table 1.1]

A small section of the survey was dedicated 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 United States units of 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 1.1.

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 (these were the nonsense forecasts mentioned above).

[Figure 1.1]

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. The Beta-PERT elicitation is nearly identical to the Beta elicitation used with United States farmers by Shaik et al. (2008). 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 three 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 potential for technological change affecting yields over time, and iii) the potential for inconsistent mapping between historical and forecasted minimum/maximum as representing comparable confidence intervals (discussed in further detail in section 1.5). 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 a later section 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. Depending on the type of forecast, we find that subjects report, on average, a forecast mean between -1% and 3% higher than their historic mode, after accounting for shift and scale adjustments. However, we find no meaningful change in goodness of fit or in the number of mixture components selected.

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, or in the optimism bias modeling of Egelkraut et al., 2006), 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 larger-scale operations.

Given this ambiguity, we are fortunate that an explicit statistical test is available: we can use the reported historical means as dependent variables, and regress them on years farming and other control covariates, including village fixed effects. If a yield trend is present, then the regression coefficient on years farming should be negative and significantly different from zero. Table 1.2 below presents the regression testing, which shows no statistical evidence of yield trends in the historic reported means. Model 1 shows a baseline specification and model 2 allows for distinct yield trends by crop. On this basis, we conclude that no yield trend adjustment is necessary for our data set. This finding is further confirmed in our later results showing the estimated optimism bias does not vary systematically with years of farming experience. Furthermore, we perform additional regression analysis and find no evidence that differences in the reported ranges (minimum yield – maximum yield) are significantly affected by years farming. This implies that the historical ranges reported are independent of the number of yields experienced, a finding which is somewhat at odds with the typical assumption of yields as following a Markov process (more sample draws should give a wider range, on average). We therefore recommend that future surveys also ask respondents to report yields in the previous

growing season in order to test for such a possibility, and therefore reduce the need for such assumptions.

[Table 1.2]

1.5 Model Selection

Figures 1.2 and 1.3 provide a graphical illustration of the misalignment between historic outcomes and future expectations. We used average minimum, maximum, and modal yield values to estimate both a triangular distribution (Figure 1.2, Panel A) and a Beta distribution (Figure 1.2, Panel B) for both historic and expected future yields. We see that on average, the farmers in our data set dramatically underestimate downside yield risk relative to what historic outcomes might suggest.

[Figure 1.2]

To identify whether a raw shift and scale model (Equation 1.1) or one incorporating a reference point (Equation 1.3) 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 1.1 alongside several variants of the model from Equation 1.3, 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 1.3. 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. The finding is also intuitively appealing in the sense that anchoring to a reference point is more meaningful: it would be difficult to interpret an arbitrary shift parameter without historical context.

[Table 1.3]

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 1.3 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 10th and 90th percentiles, as in Shaik et al., 2008), the data are still valid for our approach. However, this conclusion relies on the assumption that the minimum and maximum values being reported represent the same confidence intervals for both forecasted and historical outcomes. In practice, this may not be the case since farmers are asked to recall historical worst and best yields as specific events in specific years, whereas their forecast max and min values may not represent true extremes of the distribution. 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 1.3 using all three data points

per forecast in the regression. The new model results are shown in table 1.4: it is clear that the results are materially unchanged by the addition of mean/mode reporting.

[Table 1.4]

1.6 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 1.3, 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 1.5. 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 1.5]

[Figure 1.3]

Each GMM is fit using the well-known expectation-maximization (EM) algorithm of Dempster, Laird and Rubin (1977), an iterative 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:

$$\tau_{ic}^{t+1} = \frac{\pi_c^t \cdot f(x_i | \hat{\theta}_c^t)}{\sum_{j=1}^C \pi_j^t \cdot f(x_i | \hat{\theta}_j^t)} . \quad (1.4)$$

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:

$$\pi_c^{t+1} = \frac{1}{I} \sum_{i=1}^I \tau_{ic}^{t+1} . \quad (1.5)$$

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

$$\max_{\theta} LL(\theta; \pi) = \sum_{i=1}^I \ln \left(\sum_{c=1}^C \pi_c^{t+1} \cdot f(x_i | \theta_c) \right). \quad (1.6)$$

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 1.6. 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 1.6]

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 discuss 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 1.7.

[Table 1.7]

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 (especially in large data sets where the log-likelihood dominates the parameter penalty). 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 with outliers (due to the 4th 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 1.4.

[Figure 1.4]

[Table 1.8]

Table 1.8 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 1.9]

Table 1.9 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 1.7). The key distinction between information presented in table 1.8 versus table 1.9 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.

1.7 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 our evidence suggests they are not, in two ways. First, 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, only 65% of subjects growing both crops gave the same type of forecast for both crops. If the posterior probabilities indicate the probability that an *individual*

bears a particular classification, then this outcome is statistically impossible given the results in table 1.6 (*Pooled Data, C=4*), where 95% of forecasts had $\tau_{ic} \geq 0.99$ and 99% had $\tau_{ic} \geq 0.95$. Second, we can formally test the null hypothesis that an individual's forecast classification for one crop is independent of their forecast classification in the other crop. We would expect this null to be rejected if classifications are correlated within individuals, on average. Table 1.10 below shows the estimated cross-forecast correlation coefficients for each crop and class, for all farmers growing both crops ($N=396$). As can be clearly seen, none of the estimated correlation coefficients are statistically significant at the conventional 95% level on a two-sided test, even before making an adjustment for the multiplicity of tests. Thus, we fail to reject the null hypothesis that forecasts are, on average, independent across crops.

[Table 1.10]

We are forced to conclude that variation in forecast type, while it may be influenced by individual factors with effect sizes too small to measure, 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 1.11.

[Table 1.11]

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 1.1, the average time since a historic loss is 8.65 years with a standard deviation of 8.49 years, and with 3 and 11 years, respectively as the bounds of the interquartile range (IQR). 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 theory intuition, namely that losses are more salient than gains. Explicitly testing this hypothesis, however, is beyond the scope of both our data and this article.

As table 1.11 shows, 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 1.4, which plots the years since a historic loss in corn versus in wheat for those farmers growing both.

[Figure 1.4]

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 85% the scale of their historic yield distribution (i.e., $\hat{\beta} = 0.85$), female forecasters expect future yields to come from a distribution with roughly only 74% the scale of their historic distribution (i.e., $\hat{\beta} = 0.74$). This finding is consistent across both crops, despite the presence of within-subject heterogeneity in beliefs between crops.

Our results regarding gender differences in forecasts 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), and by white men in particular (Flynn et al., 1994). Despite race and gender being found to be strong predictive factors for risk perceptions, Finucane et al. (2000) found considerable variation across both males

and females of ethnic minorities in their replication of Flynn et al., concluding that risk is a social construct depending largely on the individual, even if its perception may be influenced by cultural and biological factors.

From the prospect theory literature, there has also been evidence of gender differences in risk preferences (as opposed to risk perceptions) 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. On the other hand, Tanaka, Camerer and Nguyen (2010) do not find significant gender differences in risk or time preferences, but Liu (2013) uses the same methodology and a larger data set and finds that females were significantly more risk averse, in the sense of the curvature (exponent) of the prospect theory value function.

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.

Persistence of Optimism Bias

As mentioned in the background section, existing evidence from Dalziel and Job (1997) suggests that optimism bias persists in the face of professional experience: they showed that even professional drivers tend to underestimate their risk of an accident. Our data support this finding, in the sense that we find no statistical evidence of changing optimism bias as a function of years of farming experience. In fact, across all models tested in table 1.11, our estimate of the marginal effect of experience is best described as being precisely estimated and near-zero in terms of economic significance while also not being statistically different from zero.

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 that 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:

$$Y = \alpha r + \beta X + \mathbf{1}_{mid} \cdot \gamma X . \quad (1.7)$$

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 1.12 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 disappearance of outlier groups is likely attributable to cleaner separation in the mixture model caused by higher dimensional data. 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 1.12]

1.8 Conclusion

Recent literature in agricultural economics has suggested that the widespread behavioral phenomenon of optimism bias may play a role in lowering crop insurance

participation in the absence of subsidies. Optimism Bias has been widely demonstrated across domains, including driver safety, job expectations, mergers and acquisitions, entrepreneurship and (non-crop) insurance purchases, yet no agreed upon model makes explicit the behavioral process of transforming historical experience into an optimistically biased forecast. To close this gap in the literature, we introduce a straightforward model in which farmers generate forecasts by shifting and scaling their historical yield distributions. We fit the model to 879 corn and wheat yield forecasts of farmers in China, and find that its predictions are robust both to varying amounts of data per forecast and also to inclusion of a shape parameter to accommodate higher moments (these turn out to be near zero). We also find that farmers anchor to past positive experience in making their forecasts.

We obtain new insight by exploring heterogeneous types of forecasts revealed by a mixture model: approximately two thirds of forecasts are found to be optimistic, understating their risks relative to historical experience, approximately 20% are realistic or neutral, and the remainder are pessimistic, predicting substantially higher risk in the future. These results are the first to give a modeled structure to repeated findings in the literature that, across domains, while a large majority of people tend to give optimistically biased forecasts, there is a consistent small minority observed who are more pessimistic than the rest. However, our results do not suggest that forecasts reflect personality types of the individual farmers, rather, they reflect crop-specific experience: in essence, forecasts seem to default towards being optimistic with recent large losses predicting a more pessimistic outlook which fades (by about 1% per year) as the losses recede into the past. We also observe statistically significant gender

effects, with female farmers reporting significantly more optimistic forecasts than their male counterparts. The gender difference is approximately equivalent to the difference between the top and bottom of the interquartile range in terms of loss experience.

Overall, our results provide a foundation for future modeling efforts to evaluate the structure of optimism bias across domains. Within agricultural economics, we expect that our efforts can form a basis for behavioral models of crop insurance demand in which optimistic, neutral or pessimistic forecasts dramatically alter the need for subsidies to achieve desired participation rates. Our results also suggest that future research is needed to estimate the degree to which optimism bias in farmer forecasts might be mitigated by interventions such as informational materials or more rigorous education or training facilitated by outreach and extension personnel. While many agricultural economists hold the perspective that crop insurance demand frictions in the United States have largely been “solved” through the use of high subsidies, our perspective is that the policy climate may change and better understanding the structure of insurance demand certainly cannot hurt. Moreover, parts of the world where generous crop insurance subsidies are not established (for example, in China, where crop insurance currently only covers production costs rather than revenue losses), understanding the role of optimism bias may be key to future advances in policy design. For example, given Sherrick’s (2002) finding that farmers are not optimistically biased with respect to weather, it may be that weather index efforts being explored in China (Turvey and Kong 2010) and elsewhere in the developing world (cf. Chantarat et al. 2007, among many) enjoy a

structural advantage over farm-level yield or revenue insurance products. Clearly, further research is needed to test whether these phenomena actually extrapolate in this manner.

A final important area for future research is reconciling the apparent phenomenon in the literature that farmers are optimistically biased, on average, with respect to both their yields and the prices they expect to receive for their harvest. In the case of our results here, a complete answer would require extending our model to a bivariate forecast and treating statistical dependence. Shaik et al. (2008) offer a potential insight in that they elicited both yield and price forecasts using a similar methodology to ours with farmers in the United States, and also elicited farmer estimates of the correlation coefficient of price and yield. On average, their farmers ($N=778$) reported price-yield correlations of -0.66 for corn and -0.43 for soybeans, indicating that their forecasts recognize the standard patterns of negative price-yield correlations driven by weather and supply-and-demand effects (see, e.g., Sproul and Kropp, 2015, among many). If farmers make optimistically biased forecasts with respect to both price and yield but they recognize the negative correlation between the two, then it will be important for future researchers to test the degree to which optimistically biased price and yield forecasts are influenced (and maybe simultaneously determined) by optimistically biased forecasts of farm income.

1.9 Works Cited

- Babcock, Linda, George Loewenstein, and Samuel Issacharoff. 1997. "Creating Convergence: Debiasing Biased Litigants." *Law and Social Inquiry* 22(4): 913-925.
- Barseghyan, Levon. 2013. "The Nature of Risk Preferences: Evidence from Insurance Choices." *The American Economic Review* 103(6): 2499-2529.
- Bartl, G., and G. Stummvoll. 2000. "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 .
- Benartzi, Shlomo, and Richard H. Thaler. 2001. "Naive Diversification Strategies in Defined Contribution Saving Plans." *American Economic Review*: 79-98.
- Bewley, J. M., M. D. Boehlje, A. W. Gray, H. Hogeveen, S. J. Kenyon, S. D. Eicher, and M. M. Schutz. 2010. "Assessing the Potential Value for An Automated Dairy Cattle Body Condition Scoring System Through Stochastic Simulation." *Agricultural Finance Review* 70(1): 126-150.
- Bhattacharya, Jayanta, David M. Cutler, Dana P. Goldman, Michael D. Hurd, Geoffrey F. Joyce, Darius N. Lakdawalla, Constantijn WA Panis, and Baoping Shang. 2004. "Disability forecasts and future Medicare costs." *Forum for Health Economics & Policy* 7.1.
- Biernacki, Christophe, Gilles Celeux, and Gérard Govaert. 2000. "Assessing a Mixture Model for Clustering With the Integrated Completed Likelihood." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(7): 719-725.

- Bruhin, Adrian, Helga Fehr-Duda, and Thomas Epper. 2010. "Risk and Rationality: Uncovering Heterogeneity in Probability Distortion." *Econometrica* 78(4): 1375-1412.
- Buckley, Neil J., Frank T. Denton, A. Leslie Robb, and Byron G. Spencer. 2004. "The Transition from Good to Poor Health: An Econometric Study of the Older Population." *Journal of Health Economics* 23(5): 1013-1034.
- Camerer, Colin, and Dan Lovallo. 1999. "Overconfidence and Excess Entry: An Experimental Approach." *The American Economic Review* 89(1): 306-318.
- Cebulla, Andreas. 1999. "A Geography of Insurance Exclusion: Perceptions of Unemployment Risk and Actuarial Risk Assessment." *AREA* 31(2): 111-121.
- Celeux, Gilles, and Gilda Soromenho. 1996. "An Entropy Criterion for Assessing the Number of Clusters in a Mixture Model." *Journal of Classification* 13(2): 195-212.
- Chantarat, S., Barrett, C.B., Mude, A.G. and Turvey, C.G., 2007. "Using weather index insurance to improve drought response for famine prevention." *American Journal of Agricultural Economics*, 89(5), pp.1262-1268.
- Chiappori, Pierre-André, Bruno Jullien, Bernard Salanié, and Francois Salanie. 2006. "Asymmetric Information in Insurance: General Testable Implications." *The RAND Journal of Economics* 37(4): 783-798.
- Coble, K.H. and B. Barnett. 2013. "Why Do We Subsidize Crop Insurance?" *American Journal of Agricultural Economics* 95(2): 498-504.
- Coble, K H., T. O. Knight, R. D. Pope, and J. R. Williams. 1997. "An Expected Indemnity Approach To the Measurement of Moral Hazard in Crop Insurance."

- American Journal of Agricultural Economics* 79(1): 216-26.
- Cooper, Arnold C., Carolyn Y. Woo, and William C. Dunkelberg. 1988.
“Entrepreneurs' Perceived Chances for Success.” *Journal of Business Venturing* 3(2): 97-108.
- Cooper, Joseph C. 2010. “Average Crop Revenue Election: A Revenue-Based Alternative To Price-Based Commodity Payment Programs.” *American Journal of Agricultural Economics*: AAQ048.
- Cutler, David M., and Richard Zeckhauser. 2004. “Extending the Theory To Meet the Practice of Insurance.” *Brookings-Wharton Papers on Financial Services* 2004(1): 1-53.
- Dalziel, James R., and Rf Soames Job. 1997. “Motor Vehicle Accidents, Fatigue and Optimism Bias in Taxi Drivers.” *Accident Analysis and Prevention* 29(4): 489-494.
- De Bondt, Werner Fm, and Richard H. Thaler. 1995. “Financial Decision-Making in Markets and Firms: A Behavioral Perspective.” *Handbooks in Operations Research and Management Science* 9: 385-410.
- Dempster, Arthur P., Nan M. Laird, and Donald B. Rubin. 1977. “Maximum Likelihood from Incomplete Data Via the EM Algorithm.” *Journal of the Royal Statistical Society. Series B B (Methodological)*: 1-38
- Eales, James S., Brian K. Engel, Robert J. Hauser, and Sarahelen R. Thompson. 1990.
“Grain Price Expectations of Illinois Farmers and Grain Merchandisers.” *American Journal of Agricultural Economics* 72(3): 701-708.
- Egelkraut, T. M., Garcia, P., Pennings, J. M., and Sherrick, B. J. 2006. “Producers'

- Yield and Yield Risk: Perceptions Versus Reality and Crop Insurance Use.”
*Agricultural and Applied Economics Association Annual Meeting, July 23-26,
 Long Beach, CA.*
- Enjolras, Geoffroy, and Patrick Sentis. 2011. “Crop Insurance Policies and Purchases
 in France.” *Agricultural Economics* 42(4): 475-486.
- Fehr-Duda, Helga, Manuele De Gennaro, and Renate Schubert. 2006. “Gender,
 Financial Risk, and Probability Weights.” *Theory and Decision* 60(2): 283-313.
- Finucane, Melissa L., Paul Slovic, Chris K. Mertz, James Flynn, and Theresa A.
 Satterfield. 2000. “Gender, Race, and Perceived Risk: The 'White Male' Effect.”
Health, Risk and Society 2(2): 159-172.
- Flynn, James, Paul Slovic, and Chris K. Mertz. 1994. “Gender, Race, and Perception
 of Environmental Health Risks.” *Risk Analysis* 14(6): 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. 1996. “Self-Preserving Assessments of
 Skill?.” *British Journal of Psychology* 87(1): 61-79.
- Güner, A. Burak, Ulrike Malmendier, and Geoffrey Tate. 2006. “The Impact of
 Boards With Financial Expertise on Corporate Policies.” *National Bureau of
 Economic Research.*
- Gutteling, Jan M., and Oene Wiegman. 1993. “Gender-Specific Reactions To
 Environmental Hazards in the Netherlands.” *Sex Roles* 28(7): 433-447

- Haigh, M. S., and List, J. A. 2005. "Do Professional Traders Exhibit Myopic Loss Aversion? An Experimental Analysis". *The Journal of Finance* 60(1): 523-534.
- Hatakka, Mika, Esko Keskinen, Niels Peter Gregersen, Alf Glad, and Kati Hernetkoski. 2002. "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): 201-215.
- Hoch, Stephen J. 1985. "Counterfactual Reasoning and Accuracy in Predicting Personal Events." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 11(4): 719.
- Hoorens, Vera. 1994. "Unrealistic Optimism in Social Comparison of Health and Safety Risks." *The Social Psychology of Health and Safety*: 153-174.
- Horowitz, John K., and Erik Lichtenberg. 1993. "Insurance, Moral Hazard, and Chemical Use in Agriculture." *American Journal of Agricultural Economics* 75(4): 926-935.
- Just, R. E., L. Calvin, and J. Quiggin. 1999. "Adverse Selection in Crop Insurance: Actuarial and Asymmetric Information Incentives." *American Journal of Agricultural Economics* 81(4):834-49.
- Kenyon, D. E. 2001, "Producer Ability to forecast Harvest Corn and Soybean Prices". *Review of Agricultural Economics* 23(1):151-162.
- Ker, Alan P., Tor N. Tolhurst, and Yong Liu. 2015. "Bayesian Estimation of Possibly Similar Yield Densities: Implications for Rating Crop Insurance Contracts." *American Journal of Agricultural Economics* 98(2): 360-382.
- Kreuter, Matthew W. and Victor J. Strecher. 1995. "Changing Inaccurate Perceptions

- of Health Risk: Results from a Randomized Trial.” *Health Psychology* 14(1): 56.
- List, J. A. “Does Market Experience Eliminate Market Anomalies?” *The Quarterly Journal of Economics* 118(1): 41-71.
- Landier, Augustin, and David Thesmar. 2009. “Financial Contracting With Optimistic Entrepreneurs.” *Review of Financial Studies* 22(1): 117-150.
- Liu, Elaine M. 2013. “Time To Change What To Sow: Risk Preferences and Technology Adoption Decisions of Cotton Farmers in China.” *Review of Economics and Statistics* 95(4): 1386-1403.
- Malcolm, D. G., J. H. Roseboom, C. E. Clark, and W. Fazar. 1959. “Application of a Technique for Research and Development Program Evaluation.” *Operations Research*: 646-669.
- Malmendier, Ulrike, and Geoffrey Tate. 2008. “Who Makes Acquisitions? Ceo Overconfidence and the Market's Reaction.” *Journal of Financial Economics* 89(1): 20-43.
- Pease, James W. 1992. “A Comparison of Subjective and Historical Crop Yield Probability Distributions.” *Southern Journal of Agricultural Economics* 24.02: 23-32.
- Pease, James W., Ernest W. Wade, Jerry S. Skees, and Chandra M. Shrestha. 1993. “Comparisons Between Subjective and Statistical forecasts of Crop Yields.” *Review of Agricultural Economics* 15(2): 339-350.
- Riley, John Michael, and John D. Anderson. 2009. “Producer Perceptions of Corn, Soybean and Cotton Price Risk.” In *Southern Agricultural Economics*

Association Annual Meeting, Atlanta, Georgia, January.

RMA United States Department of Agriculture Risk Management Agency. Trend-Adjusted Actual Production History.

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

Sandroni, Alvaro, and Francesco Squintani. 2007. "Overconfidence, Insurance, and Paternalism." *The American Economic Review* 97(5): 1994-2004.

Shaik, Saleem, Keith Coble, Thomas Knight, Alan Baquet, and George F. Patrick. 2008. "Crop Revenue and Yield Insurance Demand: A Subjective Probability Approach." *Journal of Agricultural and Applied Economics* 40(3): 757-766.

Sherrick, Bruce J. 2002. "The Accuracy of Producers' Probability Beliefs: Evidence and Implications for Insurance Valuation." *Journal of Agricultural and Resource Economics* 27(1): 77-93.

Sherrick, Bruce J., Peter J. Barry, Paul N. Ellinger, and Gary D. Schnitkey. 2004. "Factors Influencing Farmers' Crop Insurance Decisions." *American Journal of Agricultural Economics* 86(1): 103-114.

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

Skees, J., and M. Reed. 1986. "Rate Making for Farm-Level Crop Insurance: Implications for Adverse Selection." *American Journal of Agricultural Economics* 68(3): 653-59.

Slovic, Paul. *The Perception of Risk: Risk, Society and Policy*. 2000. London:

Earthscan Publications.

- Slovic, Paul, Torbjörn Malmfors, C. K. Mertz, Nancy Neil, and Iain Fh Purchase. 1997. "Evaluating Chemical Risks: Results of a Survey of the British Toxicology Society." *Human and Experimental Toxicology* 16(6): 289-304.
- Smith, V. H., and B. K Goodwin. 1996. "Crop Insurance, Moral Hazard, and Agricultural Chemical Use." *American Journal of Agricultural Economics* 78(2): 428-38.
- Spinnewijn, Johannes. 2013. "Insurance and Perceptions: How to Screen Optimists and Pessimists." *The Economic Journal* 123.569: 606-633.
- Sproul, Thomas W. and Jaclyn D. Kropp. 2015. "A general equilibrium theory of contracts in community supported agriculture." *American Journal of Agricultural Economics*, 97(5): 1345-1359.
- Sproul, Thomas W., and Clayton P. Michaud 2017. "Heterogeneity in Loss Aversion: Evidence from Field Elicitations." *Agricultural Finance Review* 77(1): 196-216.
- Statman, Meir, and Tyzoon T. Tyebjee. 1985. "Optimistic Capital Budgeting forecasts: An Experiment." *Financial Management*: 27-33.
- Stern, Paul C., Thomas Dietz, and Linda Kalof. 1993. "Value Orientations, Gender, and Environmental Concern." *Environment and Behavior* 25(5): 322-34
- Svenson, Ola. 1981. "Are We All Less Risky and More Skillful Than Our Fellow Drivers?." *Acta Psychologica* 47(2): 143-148.
- Tanaka, Tomomi, Colin F. Camerer, and Quang Nguyen. 2010. "Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam." *American Economic Review* 100(1): 557-571.

- Turvey, C. G., Gao, X., Nie, R., Wang, L., and Kong, R. 2013. "Subjective Risks, Objective Risks and the Crop Insurance Problem in Rural China." *The Geneva Papers on Risk and Insurance Issues and Practice* 38(3): 612-633.
- Turvey, C.G. and Kong, R., 2010. "Weather risk and the viability of weather insurance in China's Gansu, Shaanxi, and Henan provinces." *China Agricultural Economic Review*, 2(1), pp.5-24.
- Tversky, Amos, and Daniel Kahneman. 1992. "Advances in Prospect Theory: Cumulative Representation of Uncertainty." *Journal of Risk and Uncertainty* 5(4): 297-323.
- Umarov, Alisher, and Bruce J. Sherrick . 2005. "Farmers' Subjective Yield Distributions: Calibration and Implications for Crop Insurance Valuation." In *American Agricultural Economics Association Annual Meeting in Providence, RI, July*.
- Wang, Ming, Tao Ye and Peijun Shi. "Factors Affecting Farmers' Crop Insurance Participation in China." *Canadian Journal of Agricultural Economics* 64: 479-492.
- Weinstein, Neil D. 1980. "Unrealistic Optimism About Future Life Events." *Journal of Personality and Social Psychology* 39(5): 806.
- Wengler, Leif, and Anne-Sofie Rosen. 2000. "Measuring Optimism–Pessimism from Beliefs About Future Events." *Personality and Individual Differences* 28(4): 717-728.
- Woodard, Joshua D., and Bruce J. Sherrick. 2011. "Estimation of Mixture Models Using Cross-Validation Optimization: Implications for Crop Yield Distribution

Modeling.” *American Journal of Agricultural Economics* 93(4): 968-982.

Woodard, Joshua D., and Leslie J. Verteramo-Chiu. 2017. “Efficiency Impacts of Utilizing Soil Data in the Pricing of the Federal Crop Insurance Program.” *American Journal of Agricultural Economics* 99(3): 757-772.

1.10 Figures

Figure 1.1: 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

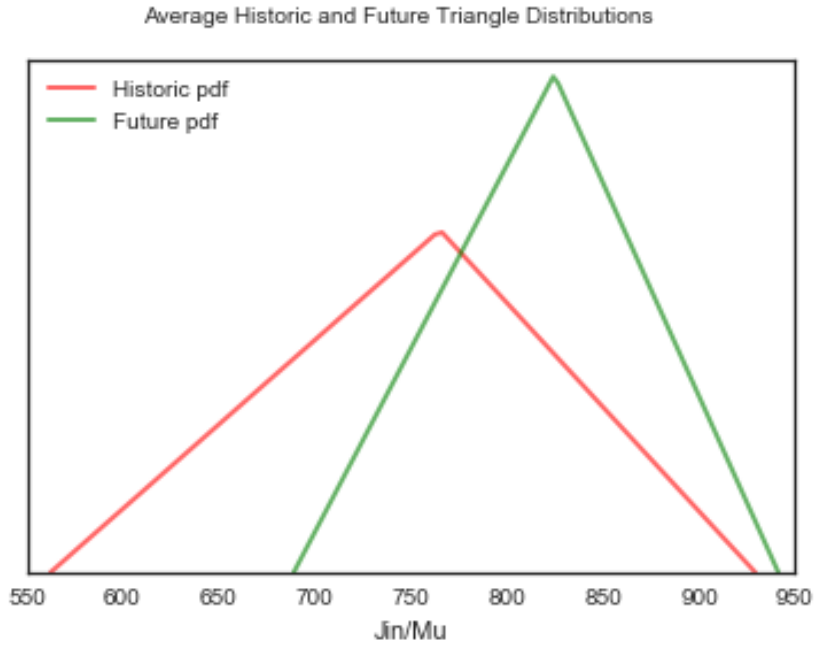
Crop	Lowest possible yield (jin/mu)	Most likely yield (jin/mu)	Highest possible yield (jin/mu)
1 Corn			
2 Wheat			

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 1.2: Elicited Historic vs. Future Yield Distributions

Panel A: Triangular Distribution



Panel B: Beta-PERT Distribution

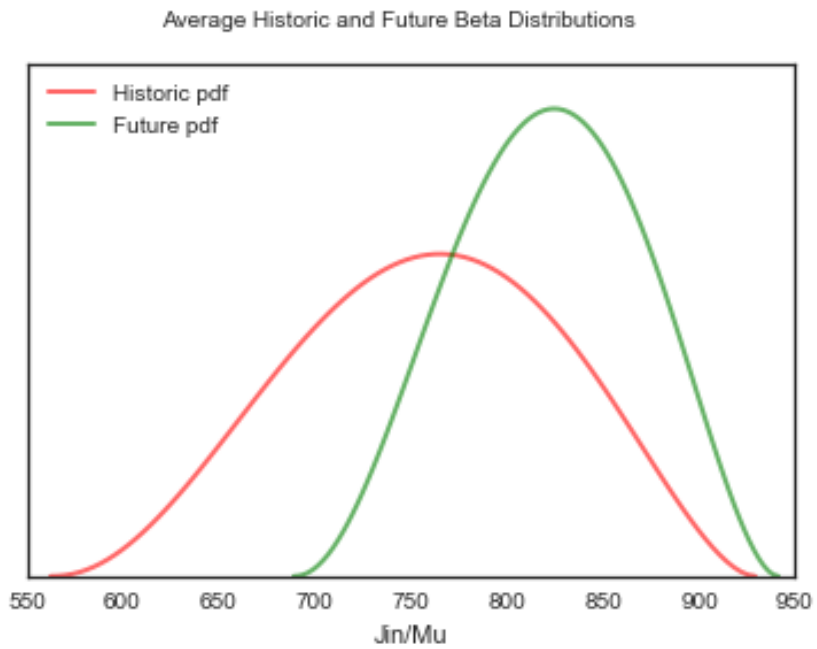


Figure 1.3: Hex Plot of Alpha and Beta with Marginal Densities

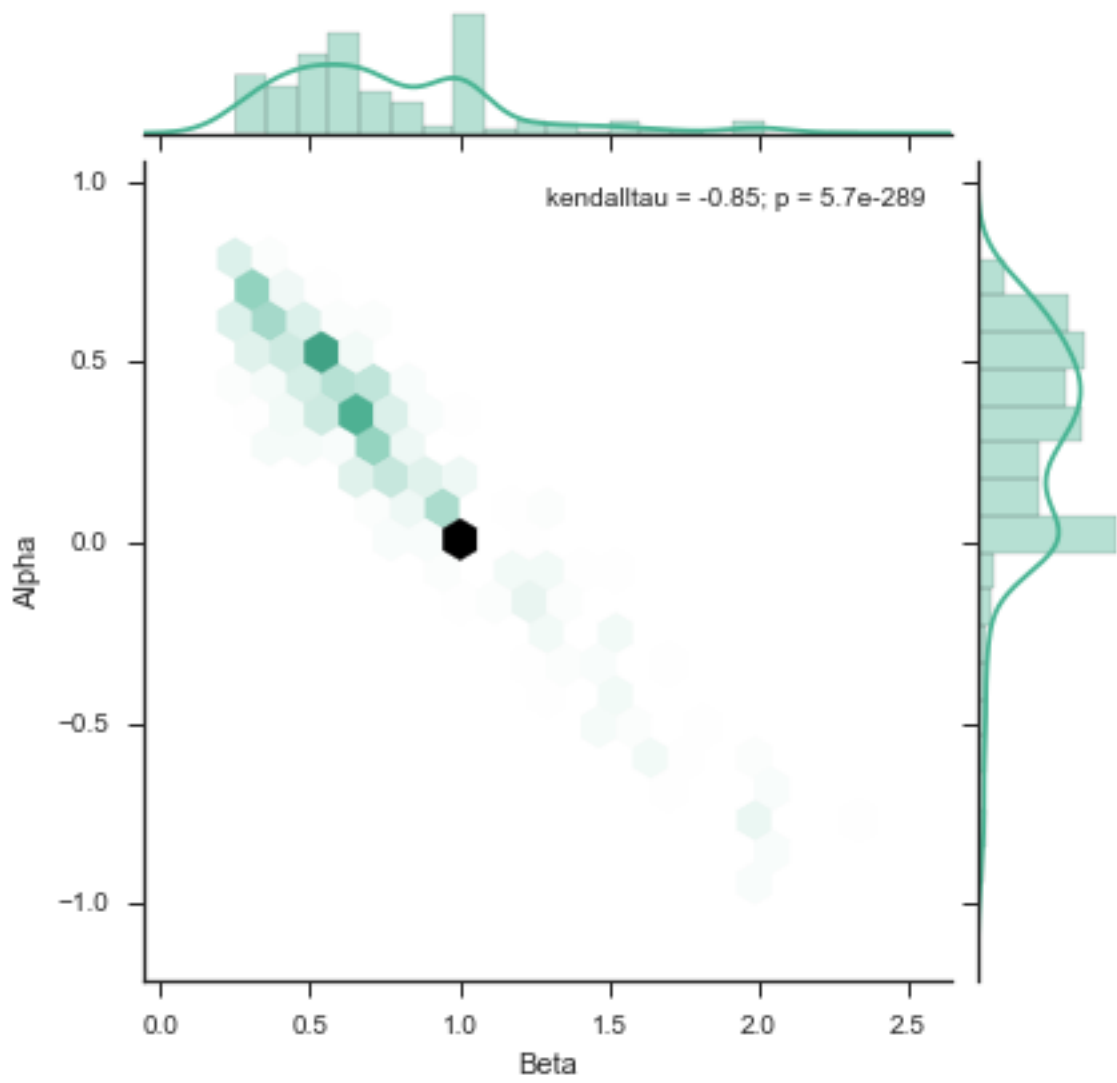
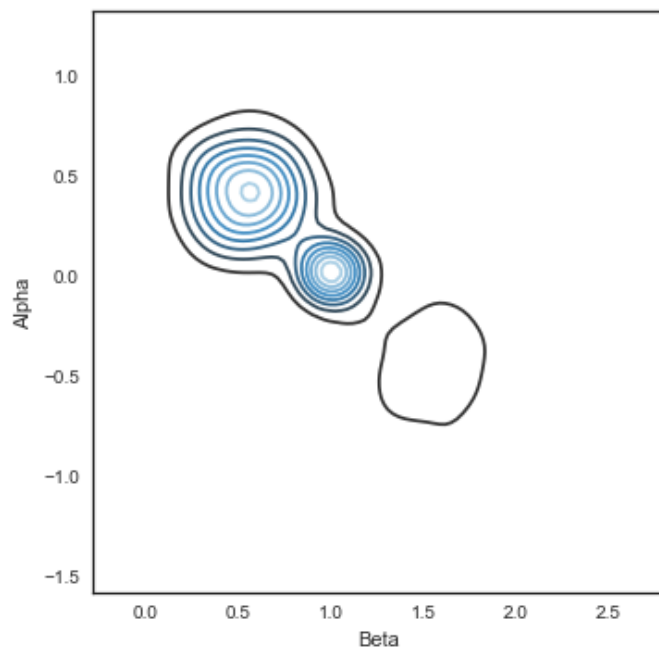


Figure 1.4: Contour Plots for Fitted 3-Mixture (Panel A) and 4-Mixture (Panel B) Models
Panel A



Panel B

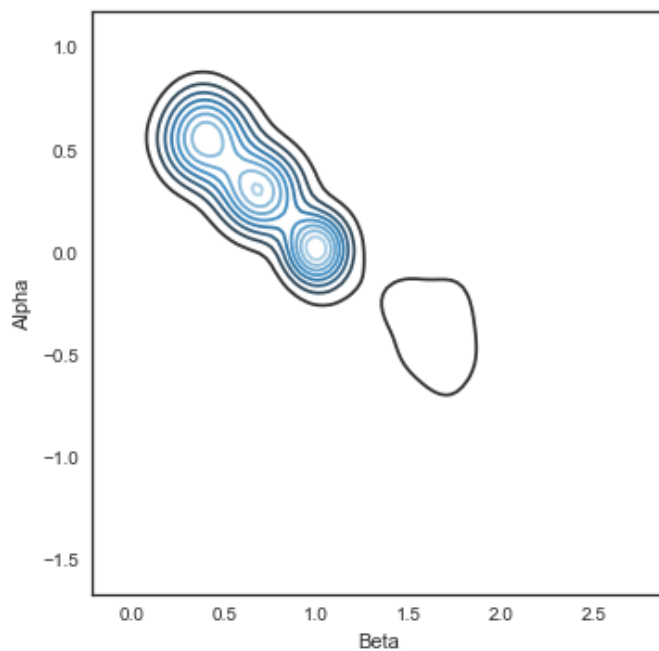
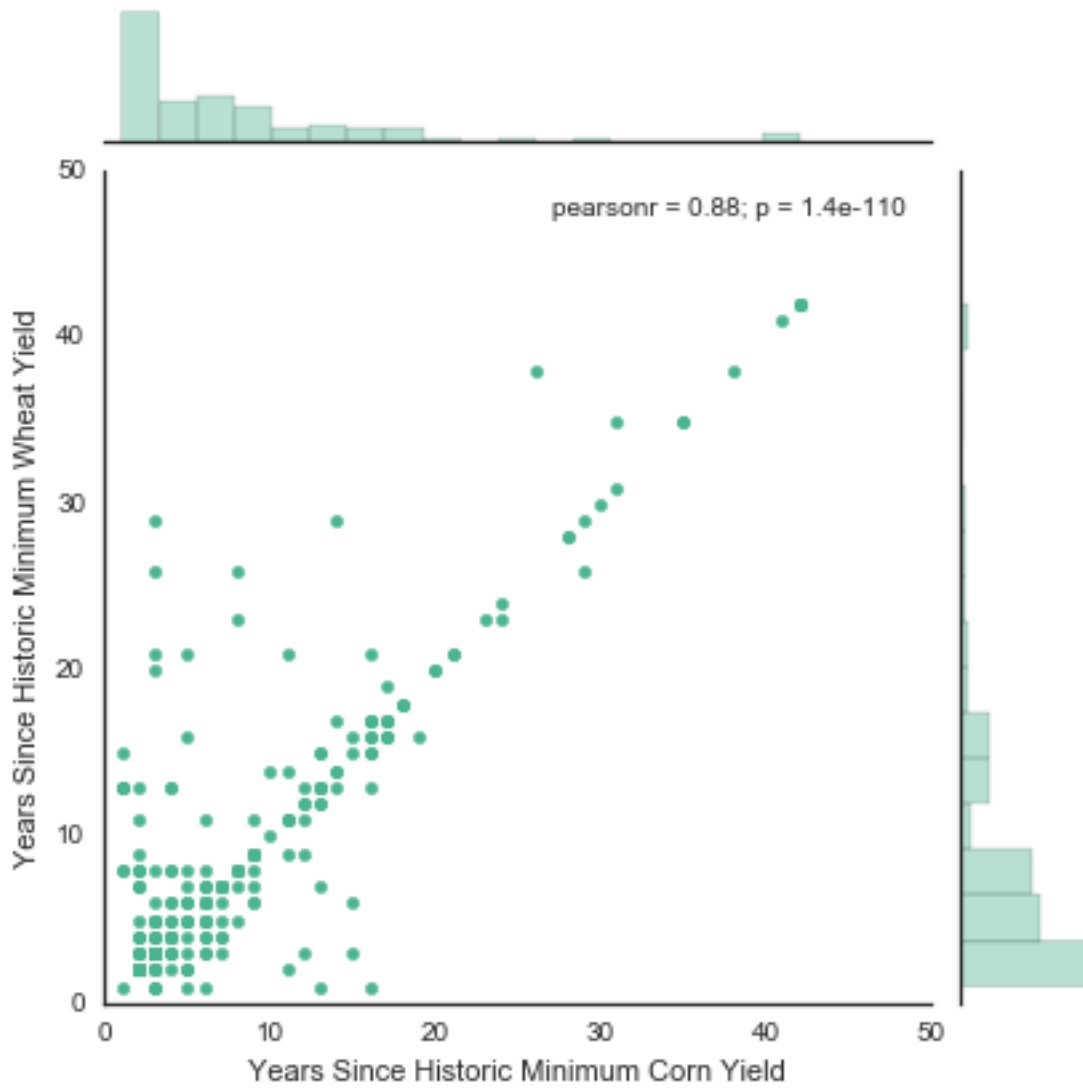


Figure 1.5: Scatter Plot: Years Since Historic Minimum Yield (Corn and Wheat)



1.11 Tables

Table 1.1: Household-Level Summary Statistics

	Mean	S.D.	Min.	25%	50%	75%	Max.
Years since Min Corn Yield ($N=432$)	8.42	8.48	1	3	6	11	42
Years since Min Wheat Yield ($N=423$)	8.88	8.52	1	3	6	12	42
Years since Max Corn Yield ($N=432$)	1.9	1.51	1	1	1	2	14
Years since Max Wheat Yield ($N=426$)	2.1	1.72	1	1	2	2	14
Gender ($I=Male$)	0.54	0.5	0	0	1	1	1
Age	48.09	10.93	20	40	50	56	72
Education	4.36	1.76	0	3	5	5	8
Years Farming	27.01	13.04	0	19	30	37	65
Percent Farm Income	41.08	26.95	0	20	35	58.1	100
Household Income (Yuan/year) ($N=479$)	23615	22316	1000	10000	20000	30000	248000

Note: $N=483$ unless otherwise noted (396 households grow both corn and wheat, 46 grow only corn and 41 grow only wheat). Not all households provided income or the years of historic gains or losses.

Table 1.2: Testing for the Presence of Yield Trends

<i>Dependent Variable: Historic Mean</i>	Model 1	Model 2
Intercept	822.849*** (106.827)	820.427*** (109.528)
Years Farming	2.6811 (3.181)	2.768 (3.296)
Years Farming ²	-0.054 (0.056)	-0.054 (0.056)
Sex (Male=1)	2.946 (25.647)	2.989 (25.665)
Age (Years)	-1.483 (1.696)	-1.485 (1.697)
Education	3.989 (7.162)	3.995 (7.167)
Crop Dummy (Corn=1)	-27.617 (21.254)	-23.107 (49.319)
Crop Dummy x Years Farming		-0.167 (1.644)
N	873	873
Adjusted R ²	0.022	0.020

Notes: Standard errors in parentheses. ***, **, and * denote statistical significance at the 99.9%, 99% and 95% levels, respectively, using a two-sided test. Village fixed effects not shown.

Table 1.3: Model Selection for Candidate Forecasting Models

	Model 1	Model 2	Model 3	Model 4
Scale	0.6378*** (0.011)	0.5821*** (0.012)	0.8457*** (0.014)	0.8926*** (0.010)
Constant	337.25*** (8.442)			
$r =$				
Historic Max		0.4053*** (0.010)		
Historic Min			0.2856*** (0.019)	
Historic Mode				0.1616*** (0.009)
Adj. R-squared	0.664	0.984	0.972	0.973
Log-Likelihood	-10736	-10687	-11189	-11083
AIC	21480	21380	22380	22170
BIC	21490	21390	22390	22180

Notes: *** Signifies statistical significance at the 99.9% level. N=879.

Table 1.4: Model Selection (2 parameters, 3 observations)

	Model 1	Model 2	Model 3	Model 4
Scale	0.3810*** (0.016)	0.6264*** (0.014)	0.5542*** (0.017)	0.8533*** (0.019)
Constant	530.68*** (12.941)			
$r =$				
Historic Max		0.3106*** (0.017)		
Historic Min			0.6597*** (0.024)	
Historic Mode				0.1551*** (0.018)
Adj. R-squared	0.174	0.937	0.914	0.893
Log-Likelihood	-17980	-17879	-18278	-18527
AIC	35960	35760	36560	37060
BIC	35980	35770	36570	37070

Notes: *** Signifies statistical significance at the 99.9% level. N=879.

Table 1.5: Summary of α and β by Gender and by Crop

<i>Corn Farmers (N=442)</i>				
	Males (N=241)		Females (N=201)	
	α	β	α	β
Mean	0.157	0.854	0.300	0.712
Std	0.595	0.638	0.331	0.371

<i>Wheat Farmers (N=437)</i>				
	Males (N=242)		Females (N=195)	
	α	β	α	β
Mean	0.170	0.851	0.259	0.762
Std	0.536	0.600	0.484	0.506

Note: Of the male farmers, 222 grow both corn and wheat. Of the female farmers, 174 grow both.

Table 1.6: Model Selection Criteria by Number of Components, C

C	N	AICc	BIC	ICL	$\tau > .99$	$\tau > .95$	$\tau > .90$
<i>Corn</i>							
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
<i>Wheat</i>							
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
<i>Pooled</i>							
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 1.7: Mixture Weights by Number of Components, C

C	Weights					
<i>Corn</i>						
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
<i>Wheat</i>						
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
<i>Pooled</i>						
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 1.8: 3-Component Mixture Model (with outliers omitted)

Corn Forecasts (N=442)

	<u>Optimistic</u>		<u>Realistic</u>		<u>Pessimistic</u>	
	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)

	<u>Optimistic</u>		<u>Realistic</u>		<u>Pessimistic</u>	
	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)

	<u>Optimists</u>		<u>Realists</u>		<u>Pessimists</u>	
	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 1.9: 4-Component Mixture Model (with outliers omitted)

Corn Forecasts (N=442)

	<u>Highly</u>		<u>Mildly</u>		<u>Realistic</u>		<u>Pessimistic</u>	
	<u>Optimistic</u>		<u>Optimistic</u>					
	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)

	<u>Highly</u>		<u>Mildly</u>		<u>Realistic</u>		<u>Pessimistic</u>	
	<u>Optimistic</u>		<u>Optimistic</u>					
	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)

	<u>Highly</u>		<u>Mildly</u>		<u>Realistic</u>		<u>Pessimistic</u>	
	<u>Optimistic</u>		<u>Optimistic</u>					
	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 1.10: Correlation Coefficients of Classes Across Forecasts

		Wheat		
		Optimistic	Neutral	Pessimistic
Corn	Optimistic	0.07	-0.05	-0.02
	Neutral	-0.05	0.05	0.01
	Pessimistic	-0.02	0.01	0.02

Notes: Only farmers growing both crops (N=396). Pearson (linear) correlation coefficients calculated with each class X crop as a dummy variable. None of the coefficients are statistically significant at the 95% level (two-sided test).

Table 1.11: Regression Results

	β			α		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Intercept	0.825*** (0.145)	0.844*** (0.146)	0.875*** (0.171)	0.191 (0.133)	0.178 (0.135)	0.199 (0.157)
Crop Dummy (Corn=1)		-0.034 (0.037)	-0.052 (0.039)		0.024 (0.034)	0.034 (0.036)
Sex (Male=1)	0.109*** (0.041)	0.109*** (0.041)	0.122*** (0.043)	-0.108*** (0.038)	-0.107*** (0.038)	-0.116*** (0.039)
Age (Years)	0.000 (0.003)	0.000 (0.003)	0.000 (0.003)	-0.000 (0.003)	-0.000 (0.003)	-0.000 (0.003)
Education	-0.004 (0.012)	-0.004 (0.012)	-0.004 (0.012)	0.003 (0.011)	0.003 (0.011)	0.004 (0.011)
Years Farming	0.001 (0.002)	0.001 (0.002)	-0.001 (0.006)	-0.001 (0.002)	-0.001 (0.002)	-0.001 (0.005)
Years Farming ²			0.000 (0.0)			0.000 (0.0)
Years since Min Yield	0.013*** (0.002)	-0.013*** (0.002)	-0.010* (0.005)	0.012*** (0.002)	0.012*** (0.002)	0.009* (0.005)
Min Yield Last Year Dummy			0.136** (0.068)			-0.136** (0.062)
Years since Max Yield	-0.003 (0.012)	-0.004 (0.012)	-0.003 (0.019)	0.003 (0.011)	0.004 (0.011)	0.005 (0.017)
Max Yield Last Year Dummy			-0.007 (0.064)			0.006 (0.059)
Min Yield Year > Max Yield Year			-0.108 (0.094)			0.063 (0.087)
Years since Min Yield, Other Crop			-0.002 (0.005)			0.002 (0.005)
Years since Max Yield, Other Crop			0.012 (0.014)			-0.008 (0.013)
Adj. R-squared	0.04	0.04	0.04	0.05	0.05	0.05
N	848	848	824	848	848	824

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

Table 1.12: Classification Results for the 3-parameter Model (Pooled Data)

Panel A: Model Selection

C	AICc	BIC	ICL	$\tau > .99$	$\tau > .95$	$\tau > .90$	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)

	<u>Optimistic</u>		
	α	β	γ
Mean	0.44	0.55	-0.01
Std. Dev.	0.166	0.171	0.087
Weight		0.669	
	<u>Realistic</u>		
	α	β	γ
Mean	0.02	1.00	0.01
Std. Dev.	0.062	0.035	0.090
Weight		0.221	
	<u>Pessimistic</u>		
	α	β	γ
Mean	-0.37	1.51	0.03
Std. Dev.	0.290	0.327	0.156
Weight		0.111	

Manuscript 2

The Effect of Optimism Bias on Cross-Coverage Level Crop Insurance Demand

This manuscript has not yet been submitted for publication

It is single authored by Clayton Michaud

Abstract

This manuscript examines the effect of producer misperceptions of yield risk on cross-coverage level demand for crop insurance. Using real-world data, we simulate both yield and revenue insurance coverage decisions for two representative farms (“low risk” and “high risk”) across four potential models of risk miscalibration and three potential models of decision-making (generalized utility models). We find that the effect of optimism bias is remarkably consistent across our 48 scenarios. Given current crop insurance subsidy levels, we find the counterintuitive result that as optimism bias regarding future yields increases, producers choose to insure their crops at decreasing levels of coverage, eventually opting not to purchase insurance at all. Additionally, we find that the magnitude of this effect is larger in the case of yield insurance than revenue insurance. We perform a further sensitivity analysis on the effects of price-yield correlation but find that our results are not highly sensitive to this parameter. Furthermore, we demonstrate that these results are driven primarily by current subsidy levels, namely the fact that lower coverage levels are awarded higher proportionate subsidies. We go on to discuss how these results can be used to develop more cost-effective crop insurance subsidy policies, and discuss avenues of future research.

2.1 Introduction

Optimism Bias is a ubiquitous and well-known phenomenon wherein individuals assume they are more likely to succeed and less likely to fail than their peers (Slovic, 1987; DeBondt and Thaler, 1995). There is a large body of evidence that such optimism bias leads to biased decision-making and has affects over a wide swath of domains ranging from health outcomes (Hoorens, 1994), legal settlements (Babcock et al., 1997), mergers and acquisitions (Malmendier and Tate, 2006, 2008), and insurance purchasing decisions (Bhattacharya et al., 2004; Barseghyan, 2013; Spinnewijn, 2013, 2015, 2017). For a more in-depth review of the optimism bias literature, see Manuscript 1 of this dissertation.

Optimism bias has also been cited as contributing to what has historically been referred to as the Crop Insurance Demand Puzzle, the observation that farmers are much more reluctant to participate in farm risk-management programs such as crop insurance than economic models would suggest, and that as a result they often require large subsidies in order to participate. A 1988 GAO Congressional Report on crop insurance participation and the costs associated with the federal program lists optimism as a possible reason why farmers were choosing not to participate in the federal crop insurance program, stating “some farmers tend to underestimate their crop losses in bad years, and expect profitable harvests in future years. Such optimism reduces the expected benefit from crop insurance”. The empirical work supports such a hypothesis. Pease (1992), Pease et al. (1993), Umarov and Sherrick, (2005) Egelkraut et al. (2006), and Turvey et al. (2013) find that on average, farmers are overly optimistic about both their yields and yield variability, expecting better than average yields and below average yield risk. A number of other studies have gone on to show how perceived yield risk

affects demand for crop insurance (Horowitz and Lichtenberg, 1993; Sherrick et al., 2004; Egelkraut et al., 2006; Shaik et al., 2008, Enjolras and Sentis, 2011; Wang, Ye and Shi, 2016).

While optimism bias has been argued as contributing to lower than expected willingness-to-pay for crop insurance, another important question regarding farmers' insurance decisions remains unanswered, the question of why farmers choose the coverage levels that they do. Expected utility theory predicts that producers will purchase crop insurance at a coverage level at least as high as the level that maximizes per-acre subsidies. In reality, however, there is no statistical evidence to support this hypothesis (Du, Feng, and Hennessy, 2016). To the best of our knowledge, no research exists looking at the effect of optimism bias on farmers' decisions of which coverage level to purchase crop insurance at. To that end, this paper simulates the effect of optimism (pessimism) bias on cross-coverage level demand for both crop yield and revenue insurance. We consider three possible models of decision-making – expected utility, cumulative prospect theory wherein decisions are framed within the broader context of farm risk-management, and a cumulative prospect theory model in which decisions/outcomes are more narrowly framed. Furthermore, we consider four possible manifestations of optimism (pessimism) bias: (1) optimism bias with regard to downside yield risk; (2) optimism (pessimism) bias regarding upside yield risk; (3) optimism bias with regard only to the variability of yields, while accurate beliefs about mean yields are maintained; (4) and optimism (pessimism) bias regarding the shape (skewness) of the yield distribution. These four models of optimism bias are mathematically equivalent to: (1) increasing the mean and decreasing the variance of the perceived yield distribution as compared to the objective yield distribution by increasing the lower-bound, while

holding the upper-bound fixed; (2) increasing (decreasing) the mean and decreasing (increasing) the variance of the perceived yield distribution as compared to the objective yield distribution by increasing (decreasing) the upper-bound, while holding the lower-bound fixed; (3) decreasing the variance of the perceived yield distribution as compared to the objective yield distribution, while holding the mean fixed; and (4) decreasing (increasing) the skewness of the perceived yield distribution as compared to the objective yield distribution such that it results in increasing (decreasing) the mean and scaling the variance.

We find that the effects of optimism bias on cross-coverage level demand for crop insurance are consistent across all 12 models of decision-making under biased perceptions of yield risk, and are as follows. (1) As optimism bias increases (decreases), farmers prefer lower (higher) levels of crop insurance coverage than they would otherwise prefer, with extreme optimism bias inducing farmers to buy no insurance at all. (2) Optimism bias affects cross-coverage demand for revenue and yield insurance similarly, however the effect is more pronounced for yield insurance. (3) The effect of optimism bias on cross-coverage demand for revenue insurance is relatively stable across changes in the correlation between prices and yields. We go on to discuss why our main result is both surprising and significant to policy design, pointing out that such a result is primarily driven by decreasing proportionate subsidies as coverage levels increase. Further simulations go on to demonstrate that under equal proportionate subsidies across coverage levels, expected utility maximizers will always choose the maximum level of coverage until optimism bias eventually induces them not to purchase insurance at all.

2.2 Modeling Optimism bias

In our model of decision-making under optimism bias we assume that farmers make their decision of which coverage level of insurance to purchase based not on $F(y, P)$, the true distribution of yields and prices³, but rather based on $F(\hat{y}, P)$, which is their forecasted or perceived distribution of yields and prices which need not be equal to the true distribution. We model forecasted or perceived yield risk as being miscalibrated from the objective yield risk such that their perceived yield distribution is a function of the true distribution of yields, and one or more miscalibration parameters such that $\hat{y}_i(y_i, \phi)$. We consider four potential models of optimism (pessimism) bias. Following standard yield modeling procedures we assume that y_i and \hat{y}_i each follow a beta distribution, $B(a, b, S_1, S_2)$ and $B(\hat{a}, \hat{b}, \hat{S}_1, \hat{S}_2)$, where a and b are the lower and upper bounds of the distribution, respectively, while S_1 and S_2 are shape parameters. Given these four parameters, the first three moments can be defined as follows:

$$E[y] = a + \frac{S_1}{S_1 + S_2} (b - a) = \frac{S_1 b + S_2 a}{S_1 + S_2},$$

$$Var[y] = \frac{S_1 S_2}{(S_1 + S_2)^2 (S_1 + S_2 + 1)} (b - a)^2,$$

$$Skew[y] = \frac{2(S_2 - S_1)\sqrt{S_1 + S_2 + 1}}{(S_1 + S_2 + 2)\sqrt{S_1 S_2}}.$$

Additionally, we assume the lower bound of the objective yield distribution, a , to be zero. This is done both because this is the case for both of the objective crop yield distributions we will be simulating outcomes from, as well as because for no crop is it

³ The distribution of yields and prices is generalized to be a joint distribution to allow for potential price-yield correlations.

ever truly the case that $\Pr(y_i = 0) = 0$. Prices are assumed to be distributed according to a log-normal distribution which is accurately perceived by the farmer, and may or may not be correlated with yields.

Optimism Bias Model 1: Underestimating Downside Yield Risk

Under this model, farmers increase the mean and decrease the variance of their perceived yield distribution relative to their objective yield distribution by increasing the lower-bound, while holding the upper-bound fixed such that

$$\hat{y}_i \sim B(\hat{a} = (1 - \beta_1)b, \hat{b} = b, \hat{S}_1 = S_1, \hat{S}_2 = S_2)$$

where β_1 is our miscalibration parameter and where optimism bias is defined as $\beta_1 < 1$.

Our resulting miscalibration function is thus $\hat{y}_i = (1 - \beta_1)b + \beta_1 y_i$. The result of this

model is that for any $\beta_1 < 1$, $E[\hat{y}] > E[y]$ and $Var[\hat{y}] < Var[y]$ or more specifically,

$E[\hat{y}] = (1 - \beta_1)b + \beta_1 E[y]$ and $Var[\hat{y}] = \beta_1 Var[y]$. This model of optimism bias can be

described as the belief that ‘the bad things that happen to others can’t happen to me’.

Optimism Bias Model 2: Over(under)estimating Upside Yield Risk

Under this model of optimism (pessimism) bias, farmers increase (decrease) the mean and decrease (increase) the variance of their perceived yield distribution relative to the objective yield distribution by increasing (decreasing) the upper-bound, while holding the lower-bound fixed. In this case,

$$\hat{y}_i \sim B(\hat{a} = a, \hat{b} = \beta_2 b, \hat{S}_1 = S_1, \hat{S}_2 = S_2)$$

where β_2 is our miscalibration parameter. Our resulting miscalibration function in this

scenario is simply $\hat{y}_i = \beta_2 y_i$. Under this model, $E[\hat{y}] = \beta_2 E[y]$, $Var[\hat{y}] = \beta_2 Var[y]$ and

optimism bias is defined as $\beta_2 > 1$, while underconfidence is defined as $\beta_2 < 1$. This

model of optimism (pessimism) bias can be described as the belief that ‘just because good (bad) things do not happen to other people does not mean they will not happen to me’.

Optimism Bias Model 3: Underestimating Yield Variability

In this model of optimism bias, farmers decrease the variance of their perceived yield distribution as compared to their objective yield distribution by increasing the lower-bound and decreasing the upper-bound, while holding the mean fixed. In this setting,

$$\hat{y}_i \sim B\left(\hat{a} = (1 - \beta_3)E[y], \hat{b} = \beta_3 b, \hat{S}_1 = S_1, \hat{S}_2 = S_2\right)$$

giving us $\hat{y}_i = (1 - \beta_3)E[y] + \beta_3 y_i$ where β_3 is our miscalibration parameter. Under this specification, optimism bias is defined as $\beta_3 < 1$ and $E[\hat{y}] = E[y]$, $Var[\hat{y}] = \beta_3 Var[y]$.

This model of optimism bias is akin to the belief that ‘really good *and* really bad things might happen to others, but *neither* will happen to me’.

Optimism Bias Model 4: Optimism Regarding the Shape of the Yield Distribution

In our fourth model, farmers decrease (increase) the skewness of their perceived yield distribution as compared to the objective yield distribution in a way that increases (decreases) the mean and scales the variance. Under this specification,

$$\hat{y}_i \sim B\left(\hat{a} = a, \hat{b} = b, \hat{S}_1 = ZS_1, \hat{S}_2 = S_2 + (1 - Z)S_1\right)$$

where Z is now our miscalibration parameter. Here, optimism bias is defined as $Z > 1$ while underconfidence is defined as $Z < 1$. By specifying the model such that

$\hat{S}_1 + \hat{S}_2 = S_1 + S_2$, we get the result that as long as $a = 0$, $E[\hat{y}] = ZE[y]$. As a result, Z is constrained such that $ZE[y] \leq b$, since the expected value of yields cannot be greater

than the maximum yield. Additionally, we get $Var[\hat{y}] = Z\left(\frac{S_2 + 2S_1 + ZS_1^2}{S_1 S_2}\right)Var[y]$, a result that

is independent of our assumption that $a = 0$. Unlike with our β s, the effect of Z on the new variance can go in either direction depending on the skewness of original objective yield distribution. More specifically, adding positive (negative) skewness to an already positively (negatively) skewed distribution will decrease the variance, while adding positive (negative) skewness to a negatively (positively) skewed distribution will decrease the variance. This model of optimism bias is analogous to the belief that ‘although the best and worst things that could happen are the same for me as for everyone else, I am more likely than others to experience good things’.

For a graphical example of how each of the four models modifies the objective yield distribution, see figure 2.1.

[Figure 2.1]

2.3 Modeling Farmer Crop Insurance Coverage Level Decision-Making

We consider three potential models of producer decision-making regarding which crop insurance cover level (if any) to purchase, specifically one based on the standard expected utility model, and two based on cumulative prospect theory, one in which the decision is framed broadly and one in which the decision is framed more narrowly.

The Objective Function

Given the choice between various insurance coverage levels, c , the farm’s objective function can be written as

$$\begin{aligned} \max_{d_c \in \mathbb{Z}^+} & \sum_{c=c_1}^C d_c V_c \\ \text{s.t.} & \sum_{c=c_1}^C d_c = 1 \end{aligned}$$

where $d_c = 1$ denotes the decision to purchase insurance at cover level c , and C represents the highest coverage level available. Given that $d_c \in \mathbb{Z}^+$, the constraint ensures that a producer must pick one and only one coverage level. Rather than allow $\sum d_c = 0$, we simply represent the choice to not participate in the insurance program as choosing a coverage level of 0. This distinction is particularly important in our setting because depending on how the outcome is framed, it is possible that $V_{c=0} < 0$.

V_c is determined according to the cumulative prospect theory model where

$$V_c = \sum_{i=-m}^n \pi(p_i)v(x_i)$$

This states that farmers value potential outcome scenarios ('prospects') by considering the weighted average of the value placed on the $M+N$ possible discrete outcomes, x_i . In accordance with cumulative prospect theory (Tversky and Kahneman, 1992), values for x_i are sorted from smallest to largest with M representing the number of losses and N representing the number of gains, and where $\pi(p_i)$ assigns a decision weight to p_i , the probability of x_i being realized. Decision weights take the form

$$\begin{aligned}\pi_i^+ &= w^+(p_i + \dots + p_n) - w^+(p_{i+1} + \dots + p_n); \\ \pi_n^+ &= w^+(p_n), \text{ and} \\ \pi_i^- &= w^-(p_m + \dots + p_i) - w^-(p_m + \dots + p_{i-1}); \\ \pi_m^- &= w^-(p_m)\end{aligned}$$

where

$$\begin{aligned}w^+(p) &= \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{1/\gamma}}; \\ w^-(p) &= \frac{p^\delta}{(p^\delta + (1-p)^\delta)^{1/\delta}}\end{aligned}$$

The value function, $v(x)$, takes the form

$$v(x) = \begin{cases} x^{1-r} & \text{if } x \geq 0 \\ -\lambda(-x)^{1-r} & \text{if } x < 0 \end{cases}$$

where λ is the coefficient of loss aversion and represents the degree with which losses are felt more strongly than corresponding gains, such that $-v(-x_0) = \lambda v(x_0)$. Figure 2.2 provides a graphical representation of the probability weighting function, while figure 2.3 illustrates the value function.

[Figure 2.2]

[Figure 2.3]

One of the major strengths of the cumulative prospect theory model is that when $\lambda, \gamma, \delta = 1$ and $x_i \in \mathbb{R}^+$, the model generalizes to that of the traditional expected utility model, where farmers are risk averse for $r > 0$, risk neutral for $r = 0$, and risk loving for $r < 0$. This allows us to use the same model even in cases where farmers are simply expected utility maximizers, rather than prospect theory value function maximizers.

The Expected Utility Model

In our first model of decision-making $\lambda, \gamma, \delta = 1$ and $x_i \in \mathbb{R}^+$ so that the cumulative prospect theory model becomes equivalent to that of the expected utility model. In the traditional expected utility model of insurance choice, where it is assumed that farmers and insurance provider have identical expectations, $x_i = P_i y_i + I_i(c) - \rho(c)(1 - s_c)$ where $I_i(c)$ represents the indemnity paid out by the insurance provider if the outcome (y_i in the case of yield insurance, and $P_i y_i$ in the case of revenue insurance) is below that of the outcome guaranteed by the insurance contract ($cE[y]$ in the case of yield insurance and $cE[y]E[P]$ for revenue insurance). $\rho(c)$ represents the fair-priced premium for insurance of a given coverage level, c , and where s_c is the corresponding proportionate

subsidy offered at that coverage level. We will layout how indemnities and premiums are calculated in more detail in the following section.

Alternatively, in our model of insurance coverage level decision-making, in which farmer expectations are potentially misaligned with what we assume are accurate expectations of insurance providers, $x_i = P_i \hat{y}_i + \hat{I}_i(c) - \rho(c)(1 - s_c)$ where \hat{y}_i represents the over(under)confident yield forecast and $\hat{I}_i(c)$ now represents the potentially inaccurate forecasted indemnity paid out by the insurance provider based on the forecasted outcome (\hat{y}_i in the case of yield insurance, and $P_i \hat{y}_i$ in the case of revenue insurance).

The Cumulative Prospect Theory Model

A key feature of prospect theory is the idea of a reference point from which outcomes above the reference point are treated as gains, and outcomes below the reference point are treated as losses. Given our specification of the value function, $v(x_i)$, x_i is calculated as being net of the reference point, R . For example, if $R = E[y]$, then $x_i(y_i, R) = y_i - E[y]$ such that $x_i(y_i < E[y])$ is valued as a loss and is thus weighted according to one's loss aversion coefficient, λ . The choice of a reference point is thus a critical part of properly specifying any model based in prospect theory. In line with Babcock (2015), we follow the suggestion of Koszegi and Rabin (2007) of using expected outcomes to guide the selection of a reference point.

Another important and related question is how outcomes are framed. Is the outcome of an insurance contract broadly framed such that it includes the entirety of the farm's financial risk, or is the outcome framed more narrowly such that only the outcome explicitly derived from the insurance is considered? In order to gain a better

understanding of how these framing decisions might change the effect of optimism bias on coverage choices, we consider two reference dependent outcomes, one in which outcomes are framed narrowly and one in which outcomes are framed more broadly.

Broadly Framed Outcomes

When outcomes are framed more broadly and $R = E[P]E[y] + \rho(c)(1 - s_c)$,

$$x_i = P_i \hat{y}_i - E(P)E(\hat{y}) + \hat{I}_i(c) - \rho(c)(1 - s_c)$$

which states that x_i is the sum of the forecasted difference between the i^{th} realization of revenue and expected revenue, and the difference between the forecasted i^{th} indemnity realization and the subsidized premium paid in order to receive that indemnity.

Narrowly Framed Outcomes

When outcomes are framed more narrowly and $R = \rho(c)(1 - s_c)$,

$$x_i = \hat{I}_i - \rho(c)(1 - s_c)$$

which states that x_i is simply the difference between the forecasted i^{th} indemnity realization and the subsidized premium paid in order to receive that indemnity. Under this model, farmers essentially treat insurance as a one-off lottery wherein they only win if $\hat{I}_i > \rho(c)(1 - s_c)$. As a result of not fully integrating the entirety of the farm's financial risk, the narrow-frame farmer becomes more likely to realize an outcome that is perceived as being a loss causing the decision to be more affected by loss aversion than it otherwise would be. As demonstrated by Babcock (2015) in a non-optimism bias setting, this model of narrow framing outperforms the other two models when it comes to explaining the seemingly 'anomalous' coverage level choices observed in the real world.

2.4 Crop Insurance Premiums and Indemnities

Calculating Premiums and Indemnities

Crop insurance is offered across eight levels in .05 increments from .50 to .85 with each coverage level having a proportionate premium subsidies, s_c , that is fixed across both crops and locations. The eight pairs of (c, s_c) are as follows: (.50, .76), (.55, .64), (.60, .64), (.65, .59), (.70, .59), (.75, .55), (.80, .48), and (.85, .38). For yield insurance, yields are guaranteed at $cE[y]$, and for revenue insurance, revenues are guaranteed at $cE[P]E[y]$, where expectations are taken at the time the insurance is contracted. The indemnity provided by yield insurance is thus $I_y = P \max[cE[y] - y, 0]$. For the majority of revenue insurance purchased in the United States, the guarantee actually increases if the actual price at harvest is greater than $E[P]$. The indemnity provided by revenue insurance is thus $I_R = \max[cE[y] \max(E[P], P) - Py, 0]$. Unsubsidized crop insurance premiums in the US are designed to be actuarially fair^{4,5}, allowing us to calculate

premiums for yield and revenue insurance as $\rho_y(c) = \int_0^b I_y dF(P, y)$ and

$$\rho_R(c) = \int_0^\infty \int_0^b I_R df(P, y), \text{ respectively.}$$

2.5 Simulating Optimal Cover Level Choices

Optimal coverage level choices under our various decision-making models cannot be solved analytically, at least not under any realistic distribution of yields and prices.

⁴ Premium rates are determined based on historical data and then loaded by a factor of 13.6% in order to account for potential future losses that are not reflected in the historical data.

⁵ Insurance companies are reimbursed by the federal government for the administrative costs of providing fair-priced crop insurance.

Therefore, we simulate the effect of changes in our optimism bias parameters on coverage level choices for both revenue and yield insurance for two representative farms – a ‘low risk’ irrigated corn farm in York County, NE with $corr[P,y] = 0$ and a ‘high risk’ wheat farm in Sumner County, KS with $corr[P,y] = -0.3$, the price-yield correlation used in determining Revenue Assurance premium rates in 2009, the year our simulation data is based on⁶. As a further sensitivity analysis, we also simulate revenue insurance coverage choices where $corr_{corn}[P,y] = -0.3$ and $corr_{wheat}[P,y] = 0$.

For each crop, we simulate 5,000 random yield and price outcomes. Yields are assumed to follow a beta distribution, while prices follow a log-normal distribution. The parameters of these distributions are taken from Babcock (2015). Babcock uses 2009 harvest price means and volatilities as used by the crop insurance program to fit the log-normal distribution, where as the parameters of the beta distribution were estimated by minimizing the sum of squared errors between target premium rates (obtained from USDA’s Risk Management Agency Premium Calculator, also for 2009) and simulated premium rates across the eight coverage levels. These values are displayed in table 2.1. Price-yield and yield distribution plots for our simulated data are shown in figure 2.4. In order to reduce any potential error causes by sampling variability, both guarantees and premiums were re-calculated based on our simulated data.

[Table 2.1]

[Figure 2.4]

For all three of our decision-making models, r , the coefficient of risk aversion, is set equal to 0.12. In our expected utility model $\lambda, \gamma, \delta = 1$, while in both of our

⁶ The price-yield correlation for corn was set to zero based on the fact that almost 90% of corn in York County is irrigated and irrigated corn yields are not highly correlated with national corn prices.

cumulative prospect theory models $\lambda = 2.25$, $\gamma = 0.61$, $\delta = 0.69$, which are the values estimated by Tversky and Kahneman (1992) based on their experimental data⁷.

Calculating Decision Weights on Simulated Outcomes

The values and signs of x_i for a given i^{th} realization vary across different coverage levels and therefore $w(p_i)$ must be calculated independently for each coverage level, c , in order to calculate V_c . For each coverage level, x_i s are sorted in ascending order from largest loss to largest gain. Let M_c be the number of losses and N_c be the number of gains. Losses are labeled such that for the largest loss, $m_c = M_c$; for the second largest loss, $m_c = M_c - 1$; ...; for the second smallest loss $m_c = 2$ and for the smallest loss $m_c = 1$. For all gains, $m_c = 0$. The same thing is done for gains such that for the largest gain, $n_c = N_c$, for the second largest gain, $n_c = N_c - 1$; ...; for the second smallest gain $n_c = 2$; and for the smallest gain $n_c = 1$. For all losses, $n_c = 0$. This set up makes calculating decision weights straightforward as each of the simulated outcomes (yield, price, and resulting indemnity) have equal probability of occurring in the simulation, $1/(M_c + N_c)$. For each coverage level, c ,

$$\pi_{n,c} = w^+[n_c/(N_c + M_c)] - w^+[(n_c - 1)/(N_c + M_c)], \text{ and}$$

$$\pi_{m,c} = w^- [m_c/(N_c + M_c)] - w^- [(m_c - 1)/(N_c + M_c)].$$

2.6 Results

Our main results are presented in tables 2.2A through 2.2D. Although we calculate significantly different optimal coverage levels across our twelve models, two crops, and two insurance types, the overall effects of optimism bias on coverage level choices are

⁷ $\lambda=2.25$ is reported as the median estimated value. Since the distribution of λ is highly non-normal and positively skewed, the median value serves as a better representative estimate than the mean value.

largely consistent across all twelve specifications for both crops and both insurance types. Tables 2.3A through 2.3D present the results of our additional sensitivity analysis with regard to correlation between prices and yields.

[Table 2.2A]

[Table 2.2B]

[Table 2.2C]

[Table 2.2D]

The Effect of Optimism (Pessimism) Bias on Cross-Coverage Level Demand

Our first result is that as optimism bias increases, the optimal coverage level selected decreases. This pattern holds true for all 24 of our yield insurance simulations, and 22 of the 24 revenue simulations. In this case of optimism bias model 3 under the broad framed prospect theory model, the optimal coverage choice stay unchanged across changes in β_3 for both crops. Furthermore, as optimism bias continues to increase, producers move towards preferring no insurance.

We find that the converse effect also holds true, where increasing pessimism bias causes producers to prefer higher levels of coverage than they would in the unbiased scenario. This holds true for all cases where the unbiased coverage choice was not already at .85, with all choices eventually approaching .85 coverage.

Further Sensitivity Analyses

The Effect of Optimism Bias Across Insurance Type

While the overall effect of optimism bias on coverage level choice is consistent across both revenue and yield insurance, we find that in general the magnitude of the effect of optimism bias on cross-coverage level demand is larger for yield insurance

than for revenue insurance. This is the case for 22 of our 24 comparisons, with 2 comparisons -wheat, expected utility, optimism bias models 1 and 2 – having identical coverage choices across both yield and revenue insurance. This effect is perhaps not surprising given that our optimism bias model assumes producers are accurate in their perceptions of price risk.

The Effect of Optimism bias Across Optimism bias Models

When comparing the effects of optimism bias on coverage level choice across our four optimism bias models we find varying results. Model 1 produces the most dramatic effects, followed by model 3. It is probably not surprising that underestimating downside risk has a more dramatic effect than simply overestimating upside risk, given that it is downside risk that insurance is protecting against.

The Effect of Optimism Bias Across Decision-Making Models

Across the parameters of our four optimism bias models it is always that case that $c_{EU}^* \leq c_{Broad}^* \geq c_{Narrow}^*$ and both $c_{Broad}^* - c_{EU}^*$ and $c_{Broad}^* - c_{Narrow}^*$ are monotonically increasing as optimism bias increases. The difference in the overall effect between Narrow-Framed CPT and Expected Utility are inconsistent though in general, but not always, $c_{Narrow}^* < c_{EU}^*$. Despite the fact that when $\hat{y}_i = y_i$, $c_{Narrow}^* \leq c_{EU}^*$, that relationship does not always hold as optimism bias increases. While it is not surprisingly that coverage choices under the narrow CPT frame are more affected than those under the broad frame given that the narrow frame is more effected by loss aversion, it is surprising that coverage choices under expected utility are more affected than the broad CPT frame given that the broad CPT frame is effected by loss aversion, whereas the expected utility choices are not.

The Effect of Optimism Bias Across Crops

Differences in the effect of optimism bias between our two crops varies across our four models. Under model 1, the effect is stronger for wheat (our ‘high risk’, positively skewed crop) than it is for corn (our ‘low risk’, negatively skewed crop). However, under model 2 we find just the opposite, with corn being more effect by optimism bias than wheat. Under model 3. We obtain conflicting results. For revenue insurance, the effect of optimism bias is more dramatic for wheat, whereas for yield insurance, the effect is more dramatic on corn. This suggests that the effect of each model depends on the skewness of the objective distribution and may be an area for future research.

The Effect of Optimism Bias Across Price-Yield Correlations

Our sensitivity analyses with regard to the correlation between prices and yields is displayed in tables 2.3A through 2.3D.

[Table 2.3A]

[Table 2.3B]

[Table 2.3C]

[Table 2.3D]

We find that results do not change dramatically based on whether or not prices and yields are negatively correlated compared to when no correlation exists. Although some coverage level choices based on whether the price-yield correlation was set to 0 or -0.3, changes were minimal overall and did not consistently change in one direction or the other.

The Role of Unequal Proportionate Subsidy Levels

It is important to point out why the finding that optimism bias causes farmers to prefer lower levels of crop insurance coverage is both surprising and significant. This result does not simply stand on its own. In fact it can be shown that in cases when subsidy levels are equivalent across coverage levels that this result no longer holds, and that instead, an expected utility maximizing farmer will purchase either maximum coverage, or no coverage at all.

This is can easily be shown mathematically for the case of a risk-neutral expected utility maximizer. Such a farmer will choose the coverage level that maximizes $E[\hat{I}_c] - (1-s_c)\rho_c$ where $\rho_c = E[I_c]$. This can be mathematically rearranged to state that such a farmer will maximize

$$s_c - (E[I_c] - E[\hat{I}_c]) / E[I_c].$$

It is indeed the case that for all four of our models of optimism bias, $(E[I_c] - E[\hat{I}_c]) / E[I_c]$ decreases as the coverage level increases. Thus when all values of s_c are equivalent, said farmer will always choose either $c = .85$, or no coverage at all. Thus, in order for the farmer to pick a coverage level below .85 it must be the case that s_c decreases faster than $(E[I_c] - E[\hat{I}_c]) / E[I_c]$ as coverage levels increase, i.e.

$$\frac{\Delta s}{\Delta c} < \frac{\Delta(E[I_c] - E[\hat{I}_c]) / E[I_c]}{\Delta c}.$$

Re-simulating coverage choices under equal proportionate subsidies ($s_c = .5$ for all c), we find that under this subsidy schedule, our mildly risk-averse EU farmer ($r = .12$) will always prefer $c = .85$ until eventually preferring $c = 0$. We find the same result for our mildly risk-averse, broad-framing CPT farmers. For our mildly risk-averse, narrow-framing CPT farmers, we find that there are still instances where the farmer

prefers an interior solution, however the number of interior solutions chosen is greatly reduced. These results are presented in table 2.4. For brevity, we only show results under optimism model 1, however they are consistent across all four models of risk misperception.

[Table 2.4]

It is thus obvious that the mechanism driving these results is the fact that proportionate subsidies increase as coverage levels decrease. Thus, the impact of optimism bias on coverage level choices could be greatly reduced by a more equivalent schedule of subsidies across coverage levels.

2.7 Conclusion

In this paper we find that although expected utility suggests that producers should purchase crop insurance at a coverage level at least as high as the level that maximizes per-acre subsidies ($c = .85$ for corn and $c = .80$ for wheat), this is no longer the case when producers have inaccurate perceptions regarding their yield risk. We find that the effect of optimistically biased misperceptions of yield risk is to reduce the level of optimal coverage in the case of both revenue and yield insurance and that this effect is strongest for yield insurance. We further find that the mechanism driving this result is the higher levels of proportionate subsidies provided for lower levels of coverage and that such a result no longer exists for our expected utility farmers or broad-framing cumulative prospect theory farmers and is greatly reduced for our narrow-framing cumulative prospect theory farmers in the case of equivalent proportionate subsidies.

These findings suggests that if a goal of policy-makers is to induce farmers to purchase lower-deductible (higher coverage) policies through subsidies in order to

reduce the risk of having to provide *ad hoc* disaster relief (Babcock, 2015), then incorporating optimism bias may allow for the design of more cost-effective subsidy schedules. Given that the effects we found were more pronounced for yield insurance relative to revenue insurance, it also suggests that in the case where yield misperceptions are heterogeneous across farmers (e.g. some are optimistic, while others are pessimistic, as was evidenced in Manuscript 1) that equivalently subsidized revenue insurance likely does a better job of minimizing deadweight loss by reducing the variance of producer surplus across farmers, relative to yield insurance. Rigorously testing these hypotheses, as well as looking at the interaction between optimism bias and yield skewness provide fruitful grounds for future research.

2.8 Work Cited

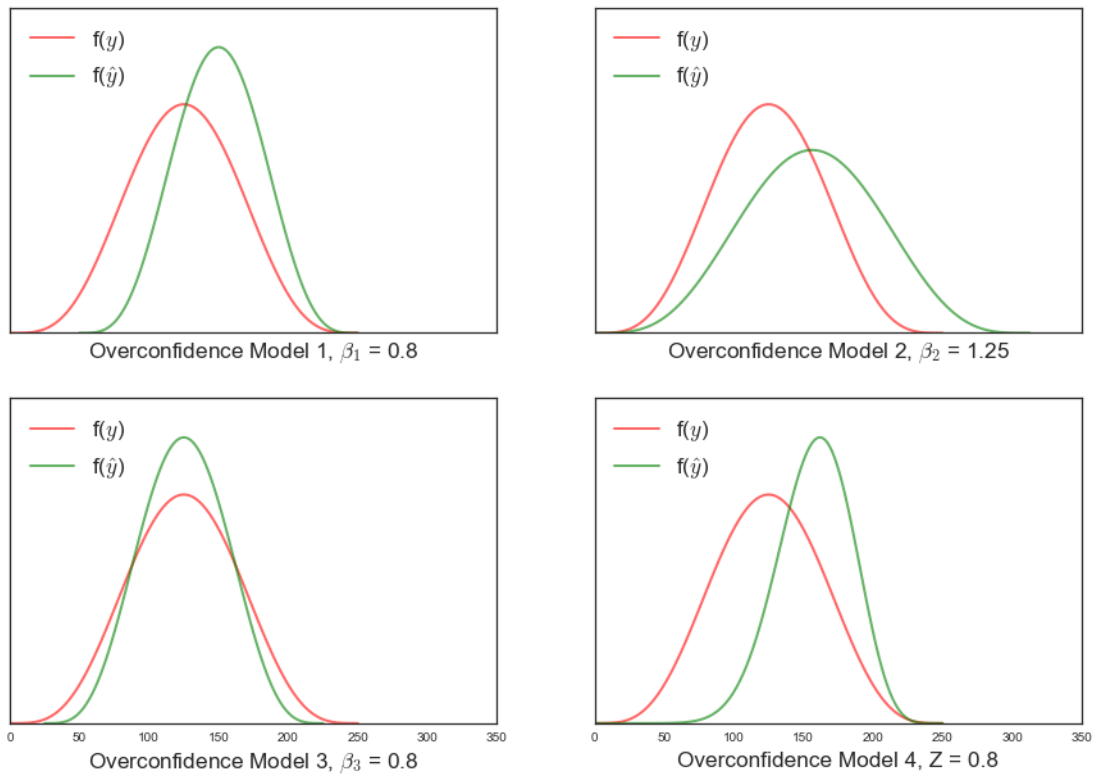
- Babcock, Bruce. 2015. "Using Cumulative Prospect Theory to Explain Anomalous Crop Insurance Coverage Choice." *American Journal of Agricultural Economics* 97(5): 1371-1384.
- Babcock, Linda, George Loewenstein, and Samuel Issacharoff. 1997. "Creating Convergence: Debiasing Biased Litigants." *Law and Social Inquiry* 22(4): 913-925.
- Barseghyan, Levon. 2013. "The Nature of Risk Preferences: Evidence from Insurance Choices." *The American Economic Review* 103(6): 2499-2529.
- Bhattacharya, Jayanta, David M. Cutler, Dana P. Goldman, Michael D. Hurd, Geoffrey F. Joyce, Darius N. Lakdawalla, Constantijn WA Panis, and Baoping Shang. 2004. "Disability forecasts and future Medicare costs." *Forum for Health Economics & Policy* 7.1.
- Kahneman, Daniel and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica* 47: 263–291.
- Kőszegi, B. and Rabin, M., 2007. Reference-dependent risk attitudes. *American Economic Review*, 97(4), pp.1047-1073.
- Egelkraut, T. M., Garcia, P., Pennings, J. M., and Sherrick, B. J. 2006. "Producers' Yield and Yield Risk: Perceptions Versus Reality and Crop Insurance Use." *Agricultural and Applied Economics Association Annual Meeting, July 23-26, Long Beach, CA.*
- Enjolras, Geoffroy, and Patrick Sentis. 2011. "Crop Insurance Policies and Purchases in France." *Agricultural Economics* 42(4): 475-486.
- Hoorens, Vera. 1994. "Unrealistic Optimism in Social Comparison of Health and Safety Risks." *The Social Psychology of Health and Safety*: 153-174.
- Malmendier, Ulrike, and Geoffrey Tate. 2008. "Who Makes Acquisitions? CEO Overconfidence and the Market's Reaction." *Journal of Financial Economics* 89(1): 20-43.
- Pease, James W. 1992. "A Comparison of Subjective and Historical Crop Yield Probability Distributions." *Southern Journal of Agricultural Economics* 24.02: 23-32.

- Pease, James W., Ernest W. Wade, Jerry S. Skees, and Chandra M. Shrestha. 1993. "Comparisons Between Subjective and Statistical Forecasts of Crop Yields." *Review of Agricultural Economics* 15(2): 339-350.
- Umarov, Alisher, and Bruce J. Sherrick. 2005. "Farmers' Subjective Yield Distributions: Calibration and Implications for Crop Insurance Valuation." *American Agricultural Economics Association Annual Meeting, July 2005, Providence, RI*.
- US General Account Office. 1988. "Crop Insurance: Participation and Costs Associated With the Federal Program." (GAO/RCED-88-171BR). Washington, D.C.: U.S. Government Printing Office.
- Wang, Ming, Tao Ye and Peijun Shi. "Factors Affecting Farmers' Crop Insurance Participation in China." *Canadian Journal of Agricultural Economics* 64: 479-492.

2.8 Figures

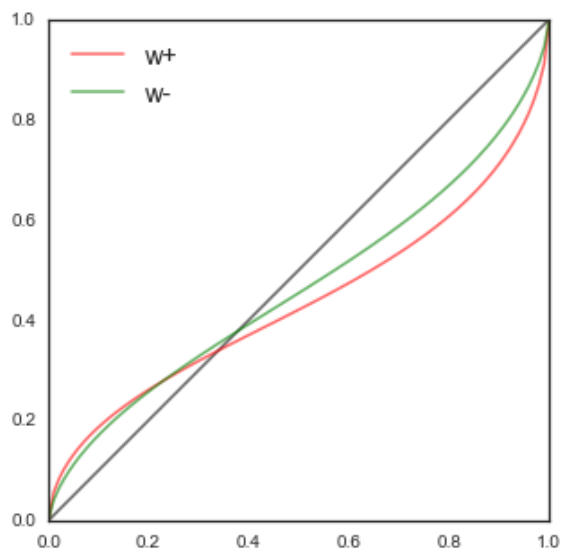
Figure 2.1: Example Comparisons of the Objective and Perceived/Forecasted Yield

Distributions According to our Four Models of Optimism Bias



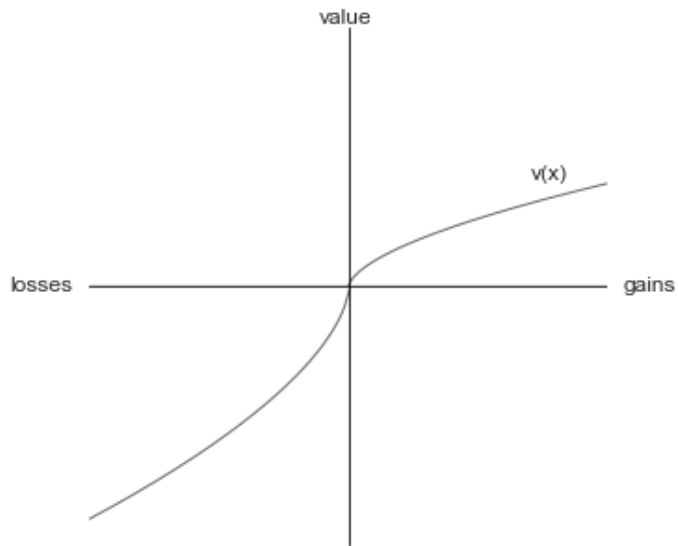
Note: Depicted above are example comparisons between a farmer's objective yield distribution (in red)- which is used to set premiums- and their misperceived subjective yield distribution (in green)- which they use to make coverage level choices- across of four models of risk misperception. Miscalibration parameters are presented below each graph.

Figure 2.2: The Probability Weighting Function ($\gamma = 0.61$, $\delta = 0.69$)



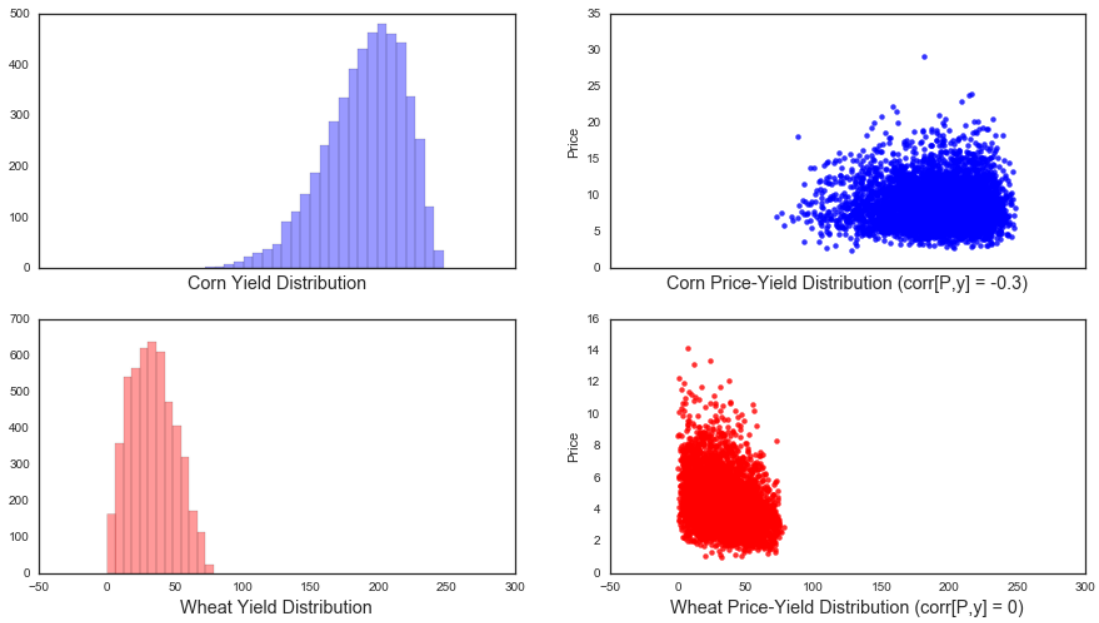
Note: Depicted above is the cumulative prospect theory probability weighting function. It shows that individuals overweight the probability of low probability events, while underweighting the probability of high probability events.

Figure 2.3: The Prospect Theory Value Function



Note: Above is a graphical representation of the cumulative prospect theory value function. It shows that individuals are risk averse in the gains domain, and risk-loving in the loss domain. It also shows that individuals are loss averse, i.e. that losses are more salient than equivalent gains.

Figure 2.4: Simulated Price-Yield and Yield Distribution Plots for Corn and Wheat



Note: On the left are displayed histograms of the yield outcomes based on the parameters of the yield distribution for corn (above) and wheat (below). On the right are scatter plots of price and yield outcomes based on the joint price-yield distribution for both corn (above) and wheat (below). For our scatter plots, yields are presented on the x-axis, while prices are presented on the y-axis.

2.10 Tables

Table 2.1: Parameters for $F(y, P)$ from Babcock (2015)

	Corn	Wheat
	York Co, NE	Sumner Co, KS
Expected Yield	190 bu/ac	33 bu/ac
Expected Price	\$4.40/bu	\$8.77/bu
Price Volatility	37%	33%
Price-Yield Correlation	0	-0.3
Yield Parameters		
Maximum	250	80
Minimum	0	0
Shape 1	9.34	1.938
Shape 2	2.949	2.76

Table 2.2A: Optimal Coverage Level Choice by Crop, Insurance Type, Decision Model: Optimism Bias Model 1

Type of Insurance	Corn (York County, NE)						Wheat (Sumner County, KS)						
	Revenue			Yield			Revenue			Yield			
	Decision Model	EU	Broad CPT	Narrow CPT	EU	Broad CPT	Narrow CPT	EU	Broad CPT	Narrow CPT	EU	Broad CPT	Narrow CPT
$\beta_1 = 0.50$	0	80	50	0	0	0	0	0	0	0	0	0	0
$\beta_1 = 0.55$	0	80	50	0	0	0	0	0	0	0	0	0	0
$\beta_1 = 0.60$	0	85	60	0	0	0	0	0	0	0	0	0	0
$\beta_1 = 0.65$	70	85	60	0	0	0	0	0	0	0	0	0	0
$\beta_1 = 0.70$	75	85	60	0	0	0	0	0	0	0	0	0	0
$\beta_1 = 0.75$	75	85	70	0	0	0	0	0	0	0	0	0	0
$\beta_1 = 0.80$	80	85	70	0	0	0	0	0	0	0	0	0	0
$\beta_1 = 0.85$	80	85	70	0	75	0	0	75	0	0	0	0	0
$\beta_1 = 0.90$	80	85	75	0	85	75	75	80	60	75	75	0	0
$\beta_1 = 0.95$	85	85	75	85	85	80	75	80	70	75	80	70	70
$\beta_1 = 1.00$	85	85	75	85	85	80	80	80	70	80	80	70	70

Note: Optimal coverage level choice under identical farmer and insurer yield expectations are highlighted in blue.

Table 2.2B: Optimal Coverage Level Choice by Crop, Insurance Type, Decision Model: Optimism Bias Model 2

Type of Insurance	Corn (York County, NE)						Wheat (Sumner County, KS)					
	Revenue			Yield			Revenue			Yield		
	Decision Model	EU	Broad CPT	Narrow CPT	EU	Broad CPT	Narrow CPT	EU	Broad CPT	Narrow CPT	EU	Broad CPT
$\beta_2 = 1.5$	0	70	0	0	0	0	70	75	60	70	75	50
$\beta_2 = 1.4$	0	75	50	0	0	0	70	75	60	70	75	60
$\beta_2 = 1.3$	0	80	50	0	0	0	75	80	60	75	75	60
$\beta_2 = 1.2$	60	80	60	0	75	70	75	80	70	75	80	60
$\beta_2 = 1.1$	80	85	70	75	80	75	75	80	70	75	80	70
$\beta_2 = 1.0$	85	85	75	85	85	80	80	80	70	80	80	70
$\beta_2 = 0.9$	85	85	80	85	85	85	80	85	75	80	85	70
$\beta_2 = 0.8$	85	85	85	85	85	85	85	85	75	85	85	75
$\beta_2 = 0.7$	85	85	85	85	85	85	85	85	80	85	85	80
$\beta_2 = 0.6$	85	85	85	85	85	85	85	85	85	85	85	80
$\beta_2 = 0.5$	85	85	85	85	85	85	85	85	85	85	85	85

Note: Optimal coverage level choice under identical farmer and insurer yield expectations are highlighted in blue.

Table 2.2C: Optimal Coverage Level Choice by Crop, Insurance Type, Decision Model: Optimism Bias Model 3

Type of Insurance	Corn (York County, NE)						Wheat (Sumner County, KS)						
	Revenue			Yield			Revenue			Yield			
	Decision Model	EU	Broad CPT	Narrow CPT	EU	Broad CPT	Narrow CPT	EU	Broad CPT	Narrow CPT	EU	Broad CPT	Narrow CPT
$\beta_3 = 0.50$	80	85	70	0	0	0	0	80	0	0	0	0	0
$\beta_3 = 0.55$	80	85	75	0	0	0	0	80	0	0	0	0	0
$\beta_3 = 0.60$	85	85	75	0	0	0	80	80	0	0	0	0	0
$\beta_3 = 0.65$	85	85	75	0	0	0	80	80	0	0	75	0	0
$\beta_3 = 0.70$	85	85	75	0	0	0	80	80	70	75	80	0	0
$\beta_3 = 0.75$	85	85	75	0	85	0	80	80	70	75	80	0	0
$\beta_3 = 0.80$	85	85	75	0	85	75	80	80	70	80	80	70	0
$\beta_3 = 0.85$	85	85	75	85	85	80	80	80	70	80	80	70	0
$\beta_3 = 0.90$	85	85	75	85	85	80	80	80	70	80	80	70	0
$\beta_3 = 0.95$	85	85	75	85	85	80	80	80	70	80	80	70	0
$\beta_3 = 1.00$	85	85	75	85	85	80	80	80	70	80	80	70	0

Note: Optimal coverage level choice under identical farmer and insurer yield expectations are highlighted in blue.

Table 2.2D: Optimal Coverage Level Choice by Crop, Insurance Type, Decision Model: Optimism Bias Model 4

Type of Insurance	Corn (York County, NE)						Wheat (Sumner County, KS)					
	Revenue			Yield			Revenue			Yield		
	Decision Model	EU	Broad CPT	Narrow CPT	EU	Broad CPT	Narrow CPT	EU	Broad CPT	Narrow CPT	EU	Broad CPT
Z = 1.5	-	-	-	-	-	-	0	60	0	0	50	0
Z = 1.4	-	-	-	-	-	-	0	70	0	0	60	0
Z = 1.3	0	75	0	0	0	0	60	75	50	50	70	50
Z = 1.2	0	80	50	0	0	0	75	80	60	75	75	60
Z = 1.1	75	85	70	0	75	0	75	80	70	75	80	60
Z = 1	85	85	75	85	85	80	80	80	70	80	80	70
Z = 0.9	85	85	80	85	85	85	80	85	75	85	85	75
Z = 0.8	85	85	85	85	85	85	85	85	75	85	85	75
Z = 0.7	85	85	85	85	85	85	85	85	80	85	85	80
Z = 0.6	85	85	85	85	85	85	85	85	80	85	85	80
Z = 0.5	85	85	85	85	85	85	85	85	85	85	85	80

Note: Optimal coverage level choice under identical farmer and insurer yield expectations are highlighted in blue. We are unable to calculate optimal coverage level choices for corn for $Z = 1.4, 1.5$ because Z cannot be greater than $b/E[y]$.

Table 2.3A: Optimal Revenue Coverage Level Choice by Crop, Decision Model, and Price-yield Correlation: Optimism Bias Model 1

Price-yield Correlation	Corn (York County, NE)						Wheat (Sumner County, KS)						
	0			-0.3			0			-0.3			
	Decision Model	EU	Broad CPT	Narrow CPT	EU	Broad CPT	Narrow CPT	EU	Broad CPT	Narrow CPT	EU	Broad CPT	Narrow CPT
$\beta_1 = 0.50$	0	80	50	0	80	0	0	0	0	0	0	0	0
$\beta_1 = 0.55$	0	80	50	0	85	50	0	0	0	0	0	0	0
$\beta_1 = 0.60$	0	85	60	0	85	60	0	0	0	0	0	0	0
$\beta_1 = 0.65$	70	85	60	75	85	60	0	0	0	0	0	0	0
$\beta_1 = 0.70$	75	85	60	75	85	70	0	0	0	0	0	0	0
$\beta_1 = 0.75$	75	85	70	80	85	70	0	0	0	0	0	0	0
$\beta_1 = 0.80$	80	85	70	80	85	70	0	75	0	0	0	0	0
$\beta_1 = 0.85$	80	85	70	80	85	70	0	75	0	0	75	0	0
$\beta_1 = 0.90$	80	85	75	80	85	70	75	80	60	75	80	60	70
$\beta_1 = 0.95$	85	85	75	85	85	75	75	80	70	75	80	70	70
$\beta_1 = 1.00$	85	85	75	85	85	75	80	80	70	80	80	70	70

Note: Optimal coverage level choice under identical farmer and insurer yield expectations are highlighted in blue.

Table 2.3B: Optimal Revenue Coverage Level Choice by Crop, Decision Model, and Price-yield Correlation: Optimism Bias Model 2

Price-yield Correlation	Corn (York County, NE)						Wheat (Sumner County, KS)					
	0			-0.3			0			-0.3		
	Decision Model	EU	Broad CPT	Narrow CPT	EU	Broad CPT	Narrow CPT	EU	Broad CPT	Narrow CPT	EU	Broad CPT
$\beta_2 = 1.5$	0	70	0	0	75	0	70	75	50	70	75	60
$\beta_2 = 1.4$	0	75	50	0	80	50	70	75	60	70	75	60
$\beta_2 = 1.3$	0	80	50	0	80	50	75	80	60	75	80	60
$\beta_2 = 1.2$	60	80	60	70	85	60	75	80	60	75	80	70
$\beta_2 = 1.1$	80	85	70	80	85	70	75	80	70	75	80	70
$\beta_2 = 1.0$	85	85	75	85	85	75	80	80	70	80	80	70
$\beta_2 = 0.9$	85	85	80	85	85	80	80	85	75	80	85	75
$\beta_2 = 0.8$	85	85	85	85	85	85	85	85	75	85	85	75
$\beta_2 = 0.7$	85	85	85	85	85	85	85	85	80	85	85	80
$\beta_2 = 0.6$	85	85	85	85	85	85	85	85	80	85	85	85
$\beta_2 = 0.5$	85	85	85	85	85	85	85	85	85	85	85	85

Note: Optimal coverage level choice under identical farmer and insurer yield expectations are highlighted in blue.

Table 2.3C: Optimal Revenue Coverage Level Choice by Crop, Decision Model, and Price-yield Correlation: Optimism Bias Model 3

Price-yield Correlation	Corn (York County, NE)						Wheat (Sumner County, KS)					
	0			-0.3			0			-0.3		
	Decision Model	EU	Broad CPT	Narrow CPT	EU	Broad CPT	Narrow CPT	EU	Broad CPT	Narrow CPT	EU	Broad CPT
$\beta_3 = 0.50$	80	85	70	80	85	75	0	80	0	0	80	0
$\beta_3 = 0.55$	80	85	75	80	85	75	75	80	0	0	80	0
$\beta_3 = 0.60$	85	85	75	80	85	75	80	80	0	80	80	0
$\beta_3 = 0.65$	85	85	75	85	85	75	80	80	0	80	80	0
$\beta_3 = 0.70$	85	85	75	85	85	75	80	80	70	80	80	70
$\beta_3 = 0.75$	85	85	75	85	85	75	80	80	70	80	80	70
$\beta_3 = 0.80$	85	85	75	85	85	75	80	80	70	80	80	70
$\beta_3 = 0.85$	85	85	75	85	85	75	80	80	70	80	80	70
$\beta_3 = 0.90$	85	85	75	85	85	75	80	80	70	80	80	70
$\beta_3 = 0.95$	85	85	75	85	85	75	80	80	70	80	80	70
$\beta_3 = 1.00$	85	85	75	85	85	75	80	80	70	80	80	70

Note: Optimal coverage level choice under identical farmer and insurer yield expectations are highlighted in blue.

Table 2.3D: Optimal Revenue Coverage Level Choice by Crop, Decision Model, and Price-yield Correlation: Optimism Bias Model 4

Price-yield Correlation	Corn (York County, NE)						Wheat (Sumner County, KS)					
	0			-0.3			0			-0.3		
	EU	Broad CPT	Narrow CPT	EU	Broad CPT	Narrow CPT	EU	Broad CPT	Narrow CPT	EU	Broad CPT	Narrow CPT
Z = 1.5	-	-	-	-	-	-	0	70	0	0	60	0
Z = 1.4	-	-	-	-	-	-	0	75	0	0	70	0
Z = 1.3	0	75	0	0	75	0	70	75	50	60	75	50
Z = 1.2	0	80	50	0	80	50	75	80	60	75	80	60
Z = 1.1	75	85	70	75	85	70	75	80	70	75	80	70
Z = 1	85	85	75	85	85	75	80	80	70	80	80	70
Z = 0.9	85	85	80	85	85	80	80	85	75	80	85	75
Z = 0.8	85	85	85	85	85	85	85	85	75	85	85	75
Z = 0.7	85	85	85	85	85	85	85	85	75	85	85	80
Z = 0.6	85	85	85	85	85	85	85	85	80	85	85	80
Z = 0.5	85	85	85	85	85	85	85	85	80	85	85	85

Note: Optimal coverage level choice under identical farmer and insurer yield expectations are highlighted in blue.

Table 2.4: Optimal Coverage Level Choice by Crop, Insurance Type, Decision Model: Optimism Bias Model 1 Under Equal Proportionate Subsidies ($s_c = .5$)

Type of Insurance	Corn (York County, NE)						Wheat (Sumner County, KS)						
	Revenue			Yield			Revenue			Yield			
	Decision Model	EU	Broad CPT	Narrow CPT	EU	Broad CPT	Narrow CPT	EU	Broad CPT	Narrow CPT	EU	Broad CPT	Narrow CPT
$\beta_1 = 0.50$	0	85	0	0	0	0	0	0	0	0	0	0	0
$\beta_1 = 0.55$	0	85	0	0	0	0	0	0	0	0	0	0	0
$\beta_1 = 0.60$	0	85	0	0	0	0	0	0	0	0	0	0	0
$\beta_1 = 0.65$	85	85	50	0	0	0	0	0	0	0	0	0	0
$\beta_1 = 0.70$	85	85	50	0	0	0	0	0	0	0	0	0	0
$\beta_1 = 0.75$	85	85	60	0	0	0	0	0	0	0	0	0	0
$\beta_1 = 0.80$	85	85	70	0	0	0	0	0	0	0	0	0	0
$\beta_1 = 0.85$	85	85	85	0	85	0	0	85	0	0	0	0	0
$\beta_1 = 0.90$	85	85	85	85	85	85	85	85	0	85	85	0	0
$\beta_1 = 0.95$	85	85	85	85	85	85	85	85	85	85	85	85	85
$\beta_1 = 1.00$	85	85	85	85	85	85	85	85	85	85	85	85	85

Note: Optimal coverage level choice under identical farmer and insurer yield expectations are highlighted in blue.

Manuscript 3

Bayesian Downscaling Methods for Aggregated Count Data

This manuscript was published in the December 2017 edition of *Agricultural and Resource Economics Review*, however this version has been revised to reflect comments from dissertation committee members.

Authors and acknowledgements as printed are:

Clayton P. Michaud, Thomas W. Sproul

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. 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

Abstract

Policy-critical, micro-level statistical data are often unavailable at the desired level of disaggregation. We present a Bayesian methodology for ‘downscaling’ aggregated count data to the micro-level, using an outside statistical sample. Our procedure combines numerical simulation with exact calculation of combinatorial probabilities. We motivate our approach with an application estimating the number of farms in a region, using count totals at higher levels of aggregation. In a simulation analysis over varying population sizes, we demonstrate both robustness to sampling variability and outperformance relative to maximum likelihood. Spatial considerations, implementation of “informative” priors, non-spatial classification problems, and best practices are discussed.

3.1 Introduction

Local economic planning often relies on micro-level data that is not always available at the desired level of disaggregation. For example, Federal government-provided economic and employment data for key industry sectors is often reported at the county level and obtaining city or ZIP-code level data may require time-consuming special requests or considerable expense, or it may be simply unavailable. In this article, we address the need for micro-level count data by developing a Bayesian methodology to ‘downscale’ aggregated count data to lower levels of aggregation using the information contained in an outside statistical sample.

Suppose a researcher knows the true size of a population (e.g., farmers, voters, customers) and would like to classify members of that population into distinct sub-groups (e.g., by farm type, county/region, political party, or demographic attributes) using independent data sampled from the full population. In this setting, we demonstrate a method for estimating the population proportion in each sub-group, in a manner that provides more stable and robust estimates than maximum likelihood estimation (MLE) in the face of sampling variability. The method consists primarily of using simulated random sampling combined with exact calculation of combinatorial probabilities to estimate the posterior distribution over combinations of counts. We leverage two key restrictions: i) the sub-group counts must add up to the population total, and ii) the sub-group counts cannot be smaller than their observed counts in the outside sample, nor larger than the population minus the sum of observed samples in the other sub-groups. This explicit handling of sampling

variability, especially in small to medium-sized samples, results in smaller normalized errors and, consequently, more reliably accurate estimates.

We are, of course, not the first to address the demand for more disaggregated data from aggregated sources. Gocht and Roder (2011), for example, employ a Bayesian procedure to downscale county-level German Agricultural Census estimates of land devoted to agricultural use. Their method incorporates land use data from GIS to facilitate micro-level environmental impact studies that would otherwise be hindered by data protection rules (i.e., censoring). Other relevant studies include Chakir (2009); Dendoncker, Bogaert, and Rounsevell (2006); Gärtner, Keller, and Schulin (2013); Howitt and Reynaud (2003); Purcell and Kish (1980); and Polasek, Llano, and Sellner (2010). These papers share a common thread of attempting to estimate land-use patterns using a variety and/or combination of methods including regression, multinomial logit, maximum entropy, cross-entropy, and various iterative fitting procedures. However, while these procedures perform well in their intended domain, they are ill-suited to solving the downscaling problem for count data. Intuitively, multinomial logit might be mapped to a count model in which sampling probabilities are estimated, but many observations and covariates are required. The methods we introduce here are designed to overcome this problem when the outside sample contains only limited categorical information.

Another popular application of downscaling involves disaggregation of global climate data (typically reported at grid levels of 100-200km) to a level of resolution more useful for decision-makers and impact assessors. Such procedures are outlined, for example, in Coelho et al. (2006); Hashmi, Shamseldin, and Melville (2009);

Fasbender and Ouarda (2010); Murphy (1999); and von Storch, Zorita, and Cubasch (1993). The goals of such estimation procedures, however, are to disaggregate weather/climate data not only spatially, but temporally as well in order to model various potential weather outcomes for use in forecasting. The procedures outlined by these studies are both unnecessarily complex given our particular problem of interest, and potentially ill-suited to the count data problem due to highly detailed data requirements in the outside sample.

In an attempt to balance precision with tractability, we develop a method that is adaptable to the data and computational resources of the applied researcher. Namely, we show that reasonable performance can be obtained using a uniform prior distribution over combinations of counts, but we also demonstrate a method for researchers to incorporate “informative” prior information generated by a simple linear regression or one of the more spatially-explicit and computationally demanding methods described above. In our simulation analysis, we demonstrate a means for testing the best performance among MLE, the uniform prior, or a more informative spatial prior, over a range of population counts and sample sizes. As might be expected, a more informative prior performs best for the smallest sample sizes and smallest population counts. However, we find that the uniform prior performs best over an unexpectedly wide range of sample size and population count combinations.

To provide context, we introduce and apply our methods in the setting of estimating spatially disaggregated farm counts by sub-region from regional data, using a sample of Rhode Island farms combined with aggregated data from the 2012 United States Department of Agriculture (USDA) Census of Agriculture (herein, “Ag

Census”). We explore both county-to-city downscaling and state-to-county downscaling, and show how spatial patterns at higher levels of aggregation might be used to construct an informative prior. We take special advantage of state-to-county downscaling as an example where the true underlying distribution is known and can be used to validate our methods. We also use published estimates of uncertainty in the Ag Census total counts to demonstrate the robustness of our methods to uncertainty in the top-level population count.

Despite the focus of much of the literature, and our own application, on spatial downscaling problems, it is important to note that there is nothing inherently spatial about the mathematics involved. Our method is equally well adapted to arbitrary classification problems in which it is desired to estimate the size of population subgroups according to a number of discrete categories. These applications might include political polling, estimation of workforce participation rates, demographic breakdowns by gender, age, race or educational attainment, or market segmentation analysis, among many others. At the same time, though our method does not require spatial information per se, it is flexible enough to incorporate arbitrarily complex spatial information as an input to the estimation procedure, by way of the informative prior.

The remainder of this article is organized as follows. The next section outlines and derives our estimation methodology, and the following section discusses selection of a prior. The fourth and fifth sections outline our sample data and methods, and the sixth section covers the results. The next section discusses

applications of our findings and areas for future research, and the final section concludes.

3.2 Bayesian Downscaling of Aggregated Count Data

Consider a source of aggregated count data for which estimated count data are required at the sub-population (e.g. sub-region, demographic classifications, etc.) level, for each sub-population, $s = 1, \dots, S$. Let N denote counts at the aggregate level, i.e. population size. We denote the counts to be estimated at the sub-population-level as N_s , such that the sum of sub-populations counts is equal to the total population count, $\sum N_s = N$. We supplement this population level data with an outside, independently sampled data set with sub-population counts, n_s , where $\sum n_s = n < N$. That is, the outside sample of sub-region data is a subset of the population to be estimated. The immediate impact of the outside data set is to constrain the range of eligible values, which we will denote as N'_s within each combination. Namely,

$$n_s \leq N'_s \leq N - \sum_{s' \neq s} n_{s'}. \quad (3.1)$$

Thus, we define \mathbb{C} to be the set of all valid combinations of integer-valued counts satisfying table 3.1. The cardinality of this set is denoted $|\mathbb{C}|$ and is given by:

$$|\mathbb{C}| = \binom{N - n + S - 1}{S - 1} = \frac{(N - n + S - 1)!}{(S - 1)!(N - n)!}. \quad (3.2)$$

For example, consider Bristol County, Rhode Island. Bristol County is comprised of three municipalities and is reported by the Ag Census as containing a total of 42 farms. Our sample counts for these three sub-regions are (6, 3, 2). A valid combination would therefore be any triple with each value equal to or exceeding the

sample count and with the total count equal to 42. Thus, (2, 21, 19) is not a valid combination because there are not enough farms in the first town, and (25, 12, 7) is not valid because there are too many total farms (44), but both (6, 34, 2) and (15, 15, 12) are valid combinations.

We have now developed sufficient notation to outline our estimation procedure. First, recall Bayes' Rule:

$$\Pr(A | B) = \frac{\Pr(B | A)\Pr(A)}{\Pr(B)} \propto \Pr(B | A)\Pr(A) . \quad (3.3)$$

For our purposes, $\Pr(A | B)$ in Equation 3.3 represents the probability of a specific combination of sub-population counts given the data, or $\Pr(C_i | N, n_1, \dots, n_s)$, and the other terms translate similarly. That is:

$$\begin{aligned} \Pr(C_i | N, \mathbf{n}) &\propto \Pr(N, \mathbf{n} | C_i)\Pr(C_i | C^0) \\ &= \Pr(\mathbf{n} | C_i)\Pr(C_i | C^0) \\ &\propto \Pr(\mathbf{n} | C_i) \end{aligned} \quad (3.4)$$

where (i) $\mathbf{n} = (n_1, \dots, n_s)$ denotes the vector of sub-population counts in the outside data, (ii) the equality follows from the constraint in Equation 1.1 since only combinations that sum to N are considered, and (iii) the final proportionality comparison relies on the assumption that the unconditional probability of a combination is uniform across combinations, representing the prior in our Bayesian approach. This is the simplest case of a uniform prior over combinations, which we will later generalize. Equation 3.4 therefore tells us that the posterior probability of a given combination is proportional to the probability of our outside data sample

conditional on that combination. For the case of a non-uniform prior this proportionality does not hold and the final reduction in Eq. 3.4 does not apply.

The analysis is further simplified because the conditional probability of our data given a combination, $\Pr(\mathbf{n} | C_i)$, has a closed form according to the formula for sampling without replacement. Namely,

$$\Pr(\mathbf{n} | C_i) = \prod_{s=1}^S \left(\prod_{k=0}^{k_s-1} \frac{C_{s,i} - k}{N - (s-1)n_s - k} \right). \quad (3.5)$$

Given Equation 3.5, it is theoretically possible to iterate over all eligible combinations of counts at the sub-population level and exactly calculate the posterior distribution over those counts given the outside sample data in \mathbf{n} . Unfortunately, the number of combinations given in Equation 3.2 grows astronomically large rather quickly in real-world applications. Table 3.1 provides examples for our application.

[Table 3.1]

Because it is not computationally feasible using contemporary hardware to calculate Eq. 3.5 for each possible combination, we propose a (pseudo) random sampling procedure in which valid combinations are sample uniformly from \mathbb{C} . These samples are generated by recognizing that each sub-population's count falls in a range containing $N - n + 1$ consecutive integers, whose lower bound is found in our sample for that sub-population. Revisiting our Bristol County example from earlier, wherein $N - n + 1 = 42 - 11 + 1 = 32$, it is only possible for sub-regional values to fall in the set, $\{n_s(+0), n_s + 1, \dots, n_s + 31\}$. Since each sub-region must have the same size range, the problem reduces to picking uniform integers in this range. If we offset the uniform integers by their minimum values, then all the random choices

must add up to the same total (*also* $N - n + I = 32$) to be a valid combination as described above.

This is a well-known problem whose solution is to randomly choose switching points, $s_s \in \{s_2, \dots, s_S\}$, without replacement from the set of integers, $\{1, 2, \dots, N - n\}$. The sampled combination is then derived by differencing the switching points after setting $s_1 = 0$ and $s_{S+1} = N - n + 1$. In order to handle the minimum switching interval being size 1, the resulting differences are added to the sample value, minus 1, and the sampled switching points are taken from $N - n + S - I$ candidate values. Downscaling Bristol County into *three* sub-regions provides $N - n + S - I = 42 - 11 + 3 + 1 = 35$. We would thus randomly sample $S - 1 = 2$ switching points from $\{1, \dots, 35\}$, setting $s_1 = 0$ and $s_4 = 36$.

3.3 Choosing a Prior

Researchers have two broad choices for estimating the Bayesian prior used in our estimation method: a uniform prior, or a more informative one. While in its most generalized form, our method has no requirement that sub-populations have additional characteristics from which to estimate a prior, researchers may be able to elicit a more informative prior based on additional characteristics of the sampling units. For example, in the case of classifying farms into sub-regions, this additional data may include population, land area, demographic data, etc. at the sub-region level. We now outline a rigorous procedure for eliciting an informative prior. In cases where such additional characteristic information is unavailable for whatever reason,

researchers have little choice but to assume a uniform prior across the sub-populations.

In many cases, the assumption that counted units have an equal probability of occurrence across sub-populations is potentially unrealistic, particularly in our example application of estimating farm counts. If the data available to the researchers consist of aggregate count data for multiple populations, as well as additional covariates at the sub-regional level (and can thus be summed to the regional level, e.g. population, land area, demographics, spatial information, etc.), one can test and identify potential informative priors by regressing these covariates (summed to the population level) on the population-level count data. By identifying the covariates that are most predictive of (correlated with) counts at the population level, one can use these relationships to estimate sub-regional farm counts. In this fashion, a more informative prior is elicited than the simple uniform prior. See figure 3.1 for an illustration of how the data analysis is structured.

[Figure 3.1]

In order to compare the accuracy of estimates resulting from a strong informative prior versus the uniform prior, we used the above procedure to elicit and compare various informative priors using supplemental sub-regional data obtained from the 2010 United States Census. Using county level Census data, we performed a comprehensive regression analysis, regressing various combinations of potentially relevant covariates such as population, area, and population density on our five county level farm counts. Land area was by far the most predictive covariate, being significantly positively correlated with farm counts (0.928). In calculating our prior,

we ignore the constant term in the regression to reduce bias introduced by the fact that the constant term is only meaningful at the aggregated level because we scaled up the data. Removing the constant term also imposes the intuitive constraint that a sub-region with zero land area must also contain zero farms. Because of the combinatorial nature of this problem, the regression-based priors are normalized to sum to the known aggregate counts. This makes the elicitation of the informative prior in our example equivalent to distributing N across sub-regions based on their proportional land area, such that $n_n^0 = \frac{Land\ Area_n}{Land\ Area_N}$.

In what follows, we will explicitly compare predictions of the uniform prior against those of this simple informative prior, and compare both against maximum likelihood.

Although geographic downscaling is traditionally a spatial problem, the general form of our method ignores issues of spatial dependency in favor of a more parsimonious method that requires much less data (and less technical expertise in the area of spatial modeling). However, the generalized method can be easily expanded and the use of an informative prior in our procedure makes the incorporation of features such as spatial dependence relatively straightforward. While we opt for a simple, one-parameter, area-based prior as an example here, myriad potential models for eliciting an informative prior exist, including those discussed in the introduction. Researchers with the prior belief that the posterior distribution follows a spatial dependency structure (such as spatial lag or autocorrelation) can easily incorporate such beliefs into this methodology by eliciting their priors using a spatial model such as geographically weighted regression (GWR), among many choices.

3.4 Sample Data

Rhode Island has a total of 39 municipalities grouped into five counties: Bristol, Kent, Newport, Providence, and Washington County. Counties range in size from 3 towns in Bristol County to 16 in Providence County. Our aggregated data source comes from the 2012 Ag Census, which contains farm counts by county and consequently, at the state-level. Our outside sample data comes from a survey administered by the University of Rhode Island in 2011-2012 in collaboration with local government agencies and agricultural organizations. It contains a list of addresses for 229 of the 1243 farms reported in the Ag Census. We further supplement this data with additional sub-regional data provided by the 2010 Census. Of these data, only land area was used in our final estimation procedure (indirectly, to elicit the informative prior). This information is presented in table 3.2.

[Table 3.2]

It is worth pointing out that since our posterior estimates depend on the probability of obtaining n from a random n -sized sample of C_i , the accuracy of our results depend upon the assumption that n was obtained via independent (random) sampling across s . In the case where sample data is collected via voluntary sampling, researchers must be confident in their belief that response rates are not affected by s . Similar restrictions apply in cases where sample data is collected through convenience sampling. For example, in the case where s represents a spatial classification, selection bias is likely to occur when samples are collected in person at non-random or limited locations. In cases where observations are business, such as

our farm application, unbiased sample data may be collected by searching business databases, such as those provided by the secretary of state. As is the case with most estimation procedures, any selection bias resulting from characteristics of s that cannot be controlled for will naturally induce estimation bias.

3.5 Methods

Clearly the unknowns in our data set are the city-level counts. We focus instead on the county totals, as if unknown, so that we can compare the results of our procedure against the true underlying distribution. By aggregating our regional counts to the state level and aggregating our sample data to the county level, we can compare the accuracy of our estimates using i) a uniform prior, ii) a simple spatial prior, and iii) maximum likelihood. The maximum likelihood estimates (MLE) differ from those of the Bayesian estimates under the uniform prior since because N and n are known,

MLE estimates for N_s simplify to $\hat{N}_s^{MLE} = \frac{n_s}{n} N$.

For each methodology, we evaluate the normalized root mean squared error (NRMSE) of our posterior point estimates relative to the values reported in the Ag Census. The NRMSE is simply the familiar root mean squared error (RMSE), scaled by the average region-size, $\bar{N}_s = N / S$. Using NRMSE supports our goals of estimating the effects of both population size, N , and sample size, n / N , on the relative performance of each method.

To obtain estimates for the Bayesian methods, we use the sampling procedure described in Section 3.2 to calculate estimates from 100,000 sample combinations. We report as point estimates the means of the posterior count distributions, which is

why decimal values are observed in the estimates despite only considering integer-valued combinations. We repeat the procedure 200 times for each estimate to obtain standard errors of our estimates and calculate the NRMSE of our posterior estimates relative to the values reported in the Ag Census.

To test the effect of population size, N , on estimation performance, we simulate five sub-regions with our county-size proportions and sample proportion (approx. 0.2), at varying population sizes from $N = 50$ up to $N = 5,000$. To test the effect of sample size on estimation performance, we again simulate five sub-regions with our county-size proportions, this time with a fixed population size of $N = 1,250$ and with varying sample sizes from $n / N = 0.01$ up to $n / N = 0.50$.

For each comparison, we conduct two simulations, one with observations bootstrapped from our observed Rhode Island sample, and another randomly sampled from a multinomial distribution taken only from the population parameters. The results are nearly identical across the paired simulations, indicating that the sample data we collected do not contain extreme deviations from the projected sampling distribution.

3.6 Results

The mean NRMSE and standard error for each method are presented in table 3.3, as a function of varying sample size. For a population of 1,250, the Bayesian methods consistently outperform MLE for all sample sizes up to half of the population. Among the Bayesian methods, the simple, area-based, informative prior was best for small samples (due to greater sampling variability), but the uniform prior was best for

samples comprising at least 5% of the population. These results suggest a somewhat counterintuitive finding: namely, that even in cases where detailed spatial information is available, many applied problems will get more accurate results using the uniform prior, even when it might not seem realistic to the application at hand. The reasoning is that even relatively small samples will quickly become more representative of the underlying population than a good informative prior, but not so representative as to obviate the need for a Bayesian approach over MLE.

[Table 3.3]

When examining table 3.3, it is important to note that the conventional wisdom regarding standard errors and statistical significance does not hold because of correlated testing. That is, we are not testing whether one measure produces a lower NRMSE on average over independent tests, rather, we are testing whether one measure produced a statistically significantly lower NRMSE than the others across the same simulation tests. So, we do not present p -values or a fully developed hypothesis testing framework with our results. Rather, the “winning method” which is bold-faced in each column of table 3.3, is determined according to a Single Transferrable Vote system as outlined in Tideman (1995) based on the percentage of “wins” (lower observed NRMSE) out of 200 trials. In essence, the winning model is chosen according to winning a simple majority of trials outright, or else winning a “runoff” between the top two candidates.

For the interested reader, pairwise tests of statistically significant better performance between two estimators can be determined as follows. The null hypothesis for pairwise comparison of two identical estimators is that the number of

wins for each estimator follows the binomial distribution with $n = 200$ draws and $p = 0.5$. So, for individual, pairwise tests of performance, the threshold for statistical significance at the 95% and 99% levels are 113 wins (56.5%) and 117 wins (58.5%), respectively. Clearly, to evaluate every possible pairwise test in Table 3.3 involves many hypotheses, so p -values would need to be adjusted using either a Bonferroni correction or stepdown methods to control the family-wise error rate (e.g., as in Romano and Wolf, 2005). Explicit testing of multiple hypotheses in this fashion is beyond the scope of this paper.

[Table 3.4]

Table 3.4 is structured similarly, but show the effect of varying population size given the sample size held fixed at $n / N = 20\%$ of the population. The table also shows the Bayesian methods consistently outperforming MLE, but show subtly different patterns of performance of the informative prior against the uniform prior. With the sample size held at a fixed percentage, the informative prior outperforms for populations smaller than 500, while the uniform outperforms for larger populations. For populations of exactly 500, the performance of the two priors is not statistically different at conventional levels.

3.7 Discussion

The above results are primarily focused on evaluating the performance of our Bayesian methods for a case where the underlying distribution is known. However, our method is only designed to be useful in cases where this information is unavailable. Furthermore, applications of our procedure to new problems will likely

involve variation in population size, sample size(s) and availability of an informative prior, distinct from the permutations described here. In this section, we consider the possibility that future researchers have access to data at higher levels of aggregation, similar to how we have both state-level and county-level farm counts for Rhode Island from the Ag Census, and county-level land area data from the US Census.

If it can be assumed that spatial (or other group-wise dependence) patterns are likely to hold at higher levels of aggregation, then an informative prior can be calibrated from that data and applied in the downscaling problem. In our application, that would mean calibrating the land-area prior from county-level data and then applying it to the city-level downscaling problem. Depending on the application, however, this assumption may not be palatable. Spatial econometric models can be conceptualized as having a direct effect from the covariates and an indirect effect from the spatial dependence structure. If this indirect effect is relatively smaller at higher levels of aggregation, then calibrating the prior at higher levels will cause it to appear more informative than it actually will be in the downscaled analysis. Identifying when this problem materially affects the analysis is an area for future research. That said, there is no reason why spatial dependence observed in an econometric model would predict non-random sampling, so whenever the spatial prior is suspect, researchers can always default to the uniform prior for reasonable performance.

It is important to note that in an attempt to make our estimation procedure generalizable to non-spatial applications, as well as to reduce its dependence on additional outside data sources, we have intentionally ignored the explicit

incorporation of spatial autocorrelation. Although some degree of autocorrelation is likely captured implicitly via our sample, estimates might be further improved through incorporating such autocorrelation into our conditional probability estimates when possible. That said, it is beyond the scope of this chapter to quantify these tradeoffs, or to demonstrate the information loss associated with discarding spatial elements in the analysis.

Beyond spatial dependence defined econometrically, there is also the possibility that the outside data sample is non-random, in the sense that spatial factors influence response rates. At higher levels of aggregation this can be tested simply by evaluating the degree to which the sample contains outliers relative to a typical sample from a multinomial distribution. A further verification step is possible using the simulation methods outlined above at higher levels of aggregation. Namely, the bootstrapped analysis can be replicated with counts drawn directly from a multinomial distribution instead of from the sample data. Below, we give an example of simulation results obtained in this fashion in Tables 3.5 and 3.6, which replicate our Tables 3.3 and 3.4 but do *not* use our sample data. For our specific application, it can be observed that the results are nearly identical, the desired outcome indicating that systematic sampling bias is unlikely to be a problem in our application.

[Table 3.5]

[Table 3.6]

If we consider the city-level downscaling problem in our application, the above procedures indicate that every county in Rhode Island should be estimated using the informative, area-based prior. Clearly, we do not have the underlying, true

distribution of city-level farm counts for verification, so we include this observation only for completeness.

Two issues not previously addressed are i) the effects of uncertainty in the top-level population estimates, and ii) the scenarios in which MLE does outperform the Bayesian estimators, according to conventional wisdom based on asymptotic results. While our procedure is designed to mitigate potential estimation error resulting from the increased sampling variability inherent in relatively small samples, it does not account for potential error in the aggregate count data. In our application, for example, the Ag Census farm counts for Rhode Island are reported as 1,243 total farms with a standard error of 236 (USDA 2014). The analysis thus far suggests that incorporating an error term on the total count may have non-linear effects because of simultaneous changes both in the population size, N , and in the sample proportion, n/N .

To address this concern, we repeated the simulation analysis using the uniform prior, with each replication using a different total farm count drawn from a normal distribution with mean and standard deviation according to the reported Ag Census mean and standard error. The mean estimated farm counts arising from this procedure were within 1% of the values estimated with $N = 1,243$. This suggests that errors in top-level counts are less of a concern, as long as i) it is recognized that the division of the population into groups will necessarily result in estimates that are proportional to the total used, and ii) that the estimation error in the total count is not so large as to make the collected sample size unlikely or impossible.

Additionally, it is important to give proper context to our finding that these Bayesian methods outperform MLE. Clearly, this is a finite sample result since, asymptotically, Bayesian updating with a uniform prior converges to MLE, whereas in small samples MLE is equivalent to Bayesian updating with zero sample weight on the prior. Also, it may not be immediately obvious, but our application data set includes considerable variation in the group sizes to be estimated: 42, 126, 214, 425, and 436 (from table 3.2). Having extremes in the group sizes, especially on the small end, leads to inherently noisier sampling of the smaller groups. This problem can be conceptualized as arising from the probability that a given sample will be representative of the population conditional on population size and sample size.

To show how variation across group sizes affects the performance of MLE relative to the Bayesian methods described here, we ran some preliminary simulations. The simulated group sizes were all drawn IID from a normal distribution with sigma given by a fraction of the mean value, and samples were then drawn from the resulting multinomial distribution. Our sample data had a standard deviation of 71% of the mean count, and MLE did not outperform the Bayesian methods for any population of $N = 5,000$ or below. Reducing the standard deviation to 50% of the mean count, we found that MLE was statistically significantly best (lowest NRMSE) for populations above 2,000 (as might appear in table 3.6). These preliminary results suggest that the efficacy of MLE relative to the Bayesian methods is not only a function of population size and sample size, but also of the degree of heterogeneity in the sub-population counts to be estimated. Naturally, however, the true distribution of

sub-population counts is inherently unknown a priori. We leave exact quantification of these tradeoffs as an area for future research.

Lastly, one might wonder if it would be possible to apply these methods to the estimation procedure performed in Manuscript 1. There are two reasons why such an application would be inappropriate. Firstly, the classification estimates performed in Manuscript 1 are not explicit or ‘known’ classifications, but rather they are probabilistic classifications with respect to a distribution, where the probability of belonging to each class (or distribution) is calculated for each observation. Secondly, we do not possess the required independent sample. Rather, we are trying to estimate the distribution of classes across a sample of farmers, rather than the entire population, making claims about which population this distribution belongs to. However, if one was trying to estimate the proportion of optimistically biased, pessimistically biased, and unbiased individuals across a population of known size using sample data, our distributional estimates from Manuscript 1 would likely serve as a good informative prior.

3.8 Conclusion

Micro-level statistical data are often unavailable at the desired level of disaggregation, despite their critical importance for applied policy research. Herein, we present a Bayesian methodology for ‘downscaling’ aggregated count data to the micro-level, using an outside statistical sample. Our procedure combines numerical simulation with exact calculation of combinatorial probabilities. We motivate our approach with an application estimating the number of farms in a region, using count

totals at higher levels of aggregation, and data sourced from the 2012 USDA Ag Census. In a simulation analysis over varying population sizes, we demonstrate both robustness to sampling variability and outperformance relative to maximum likelihood. Our results show that Bayesian methods have better finite sample performance than MLE in many cases relevant to applied research, especially for relatively small populations ($N < 5,000$).

We develop a number of methods for applied researchers to calibrate informative prior probabilities, and to estimate whether the combination of sample size and population size in their application will perform best with their more informative prior, or a simple uniform prior. In many cases, the uniform prior performs reasonably well and can be used as a default in cases where a more informative prior is unavailable, or cannot be reasonably calibrated due to spatial considerations. We also show how the process of calibrating the priors can be simulated to verify that they are not being affected adversely by outside sample data that contains too many outliers. Our methods appear to be robust both to sampling variability in the outside data sample, and also to uncertainty in the top-level population counts. An area for future research is determining the effects of heterogeneity in sub-population sizes on the relative performance of maximum likelihood estimates in smaller populations.

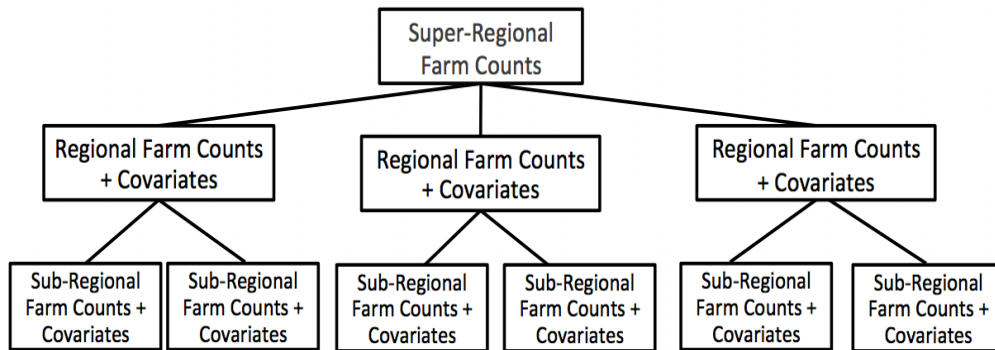
3.9 Work Cited

- Chakir, R. 2009. Spatial Downscaling of Agricultural Land-Use Data: An Econometric Approach Using Cross Entropy. *Land Economics*, 85(2), 238-251.
- Coelho, C. A. S., D. B. Stephenson, F. J. Doblas-Reyes, M. Balmaseda, A. Guetter, and G. J. Van Oldenborgh. 2006. A Bayesian Approach For Multi-Model Downscaling: Seasonal Forecasting of Regional Rainfall and River Flows In South America. *Meteorological Applications*, 13(01), 73-82.
- Dendoncker, N., P. Bogaert, , and M. Rounsevell. 2006. A Statistical Method to Downscale Aggregated Land Use Data and Scenarios. *Journal of Land Use Science*, 1(2-4), 63-82.
- Fasbender, Dominique, and Taha BMJ Ouarda. 2010. Spatial Bayesian Model For Statistical Downscaling of AOGCM to Minimum and Maximum Daily Temperatures. *Journal of Climate* 23, No. 19: 5222-5242.
- Gärtner, D., A. Keller, and R. Schulin. 2013. “A Simple Regional Downscaling Approach For Spatially Distributing Land Use Types For Agricultural Land.” *Agricultural Systems*, 120, 10-19.
- Gocht, Alexander, and Norbert Roder. 2011. “Salvage The Treasure of Geographic Information In Farm Census Data.” In *Proc Int Congress European Association of Agricultural Economists*.
- Howitt, R., and A. Reynaud. 2003. “Spatial Disaggregation of Agricultural Production Data Using Maximum Entropy.” *European Review of Agricultural Economics*, 30(3), 359-387.

- Hashmi, M. Z., A. Y. Shamseldin, and B. W. Melville. 2009. "Statistical Downscaling of Precipitation: State-of-The-Art and Application of Bayesian Multi-Model Approach For Uncertainty Assessment." *Hydrology and Earth System Sciences Discussions*, 6(5), 6535-6579.
- Murphy, J. 1999. "An Evaluation of Statistical and Dynamical Techniques For Downscaling Local Climate." *Journal of Climate*, 12(8), 2256-2284.
- Polasek, W., C. Llano, and R. Sellner. 2010. "Bayesian Methods For Completing Data In Spatial Models." *Review of Economic Analysis*, 2(2), 194-214.
- Purcell, Noel J., and L. Kish. 1980. "Postcensal Estimates For Local Areas (Or Domains)." *International Statistical Review/Revue Internationale De Statistique*: 3-18.
- Romano, J. P., and M. Wolf. 2005. "Exact and Approximate Stepdown Methods For Multiple Hypothesis Testing." *Journal of The American Statistical Association*, 100(469), 94-108.
- Tideman, N. 1995. "The Single Transferable Vote." *The Journal of Economic Perspectives*, 9(1), 27-38.
- United States Department of Agriculture. 2014. *2012 Census of Agriculture, Appendix A: Census of Agriculture Methodology*. Washington, DC: U.S. Government Printing office.
- von Storch, H., E. Zorita., and U. Cubasch. 1993. "Downscaling of Global Climate Change Estimates to Regional Scales: An Application to Iberian Rainfall In Wintertime." *Journal of Climate*, 6(6), 1161-1171.

3.10 Figures

Figure 3.1: Nested Data Structure



3.11 Tables

Table 3.1: Number of Eligible Combinations per County

County	Farms	Towns	Sample	Number of Eligible Combinations
Bristol	42	3	11	55,278
Kent	126	5	27	4,421,275
Newport	214	6	44	1,291,150,035
Providence	425	16	82	115,508,396,906,738,000,000,000,000
Washington	436	9	65	9,801,540,147,002,170

Table 3.2: Farms, Municipalities,
and Land Area (mi²) per County

County	Farms	Land Area (mi ²)
Bristol	42	24.16
Kent	126	168.53
Newport	214	102.39
Providence	425	409.5
Washington	436	329.23

Note: Pearson's Rho = 0.928.

Table 3.3: Mean Normalized Root Mean Square Error by Estimate Type:
Fixed Population Size (1,250) with Increasing Sample Size

Sample Proportion (n/N)		0.01	0.05	0.10	0.20	0.33	0.50
Uniform Prior	Mean	0.393	0.207	0.143	0.1	0.077	0.055
	Std. Dev.	(0.163)	(0.088)	(0.058)	(0.042)	(0.029)	(0.022)
	Overall Win %	0.185	0.535	0.61	0.64	0.495	0.465
	Win % vs. Informative	0.19	0.59	0.74	0.805	0.785	0.745
	Win % vs. MLE	0.95	0.69	0.815	0.79	0.61	0.585
Informative Prior	Mean	0.245	0.207	0.173	0.133	0.101	0.069
	Std. Dev.	(0.014)	(0.024)	(0.025)	(0.026)	(0.028)	(0.024)
	Overall Wins	0.795	0.405	0.25	0.19	0.185	0.18
	Wins vs. Uniform	0.81	0.41	0.26	0.195	0.215	0.255
	Wins vs. MLE	0.96	0.69	0.505	0.315	0.31	0.335
Maximum Likelihood	Mean	0.57	0.265	0.176	0.119	0.086	0.059
	Std. Dev.	(0.209)	(0.098)	(0.063)	(0.044)	(0.029)	(0.021)
	Overall Wins	0.02	0.06	0.14	0.17	0.32	0.355
	Wins vs. Uniform	0.05	0.115	0.185	0.21	0.39	0.415
	Wins vs. Informative	0.04	0.31	0.495	0.685	0.69	0.665

Table 3.4: Mean Normalized Root Mean Square Error by Estimate Type:
Fixed Sample Proportion (0.20) and Increasing Population Size

Total Farm Count (N)	50	100	250	500	750	1,250	2,000	5,000	
Uniform Prior	Mean	0.424	0.315	0.194	0.147	0.124	0.1	0.079	0.053
	Std. Dev.	(0.168)	(0.129)	(0.081)	(0.06)	(0.052)	(0.042)	(0.033)	(0.021)
	Overall Wins	0.06	0.105	0.255	0.42	0.605	0.64	0.63	0.635
	Wins vs. Informative	0.06	0.11	0.285	0.455	0.665	0.805	0.85	0.985
	Wins vs. MLE	0.965	0.905	0.88	0.865	0.635	0.79	0.755	0.635
Informative Prior	Mean	0.258	0.203	0.159	0.138	0.135	0.133	0.129	0.129
	Std. Dev.	(0.097)	(0.075)	(0.045)	(0.038)	(0.034)	(0.026)	(0.022)	(0.012)
	Overall Wins	0.94	0.885	0.705	0.525	0.335	0.19	0.13	0.005
	Wins vs. Uniform	0.94	0.89	0.715	0.545	0.335	0.195	0.15	0.015
	Wins vs. MLE	1	0.975	0.885	0.7	0.62	0.315	0.22	0.015
Maximum Likelihood	Mean	0.605	0.42	0.252	0.182	0.15	0.119	0.094	0.061
	Std. Dev.	(0.225)	(0.142)	(0.103)	(0.071)	(0.055)	(0.044)	(0.036)	(0.022)
	Overall Wins	0	0.01	0.04	0.055	0.06	0.17	0.24	0.36
	Wins vs. Uniform	0.035	0.095	0.12	0.135	0.105	0.21	0.245	0.365
	Wins vs. Informative	0	0.025	0.115	0.3	0.38	0.685	0.78	0.985

Table 3.5: Mean Normalized Root Mean Square Error by Estimate Type using Simulated Sample: Fixed Population Size (1,250) with Increasing Sample Size

Sample Proportion (n/N)		0.01	0.05	0.10	0.20	0.33	0.50
Uniform Prior	Mean	0.409	0.213	0.151	0.103	0.077	0.055
	Std. Dev.	(0.168)	(0.076)	(0.062)	(0.049)	(0.032)	(0.022)
	Overall Wins	0.17	0.49	0.58	0.64	0.485	0.47
	Wins vs. Informative	0.175	0.57	0.7	0.805	0.76	0.725
	Wins vs. MLE	0.945	0.86	0.795	0.78	0.63	0.585
Informative Prior	Mean	0.245	0.208	0.174	0.132	0.098	0.069
	Std. Dev.	(0.015)	(0.024)	(0.026)	(0.027)	(0.029)	(0.026)
	Overall Wins	0.815	0.43	0.3	0.18	0.21	0.225
	Wins vs. Uniform	0.825	0.43	0.3	0.195	0.24	0.275
	Wins vs. MLE	0.975	0.66	0.5	0.34	0.37	0.365
Maximum Likelihood	Mean	0.604	0.255	0.181	0.122	0.089	0.061
	Std. Dev.	(0.228)	(0.086)	(0.066)	(0.044)	(0.034)	(0.023)
	Overall Wins	0.015	0.08	0.12	0.18	0.305	0.305
	Wins vs. Uniform	0.055	0.14	0.205	0.22	0.37	0.415
	Wins vs. Informative	0.025	0.34	0.5	0.66	0.63	0.635

Table 3.6: Mean Normalized Root Mean Square Error by Estimate Type using Simulated Sample Fixed Sample Proportion (0.20) and Increasing Population Size

Total Farm Count (N)		50	100	250	500	750	1,250	2,000	5,000
Uniform Prior	Mean	0.394	0.277	0.214	0.15	0.126	0.103	0.077	0.053
	Std. Dev.	(0.173)	(0.111)	(0.09)	(0.063)	(0.053)	(0.049)	(0.031)	(0.022)
	Overall Wins	0.05	0.105	0.245	0.41	0.54	0.64	0.72	0.675
	Wins vs. Informative	0.05	0.115	0.265	0.46	0.635	0.805	0.93	0.99
	Wins vs. MLE	1	0.98	0.93	0.725	0.835	0.78	0.78	0.675
Informative Prior	Mean	0.242	0.193	0.153	0.142	0.134	0.132	0.131	0.126
	Std. Dev.	(0.095)	(0.063)	(0.05)	(0.036)	(0.035)	(0.027)	(0.023)	(0.018)
	Overall Wins	0.95	0.88	0.725	0.525	0.35	0.18	0.07	0.01
	Wins vs. Uniform	0.95	0.885	0.735	0.54	0.365	0.195	0.07	0.01
	Wins vs. MLE	1	0.98	0.93	0.725	0.57	0.34	0.155	0.015
Maximum Likelihood	Mean	0.561	0.397	0.265	0.18	0.148	0.122	0.092	0.059
	Std. Dev.	(0.198)	(0.165)	(0.096)	(0.067)	(0.055)	(0.044)	(0.037)	(0.021)
	Overall Wins	0	0.015	0.03	0.065	0.11	0.18	0.21	0.315
	Wins vs. Uniform	0.01	0.075	0.11	0.14	0.165	0.22	0.22	0.325
	Wins vs. Informative	0	0.02	0.07	0.275	0.43	0.66	0.845	0.985