Accounting for Weather Probabilities in Crop Insurance Rating

Keith H. Coble*, Mary Frances Miller†, Roderick M. Rejesus‡, Ryan Boyles‡, Barry K. Goodwin‡, Thomas O. Knight§

*Miss. State, †Select Actuarial, ‡NC State, § Texas Tech

2013 Crop Insurance and the Farm Bill Symposium
Louisville, KY (October 8-9, 2013)

Introduction

- Historical loss experience is the foundation of US crop insurance premium rating
- Previously used simple average of equally-weighted historical loss cost data from 1975 onwards
- Does this series capture the “longer” term weather experience needed to accurately estimate premium rates?

Introduction

- 38 years of loss history (1975-2012) may still not accurately reflect the long-term probabilities of weather events
- With simple averaging, 2012 drought year is given 1/38 weight
- But this drought may be a 1 in 20 year event (need larger weight) or a 1 in 50 year event (need smaller weight)

Introduction

- The inherent tension of rating
  - Longer series gives more appropriate weight to random events, but
  - Longer series picks up:
    - Changes in policy terms
    - Changes in program participation
    - Changes in risk
      - production technology
      - Climate change
    - Changes in data quality

Objective

- To develop a methodology for weighting the historical loss cost experience data based on a longer time-series weather information
  - Improve statistical validity of estimated premium rates
  - Approaches evaluated based on statistical validity, feasibility, sustainability, and balancing improvement vs. complexity

Data Issues and Conceptual Considerations

Historical Loss Cost Data and Weather Data

Historical Loss Cost Data

- Aggregate county level loss cost data is starting point for rating
- Indemnities & liabilities "normalized"
- Simple, equally-weighted average:

$$County\ Base\ Rate_i = E(LC_{ij}) = \sum_{j=1}^{n} f(LC_{ij}) \times LC_{ij} = \frac{1}{n}\sum_{j=1}^{n} LC_{ij}\ ,$$

Historical Loss Cost Data

- Catastrophic loading also imposed
  - Previously, losses above the 80th percentile are spread across all counties for a crop in a state
- Equal weighting assumes a uniform pdf, but weather distributions not necessarily uniform
- Conceptually, longer time-series weather data can augment the smaller sample LC data

Weather/Climate Data

- In developing a system to weight historical LC data with long-term weather/climate data, must consider:
  - Weather/Climate data to be utilized (i.e., how to choose)
  - Procedure for weighting each year (i.e., how to categorize year and create weight)

Weather Data Considerations

- Length of different weather/climate data available
- Degree of coverage and level of aggregation
- Availability of different weather variables
- Source of the weather/climate data and availability of the data in the future

Weather Data Choice

- Examined various weather data based on considerations above
- Choice – National Climatic Data Center's Time Bias Corrected Divisional Temperature-Precipitation-Drought Index data
- Also called Climate Division Data

Weather Data Choice

- Climate Division Data
  - Longest record with national coverage (since 1895)
  - Updates regularly provided – drought, precipitation, temperature and heat accumulation
- Only data set routinely available that provides both critical measures and long term record

Climate Division Boundaries

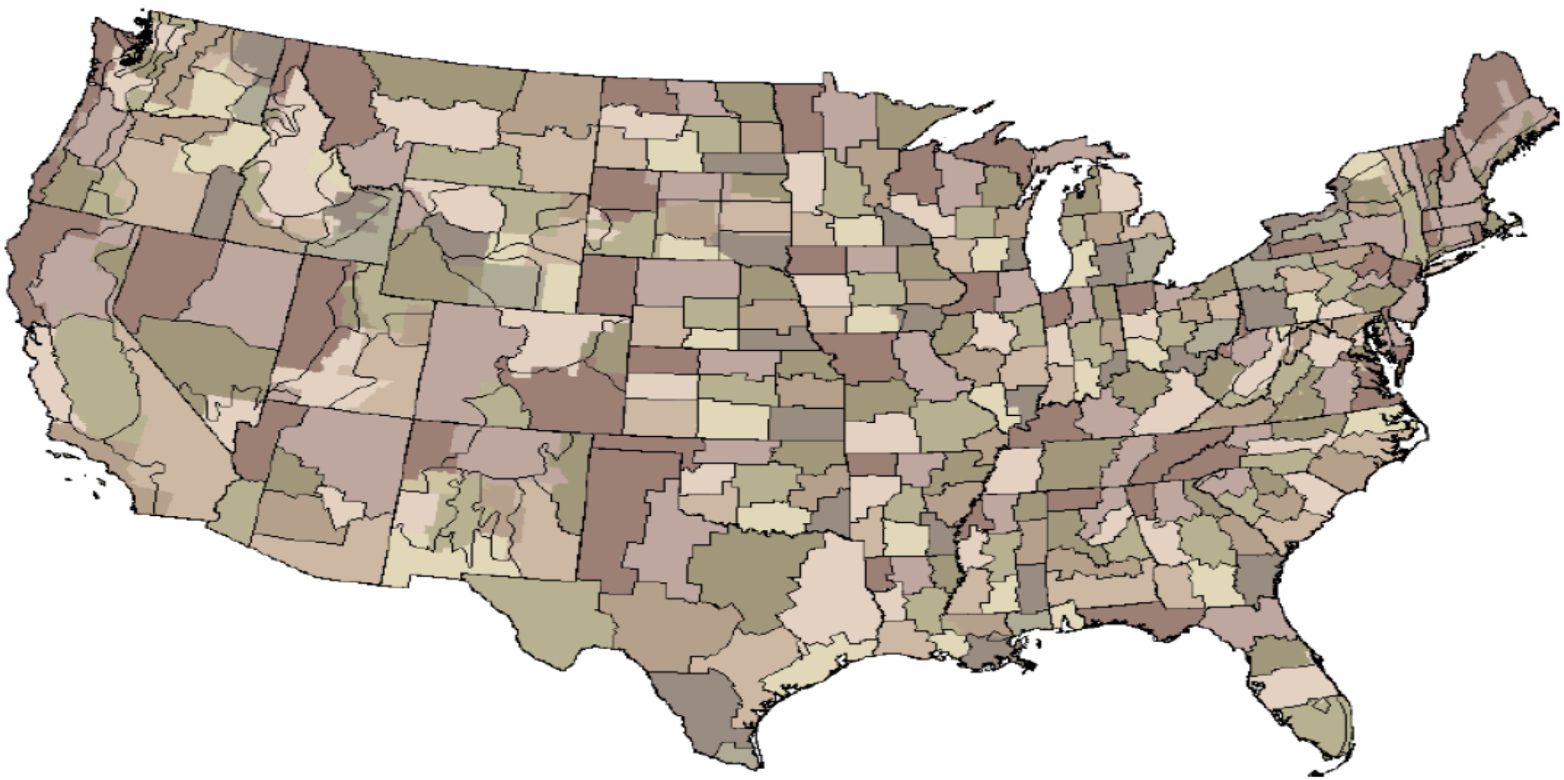

Figure 1. Climate Division Boundaries and County Assignment within Climate Divisions

Merged Loss Cost and Weather Data

- Loss cost (StatPlan) data at county level, but weather data at climate division level
- Counties within climate division has same weather data
- Merged LC and climate division data used to classify loss years, county data used to average loss cost data to calculate base rate

Empirical Approach

Weather Index Development, Loss Year Classification, Variable Bin Width Assignment & Loss Cost Averaging

Weather Index Development

- Need to choose weather variables to determine relative weights assigned to each loss year
  - Want fewest variables that explain losses
- Palmer Drought Severity Index (PDSI)
- Cooling Degree Days (CDD)
  - Also called Growing Degree Days (GDD) at base 65°F

Weather Index Development

- PDSI
  - Captures both extreme wet and dry conditions
  - Incorporates temperature, precipitation, and evaporation
- CDD
  - Captures effect of extended cold or heat events not captured by PDSI

Weather Index Development

- For corn, soybeans, cotton, sorghum, potatoes:
  - PDSI – positive and negative
    - May/June July/Aug for Midwest but adjusted by latitude & RMA dates
  - Total heat units = CDD total season
  - Extreme heat = CDD
    - June/July for Midwest but adjusted by latitude & RMA dates
- For winter cereals:
  - Drought is still an issue: root system establishes in the fall
  - PDSI – positive and negative
    - Planting months
    - Spring growth period
  - Excess heat is not a factor

Weather Index Development

- Fractional logit regression used to estimate index (due to censoring in data)
  - Climate division level adjusted loss cost as a function of weather variables
- Out-of-Sample competition for each state to determine optimal combination of weather variables that predicts loss costs
  - For more parsimonious specification

Weather Index Development

- Weather index based on predicted loss costs from fractional logit regression models
- Can use weather/climate data to "backcast" a weather index for each year from 1895 onwards
  - Relative probability of extreme loss event can be more accurately assessed
- Weights used only if weather variables are statistically significant

Loss Year Classification

- Use predicted weather index to classify a year and assign weight
- Options:
  - Standard histogram with equal bin widths and variable frequencies
  - Variable bin widths with equal probabilities
    - Variable bin width preferred – less severe "empty bin problem" and simplicity

Equal Width Bins

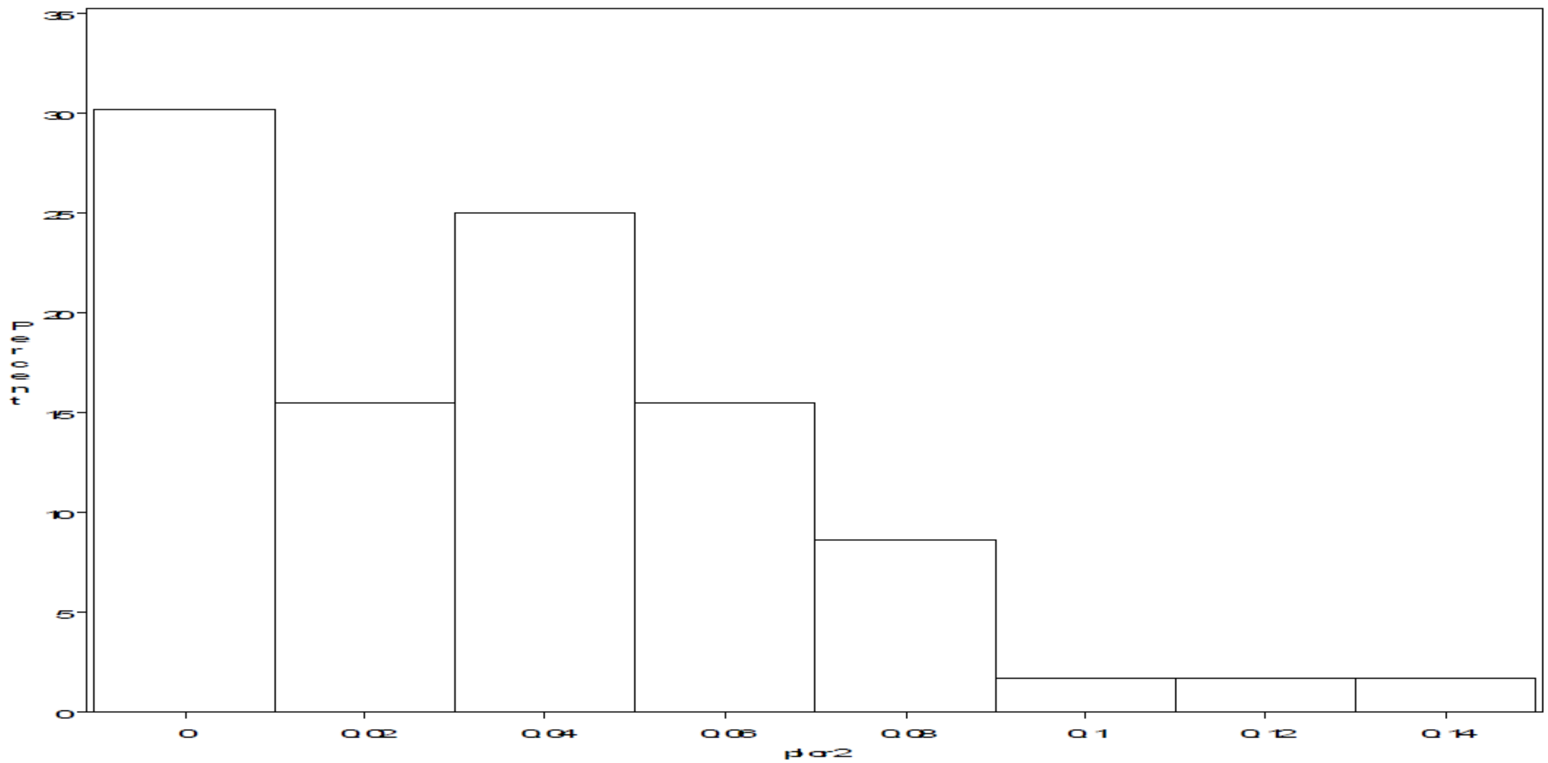

Figure 2. Example Histogram with Equal Bin Widths and Variable Probabilities for Each Bin

Variable Bin Width Assignment

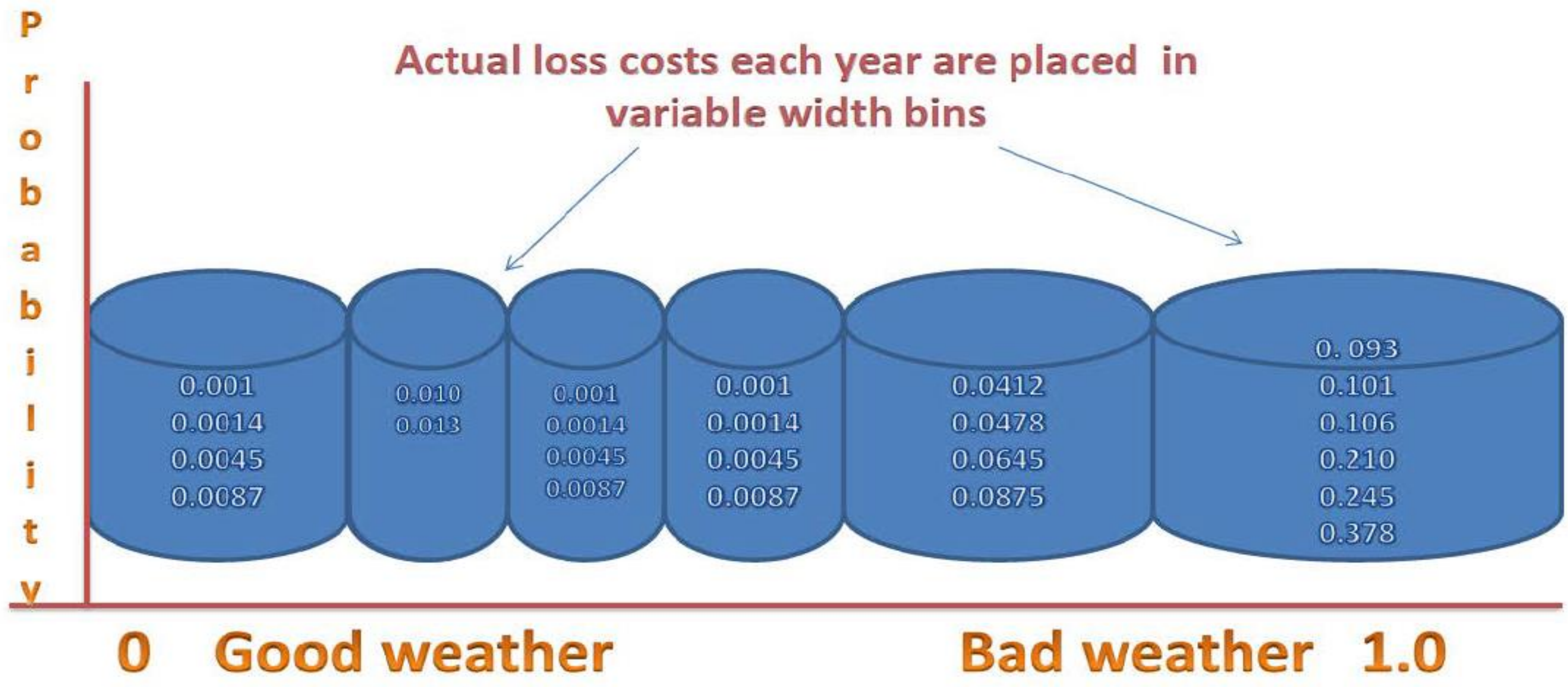

Figure 3. Example of Variable Width Bins with Equal Probability for each Bin

Variable Bin Width Assignment

- Determine number of bins
  - Choose number so that no empty bins in 1980-2009 county LC data
  - If 10 bins – find weather indexes that falls into 10^{th}, 20^{th}, 30^{th}, .. 90^{th} percentiles
- Weather index for each year can be classified and assigned to the bin in which it falls

Variable Bin Width Assignment

Table 1. Hypothetical Example of Bin Classification: Soybeans in Mississippi (State=28) climate division 1 (1980-2009).

State	Climate Division	Year	State proxy flag=1 if used state predicted values	Bin Classification	No of Bins for the Climate Division	Flag =1 if insignificant
28	1	1980	0	4	10	0
28	1	1981	0	8	10	0
28	1	1982	0	2	10	0
28	1	1983	0	5	10	0
28	1	1984	0	4	10	0
28	1	1985	0	8	10	0
28	1	1986	0	9	10	0
28	1	1987	0	1	10	0
28	1	1988	0	10	10	0
28	1	1989	0	8	10	0
28	1	1990	0	4	10	0
28	1	1991	0	6	10	0
28	1	1992	0	5	10	0
28	1	1993	0	2	10	0
28	1	1994	0	4	10	0
28	1	1995	0	1	10	0
28	1	1996	0	1	10	0
28	1	1997	0	5	10	0
28	1	1998	0	10	10	0
28	1	1999	0	5	10	0
28	1	2000	0	8	10	0
28	1	2001	0	5	10	0
28	1	2002	0	4	10	0
28	1	2003	0	5	10	0
28	1	2004	0	3	10	0
28	1	2005	0	7	10	0
28	1	2006	0	9	10	0
28	1	2007	0	8	10	0
28	1	2008	0	6	10	0
28	1	2009	0	4	10	0
28	1	2010	0	10	10	0

Note: The state proxy flag is equal to 1 if there are not enough observations (n>10) in the climate divisions to run a credible fractional regression model and calculate a predicted loss cost (weather index).

Loss Cost Averaging

- Use actual county level loss cost data (StatPlan) with results of variable bin assignment merged in
- Do Weather Weighting:
  - Take average LC within bins
  - Take “average of the average loss costs” across bins

Loss Cost Averaging

- "Recency Weighting" can be applied when averaging within bins
  - More weight to more recent data
- Can also shorten StatPlan data to use (i.e., 1990 onwards rather than 1980 onwards)
- Catastrophic loading at 80th and 90th percentile consistent with weighting approach

Results

Premium Rate Impacts

Results

Table 3. Hypothetical Example of Unweighted and Weather Weighted Loss Costs at the County-level for Boone County (county=15), Dallas County (county=49), and Grundy County (county=75), IA (State=19).

State	Climate Division	County	County Average loss costs	Flag =1 if insignificant	Weighting Type
19	5	15	0.0096378	0	1
19	5	15	0.0076921	0	2
19	5	15	0.0028386	0	3
19	5	15	0.0027737	0	4
19	5	15	0.0035587	0	5
19	5	15	0.0033862	0	6
19	5	49	0.0100697	0	1
19	5	49	0.0097928	0	2
19	5	49	0.0058953	0	3
19	5	49	0.0058029	0	4
19	5	49	0.007514	0	5
19	5	49	0.0075715	0	6
19	5	75	0.0091694	0	1
19	5	75	0.0051299	0	2
19	5	75	0.001323	0	3
19	5	75	0.0010593	0	4
19	5	75	0.0044935	0	5
19	5	75	0.0032308	0	6

Note: Weighting type = 1 if the average loss cost is calculated with no weather weighting and no censoring; Weighting type =2 if the average loss cost is calculated with weather weighting but no censoring; Weighting type = 3 if the average loss cost is calculated with censoring at the 80th percentile and no weather weighting; Weighting type = 4 if the average loss cost is calculated with censoring at the 80th percentile and with weather weighting; Weighting type = 5 if the average loss cost is calculated with censoring at the 90th percentile and no weather weighting; Weighting type = 6 if the average loss cost is calculated with censoring at the 90th percentile and with weather weighting.

Results

Table 4. Liability Weighted National Average (across counties) of Unweighted and Weather Weighted Average Loss Costs for Apples, Barley, Corn, Cotton, Potatoes, Rice, Sorghum, Soybeans, Spring Wheat and Winter Wheat. – table 4.9

Crop	No. of Counties	Unweighted loss costs (no censoring)	Weather weighted loss costs (no censoring)	Unweighted loss costs (censoring at 80th)	Weather weighted loss costs (censoring at 80th)	Unweighted loss costs (censoring at 90th)	Weather weighted loss costs (censoring at 90th)
apples	140	0.1839529	0.1756118	0.1509251	0.1458255	0.1722479	0.1649113
barley	646	0.1033683	0.0952631	0.071994	0.0677116	0.088203	0.0820236
corn	1930	0.0505333	0.0525652	0.028726	0.0293841	0.0394102	0.0409063
cotton	437	0.143511	0.1459077	0.1103868	0.1110684	0.1292813	0.1305584
potatoes	128	0.083174	0.0807186	0.0659818	0.0646853	0.0752233	0.0730846
rice	84	0.0263574	0.0251909	0.015527	0.0148564	0.0203618	0.0193536
sorghum	750	0.1208383	0.1317581	0.0887164	0.09226	0.1079448	0.1140653
soybeans	1523	0.0542112	0.0538458	0.0384229	0.0379807	0.0467105	0.0460899
spring wheat	244	0.1218715	0.1171909	0.0887732	0.0872793	0.1094074	0.1063092
winter wheat	951	0.0982152	0.0852073	0.0719574	0.065563	0.0851164	0.0759965

Note: These are the national average loss costs across all counties (i.e., liability weighted average) where the insignificance flags and state proxy flags are not equal to one. All weighted and unweighted loss costs for each county is available from the authors upon request.

Results

- Rates gets adjusted for weather in both directions (positive & negative)
- For apples, barley, cotton, potatoes, rice, and spring/winter wheat, the weather weighted average loss costs (at the national level) tend to be smaller
- For corn, cotton, sorghum, and soybeans the weather weighted average loss costs (at the national level) tend to be larger.

Results

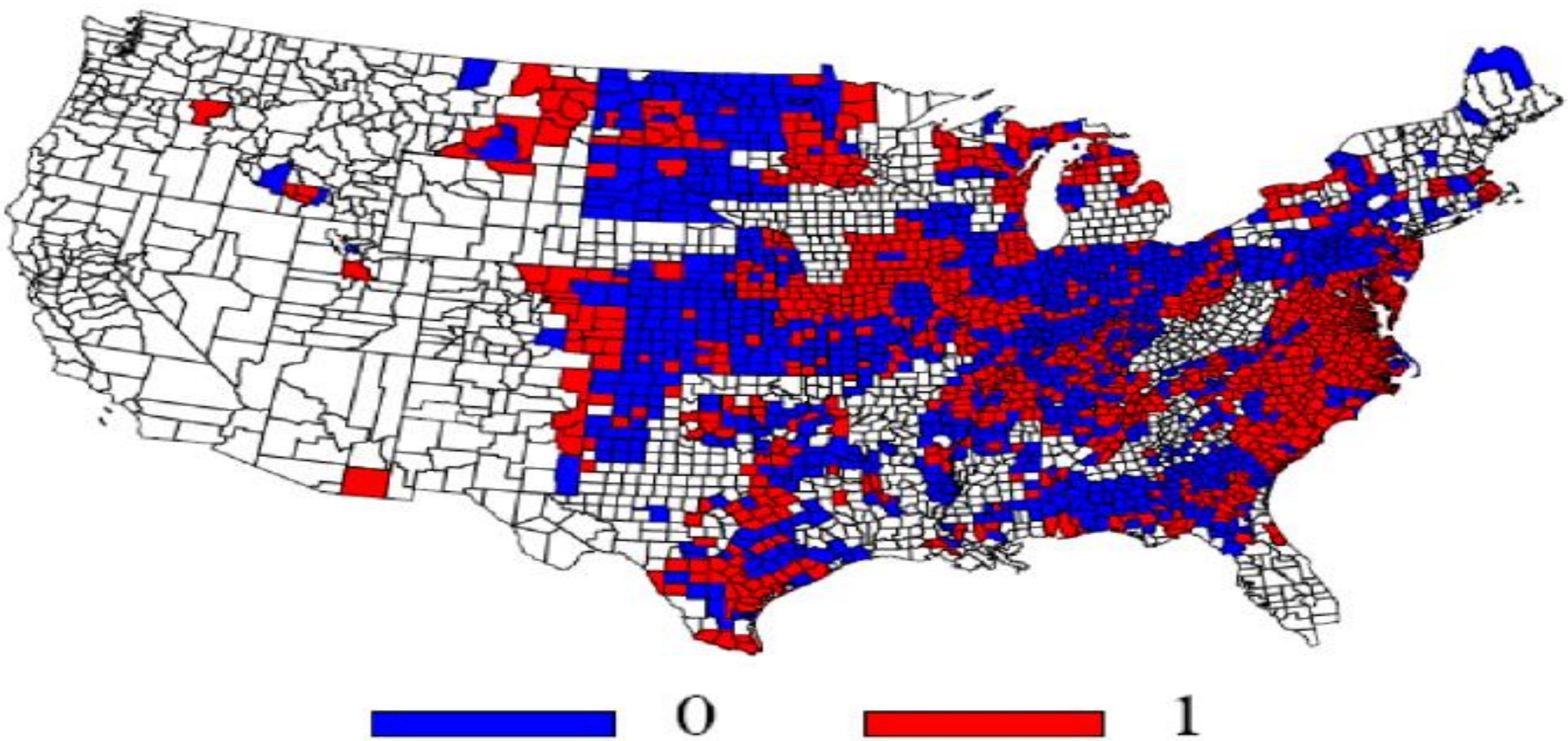

Figure 4. Map of the Difference between the Unweighted Average Loss Cost and the Weather Weighted Loss Costs for Corn [Note: negative difference (e.g., weather weighted < unweighted) is in blue (0) and positive difference (e.g., weather weighted > unweighted) is in red (1).]

Conclusions

Implications

Conclusions

- Idea is to utilize longer time-series information about weather to augment shorter historical county loss cost data used for rating and better account for weather probabilities
- This study shows that a weather weighting approach can indeed be feasibly implemented within the context of the US crop insurance program

Conclusions

- "Hidden" weather probability information, not embedded in shorter historical loss cost data, can now be utilized from long-term weather data
- Allows for better characterization of county level risk and consequently reduce asymmetric information problems

Questions/Concerns?

THANK YOU!