Risky Business:
Using Machine Learning to Predict
Insurance Premium Rates for Specialty Crops

Jonathon Siegle, The Ohio State University, siegle.14@osu.edu
Joseph Cooper, Office of the Chief Economist USDA, joseph.cooper@usda.gov
H. Allen Klaiber, The Ohio State University, klaiber.16@osu.edu

Selected Poster prepared for presentation at the 2022 Agricultural & Applied Economics Association Annual Meeting, Anaheim, CA; July 31-August 2

The findings and conclusions in this poster are those of the authors and should not be construed to represent any official USDA or U.S. Government determination or policy.

The Ohio State University
College of Food, Agricultural, and Environmental Sciences

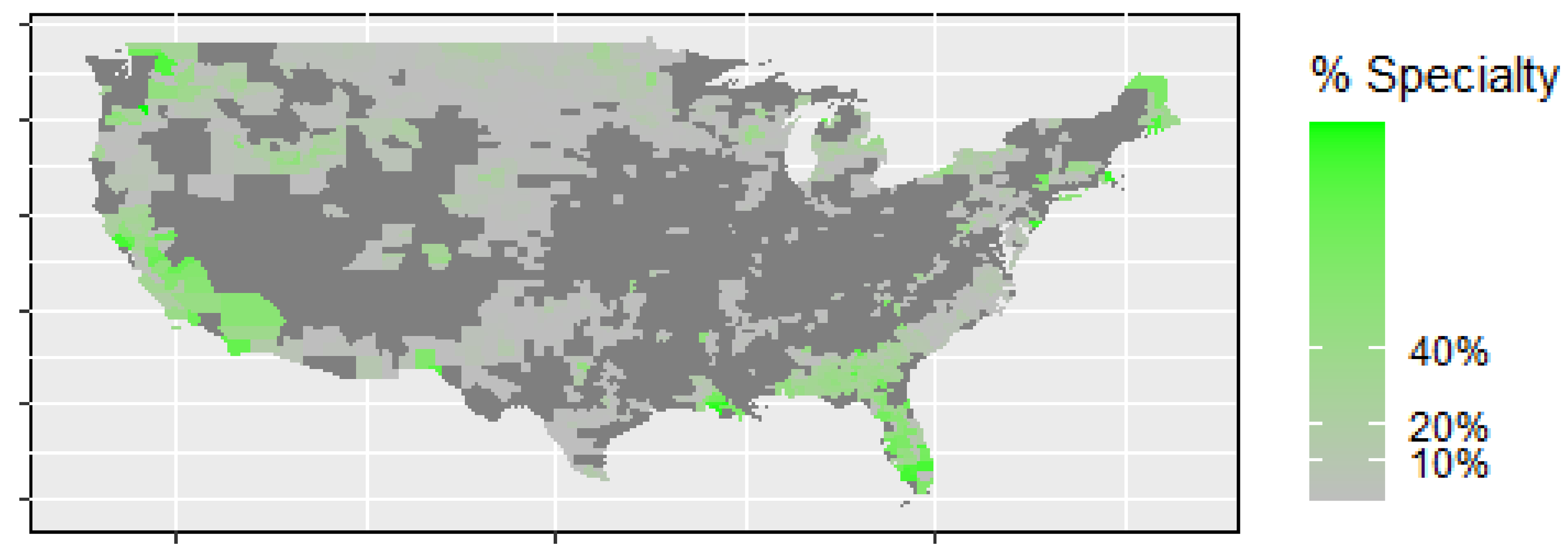
Risky Business:
Using Machine Learning to Predict Insurance Premium Rates for Specialty Crops

Jonathon Siegle[1], Joseph Cooper[2] and H. Allen Klaiber[1]
1 Department of Agricultural, Environmental and Development Economics
2 Office of the Chief Economist, United States Department of Agriculture

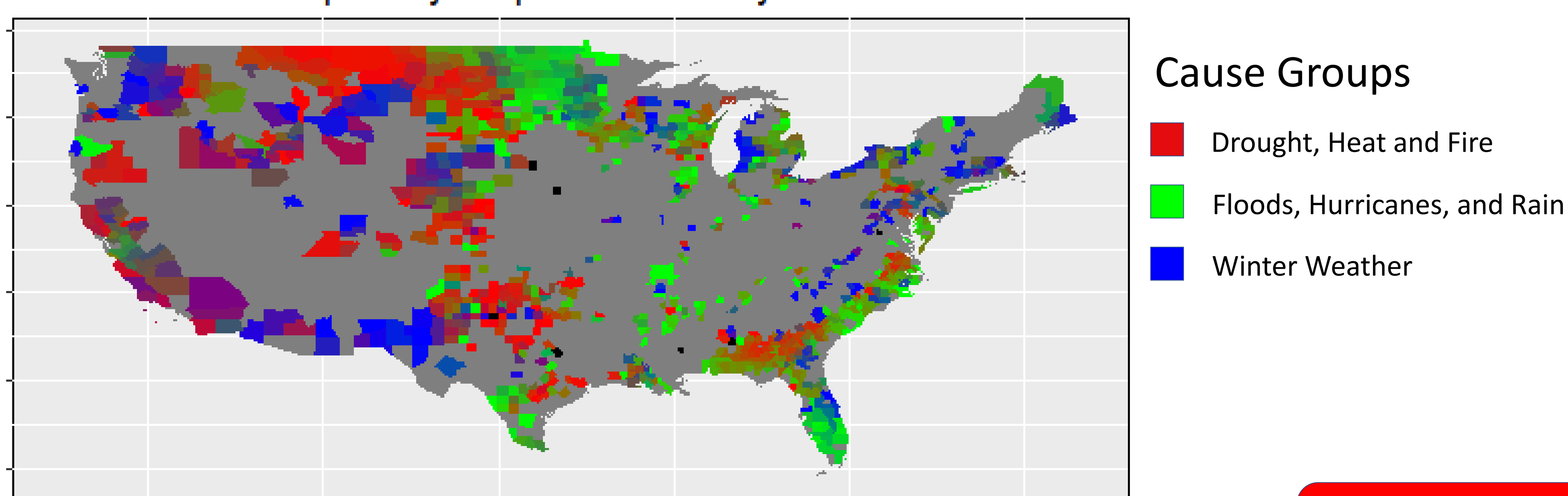
Introduction and Objectives

- Not all specialty crops are covered by the popular Federal Crop Insurance Program (FCIP)
  - Commodities with relatively low acreages and little production histories face delays getting covered by FCIP
  - Expanding the FCIP requires an intensive process to estimate the degree of risk the insurance offering would face
- The relative importance of a risk factor depends on region, commodity, and interaction effects with weather
- We use machine learning to identify key aspects of risk prediction functions without assuming simplistic functional forms *a-priori*
  - Project Base Premium Rates (BPR) for specialty crops not currently covered by FCIP or not covered in certain counties
  - Use derived functions to project future risk to illustrate heterogeneity in risk profiles for specialty crops
  - If successful, these models could serve as a baseline for estimating future USDA offerings

Specialty Crop Percentage of Insured Acres 2000-2021



Relative Share of Specialty Crop Indemnities by Cause



Data

- Create a dataset of BPR's by specialty crop policy with their associated local history of loss
  - Combining USDA data on recorded FCIP indemnities with actuarial records
  - Data is specific to specialty crop 'buckets', such as Citrus Fruit or Root Vegetables
- Describe loss by the frequency and average severity losses over the past decade
  - Use groups of causes, e.g. "Drought" or "Pests/Disease"

Machine Learning: Reducing Dimensionality

- Utilize a LASSO regression algorithm to eliminate unnecessary variables from the many created above
- Lambda, the punishment parameter eliminates irrelevant variables, determined by 10-fold cross validation
- After removing the variables whose coefficients are forced to zero, we have a more parsimonious regression function predicting aggregate risk from local history with risk factors

Types of Specialty Crops
Citrus
Berries
Root Vegetables
…

Actuarial Data
Base Premium Rate
Coverage Level
Location
…

Loss Frequency	Loss Severity	Climate Variables
# of Droughts	% Acres Lost to Drought	Rainfall
# of Freezes	% Acres Lost to Freeze	Temperature
…	…	…
Interactions		

LASSO

$$\sum_{i=1}^{n}\Big(y_i - \sum_{j} x_{ij}\beta_j\Big)^2 + \lambda\sum_{j=1}^{p}|\beta_j|$$

Tibshirani R. 1996. Regression shrinkage and selection via the lasso. J. R. Stat. Soc. B 58:267–88

Preserved Drivers of Base Premium Rates

Citrus BPR	# of Hurricanes	% Acres Lost to Wind	…

Berries BPR	% Lost to Pests	# of Pests Outbreaks	…

Root Vegetable BPR	% Acres Lost to Winter Weather	% Acres Lost to Flood	…

Next Steps:

- Evaluate model fit
  - Can training sets precisely estimate BPR out of sample?
    - Across crops, across space, or both
- Model Extensions
  - Include PRISM weather data, FEMA disaster declarations, etc.
  - Expand to additional loss causes and explore grouping
- Contrast the accuracy and precision of our LASSO regression functions against simple, 'rule of thumb' methods