Crop Production and Climate Change: The Importance of Temperature Variability

David Wigglesworth

University of Pennsylvania

2019-10-19

Introduction and Motivation

- **Introduction and Motivation**
- Relevant Literature
- Data
- Causal Models
- Forecasting Models
- Discussion

Climate Change and Crop Production

- Climate change is the most pressing issue of our time.
- Agriculture is uniquely affected.
- Consequences for food security, migration, international, trade, and political stability.

Motivation: Limitations of Mean Temperature

- T_{mean}: the preeminent measure of climate change.

$$T_{mean} = \frac{T_{min} + T_{max}}{2}$$

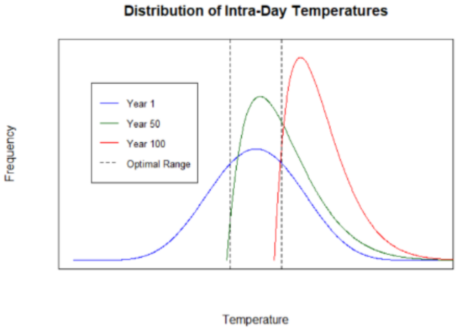
- Limitations
  - Climate change is a multivariate shift in the DGP of the environment (Hsiang and Coop, 2018).
  - Crops grow healthily within a range of temperatures (Schlencker and Roberts, 2008).

Motivation: Importance of Temperature Variability

- Sole focus on T_{mean} obscures other information.
- Measures of variability may help:
  - $T_{range} = T_{max} - T_{min}$. Increasing.
  - T_{min} itself. Increasing.
  - T_{max} itself. Increasing slowly or stagnant.

Putting The Trends Together

- T_{mean} is increasing.
- T_{range} is decreasing.
- T_{min} is increasing faster than T_{max} is increasing.



Relevant Literature

Two General Approaches

- Causal Effects Approach
  - Often looking to the past.
  - Best estimate of $\frac{\partial Y}{\partial X}$ or $E(Y|X_{treat}) - E(Y|X_{control})$.
- Forecasting Approach
  - Often looking to the future.
  - Minimum-error out-of-sample approximation of $Y = f(X, \varepsilon)$.
- How is temperature variabilty important?

Causal Effects Literature

- Popular methods:
  - Explanatory regressions.
  - Panel regressions with direct inference.
  - Simulation methods for “what-if” scenarios.
- Results:
  - Generally mixed, becoming more negative over time.
- Presence of temperature variability?
  - Few consider T_{min} and T_{max}.
  - Even fewer consider T_{range}.

Forecasting Literature

- Popular methods:
  - Extrapolating marginal effects.
  - Riccardian hedonic models.
  - Box-Jenkins Type Models: ARIMA, ARDL, VAR.
  - Machine learning methods.
- Results:
  - Optimistic going into the early 2000s.
  - Pessimistic going into 2050 and beyond.
- Presence of temperature variability?
  - Few papers have considered T_{min} and T_{max}.

Contributing to the Literature

- Is temperature variability important?
  - Consequences for future models.
  - Changes the "story."
- Make analysis general to avoid to spurious conclusions.
  - Standard models.
  - Aggregate production.

Data

- Introduction and Motivation
- Relevant Literature
- **Data**
- Causal Models
- Forecasting Models
- Discussion

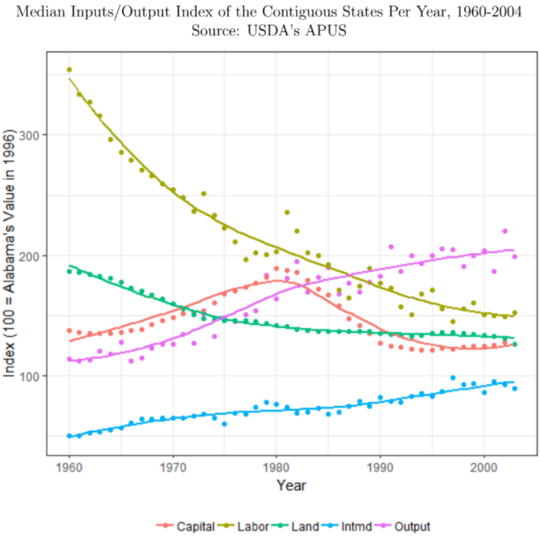
“Agricultural Productivity in the US” (APUS)

- Want to find aggregate crop production data.
- United States Department of Agriculture
  - “Agricultural Productivity in the US” Data Product.
  - Panel of 50 states within 1960-2004.
  - State-level aggregate crop production.
  - State-level input data.

APUS Metrics

- **Crop Output**: Total number of crop bushels sold and added to inventories.
- **Capital**: The number of appreciable assets via the perpetual inventory method from balance sheet data plus inventories derived as implicit quantities from balance sheet data weighted by rental prices.
- **Labor**: The number of labor hours weighted by the real wage constructed by labor accounts.
- **Land**: The ratio of the land of farms to an intertemporal price index constructed by hedonic regressions intended to reflect the value of land.
- **Intermediate Goods**: Aggregates the use of seeds, energy, and chemicals weighted by their implicit prices.

APUS Metrics Over Time



PRISM Weather Data

- Want to find corresponding weather data.
- The Parametric-elevation Regression on Independent Slopes Model (PRISM) at Oregon State University.
  - Uses historical data from weather station to interpolate conditions in a given 4 mile $\times$ 4 mile block (within the continental US) at a given time.
  - Objects of interest T_{mean}, T_{range}, T_{min}, T_{max}, and precipitation Pcp.

Constructing A Panel

- Want to combine two datasets into one.
- Intersection of the datasets:
  - 48 contiguous states.
  - 45 years (1960-2004).
- **Input Data:** capital, labor, land, intermediate goods.
- **Weather Data:** state-year-represented temperature and precipitation.

Causal Models

- Introduction and Motivation
- Relevant Literature
- Data
- **Causal Models**
- Forecasting Models
- Discussion

Strategy

- Motivated by Lobell et al. (2011).
- Estimate Cobb-Douglas production function for state-level aggregate crop production with inputs I, weather W, and spatio-temporal effects X:

$$Y_{it} = I_{it}^{\alpha} \cdot W_{it}^{\beta} \cdot e^{\gamma X_{it}} \cdot e^{\varepsilon_{it}}$$

- This fitted values estimate the conditional expectation function of crop production in the world of climate change: $E(Y_{it}|I_{it}, W_{it}, X_{it})$.
- Substitute detrended weather data $\tilde{W}$ into the production function to obtain estimate of $E(Y_{it}|I_{it}, \tilde{W}_{it}, X_{it})$.

$$\Delta_{cc} Y_{it} = E(Y_{it}|I_{it}, W_{it}, X_{it}) - E(Y_{it}|I_{it}, \tilde{W}_{it}, X_{it})$$

- Significance via bootstrap.

Estimation

- Estimate fixed effects model via OLS.

$$Y_{it} = I_{it}^{\alpha} \cdot W_{it}^{\beta} \cdot e^{\gamma X_{it}} \cdot e^{\varepsilon_{it}}$$

$$log(Y_{it}) = \alpha \cdot log(I_{it}) + \beta \cdot log(W_{it}) + \gamma \cdot X_{it} + \varepsilon_{it}$$

- Consider specifications with and without temperature variability.

$$I_{it} = \{Capital, Labor, Land, Intmd\}$$

$$W^{I} = \{T_{mean}, Pcp\}$$

$$W^{II} = \{T_{mean}, T_{range}, Pcp\}$$

$$W^{III} = \{T_{min}, T_{max}, Pcp\}$$

$$X^{I} = \{Trend, State, Trend \cdot State\}$$

$$X^{II} = \{Year\ FEs, State, Trend \cdot State\}$$

Production Function Table

Estimated Production Functions
Source: Own calculations using data from USDA/PRISM

Y : log(Output)	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
log(Capital)	0.206^{*}	0.073	0.213^{***}	0.074	0.212^{***}	0.075
	(0.034)	(0.107)	(0.034)	(0.107)	(0.034)	(0.107)
log(Labor)	0.064^{***}	0.022	0.063^{*}	0.022	0.064^{*}	0.023
	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)
log(Land)	0.128	0.134	0.127	0.133	0.127	0.134
	(0.072)	(0.081)	(0.072)	(0.081)	(0.071)	(0.081)
log(Intmd)	0.204^{***}	0.278^{***}	0.203^{***}	0.274^{***}	0.203^{***}	0.275^{***}
	(0.057)	(0.063)	(0.057)	(0.064)	(0.058)	(0.064)
$\log(T_{mean})$	-0.756^{***}	-0.785^{***}	-0.682^{***}	-0.722^{***}		
	(0.160)	(0.184)	(0.165)	(0.192)		
$\log(Pcp)$	-0.064^{***}	0.052^{**}	0.036	0.032	0.036	0.034
	(0.018)	(0.018)	(0.020)	(0.020)	(0.020)	(0.020)
$\log(T_{range})$			-0.246^{**}	-0.195^{*}		
			(0.088)	(0.094)		
$\log(T_{min})$					0.225	0.049
					(0.195)	(0.207)
$\log(T_{max})$					-1.163^{***}	-0.967^{***}
					(0.249)	(0.250)
Time Control	Trend	Year FEs	Trend	Year FEs	Trend	Year FEs
R^2	0.962	0.994	0.993	0.995	0.993	0.995
Adj. R^2	0.961	0.994	0.993	0.994	0.993	0.994
Durbin-Watson	1.913	1.954	1.949	1.949	1.947	1.947
BIC	-2679.601	-2765.603	-2685.462	-2768.729	-2585.307	-2767.420

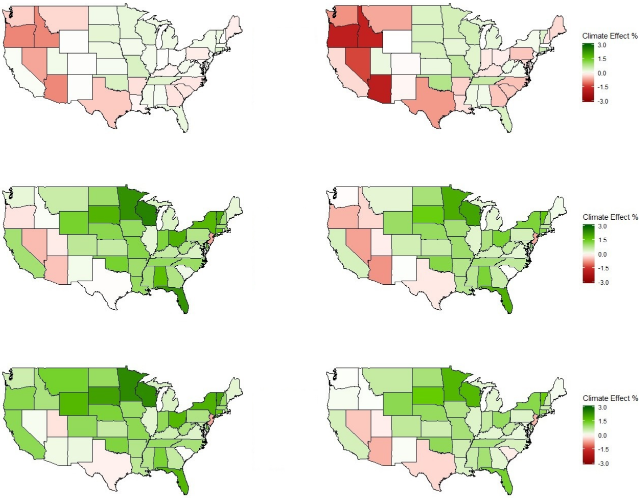
Stata State-Clustered Standard Errors $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

National-Level Impact

Model	T_1	T_2	Time	National Impact
Model 1	T_{mean}	-	Trend	$+0.04\%$
Model 2	T_{mean}	-	Year FE	-0.05%
Model 3	T_{mean}	T_{range}	Trend	$+1.01\%^*$
Model 4	T_{mean}	T_{range}	Year FE	$+0.74\%^*$
Model 5	T_{min}	T_{max}	Trend	$+1.07\%^*$
Model 6	T_{min}	T_{max}	Year FE	$+0.66\%$

* significant at the $\alpha = 0.05$ level

State-Level Impacts



Row 1 - Models 1-2 ; Row 2 - Models 3-4 ; Row 3 - Models 5-6

Forecasting Models

Standard Specification

- Outline a standard forecasting model (Lobell and Burke, 2010).
- Log-Quadratic model.

$$log(Y_{it}) = \alpha_i + \sum_{k=1}^{2} \beta_k \cdot T_{mean,it}^k + \sum_{k=1}^{2} \delta_k \cdot Pcp_{it}^k + \varepsilon_{it}$$

- Once equation is estimated, substitute in some forecasted $T_{mean,it}$ to obtain forecasted Y_{it}.

Specifications with Variability

- Incorporate T_{range} and T_{min}/T_{max}:

$$log(Y_{it}) = \alpha_i + \sum_{k=1}^{2} \beta_k \cdot T_{mean,it}^k + \sum_{k=1}^{2} \gamma_k \cdot T_{range,it}^k + \sum_{k=1}^{2} \delta_k Pcp_{it}^k + \varepsilon_{it}$$

$$log(Y_{it}) = \alpha_i + \sum_{k=1}^{2} \beta_k \cdot T_{min,it}^k + \sum_{k=1}^{2} \gamma_k \cdot T_{max,it}^k + \sum_{k=1}^{2} \delta_k Pcp_{it}^k + \varepsilon_{it}$$

.

- Selection by BIC.

Comparisons

- Comparing T_{mean}-only/variability model performance.
  - F-Test.
  - BIC.
- Comparing forecasts made by these models.
  - Diebold-Mariano Test.

Forecasting Models Table

Estimated Forecasting Models
Source: Own calculations using data from USDA/PRISM

Y : log(Output)	Model 7	Model 8	Model 9	Model 10[1]
Pcp	0.019^{***}	-0.015^{**}	-0.013^{*}	-0.006^{***}
	(0.006)	(0.006)	(0.006)	(0.001)
Pcp^2	-0.000^{*}	0.000	0.000	
	(0.000)	(0.000)	(0.000)	
T_{mean}	0.018	0.077		
	(0.053)	(0.046)		
T^2_{mean}	−0.000	−0.000		
	(0.000)	(0.000)		
T_{range}		−0.013		
		(0.040)		
T^2_{range}		-0.002^{*}		-0.002^{***}
		(0.001)		(0.000)
T_{min}			0.181^{***}	0.022^{***}
			(0.045)	(0.005)
T^2_{min}			−0.001	
			(0.000)	
T_{max}			-0.111^{*}	
			(0.056)	
T^2_{max}			0.000	
			(0.000)	
R^2	0.962	0.968	0.968	0.968
Adj. R^2	0.961	0.967	0.967	0.967
RMSE	0.255	0.232	0.232	0.232
Durbin-Watson	0.751	1.052	1.063	1.676
BIC	625.983	240.290	246.604	224.609

Newey-West HAC Standard Errors $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

1. Model chosen by BIC

F-Tests and DM-Tests

F-Tests Comparing Forecasting Panel Models
Source: Own calculations using data from USDA/PRISM

T_{mean} Model	Variability Models	F-Test
	Model 8	$p < 2.2 \times 10^{-16}$
Model 7	Model 9	$p < 2.2 \times 10^{-16}$
	Model 10	$p < 2.2 \times 10^{-16}$

DM Test Comparing Errors of Forecasting Models
Source: Own calculations using data from USDA/PRISM

T_{mean} Errors	T_{mean} MSE	Variability Errors	Variability MSE	DM Test, $h = 1$	DM Test, $h = 10$
		Model 8	0.0537	$p < 0.001$	$p < 0.001$
Model 7	0.0650	Model 9	0.0539	$p < 0.001$	$p < 0.001$
		Model 10	0.0539	$p < 0.001$	$p < 0.001$

Discussion

Shortcomings

- Problems with aggregating crops together.
  - **Advantage:** tacitly adjusts for crop rotation.
  - **Disadvantage:** the number of crops may increase or decrease but volume might change and obscure the real movements.
- Picking weather data to be representative of a state.
- Endogeneity: weather begets crops begets weather.

Importance of Temperature Variability

- Measures of temperature variability change our causal story concerning the relationship between climate change and crop production.
- Measures of temperature variability augment our ability to predict crop production under climate change.
- These two points suggest that temperature variability contains important information that is not always being exploited, but likely should.

Citations

- Hsiang, S., Kopp, R. E. (2018). An economist's guide to climate change science. Journal of Economic Perspectives, 32(4), 3-32.
- Lobell, D. B., Burke, M. B. (2010). On the use of statistical models to predict crop yield responses to climate change. Agricultural and forest meteorology, 150(11), 1443-1452.
- Lobell, D. B., Schlenker, W., Costa-Roberts, J. (2011). Climate trends and global crop production since 1980. Science, 333(6042), 616-620.
- Schlenker, W., Roberts, M. J. (2008). Estimating the impact of climate change on crop yields: The importance of nonlinear temperature effects (No. w13799). National Bureau of Economic Research.