Learning from Climate Big Data: the Case of Climate Impacts on US Agriculture.

Emanuele Massetti

Georgia Institute of Technology

Rimini Conference on the Economics of Climate Change
April 29, 2017

Outline

Outline

Introduction

- The econometric literature on climate change impacts has used very simplistic characterization of climate.
  - Mostly temperature and precipitations.
  - Very limited work on extreme events.

- Partial representation of climate impacts.
- Possibility of omitted-variables as climate variable are correlated.
  - In both cross-section and panel methods with fixed effects.

Introduction

- The econometric literature on climate change impacts has used very simplistic characterization of climate.
  - Mostly temperature and precipitations.
  - Very limited work on extreme events.

- Partial representation of climate impacts.
- Possibility of omitted-variables as climate variable are correlated.
  - In both cross-section and panel methods with fixed effects.

Introduction

- The econometric literature on climate change impacts has used very simplistic characterization of climate.
  - Mostly temperature and precipitations.
  - Very limited work on extreme events.

- Partial representation of climate impacts.
- Possibility of omitted-variables as climate variable are correlated.
  - In both cross-section and panel methods with fixed effects.

Examples

- Heat waves are not random.
- Omission of humidity, wind and other variables biases temperature coefficients (Zhang et al., 2017).

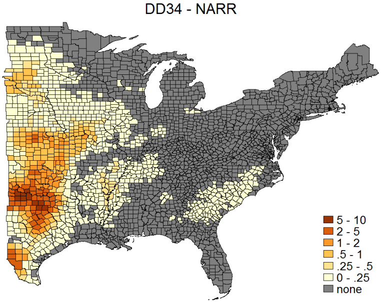


Figure 1: Extreme Heat Days

Characterizing Climate

- In previous work I have used observational data on extreme weather events
  - No significant effect on agricultural land values.
  - Significant effects on crop yields.
- In this paper I use a large set of raw climate variables to study agricultural land values.
- Goals
  - Better characterization of climate-agricultural productivity relationship.
  - Exploratory analysis of "big data" methods to climate change research.

Characterizing Climate

- In previous work I have used observational data on extreme weather events
  - No significant effect on agricultural land values.
  - Significant effects on crop yields.

- In this paper I use a large set of raw climate variables to study agricultural land values.
- Goals
  - Better characterization of climate-agricultural productivity relationship.
  - Exploratory analysis of "big data" methods to climate change research.

Outline

OLS Regression

- Standard pooled panel Ricardian regression
- Log of land value per hectare in county i at time t regressed of climate and other control variables:

$$y_{it} = \beta_0 + \boldsymbol{\beta}_C \mathbf{C}_i + \boldsymbol{\beta}_G \mathbf{G}_i + \boldsymbol{\beta}_Z \mathbf{Z}_{it} + \sum_{t=1}^{T} d_t YEAR + \sum_{s=1}^{S} d_s STATE + u_{it}$$

- $\mathbf{C}_i$: vector of climate variables;
- $\mathbf{G}_i$: vector of geographic and soil characteristics;
- $\mathbf{Z}_{it}$: vector of time-varying socio-economic variables;
- d_t: time dummies;
- d_s: state dummies.

OLS with many variables

- If model is well-specified
  - OLS estimates have low bias
  - If $n >> k$ OLS also has low variance.
- As k increases, OLS regression leads to *overfitting*, with high variance and poor out-of-sample accuracy.
  - A small change in the data used for the regression leads to a large change in the coefficients.
- If $k > n$ there is not a unique set of coefficient: the variance is infinite.

Shrinkage Methods

- Some methods allow to *constraint* or *shrink* the estimated coefficients with little increase in bias and large reductions in variance.
- Some variable are irrelevant: *variable selection*.

- Subset selection
  - Select a subset of the p predictors, then use LS.
- Shrinkage
  - Use all p regressors, but irrelevant regressors are shrunken towards zero, or to zero (variable selection).
- Dimension reduction
  - Project p predictors into a M-dimensional subspace, where $M < p$.

Shrinkage Methods

- Some methods allow to *constraint* or *shrink* the estimated coefficients with little increase in bias and large reductions in variance.
- Some variable are irrelevant: *variable selection*.

- Subset selection
  - Select a subset of the p predictors, then use LS.
- Shrinkage
  - Use all p regressors, but irrelevant regressors are shrunken towards zero, or to zero (variable selection).
- Dimension reduction
  - Project p predictors into a M-dimensional subspace, where $M < p$.

Shrinkage Methods

- Some methods allow to *constraint* or *shrink* the estimated coefficients with little increase in bias and large reductions in variance.
- Some variable are irrelevant: *variable selection*.

- Subset selection
  - Select a subset of the p predictors, then use LS.
- Shrinkage
  - Use all p regressors, but irrelevant regressors are shrunken towards zero, or to zero (variable selection).
- Dimension reduction
  - Project p predictors into a M-dimensional subspace, where $M < p$.

The Lasso

- Shrinkage and variable selection.
- The lasso coefficients $\widehat{\beta}_{\lambda}^{L}$ minimizes the quantity

$$\sum_{i=1}^{N}\left(y_i - \beta_0 - \sum_{j=1}^{k}\beta_j x_{ij}\right)^2 + \lambda\sum_{j=1}^{k}|\beta_j| = \mathrm{RSS} + \lambda\sum_{j=1}^{k}|\beta_j|\,.$$

- Intuition:
  - Minimize RSS given constraint on coefficients.
  - Variables that contribute little or nothing to explaining the dependent variable are dropped.
- Relationship with LS
  - With $\lambda = 0$: LS
  - As λ increases, the model starts shrinking coefficient: variance declines while bias increases.
  - If λ is sufficiently large some coefficients are set to zero.

The Lasso

- Shrinkage and variable selection.
- The lasso coefficients $\widehat{\beta}_\lambda^L$ minimizes the quantity

$$\sum_{i=1}^{N}\left(y_i - \beta_0 - \sum_{j=1}^{k}\beta_j x_{ij}\right)^2 + \lambda\sum_{j=1}^{k}|\beta_j| = \text{RSS} + \lambda\sum_{j=1}^{k}|\beta_j|\,.$$

- Intuition:
  - Minimize RSS given constraint on coefficients.
  - Variables that contribute little or nothing to explaining the dependent variable are dropped.
- Relationship with LS
  - With $\lambda = 0$: LS
  - As λ increases, the model starts shrinking coefficient: variance declines while bias increases.
  - If λ is sufficiently large some coefficients are set to zero.

The Lasso

- Shrinkage and variable selection.
- The lasso coefficients $\widehat{\beta}_{\lambda}^{L}$ minimizes the quantity

$$\sum_{i=1}^{N}\left(y_i-\beta_0-\sum_{j=1}^{k}\beta_j x_{ij}\right)^2+\lambda\sum_{j=1}^{k}|\beta_j| = \text{RSS}+\lambda\sum_{j=1}^{k}|\beta_j|\,.$$

- Intuition:
  - Minimize RSS given constraint on coefficients.
  - Variables that contribute little or nothing to explaining the dependent variable are dropped.
- Relationship with LS
  - With $\lambda = 0$: LS
  - As λ increases, the model starts shrinking coefficient: variance declines while bias increases.
  - If λ is sufficiently large some coefficients are set to zero.

Outline

Agricultural Data

- Agricultural data from US Census of Agriculture
- Socio-economic data from US Census Bureau and other sources
  - As in Massetti, Mendelsohn and Chonabayashi (2016).
- Climate data from North American Regional Reanalysis
  - 1979-2011 reanalysis data
  - 3-hour time step
  - 32 x 32 Km grid (average over counties)

Climate Variables

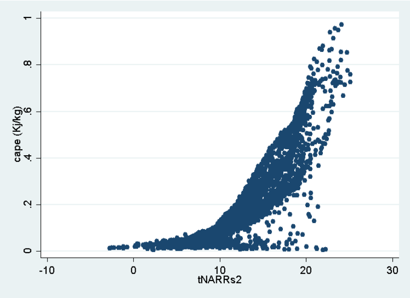
- Total Precipitation (kg/m^2)
- Convective Available Potential Energy (J/kg)
- Categorical Freezing Rain ([yes=1, no=0])
- Categorical Snow ([yes=1, no=0])
- Downward Longwave Radiation Flux (W/m^2)
- Downward Shortwave Radiation Flux (W/m^2)
- High Level Cloud Cover (%)
- Storm Relative Helicity (m^2/s^2)
- Low Level Cloud Cover (%)
- Mid Level Cloud Cover (%)
- Mean Sea Level Pressure (ETA model) (Pa)
- Precipitation Rate ($kg/m^2/s$)
- Surface Pressure (Pa)
- Tropopause Pressure (Pa)
- Pressure Reduced to MSL (Pa)
- Relative Humidity (%)
- Snow Depth (m)
- Snow Cover (%)
- Soil Moisture Content (kg/m^2)
- Specific Humidity (kg/kg)
- 2-m Temperature ($^\circ C$)
- Surface Temperature ($^\circ C$)
- Upward Longwave Radiation Flux (W/m^2)
- U-component of Storm Motion (m/s)
- Upward Shortwave Radiation Flux (W/m^2)
- Vertical Speed Shear (1/s)

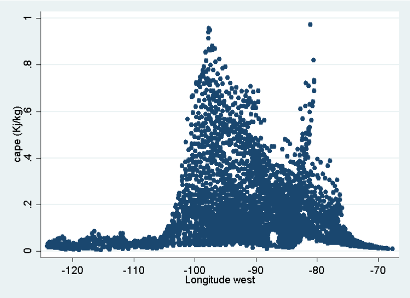
Correlations

	Y	Temp	Precip
Temperature	0.016		
Precipitations	0.547	0.4	
Convective Available Potential Energy	0.003	0.841	0.471
Categorical Freezing Rain	0.314	-0.481	0.098
Categorical Snow	0.004	-0.767	-0.28
Surface Pressure	0.431	0.465	0.628
Tropopause Pressure	-0.019	-0.948	-0.41
Downward Longwave Radiation Flux	0.244	0.892	0.648
Storm Relative Helicity	-0.198	-0.321	-0.213
Pressure Reduced to MSL	0.573	0.275	0.787
Relative Humidity	0.597	0.002	0.734
Snow Depth	-0.076	-0.725	-0.357
Snow Cover	-0.08	-0.74	-0.437
Upward Longwave Radiation Flux	-0.068	0.991	0.302
U-component of Storm Motion	0.258	-0.722	-0.039
Vertical Speed Shear	0.534	-0.098	0.443

Climate Variables - CAPE

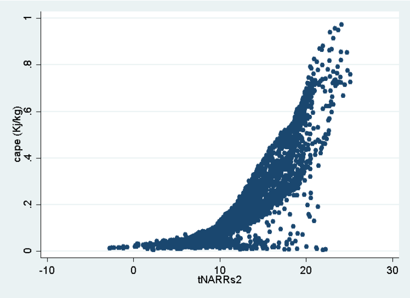
- **CAPE** - Convective Available Potential Energy
  - An indicator of atmospheric instability, which makes it very valuable in predicting severe weather.
  - Extreme CAPE can result in explosive thunderstorm development.

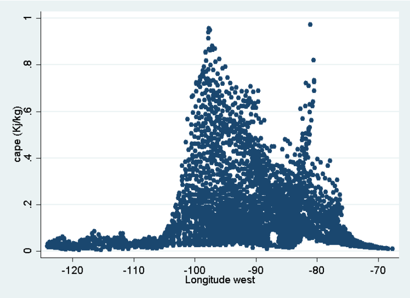




Climate Variables - CAPE

- **CAPE** - Convective Available Potential Energy
  - An indicator of atmospheric instability, which makes it very valuable in predicting severe weather.
  - Extreme CAPE can result in explosive thunderstorm development.





Storm Relative Elicity - SRH

- **SRH** - Storm Relative Elicity.
  - A measure of the potential for cyclonic updraft rotation in right-moving supercells.
  - More than likely become a supercell and possibly spawn one or more tornadoes.
  - There is no clear threshold value for SRH when forecasting supercells.

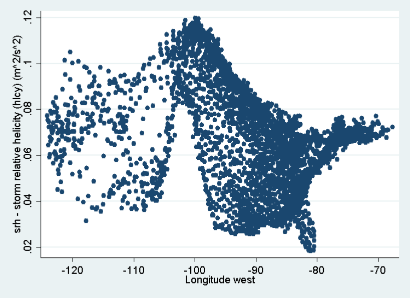
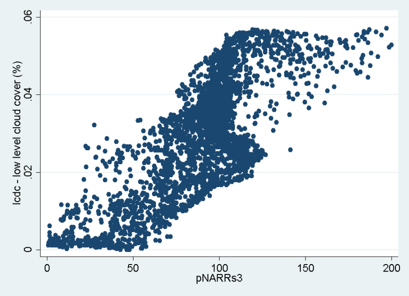
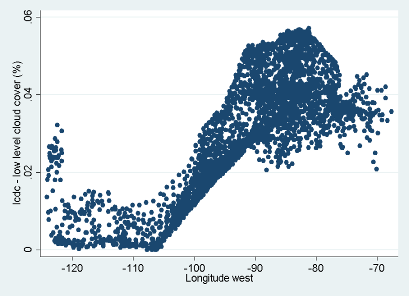


Figure 2: SRH in Spring and Longitude.

Cloud Cover - Low, Middle and High Elevation

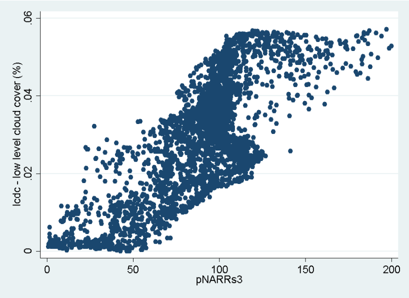
- **Cloud Cover** - Low, Middle and High Elevation.
  - Correlated with precipitations (+) and temperature (-).
  - Very strong regional patterns.

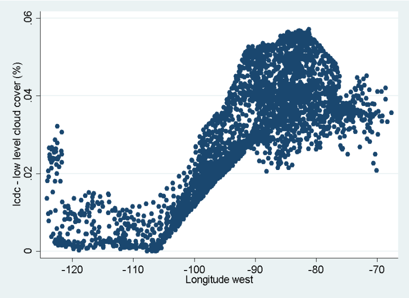




Cloud Cover - Low, Middle and High Elevation

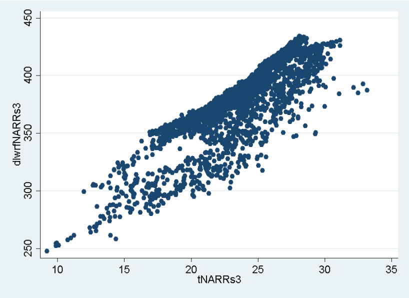
- **Cloud Cover** - Low, Middle and High Elevation.
  - Correlated with precipitations (+) and temperature (-).
  - Very strong regional patterns.

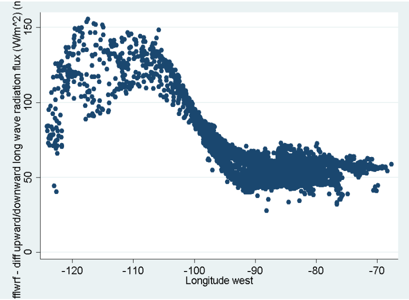




Radiative Flux

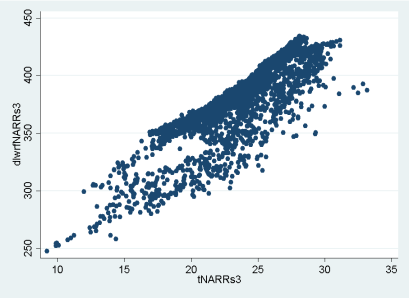
- **Radiative Flux**:
  - Short wave: diffuse reflection of incident shortwave radiation by the underlying surface.
  - Long wave (upward and downward): explains temperature inversion and fog formation (enters independently and as difference).

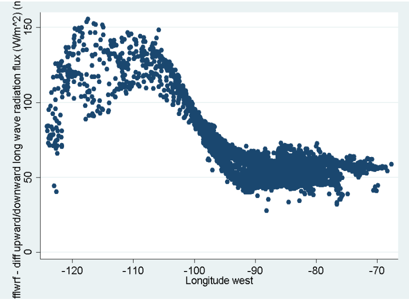




Radiative Flux

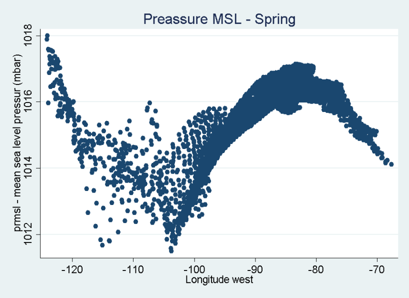
- **Radiative Flux**:
  - Short wave: diffuse reflection of incident shortwave radiation by the underlying surface.
  - Long wave (upward and downward): explains temperature inversion and fog formation (enters independently and as difference).

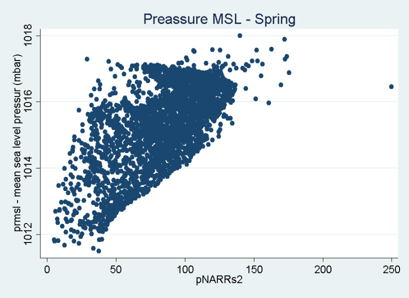




Mean Sea Level Pressure

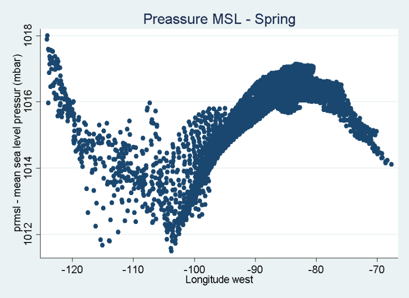
- The mean sea level pressure (MSLP) is the average atmospheric pressure at sea level.
- This is the atmospheric pressure normally given in weather reports.
- Pressure systems cause weather experienced locally.
  - Low-pressure systems are associated with clouds and precipitation that minimize temperature changes through the day.
  - High-pressure systems normally associated with dry weather and mostly clear skies with larger diurnal temperature changes.

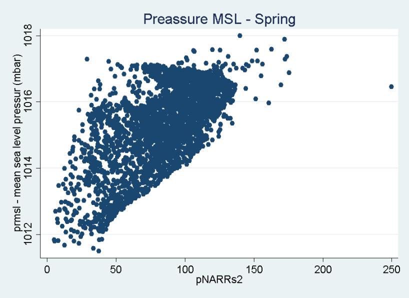




Mean Sea Level Pressure

- The mean sea level pressure (MSLP) is the average atmospheric pressure at sea level.
- This is the atmospheric pressure normally given in weather reports.
- Pressure systems cause weather experienced locally.
  - Low-pressure systems are associated with clouds and precipitation that minimize temperature changes through the day.
  - High-pressure systems normally associated with dry weather and mostly clear skies with larger diurnal temperature changes.





Outline

- Introduction
- Methods
- Data
- Preliminary Results
- Conclusions

Importance of Additional Climate Variables in OLS

	Standard Model		Climate Enhanced Model	
	Marginal	95% CI	Marginal	95% CI
Temperature (°C)				
Winter	-0.216	[-0.272 , -0.161]	0.104	[-0.01 , 0.218]
Spring	0.125	[0.068 , 0.182]	-0.097	[-0.26 , 0.065]
Summer	-0.307	[-0.356 , -0.259]	-0.026	[-0.22 , 0.168]
Fall	0.347	[0.27 , 0.424]	0.083	[-0.099 , 0.265]
Annual	-0.052	[-0.085 , -0.019]	0.063	[-0.01 , 0.137]
Precipitation (cm)				
Winter	0.036	[0.01 , 0.063]	0.050	[0.009 , 0.09]
Spring	0.052	[0.019 , 0.084]	0.066	[0.023 , 0.109]
Summer	-0.047	[-0.069 , -0.025]	-0.048	[-0.077 , -0.018]
Fall	-0.043	[-0.071 , -0.016]	-0.075	[-0.125 , -0.024]
Annual	-0.002	[-0.023 , 0.018]	-0.006	[-0.044 , 0.031]

Notes: Marginal effects at average temperature and precipitation east of the 100th meridian

OLS vs Lasso: Coefficients

	OLS			Lasso
	Coef	95% CI		Coef
T win	-0.034	-0.126	0.058	-0.011
T win sq	0.003	0.000	0.007	0.003
T spr	0.223	0.017	0.429	0.133
T spr sq	-0.009	-0.014	-0.004	-0.008
T sum	-0.128	-0.442	0.186	
T sum sq	0.001	-0.004	0.007	
T aut	0.219	-0.037	0.474	0.134
T aut sq	-0.002	-0.010	0.007	
P win	-0.022	-0.090	0.046	
P win sq	0.005	0.001	0.008	0.003
P spr	0.229	0.106	0.352	0.192
P spr sq	-0.008	-0.013	-0.002	-0.006
P sum	-0.081	-0.162	-0.001	-0.074
P sum sq	0.002	-0.001	0.005	0.002
P aut	-0.053	-0.168	0.062	-0.029
P aut sq	0.000	-0.006	0.006	-0.001

	OLS	Lasso
Temperature Marginal (°C)		
Winter	-0.023	-0.001
Spring	-0.023	-0.084
Summer	-0.065	--
Fall	0.167	0.134
Annual	0.057	0.048
Precipitation Marginal (cm)		
Winter	0.041	0.048
Spring	0.086	0.077
Summer	-0.044	-0.043
Fall	-0.094	-0.220
Annual	-0.012	-0.139

OLS vs Lasso: Out-of-sample Forecasting Accuracy

- 120 Random samples of 50% of counties to train the model
- Prediction on the remaining 50%
- Out-of-sample RMSE

Model	Mean	St. Dev.	Min	Max
OLS	0.275	0.024	0.2534	0.372
Lasso	0.268	0.014	0.2531	0.339

Outline

Conclusions

- Omitted climate variables correlated with included climate variables and land values.
- Complex modeling choices, many variables, interactions.
- Methods for selection of variables.
- Preliminary results suggest
  - Temperature and precipitation coefficients may be biased by omitted climate variables.
  - Lasso coefficients different from OLS coefficients.
  - Lasso has lower out-of-sample forecasting RMSE than OLS.