NBER WORKING PAPER SERIES

ESTIMATING THE IMPACT OF CLIMATE CHANGE ON CROP YIELDS: THE IMPORTANCE OF NONLINEAR TEMPERATURE EFFECTS

Wolfram Schlenker
Michael J. Roberts

Working Paper 13799
http://www.nber.org/papers/w13799

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2008

We would like to thank Spencer Banzhaf, Larry Goulder, Jim MacDonald, Mitch Renkow, Bernard Salanie, Kerry Smith, Wally Thurman as well as seminar participants at Arizona State, Dartmouth, Georgia State, Harvard, NC State, Stanford, UC Berkeley, UC Davis, University of Illinois, University of Maryland, University of Nebraska, University of Wyoming for useful comments. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research or the U.S. Department of Agriculture.

Estimating the Impact of Climate Change on Crop Yields: The Importance of Nonlinear Temperature Effects
Wolfram Schlenker and Michael Roberts
NBER Working Paper No. 13799
February 2008
JEL No. C23,Q54

ABSTRACT

The United States produces 41% of the world's corn and 38% of the world's soybeans, so any impact on US crop yields will have implications for world food supply. We pair a panel of county-level crop yields in the US with a fine-scale weather data set that incorporates the whole distribution of temperatures between the minimum and maximum within each day and across all days in the growing season. Yields increase in temperature until about 29C for corn, 30C for soybeans, and 32C for cotton, but temperatures above these thresholds become very harmful. The slope of the decline above the optimum is significantly steeper than the incline below it. The same nonlinear and asymmetric relationship is found whether we consider time series or cross-sectional variation in weather and yields. This suggests limited potential for adaptation within crop species because the latter includes farmers' adaptations to warmer climates and the former does not. Area-weighted average yields given current growing regions are predicted to decrease by 31-43% under the slowest warming scenario and 67-79% under the most rapid warming scenario by the end of the century.

Wolfram Schlenker
Department of Economics
School of International and Public Affairs
Columbia University
420 West 118th Street, MC 3323
New York, NY 10027
and NBER
wolfram.schlenker@columbia.edu

Michael Roberts
Economic Research Service
United States Department of Agriculture
1800 M Street NW
Washington, DC 20036
mroberts@ers.usda.gov

With accumulating evidence that greenhouse gas concentrations are warming the world's climate, research has increasingly focused on estimating the impacts that are likely to occur under different warming scenarios as well as how economies might adapt to a change in climatic conditions. Many impact studies focus on the agricultural sector for several reasons. The first reason is that agricultural production is directly exposed to changes in temperatures and precipitation. The second reason is that agricultural production and consumption still comprise a large share of income in poorer developing economies. And while agriculture comprises a smaller share of GDP in the United States, the U.S. is still the world's largest agricultural producer and exporter of agricultural commodities. The US produces 41% of the world's corn and 38% of the world's soybeans[1], so substantial climate impacts on U.S. agriculture would have broad implications for food supply and prices worldwide. At the same time, there continues to be a debate whether warming will be a net gain or loss for agriculture in the more temperate climates like that in the United States (Mendelsohn et al. 1994, Darwin 1999, Schlenker et al. 2006, Kelly et al. 2005, Timmins 2006, Ashenfelter and Storchmann 2006, Deschenes and Greenstone 2007).

In this paper we develop novel estimates of the link between weather and yields for the three most valuable crops grown in the United States: corn, soybeans, and cotton. Corn and soybeans are the nation's most prevalent crops and are the predominant source of feed grains in cattle, dairy, poultry, and hog production. Cotton is the fourth largest in acres planted, but more valuable on a per-acre basis and more suited to warmer climates than corn and soybeans. Estimating the correct relationship between weather and yields for these major crops is a critical first step before more elaborate models can be used to estimate how crop choices, food supply, and prices might shift in response to climate change. These models will give biased results if the underlying relationship between weather and yields is modeled incorrectly.

In this paper we pair yields for these three crops with a newly constructed fine-scale weather data set resulting in a large panel that spans most U.S. counties from 1950 to 2005. The new weather data includes the length of time a crop is exposed to each 1-degree Celsius temperature interval in each day of the growing season. We estimate these times for the specific locations within each county where crops are grown.

The new fine-scale weather data facilitate estimation of a flexible model in order to identify nonlinearities and breakpoints in the effect of temperature on yield. If the true underlying relationship is nonlinear (e.g., increasing and then decreasing in temperature),

[1] Foreign Agricultural Statistical Service data for 2005/2006: http://www.fas.usda.gov/psdonline/psdHome.aspx

averaging over time or space dilutes effects of extreme outcomes. Yet extreme temperatures are critical for crop yields (Tubiello et al. 2007). Accurate estimation of nonlinear effects is particularly important when considering large non-marginal changes in temperatures now expected with climate change.

We find a robust nonlinear relationship between weather and yields that is consistent across space, time, and crops: plant growth increases approximately linearly in temperature up to a point where additional heat becomes harmful (See Figure 2 below). The nonlinear relationship is starkly asymmetric, with the slope of the decline above the optimal temperature being much steeper than the slope of the incline below the optimal temperature. Our flexible specification is superior in explaining yields and results in different predicted climate change impacts due to the sharp nonlinearity. Despite significant technological progress over our 56-year sample period, we find no evidence that crops have become better at withstanding extreme heat above the optimal temperature. Moreover, warmer southern states exhibit the same threshold as cooler states in the north. This robustness and consistency across time, space, and sources of identification suggests the links are causal and the potential for adaptation is limited.

The nonlinear and asymmetric relationship between temperature and yields is confirmed in an analysis of the futures market. Weekly corn futures prices increase significantly in response to extremely high temperatures, while there is no statistically significant relationship with average temperatures.

The sharply negative effects of temperatures above the critical temperature threshold hold powerful implications for climate change. If climate change shifts the temperature distribution such that a significantly larger portion of it exceeds the threshold, overall impacts are substantial. Indeed, under the latest warming predictions, the high-end of the temperature distribution shifts upward enough so that damaging heat waves are observed more frequently. As a result, yields at the end of the century are predicted to decrease by 43% for corn, 36% for soybeans, and 31% for cotton under a slow warming scenario (B1) and 79%, 74%, and 67%, respectively, under a fast warming scenario (A1FI). These predictions are highly significant and consistent across alternative model specifications. These scenarios, however, assume unaltered crop choices, technologies, and no effects from possible CO_2 fertilization, and so likely overstate true potential climate impacts. The nonlinear relationship between temperatures and yields can consecutively be used in more structural models to estimate crop switching and other farmer responses.

1 Literature Review

Many earlier studies have linked weather and climate to outcomes such as yields, land values, and farm profits. These studies span several disciplines and methods. Agronomic studies focus on yields and emphasize the dynamic physiological process of plant growth and seed formation. This process is understood to be quite complex and dynamic in nature and thus not easily molded into a regression framework. Instead, these studies use a rich theoretical model to simulate yields given daily and sub-daily weather inputs, nutrient applications, and initial soil conditions. In some cases, simulated yields are compared to observed yields with some success. But we are not aware of any study that has tested a simulation model using data besides that used to calibrate it. Current versions of models developed for many crops are maintained by the Decision Support System for Agrotechnology Transfer (http://www.icasa.net/dssat/).

A clear strength of simulation models is the way they incorporate the whole distribution of weather outcomes over the growing season. This differs from regression-based approaches that typically use average weather outcomes or averages from particular months. A weakness of the approach is uncertainty about the physiological process (functional form) and the sheer number of parameters in these dynamic and highly nonlinear models. Some agronomists seem to worry about possible misspecification and omitted variables biases (Sinclair and Seligman 1996, Sinclair and Seligman 2000, Long et al. 2005). These models also take production systems and nutrient applications as exogenous: there is no account for behavioral response on behalf of farmers. Nevertheless, these models are the predominant tool used to evaluate likely effects from climate change on crop yields. Examples include Black and Thompson (1978), Adams et al. (1995), Brown and Rosenberg (1999), Mearns et al. (2001), and Stockle et al. (2003), but there are many others.

Several economic studies use hedonic models to link land values to land characteristics, including climate, using reduced-form linear regression models (e.g., Mendelsohn et al. (1994); Schlenker et al. (2006); Ashenfelter and Storchmann (2006)). One strength of the approach is that, unlike crop simulation models, it can account for the whole agricultural sector rather than a single crop at a time. It can also account for behavioral response or adaptation. Cooler areas are likely to become more like warmer areas, with crops choice, management, and land values changing in accordance with the cross-section of climate.

The overarching concern with the hedonic approach or any other cross sectional study is omitted variables bias. Climate variables (e.g., average temperature) and other critical variables, such as soil types, distance to cities, and irrigation, are all spatially correlated. If critical variables correlated with climate are omitted from the regression model, the climate variables may pick up effects of variables besides climate and lead to biased estimates and predictions. This has been a concern since the early part of the last century and the birth of modern statistics when Ronald Fisher wrote "Studies in Crop Variation I-VI." Indeed, earlier work shows how omission of irrigation critically influences predicted climate impacts (Schlenker et al. 2005).

Most recently Deschenes and Greenstone (2007) (DG, henceforth) use year-to-year weather variation as a source of identification when they link agricultural profits to weather using county fixed effects to capture time-invariant factors like soil quality. DG argue that their measure overstates any adverse impact from climate change because it reflects the short-run response to weather fluctuations and does not allow for long-run adaptation. A problem with this argument is that many important time-varying factors, such as storage, irrigation, and price effects, embody short-run weather responses that are not available in long run.[2] For example, a one time heat wave might be mitigated by applying more groundwater, but it might not be feasible to sustain an increased use of groundwater on a continued basis.

Kelly et al. (2005), another recent study employing panel data, examine county-level profits for Midwestern states in relation to both climate and weather. The authors include historic mean climate and climate variability (standard deviation between years) as well as yearly weather shocks. Since the authors include climate averages (which are constant over time in the cross-section), they cannot include county fixed effects as they would be perfectly collinear with climate. Omitted variables hence remain a possible concern, as do time-varying factors associated with both weather and reported profits.

While our model is simpler than crop simulation models, it shares the feature of incorporating the whole distribution of weather outcomes. And like DG, we consider specifications with county fixed effects that narrow the source of identification to arguably random year-to-year weather variation. However, we do find that omitting county fixed effects does not significantly alter our results. Since we focus on yields rather than profits (which rely on sales in a given year), storage and price effects are not a concern. We also consider specifications

[2] This issue, and others, are considered thoroughly by Fisher et al. (2007). Most notably, DG's profit measure relies on agricultural sales in a given year, yet sales in a given year omit storage between periods. Analogous to the permanent income hypothesis that stresses that yearly consumption is a bad proxy for income, sales in a given year only partially reflect economic profit in the same year if a commodity can be stored. Second, demand for agricultural goods is highly inelastic in the short-run and hence reductions in output might be offset by large price increases, limiting the effect on profits. Third, DG also present one regression using yields instead of profit, but the model relies on average daily temperatures and does not account separately for extreme temperatures which are shown to critically influence yields in this study.

based on the cross-section of average yields, akin to hedonic models where *climate* variation serve as the source of identification rather than weather, again with similar results.

2 Model

Our objective is to discern the effect of weather, particularly heat, on crop yields using a new and rich data set and a novel approach that allows us to estimate nonlinear effects of heat over the growing season. By using the whole distribution of temperature outcomes, which is critical for estimating nonlinear effects, we depart from earlier cross-sectional and panel data studies and share a common thread with the agronomic literature that employs crop simulation models. However, unlike agronomic literature, we incorporate our model into a statistical regression framework. We focus on yields because yields, unlike profits, are linked to the year a plant is grown. Here we describe our model, discuss its assumptions, and consider various sources of temperature variation used for identification.

We postulate that the effect of heat on relative plant growth is cumulative over time and that yield is proportional to total growth. This assumes temperature effects are additively substitutable over time. Specifically, plant growth $g(h)$ depends nonlinearly on heat h and log yield, y_{it}, in county i and year t is

$$y_{it} = \int_{\underline{h}}^{\overline{h}} g(h)\phi_{it}(h)dh + \mathbf{z}_{it}\boldsymbol{\delta} + c_i + \epsilon_{it} \tag{1}$$

where $\phi_{it}(h)$ is the time distribution of heat over the growing season in county i and year t. We fix the growing season to months March through August for corn and soybeans and the months April through October for cotton. Observed temperatures during this time period range between the lower bound $\underline{h}$ and the upper bound $\overline{h}$. Other factors, such as precipitation and technological change, are denoted $\mathbf{z}_{it}$, and c_i is a time-invariant county fixed effect to control for time-invariant heterogeneity, such as soil quality.

While time separability is partially rooted in agronomy,[3] we implicitly validate this assumption by showing a statistically significant relationship between the cumulative distribution of temperatures and yields. We would not observe this if time separability were not appropriate, because random pairing of various temperatures over a season and between years

[3] This assumption underlies the concept of degree days by which many corn varieties are classified, i.e., farmers count the additive number of daily temperatures above a baseline a specific crop variety requires to mature.

would not provide clear identification. In the empirical section we also split the six-month growing season for corn into two three-month intervals and find comparable estimates for both subintervals. That is, temperature effects in the earlier and later parts of the growing season are similar.

A special case of time-separable growth is the concept of growing degree days, typically defined as the sum of truncated degrees between two bounds. For example, Ritchie and NeSmith (1991) suggests bounds of 8°C and 32°C for "beneficial heat". A day of 9°C hence contributes 1 degree day, a day of 10°C contributes 2 degree days, up to a temperature of 32°C, which contributes 24 degree days. All temperatures above 32°C also contribute 24 degree days. Degree days are then summed over the entire season. Temperatures above 34°C are included as a separate variable and speculated to be harmful. These particular bounds have been implemented in a cross-sectional analysis by Schlenker et al. (2006). Thus, growing degree days are the special case of our model where (using the above bounds as an example)

$$g(h) = \begin{cases} 0 & \text{if } h \leq 8 \\ h - 8 & \text{if } 8 < h < 32 \\ 24 & \text{if } 32 \leq h \end{cases}$$

The appropriate bounds for growing degree days are still debated, partly because earlier studies use a limited number of observations from field experiments to identify them. There is also uncertainty about temperature effects above the upper bound. While some speculate that high temperatures are harmful, the critical temperature and severity of damages remain uncertain.

In the data section we explain how we derive the amount of time a plant is exposed to each 1-degree Celsius interval. With these data, we approximate the integral over temperature with

$$y_{it} = \sum_{h=-5}^{49} g(h+0.5)[\Phi_{it}(h+1) - \Phi_{it}(h)] + \mathbf{z}_{it}\boldsymbol{\delta} + c_i + \epsilon_{it} \tag{2}$$

where $\Phi_{it}(h)$ is the cumulative distribution function of heat in county i and year t. We consider two specifications of this model.

First, we approximate $g(h)$ using dummy variables for each three-degree temperature interval.[4] The dummy-variable model effectively regresses yield on season-total time within each temperature interval. Because temperatures rarely exceed 39°C (102 degrees Fahrenheit) we lump all time a plant is exposed to a temperature above 39°C into one category.

[4]We obtain similar results when estimating even more flexible models with dummy variables for each one-degree interval. We report results for three-degree intervals in order to make figures easy to interpret.

Similarly, we lump all times temperatures are below freezing into the interval $[-1;0]$. The existing temperature distribution is displayed in the left column of Figure 1. The model becomes[5]

$$y_{it} = \sum_{j=0,3,6,9,\ldots}^{39} \gamma_j \underbrace{\left[\Phi_{it}(h+3) - \Phi_{it}(h)\right]}_{x_{it,j}} + \mathbf{z}_{it}\boldsymbol{\delta} + c_i + \epsilon_{it}. \tag{3}$$

County fixed effects c_i account for time-invariant factors in the cross-section. The error terms, however, remain spatially correlated within each year. The non-parametric routine by Conley (1999) is used to adjust the variance-covariance matrix for spatial correlation.

Second, we model the function $g(h)$ as a m-th order Chebychev polynomial of the form $g(h) = \sum_{j=1}^{m} \gamma_j T_j(h)$, where $T_j()$ is the $j-th$ order Chebyshev polynomial. Chebyshev polynomials are a relatively parsimonious approximation for the function $g(h)$, assuming it is smooth.

By interchanging the sum we obtain

$$\begin{aligned} y_{it} &= \sum_{h=-1}^{39} \sum_{j=1}^{m} \gamma_j T_j\left(h+0.5\right)\left[\Phi_{it}\left(h+1\right) - \Phi_{it}\left(h\right)\right] + \mathbf{z}_{it}\boldsymbol{\delta} + c_i + \epsilon_{it} \\ &= \sum_{j=1}^{m} \gamma_j \underbrace{\sum_{h=-1}^{39} T_j\left(h+0.5\right)\left[\Phi_{it}\left(h+1\right) - \Phi_{it}\left(h\right)\right]}_{x_{it,j}} + \mathbf{z}_{it}\boldsymbol{\delta} + c_i + \epsilon_{it} \end{aligned} \tag{4}$$

where $x_{ij,t}$ is the exogenous variable obtained by summing the $j-th$ Chebyshev polynomial evaluated at each temperature interval midpoint, multiplied by the time spent in each temperature interval. Successively higher-order polynomials were estimated until the relationship appeared stable.

While equations (3) and (4) specify our main two models, the concept of degree days is a special case where the function $g(h)$ is piecewise linear. In a sensitivity check we therefore estimate a piecewise-linear model, i.e., growth is forced to increase linearly in temperature up to an endogenous threshold and then forced to decrease linearly above the threshold. Since our data is aggregated by 1-degree Celsius intervals, we loop over possible combinations of bounds and pick the ones with the least sum of squared residuals.

[5] The omitted category is the time temperatures are below 0°C.

3 Data

Dependent Variable

Yields for corn, soybeans, and cotton for the years 1950-2005 are reported by the U.S. Department of Agriculture's National Agricultural Statistical Service (USDA-NASS).[6] These yields equal total county-level production divided by acres harvested.[7] We limit the analysis to counties east of the 100 degree meridian for corn and soybeans (as these counties are primarily nonirrigated), but use all counties that report cotton yields.[8]

In a sensitivity check we also examine changes in futures prices. We collect daily closing prices of futures with a delivery date of September for the years 1950-2005 from the Chicago Board of Trade.

Weather Variables

Earlier statistical studies have examined average temperatures over a longer time horizon (e.g., an entire season, month, or day), which can hide extreme events like high temperatures that occur during a fraction of the day. Our fine-scale weather aids identification of these effects which are diluted when weather outcomes are averaged over time or space. Construction of these data is briefly described here and in more detail in Schlenker and Roberts (2006).

The basic steps are as follows. We first develop daily predictions of minimum and maximum temperature on a 2.5x2.5mile grid for the entire United States. We then derive the time a crop is exposed to each 1 degree Celsius interval in each grid cell. These predictions are merged with a satellite scan that allows us to select only those grid cells with cropland. We then aggregate the whole distribution of outcomes for all days in the growing season in each county. Since our study emphasizes nonlinearities, it is important to derive the time each grid cell is exposed to each 1 degree Celsius interval before aggregating to obtain the county-level distribution. This preserves within-county variation in temperatures in our county-level distribution estimates.[9]

[6]We include all reported yields, even though some appear artificially low. These few outliers are very infrequent. If we drop them, the results do not change, but the cutoff point becomes somewhat arbitrary (The outliers have little influence because we use log yield, so they have small errors).

[7]For about 80% of the observations, NASS reports planted acres. As a sensitivity check we derive our own yield measure for these observations by taking total production over total acres planted. The results do not change significantly. Since the area planted is not reported in all areas and years, our analysis focuses on the larger sample of output per acre harvested that is the standard USDA definition of yield.

[8]This gives us 105,981 observations with corn yields, 82,385 observations with soybeans yields, and 31,540 observations with cotton yields.

[9]Thom (1966) develops an alternative method to approximate the distribution of daily temperatures

More specifically, we use the Parameter-elevation Regressions on Independent Slopes Model (PRISM), widely regarded as one of the best geographic interpolation procedures (http://www.ocs.orst.edu/prism/). It accounts for elevation and prevailing winds to predict weather outcomes on 2.5x2.5 mile grid across the contiguous United States. However, the PRISM data are on a *monthly* time scale. We therefore combine the advantages of the PRISM model (good spatial interpolation) with better temporal coverage of individual weather stations (daily instead of monthly values). We do this by pairing each of the 259,287 PRISM grid cells that cover agricultural area in a LandSat satellite scan with the closest seven weather stations having a continuous record of daily observations. We then estimate a separate regression for each grid cell, where the dependent variable is the monthly PRISM grid cell estimate and the explanatory variables are the monthly averages at each of the seven closest weather stations, plus fixed effects for each month. The R-squares are usually in excess of 0.999. The derived relationship between *monthly* PRISM grid cell averages and *monthly* averages at each of the seven closest stations is then used to predict *daily* records at each PRISM grid cell from the *daily* records at the seven closest weather stations.

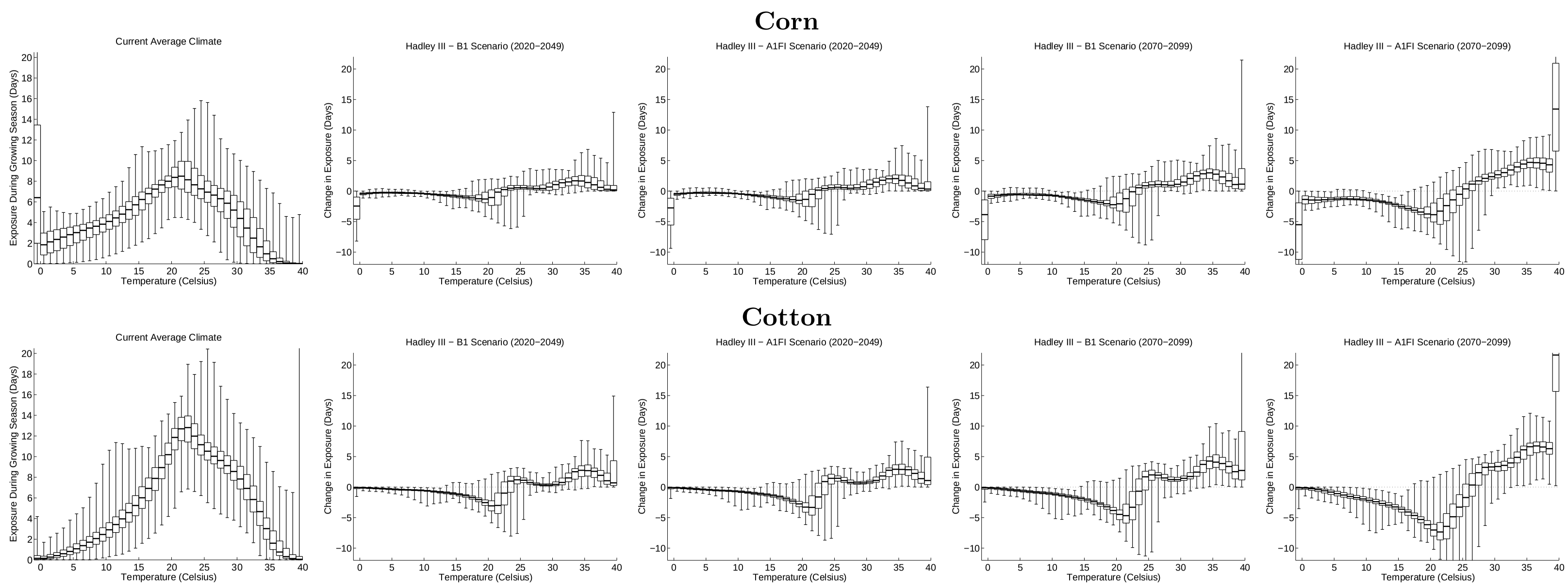
A cross-validation exercise is used to test the accuracy of the daily weather predictions. Specifically, we construct a daily weather record at each PRISM cell that harbors a weather station without using that weather station in the interpolation procedure. We then compare predicted daily outcomes at the PRISM cell with a weather station to actual outcomes recorded at the weather station in the grid cell. The mean absolute error is 1.36°C for minimum temperature and 1.49°C for maximum temperature. Due to the law of large numbers, our county-level distribution estimates contain less error, since they average errors over all grid cells in each county and all days of the growing season.

We approximate the distribution of temperatures within a day with a sinusoidal curve between minimum and maximum temperatures (Snyder 1985).[10] We derive the time spent in each 1°C-degree temperature interval between -5°C and $+50$°C. Finally, we construct the area-weighted averages over all PRISM grid cells in a county. The agricultural area in

from the distribution of average monthly temperatures. This method appears appropriate for predicting the *average* frequency that a certain weather outcome will be realized, but less appropriate in predicting a *specific* frequency of a weather outcome in a particular year. As a result, these methods work well in a forward-looking cross-sectional analysis where the dependent variable is tied to weather expected outcomes rather than realized outcomes (for example, the link between land values and climate), but less well in our analysis where the dependent variable (yield) linked to specific weather outcomes. Obtaining daily values on a small scale requires a spatial interpolation procedure to approximate daily weather outcomes between individual weather stations.

[10] In a sensitivity check we instead use a linear interpolation between minimum and maximum temperature. Both methods give similar results.

Figure 1: Descriptive Weather Statistics



Notes: Graphs display the amount of time a crop is exposed to each 1°C interval during the 183-day growing season. The lowest interval has no lower bound and includes the time temperatures fall below 0°C . The topmost interval has no upper bound and includes the time temperatures are above 39°C. For each interval, the range between minimum and maximum among counties is shown by whiskers, the 25%-75% percentile range is outlined by a box, and the median is added as a solid bold line. The top row displays the results for the 2277 eastern counties that grow corn or soybeans while the second row displays the results for the 983 counties that grow cotton. The first column displays current average climate conditions, while rows two through five display climate change predictions.

each cell was obtained from LandSat satellite images.[11] Boxplots in the left column of Figure 1 summarize the average historical weather distribution and its variability across space. Whiskers indicate the minimum and maximum average exposure to a certain temperature range among counties. The box marks the 25%-75% range, while the middle line within each box is the median. The weather variables are summed over the six-month period from March through August for corn and soybeans, and the seven-month period April through October for cotton.

We divide the United States into three regions to see whether warmer regions have adapted to higher temperatures and show a different relationship between yields and temperatures: northern, interior, and southern counties east of the 100 degree meridian.[12] The default data set for corn and soybeans is the union of northern, interior, and southern states–what we label eastern counties. We exclude counties in the Western United States and Florida because agricultural production in these areas relies on heavily subsidized access to irrigation water. Since the access to subsidized water rights is correlated with climate, omitting these variables, which vary on the sub-county level of irrigation districts, will result in biased coefficient estimates on the climatic variables in a cross-sectional analysis (Schlenker et al. 2005). Moreover, the response to temperatures is assumed to be different in these highly irrigated areas. Cotton is predominantly grown in the south and west, and we include all states in the analysis to obtain a larger sample of counties.

Climate Change Scenarios

Climate change predictions are drawn from the Hadley 3 model.[13] This major climate change model forms the basis for the report by the Intergovernmental Panel on Climate Change (IPCC). We obtain monthly model output for both minimum and maximum temperatures under four major emissions scenarios (A1FI, A2, B1, and B2) for the years 1960-2099. Each emission scenario rests on a different assumption about population growth and availability of alternative fuels, among other factors (Nakicenovic, ed 2000). The model run B1 assumes the slowest rate of warming over the next century, while model run A1FI assumes continued use

[11] Vince Breneman and Shawn Bucholtz at the Economic Research Service were kind enough to provide us with the agricultural area in each PRISM grid cell. Since we use the LandSat scan of a given year, we are not able to pick up shifts in growing regions.

[12] The northern subset includes counties in Illinois, Indiana, Iowa, Michigan, Minnesota, New Jersey, New York, North Dakota, Ohio, Pennsylvania, South Dakota, and Wisconsin that lie east of the 100 degree meridian. Interior counties are in Delaware, Kansas, Kentucky, Maryland, Missouri, Nebraska, Virginia, and West Virginia. Finally, southern counties are in Alabama, Arkansas, Georgia, Louisiana, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, and Texas.

[13] http://www.metoffice.com/research/hadleycentre/

of fossil fuels, which results in the largest increase in CO_2-concentrations and temperatures. We choose the two extreme scenarios, B1 (slowest increase) and A1FI (largest increase), to derive the range of possible climate change scenarios. In an appendix available upon request, we consider the effects for a range of uniform temperature increases.

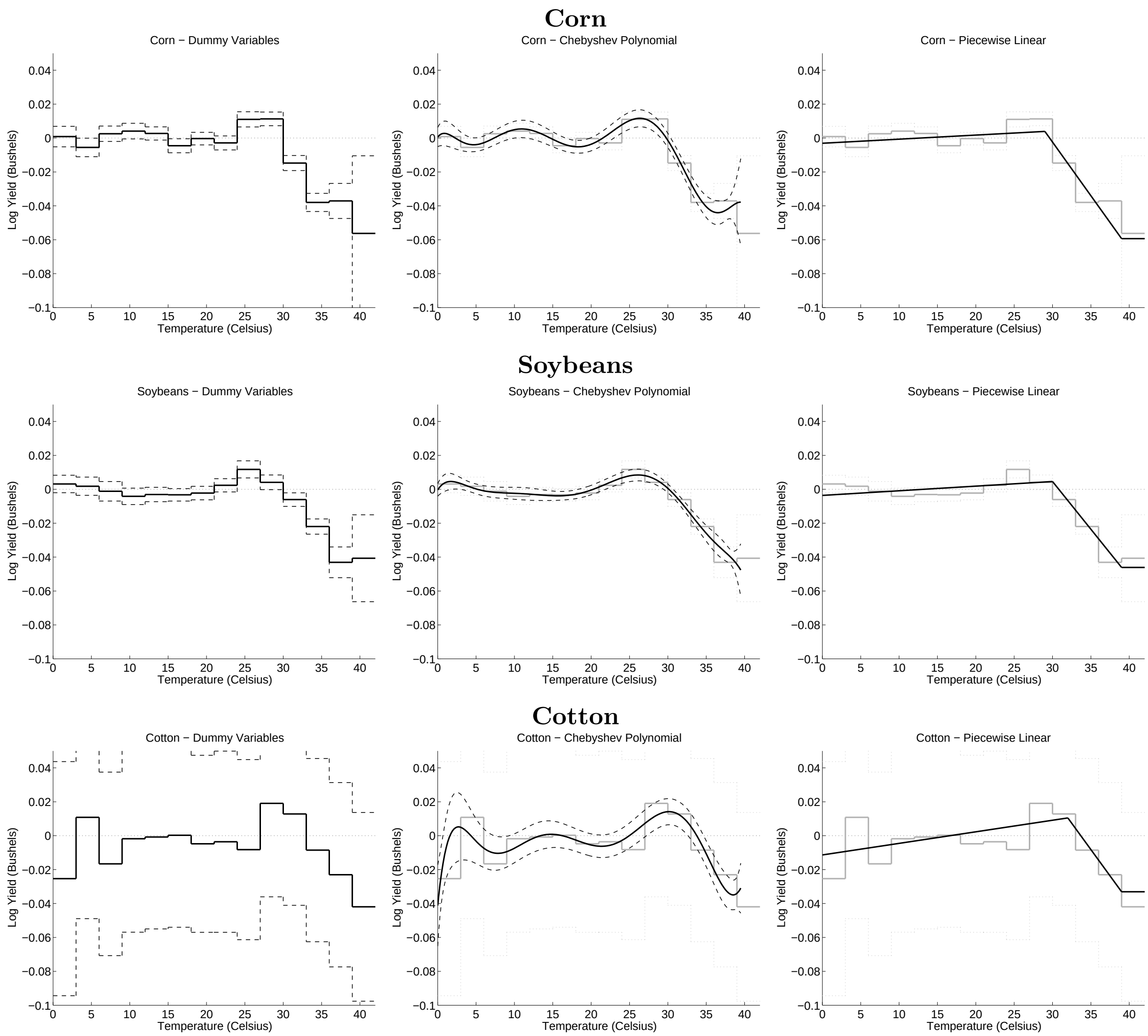
Predicted weather under climate change is derived as follows. At each of 216 Hadley grid nodes covering the United States we find the predicted difference in monthly mean temperature for 2020-2049 (medium-term), 2070-2099 (long-term), and historic averages (1960-1989). Next, predicted changes in monthly minimum and maximum temperature at each 2.5x2.5 mile PRISM grid are calculated as the weighted average of the monthly mean change in the four surrounding Hadley grid points, where the weights are proportional to the inverse squared distance and forced to sum to one. In a final step, we add the predicted *absolute* changes in monthly minimum and maximum temperatures at each PRISM grid to observed daily time series from 1960 to 1989. In other words, we shift the historical distribution mean for each climate scenario. An analogous approach was used for precipitation, except that we use the *relative* ratio of future predicted rainfall to historic rainfall instead of absolute changes. Each county's weather outcomes in a climate scenario are the area-weighted averages of all PRISM grids that cover farmland.

The last four columns of Figure 1 show the shift in the temperature distribution under the B1 and A1FI scenarios in the medium-term (2020-2049) and long-term (2070-2099), with separate plots for eastern counties that grow corn or soybeans (top row) as well as all counties that grow cotton (bottom row). Each figure shows a series of box plots, one for each degree Celsius. Each boxplot summarizes the predicted change in the frequency of that specific temperature across all counties growing that crop. Generally, temperatures below 22°C become less frequent in corn and soybeans counties, as well as temperatures below 25°C in cotton counties. Temperatures above these levels generally become more frequent.

4 Estimation Results

Estimates and standard errors of each model's temperature effects are displayed in Figure 2. The figure has nine panels, where each row represents one of three crops, and each column uses a different specification of the function $g(h)$. The left column uses the most flexible dummy-variable specification (equation (3) above); the middle column uses Chebyshev polynomials (equation (4) above); and the third column uses a piecewise linear specification. Each specification shows the same characteristic shape, increasing modestly up to a critical

Figure 2: Nonlinear Relation Between Temperature and Corn, Soybean, and Cotton Yields



Notes: Graphs show the impact of a given temperature for one day of the growing season on yearly log yields. The first row use corn yields, the second soybean yields, and the last cotton yields. The left graphs use dummy variables for each 3°C interval (which are added in grey to the middle and right graphs), the middle graphs use an 8th-order Chebyshev polynomial, and the plots on the right use a piecewise-linear function. Curves are centered so the exposure-weighted impact is zero. The lower bounds for the piecewise linear function were fixed at 0°C, but the optimal breakpoint was estimated.

temperature and then decreasing sharply. For corn the critical temperature is 29°C; for soybeans it is 30°C; and for cotton it is 32°C.

The vertical axis in each figure marks the log of yield in bushels per acre with the exposure-weighted average predicted yield normalized to zero. Thus, in comparing two points on any curve, a vertical difference of 0.01 indicates approximately a 1% difference in average yield growth. For example, on the top-left panel (the dummy variables model for corn) substituting a full day (24 hours) at 29°C temperature with a full day at 40°C temperature results in a predicted yield decline of approximately 7 percent, holding all else the same.

For brevity, other explanatory variables (precipitation, squared precipitation, county fixed effects, and state-specific quadratic time trends) are not reported. Precipitation has a statistically significant inverted-U shape with an estimated yield-maximizing level of 25.0 inches for corn and 27.2 inches for soybeans in the dummy-variable specification in the left column of Figure 2. The precipitation variables are not statistically significant for cotton, which is not surprising given that it is highly irrigated. The fixed effects and trends control for time-invariant heterogeneity and technological change and are of little interest by themselves. Given the wide geographic variation in yields and three-fold increase in yields over the sample period, these controls have strong statistical significance. Interestingly, however, the temperature effects are similar whether or not the controls are included in the regressions. In alternative specifications (not reported) we also find the estimated temperature effects to be similar if we instead control for technology and time effects using year fixed effects or state-by-year fixed effects rather than state-level trends.

Table 1 reports encompassing tests that compare our new model and approach to others in the literature. Comparisons are based on out-of-sample forecasts. Each model is estimated using 85 percent of the sample (randomly selected) and performance is measured according to the accuracy of the estimated model's prediction for the omitted 15 percent of the sample. Models compared include our own three specifications of temperature effects (dummy variables, Chebychev polynomial, and piecewise linear), a model with average temperatures for each of four months (Mendelsohn et al. 1994), an approximation of growing-degree days based on monthly average temperatures (Thom's formula) used in Schlenker et al. (2006), and a measure of growing degree days that is calculated using daily mean temperatures used by Deschenes and Greenstone (2007).[14] As a baseline, we also report a model with county fixed effects and no weather effects.

[14]We use degree days bounds of each study, but the ranking of models does not change if we were to use the bounds of this study instead.

Table 1: Model Comparison Test for Out-Of-Sample Prediction Accuracy

	Corn			Soybeans			Cotton		
	RMS	**GW**	**MGN**	**RMS**	**GW**	**MGN**	**RMS**	**GW**	**MGN**
Dummy Variables	0.2179			0.1899			0.3130		
Chebyshev Polynomials	0.2179	0.5028	0.03	0.1900	0.9697	1.86	0.3135	0.8441	1.66
Piecewise Linear	0.2199	0.9858	8.60	0.1921	1.0202	8.71	0.3150	0.8929	3.43
Monthly Averages	0.2289	0.7113	13.33	0.2008	0.7480	13.57	0.3162	0.5925	2.14
Degree Days 8-32°C, >34°C (Thom)	0.2398	0.9935	28.81	0.2027	0.8952	18.49	0.3220	0.8879	7.33
Degree Days 8-32°C (Daily Mean)	0.2436	0.9763	30.76	0.2083	0.9191	22.97	0.3272	0.9153	9.55
County-Fixed Effects (No Weather)	0.2598			0.2211			0.3323		

Notes: Table compares various temperature specifications for corn, soybeans, and cotton according to three out-of-sample criteria: (i) **RMS** is the root mean squared out-of sample prediction error; (ii) **GW** gives the Granger weight on the dummy variable regression of the optimal convex combination between the dummy variables regression and the model listed in the row; (iii) **MGN** is the normally distributed Morgan-Newbold-Granger statistic of equal forecasting accuracy.
Each model is estimated using the same 85% of the data (randomly selected) and yields are forecasted out-of-sample for the omitted 15%. **Dummy variables Chebyshev Polynomials**, and **Piecewise Linear** are the models developed in this paper; **Monthly Averages** uses a quadratic specification in both average temperature and total precipitation for the months January, April, July, and October (Mendelsohn et al. 1994); **Degree Days Thom** uses Thom's formula to extrapolate degree days (which are based on daily data) from monthly average temperature data (Schlenker et al. 2006); **Degree Days (Daily Mean)** first derive the average temperature for each day from daily temperature readings and then construct degree days from this average (Deschenes and Greenstone 2007); **County fixed Effects (No weather)** uses *only* fixed effects but no weather measure at all.

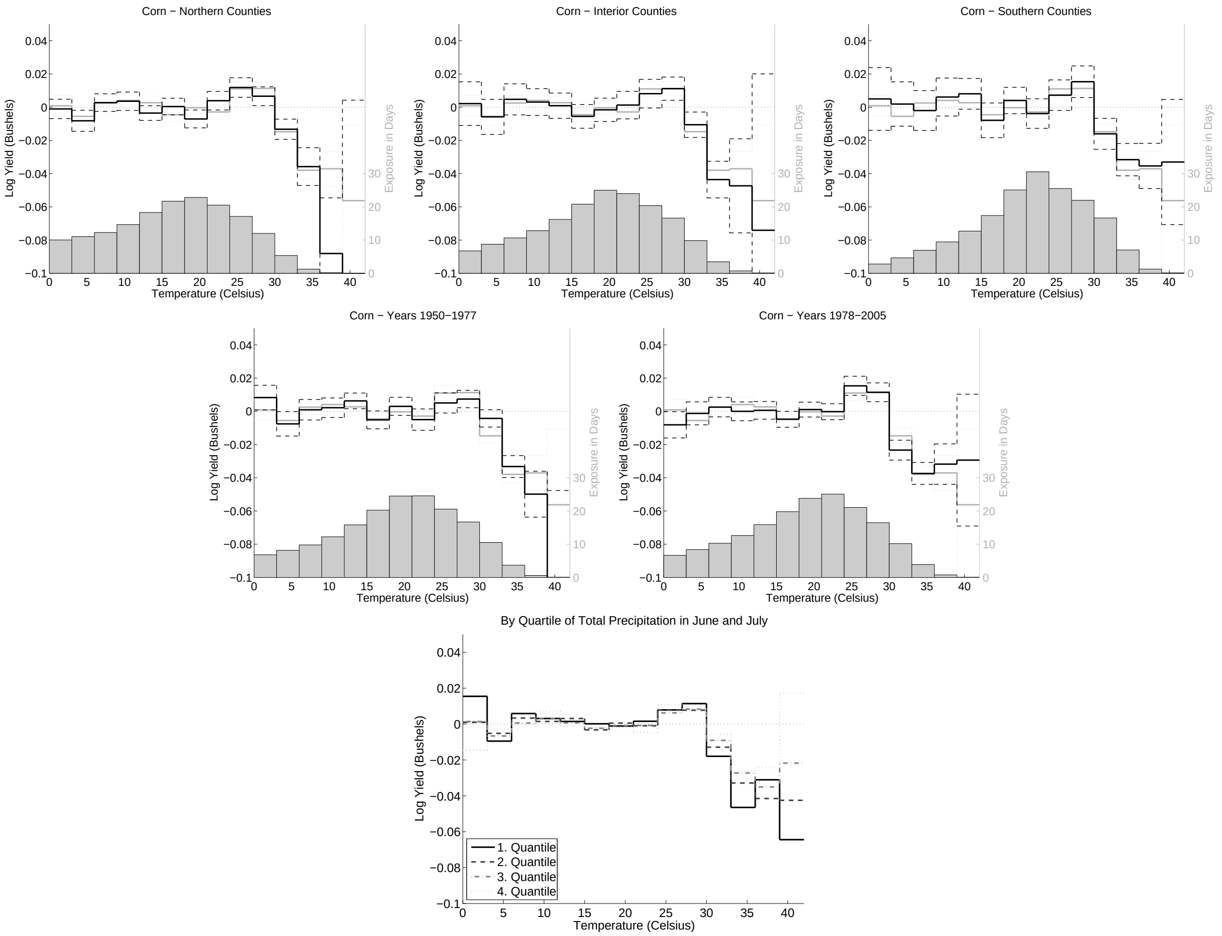
The table reports the root-mean squared prediction error (RMS) and two statistics that facilitate comparison of the best-predicting dummy-variable model to each of the other models (Diebold and Mariano 1995). The first statistic is the Granger weight, which is the weighted average of two forecasts where the weights are forced to sum to one. We report the weight on the dummy-variable model. If both models forecast equally well, each receives a weight of 0.5. If one is superior to the other, it receive a weight greater than 0.5. The second statistic is the normal-distributed Morgan-Newbold-Granger statistic against the null hypothesis of equal forecasting ability between the dummy-variable model and the comparison model. The statistics show little difference in forecasting ability between our dummy-variable approach and the smoothed Chebyshev polynomials, but large and statistically significant differences between the dummy-variable model and other models in the literature. Models that average temperatures over time or space have significantly inferior out-of-sample predictions relative to our new approach. Starting from a baseline model without weather variables, the new model reduces the root mean squared prediction error nearly three times as much as a model that uses daily temperature averages.

We explore the robustness of the preferred dummy variable model over various subsets of the corn panel data set in Figure 3. We focus on corn because it is grown over the widest geographic area and has been by far the most valuable crop in the United States. The first three panels of Figure 3 (top row) show results for each of three mutually exclusive subsets of counties corresponding to the most northern (and coolest) states, the most southern (and warmest) states, and those in the middle. In all cases we consider the more flexible dummy-variable specification of the temperature function. Estimates for the pooled sample from Figure 2 are plotted in grey for comparison. Each plot also includes the empirical distribution of temperatures within each subregion as grey histogram. These show how much warmer the southern counties are in comparison to the northern counties. The interesting and notable feature is the stability of the estimated temperature relationship across the three subregions.

The next two panels of Figure 3 (middle row) divide the sample into two time periods, 1950-1977 and 1978-2005. Although average yields in the more recent panel are substantially greater than those in the earlier period, the temperature relationships are similar to each other and to the pooled sample.

The last panel of Figure 3 overlays estimates from four regression models that divide the sample by quartiles of total precipitation in June and July. These estimates have a similar shape to that of the pooled sample up to the critical temperature of 30°C. The decline above the threshold, however, is less steep for subsamples with greater precipitation. Thus, there is

Figure 3: Nonlinear Relation Between Temperature and Corn Yields for Subsets of Counties or Years

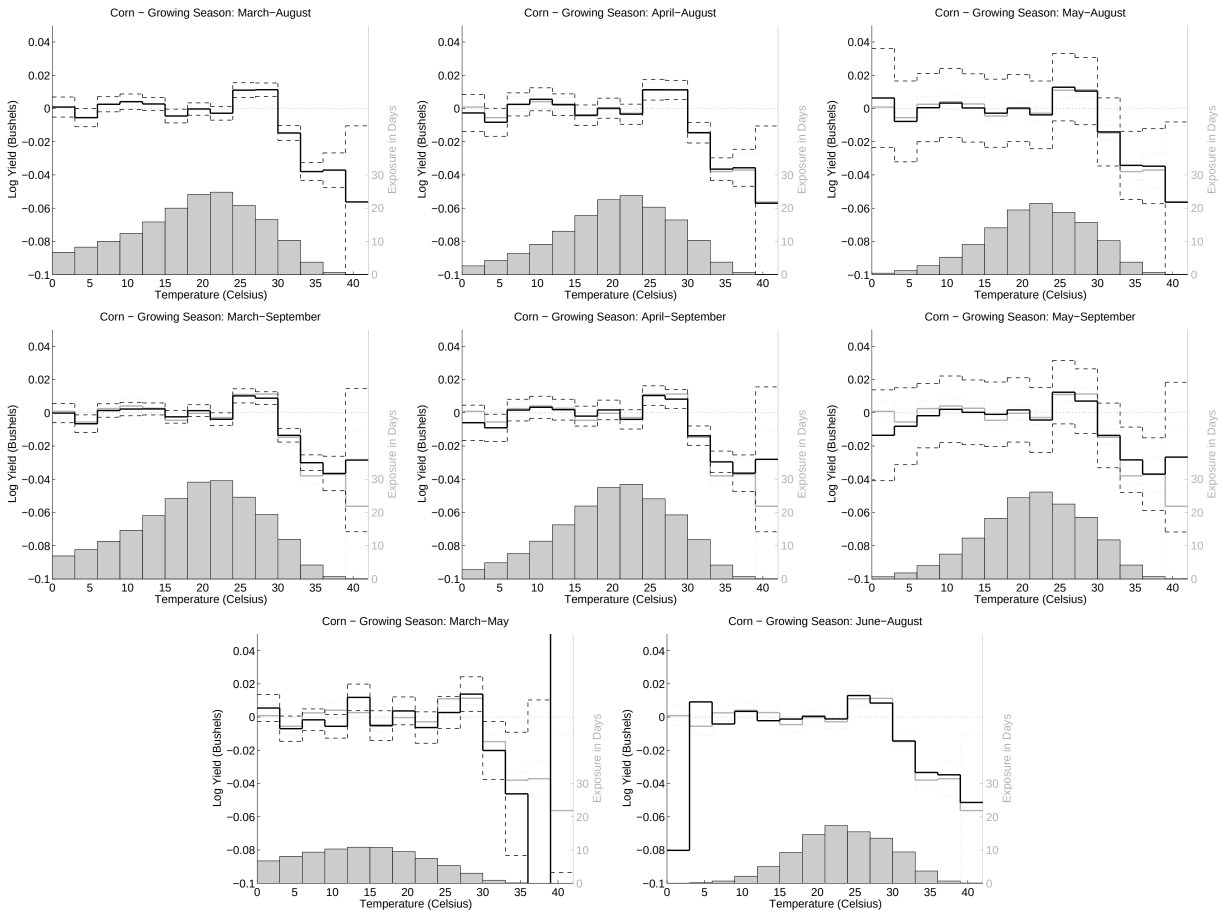


Notes: Graphs display changes in annual log yield if the crop is exposed to a particular temperature for one day. Grey histograms display average weather outcomes in the sample. The top row limits the analysis to various geographic subsets; the middle row considers temporal subsets; and the bottom row splits the data into quartiles based on the total precipitation in the months of June and July. Curves are centered so the exposure-weighted impact is zero. Results from the pooled model are added in grey for comparison in the top two rows. The 95% confidence band, after adjusting for spatial correlation, is added as dashed lines.

some evidence that precipitation partly mitigates the damages from extreme temperatures.[15] Since we do not find a significant correlation between temperatures and rainfall in the raw

[15]We did estimate models with richer interactions between temperature and rainfall, but these models do not predict out-of-sample significantly better than additively separable model reported above. It is possible that the relatively poor predictive power of precipitation in comparison to temperature stems from greater measurement error in the precipitation variable as spatial smoothing is more difficult for the latter.

Figure 4: Nonlinear Relation Between Temperature and Corn Yields for Different Definitions of the Growing Season



Notes: Graphs display changes in annual log yield if the crop is exposed to a particular temperature for one day. Grey histograms display average weather outcomes in the sample. The top left panel is the baseline model. All other graphs use different definitions of the growing season: The growing season ends in August in the first row and September in the second row, while it starts in April, March, and May in the three columns of the first two rows. The last row breaks the six-month growing season into two three-month periods. Curves are centered so the exposure-weighted impact is zero. The 95% confidence band, after adjusting for spatial correlation, is added using dashed lines.

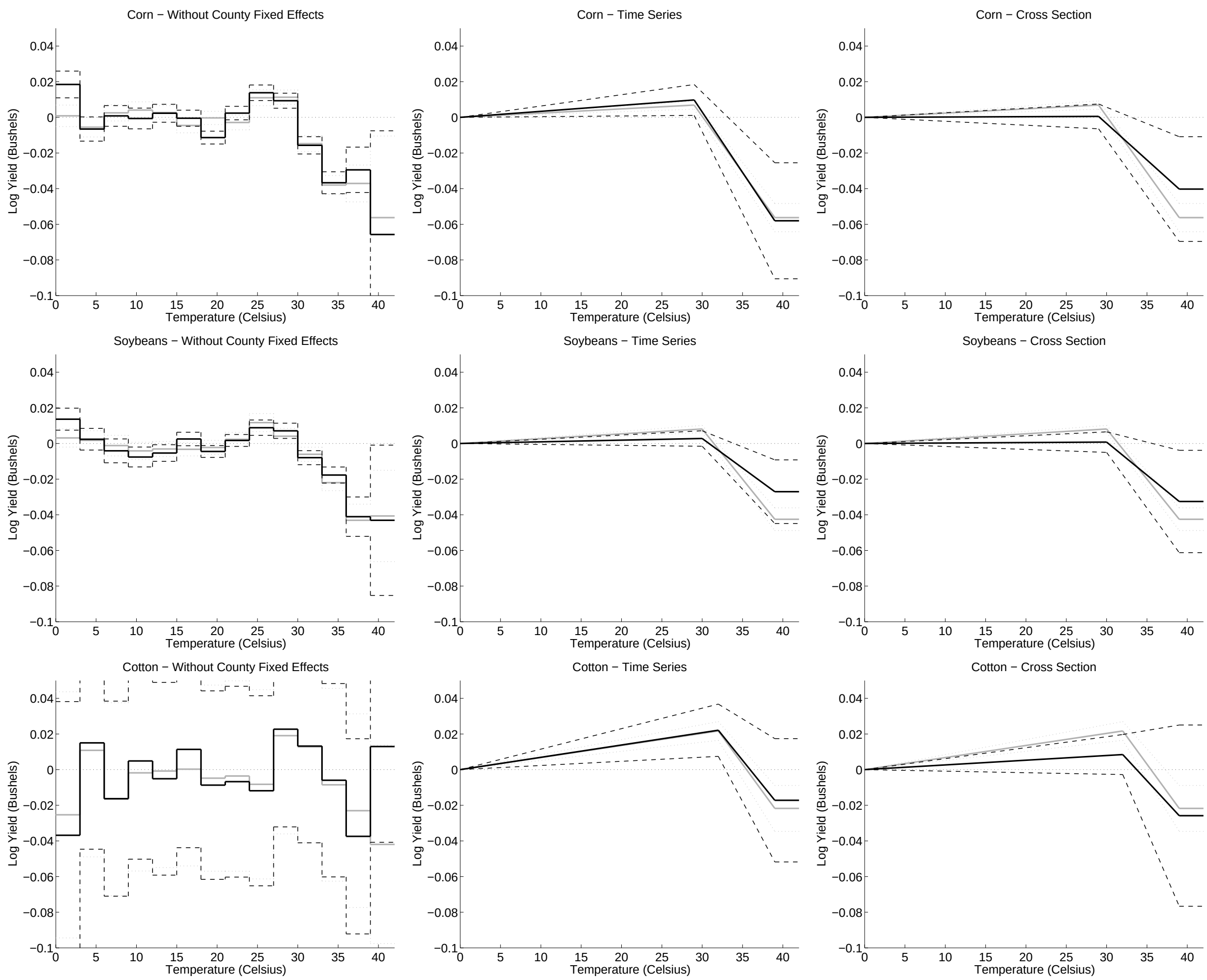
daily data, omitting temperature-rainfall interactions will not bias our predictions and give the right average effects of temperature and rainfall.

An important assumption of the empirical model is the additive separability of temperature effects over time. We fix the growing season to the months March through August for corn and soybeans, even though northern regions tend to plant later than southern regions, and planting dates may vary from year-to-year depending on weather conditions. We explore the sensitivity of the results to various definitions of the growing season in Figure 4. The figure shows seven alternative specifications of the growing season together with the baseline (top left). Again, the estimated temperature effects appear similar regardless of how we shift the growing season. The first two rows vary the start date (March in the first column, April in the second, and May in the third) as well as the end date (August in the top row and September in the second row). The third row breaks the six-month growing season into two three-month periods and still obtains similar results. This lends support to the assumption of additive separability.

Another way to consider endogenous grower responses to a permanent shift in climate is to compare regression results from a panel-data analysis to those from a cross sectional analysis. A panel data analysis with county fixed effects is identified from arguably random time-series variation in weather, which accounts for little grower adaptation to weather. In contrast, a cross-sectional analysis compares yields and grower management choices across areas with different weather expectations (i.e., climates). Much like the hedonic model, these comparisons therefore embody grower adaptations to weather, not just the direct effects of weather. These comparisons are presented in Figure 5. In all plots, results from a panel with fixed effects are displayed in grey for comparison. Plots on the left replicate the panel analysis without the use of county-fixed effects. It uses both cross-sectional (climate) and time-series (weather) variation. Results are very similar to the model with county fixed effects. The middle and right plots use, exclusively, the aggregate time series and cross-sectional variation, respectively. For the time-series we derive the average national yield and regress it on the area-weighted average weather outcome in a given year.[16] Since our panel includes 56 years, the sample size in the pure time series reduces to 56 observations. This makes estimation of a dummy-variables approach questionable due to insufficient degrees of freedom. Accordingly, we estimate a piecewise linear function which only has two temperature variables and two precipitation variables. For the cross-section we regress the average difference between the county yield and the nationwide average yield on the average climate in a county. Both the cross-section and the time series give us comparable results, except that the standard errors become larger in the case of cotton.

[16]To adjust for technological progress the dependent variable is the log of the average yield in a year minus a linear time trend of log yields.

Figure 5: Nonlinear Relation Between Temperature and Corn Yields Using Various Sources of Identification



Notes: Graphs display changes in annual log yield if the crop is exposed for one day to a particular temperature. The grey lines are the baseline model (first column uses the first column of Figure 2, while the second and third column use the third column of Figure 2). The black lines are sensitivity checks: The first column replicates the model without county fixed effects; the second column uses the aggregate time-series of 56 annual averages; the third column uses county average yields in the cross-section. Curves are centered so the exposure-weighted impact is zero in the left column and fixed at 0 in the right two columns. The 95% confidence band, after adjusting for spatial correlation, is added as dashed lines.

Market Assessment of Extreme Weather Events

Our preceding analysis reveals that temperatures above an upper threshold result in substantial yield reductions. In the following we briefly examine how futures market assess such extreme temperature events.[17] An efficient futures market immediately incorporates news about impending shifts in commodity supply and demand. Thus, if extreme temperatures harm yields, news about current or impending extreme temperatures signal an impending inward shift in supply, and causes an increase in futures prices. Since it is impossible to determine precisely when expectations are formed, we focus on weekly changes in future prices rather than daily values.[18] We link percent changes in futures price to weather variables in both the current and the subsequent week, since weather outcomes one week out might be forecastable.[19] We consider weekly price changes for the months May-July for future contracts that expire in September, by which time most uncertainty about nationwide yield has been resolved.[20]

Results are reported in Table 2. The first column relates futures price changes to degree days 8-29°C and degree days above 29°C, the same variables as the right column of Figure 2. The second column uses the more conventional temperature measures, average temperature and average temperature-squared. Both specifications include a quadratic in precipitation and fixed effects for each week so identification comes from deviations from predictable seasonal averages. Futures prices are statistically significantly decreasing in degree days below 29°C and increasing in degree days above 29°C.[21] The magnitude of the coefficient on degree days above 29°C is much larger than the one on temperatures below this threshold. This sharp asymmetry is consistent with our finding from the yield regression.[22] The coefficients indicate that one additional day at 40°C instead of 29°C increases future prices by 4.4 percent. The estimated price impact of extreme heat is substantial, particularly given storage tends to buffer the price effects of yield shocks. In contrast, average temperature and average temperature squared are not statistically significant. And while the two specifications have

[17] A more detailed analysis is given in a separate paper.

[18] We calculate the percent change in closing prices on Friday compared to the previous Friday. Weather variables are the corn-area weighted average of all counties east of the 100 degree meridian for the week in question.

[19] Campbell and Diebold (2005) show that an autoregressive process predicts average temperatures as well as a professional weather forecast for time periods more than 5-8 days into the future.

[20] We exclude weeks before May as markets are less liquid before this period: average trade volume is less than 20% of the weekly volume in the peak season. We also exclude August when USDA publishes its first yield forecasts, which also influences futures markets.

[21] Recall that a reduction in quantity implies an increase in price and vice versa.

[22] Precipitation peaks at 3.06 cm or 1.2 inches. Since we are looking at weekly intervals this translates into 31.5 inches for the 183 growing season, again comparable to our yield regression.

Table 2: Impact of Extreme Heat on Corn Futures Prices

	Coeff.	t-val	Coeff.	t-val
Weather in current week				
Degree Days 8-29°C	-0.0296	(2.40)		
Degree Days >29°C	0.2241	(2.30)		
Average Temperature			-0.3374	(0.88)
Average Temperature Squared			0.0072	(0.75)
Precipitation	-2.2006	(4.38)	-2.4595	(4.93)
Precipitation squared	0.3600	(3.94)	0.3959	(4.34)
Weather is subsequent week				
Degree Days 8-29°C	0.0168	(1.23)		
Degree Days >29°C	0.1783	(1.77)		
Average Temperature			-0.6856	(1.47)
Average Temperature Squared			0.0220	(1.95)
Precipitation	-0.6988	(1.34)	-0.7703	(1.49)
Precipitation squared	0.1147	(1.20)	0.1206	(1.27)
Observations	698		698	
R-squared	0.1006		0.0908	
Durbin-Watson statistic	1.71		1.70	
Week fixed effects	yes		yes	

Notes: Table lists regression results when weekly percent changes in futures prices are regressed on weather variables for the same week as well as the subsequent week. We include the subsequent week as weather can be forecasted and hence anticipated in advance.

the same dependent variable and identical degrees of freedom, the R^2 of the first regression is 10 percent higher than the second. This is additional evidence that the frequency of very warm temperatures is especially influential for yields.

5 Climate Change Impacts

Yield predictions under climate change are summarized in Table 3. The table reports nationwide area-weighted impacts and summary statistics for the predicted impacts across counties under each of the climate scenarios both over medium-term (2020-2049) and long-term (2070-2099) horizons. All predictions in Table 3 use the most flexible dummy-variable

Table 3: Predicted Impacts of Global Warming on Crop Yields (Percent)

	Medium-term 2020-2049						Long-term 2070-2099					
	Area-weighted		Impact by County				Area-weighted		Impact by County			
Variable	**Impact**	**(t-val)**	**Mean**	**Min**	**Max**	**Std**	**Impact**	**(t-val)**	**Mean**	**Min**	**Max**	**Std**
						Corn						
HCM3 - B1	-22.34	(21.03)	-28.32	-63.67	11.70	17.78	-43.16	(19.50)	-45.70	-83.76	18.11	18.18
HCM3 - B2	-23.02	(22.70)	-29.43	-70.01	11.08	17.09	-50.66	(21.24)	-53.51	-90.03	18.16	18.08
HCM3 - A2	-27.62	(23.29)	-32.55	-68.99	14.39	17.09	-69.71	(16.07)	-71.07	-96.34	4.27	16.33
HCM3 - A1FI	-28.54	(21.14)	-32.26	-68.95	11.55	17.19	-78.59	(14.75)	-79.83	-98.45	-7.70	14.35
						Soybeans						
HCM3 - B1	-18.62	(21.10)	-19.39	-62.24	16.49	17.10	-36.10	(22.94)	-34.27	-82.53	25.01	19.61
HCM3 - B2	-19.50	(22.37)	-20.24	-67.21	17.49	16.55	-43.73	(25.04)	-42.15	-87.53	26.09	20.42
HCM3 - A2	-23.11	(23.43)	-23.02	-67.71	20.08	16.78	-63.72	(20.87)	-61.33	-94.56	19.72	19.54
HCM3 - A1FI	-23.04	(21.76)	-22.72	-67.82	16.61	17.11	-73.64	(19.53)	-71.36	-96.79	11.87	17.32
						Cotton						
HCM3 - B1	-21.71	(6.58)	-15.39	-47.37	21.82	14.53	-31.08	(5.59)	-22.37	-66.83	31.24	18.20
HCM3 - B2	-20.98	(5.30)	-14.54	-56.40	25.98	15.01	-40.42	(6.21)	-31.45	-73.82	32.48	18.60
HCM3 - A2	-22.27	(5.81)	-15.41	-53.98	30.15	15.70	-56.99	(7.10)	-49.26	-86.22	42.03	18.93
HCM3 - A1FI	-21.59	(5.53)	-14.67	-51.13	23.18	14.16	-67.18	(7.97)	-58.79	-91.95	50.78	19.43

Notes: Table lists percent changes in corn, soybeans, and cotton yields predicted in the medium term 2020-2049 (first six columns) and long term 2070-2099 (last six columns) under four emission scenarios. The first two columns show the area-weighted impact including t-values, while the next four columns give the distribution of impacts among counties. Standard errors are adjusted for spatial correlation.

model. Across all scenarios and crops, some counties see yield gains and some see losses, but the nationwide impacts all show marked declines, ranging from about -19 to -29 percent in the medium-term and from about -31 to -79 percent in the long term. The driving force behind these large and significant impacts is the increased frequency of extremely warm temperatures that sharply reduce yields. While the previous section has shown that a model that accounts for the effect of extreme heat is better at explaining yields, it also gives significantly different impacts than traditional models that do not adequately model the nonlinearity.[23]

Figure 6 shows a map of the predicted impacts for corn across counties under the slow-warming scenario (B1) in the top row as well as the fast-warming scenario (A1FI) in the bottom row. Impacts are comparable in the medium-term (left column), but start to diverge towards the end of the century (right column). The geographic distribution of impacts is intuitive, with warmer southern areas seeing much larger declines than cooler northern areas. The exception is the Appalachian mountain range where temperatures are cooler compared to neighboring counties of comparable latitude. The central Midwestern region, which possesses among the world's best soils, sees substantial declines, and this is critical for nationwide impacts. A similar pattern is observed for soybeans in Figure 7. Corn is only grown in the south and we display the geographical distribution of impacts in Figure 8.

While nationwide predictions are quiet large, the predicted impacts keep planting dates as well as growing areas fixed.[24] In a sensitivity check we shift the growing season one month forward, i.e., we assume corn is grown from February through July instead of March to August.[25] We still use the coefficient estimates from our baseline dummy variable model for the months March-August, but derive the predicted change in each 3°C interval for the months February-July. Predicted yield impacts by the end of the century reduce from -43% to -31% under the B1 scenario and from -79% to -64% under the A1FI scenario. Damages decline as extremely warm temperatures are observed less frequently in February than in August. With regards to planting locations, Figures 6-8 show that some areas get hit less severely and shifts in the growing area could mitigate some of the negative impacts.

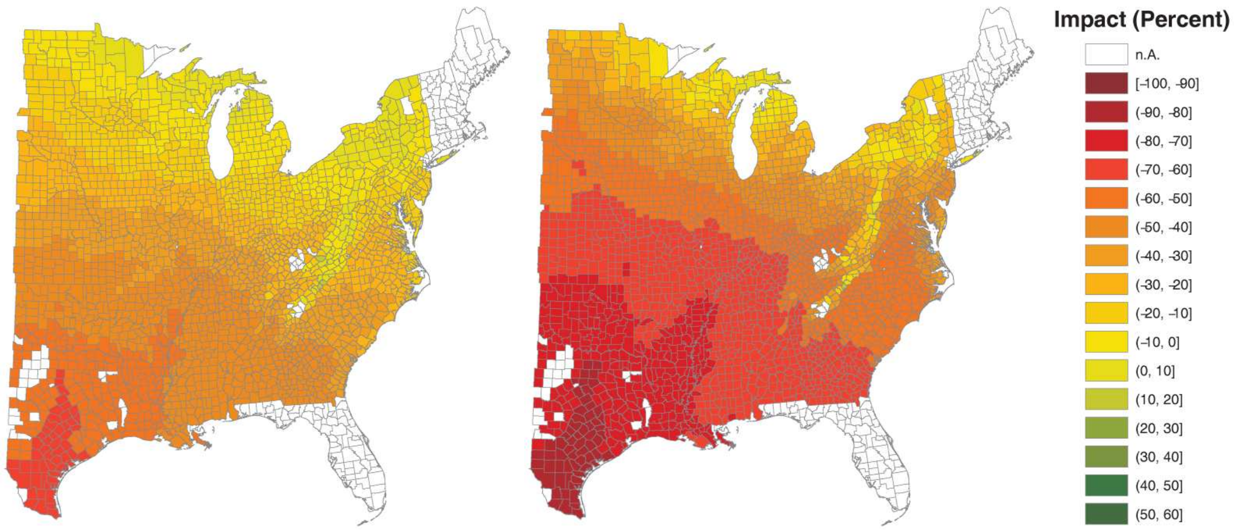
[23] If we compare our preferred model using dummy variables to a model using (i) monthly average temperatures; (ii) degree days derived using Thom's formula; and (iii) degree days derived using daily means (models described in the bottom three rows of Table 1) the root mean squared difference in predicted corn yields by the end of the century under the A1FI scenario is 10.3%, 7.4%, and 18.23%, respectively.

[24] Note that predictions are relative to what would have been realized without climate change, not relative to current production levels. These baselines differ because yields have been trending up with technological change over the last 50-plus years, and are likely to increase in the future as well.

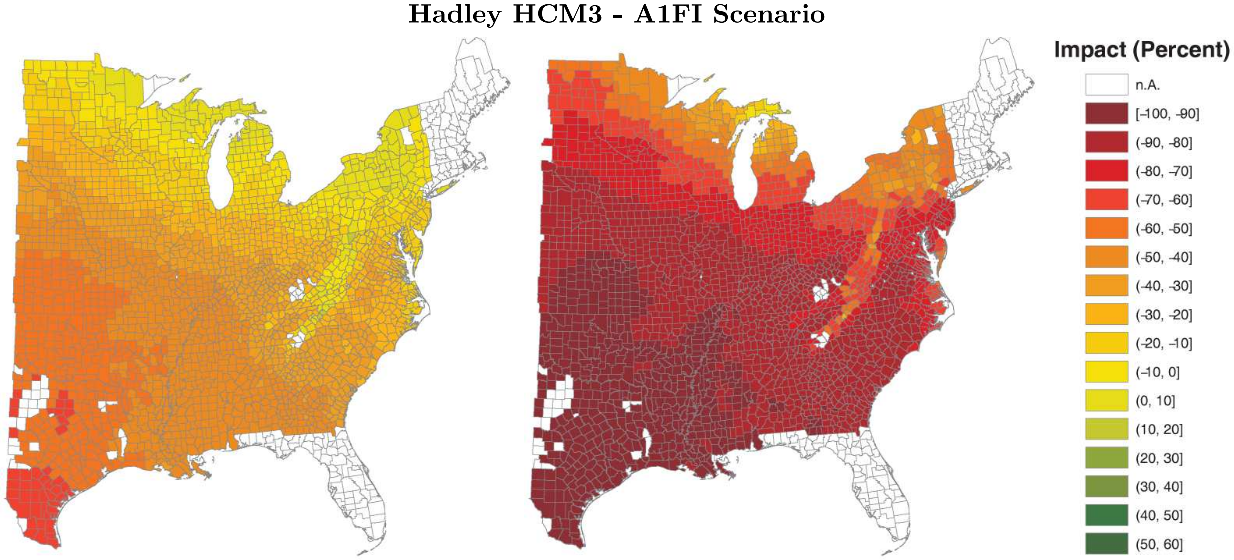
[25] A further forward shift seems unlikely as it simultaneously reduces available sunlight. Solar radiation is lower during winter months.

Figure 6: Predicted Changes for Corn Yields in the Eastern United States (Percent)

Hadley HCM3 - B1 Scenario

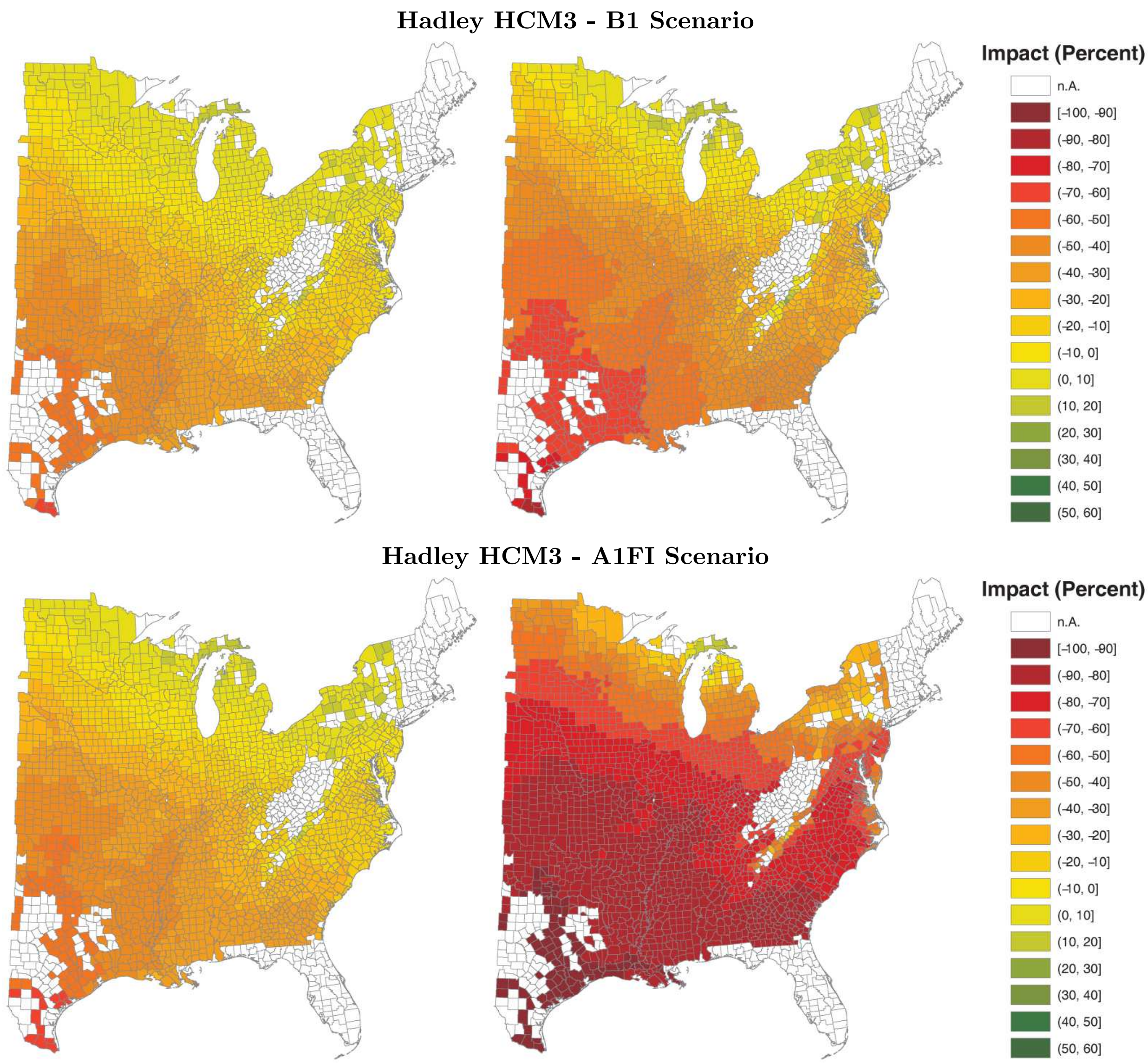


Hadley HCM3 - A1FI Scenario



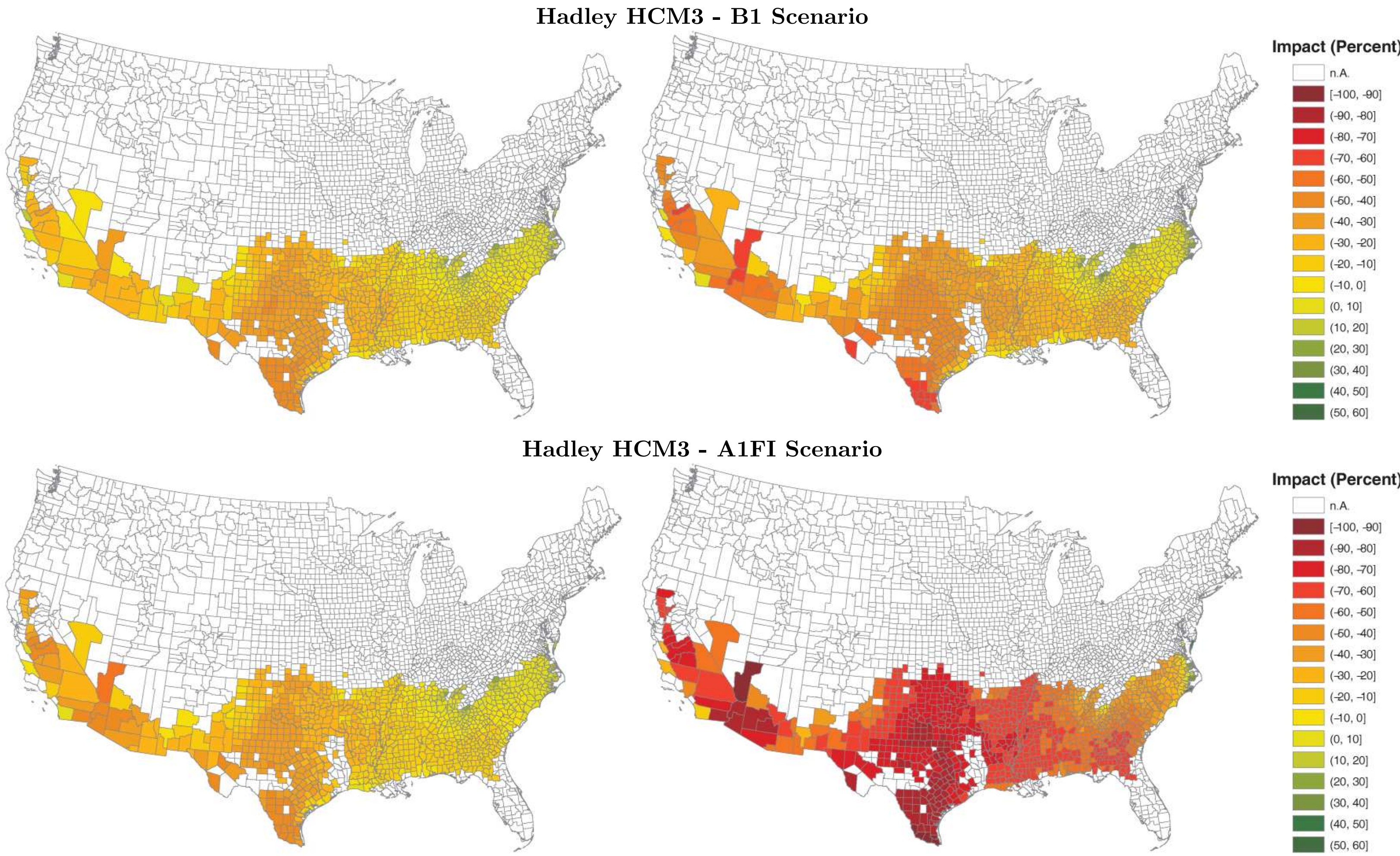
Notes: Graphs display predicted changes in corn yields under the slow warming B1 scenario (top row) and fast warming A1FI scenario (bottom row). The left column shows predicted changes in the climatic variables for 2020-2049, while the right column shows predicted changes for 2070-2099. Impacts are evaluated using the dummy variable regression in the top left panel of Figure 2.

Figure 7: Predicted Changes for Soybeans Yields in the Eastern United States (Percent)



Notes: Graphs display predicted changes in soybeans yields under the slow warming B1 scenario (top row) and fast warming A1FI scenario (bottom row). The left column shows predicted changes in the climatic variables for 2020-2049, while the right column shows predicted changes for 2070-2099. Impacts are evaluated using the dummy variable regression in the middle left panel of Figure 2.

Figure 8: Predicted Changes for Cotton Yields (Percent)



Notes: Graphs display predicted changes in cotton yields under the slow warming B1 scenario (top row) and fast warming A1FI scenario (bottom row). The left column shows predicted changes in the climatic variables for 2020-2049, while the right column shows predicted changes for 2070-2099. Impacts are evaluated using the dummy variable regression in the bottom left panel of Figure 2.

However, even cooler areas are predicted to be impacted quiet substantially. Furthermore, shifts in growing areas are limited by the availability of the right soil.

We further investigate the issue of adaptation in Table 4 by comparing the predicted nationwide climate impacts derived using different sources of identification. The first two rows use the full panel data set to estimate the flexible dummy variable model and the piecewise-linear model.[26] The third row of each crop gives the predicted impact if we only use the time-series of 56 aggregate yields, while the last two use the cross-section of average yields across counties.[27] The underlying regression results are displayed in Figure 5. As outlined above, the cross section should capture adaptation within a crop species as farmers with a permanently warmer climate have an incentive to adapt to these warmer climates. However, the predicted climate impacts are comparable if we look at the panel, the cross-section, or the time series. However, the confidence intervals increase significantly for the cotton specification.

Table 5 reports further sensitivity checks for corn. We focus on corn because it is grown over the largest geographic area. We estimate the regression coefficients using one of three geographic subsets in the estimation but then evaluate the impact for all counties. If southern counties had successfully adapted to their warmer climate one would expect extreme temperatures to have a less harmful effect on yields. Again, the results are rather robust to which subset of counties is used in the estimation, suggesting limited potential for adaptation. The same holds true if the estimation is done using different time periods. This suggests that corn has not developed improved heat tolerance over time. We also consider an alternative interpolation procedure for estimating the temperature distribution within each day. In our baseline model we follow the natural science literature and fit a sinusoidal curve between minimum and maximum temperature. Table 5 shows that a linear interpolation between minimum and maximum gives comparable climate impacts. Finally, we evaluate predictions when yields are derived as total production divided by the land area planted instead of the land area harvested and we again obtain similar predictions.

[26] As mentioned above, the time-series regression has only 56 observations, which impedes estimation of the dummy-variables model due to insufficient degrees of freedom. We therefore consider the piecewise linear model with only two temperature variables.

[27] In the cross-section, controls for soil quality include water capacity, percent clay, permeability, soil erodibility (k-factor) and the fraction of soil classified as high quality.

Table 4: Predicted Climate Impacts Using Panel, Cross-section, and Time Series (Percent)

	Medium-term (2020-2049)				**Long-term (2070-2099)**			
	B1	**(t-val)**	**A1FI**	**(t-val)**	**B1**	**(t-val)**	**A1FI**	**(t-val)**
			Corn					
Dummy Variables (Baseline)	-22.34	(21.03)	-28.54	(21.14)	-43.16	(19.50)	-78.59	(14.75)
Piecewise-linear	-22.84	(21.57)	-29.34	(23.19)	-45.06	(27.18)	-81.87	(57.91)
Piecewise-linear (Time Series)	-23.38	(6.88)	-30.13	(7.23)	-45.85	(8.31)	-82.99	(16.27)
Piecewise-linear (Cross Section)	-19.58	(7.51)	-24.67	(7.00)	-37.88	(7.57)	-72.12	(9.83)
Piecewise-linear (Cross Section + Soil)	-19.33	(8.36)	-24.29	(7.97)	-37.61	(8.75)	-72.05	(12.40)
			Soybeans					
Dummy Variables (Baseline)	-18.62	(21.10)	-23.04	(21.76)	-36.10	(22.94)	-73.64	(19.53)
Piecewise-linear	-19.67	(22.20)	-24.16	(22.94)	-37.33	(25.88)	-74.50	(48.52)
Piecewise-linear (Time Series)	-14.12	(5.32)	-17.24	(5.35)	-27.31	(5.75)	-59.18	(7.72)
Piecewise-linear (Cross Section)	-17.13	(4.84)	-20.89	(4.71)	-32.33	(4.99)	-65.38	(6.14)
Piecewise-linear (Cross Section + Soil)	-17.98	(6.87)	-21.91	(6.71)	-33.93	(7.31)	-68.18	(10.01)
			Cotton					
Dummy Variables (Baseline)	-21.71	(6.58)	-21.59	(5.53)	-31.08	(5.59)	-67.18	(7.97)
Piecewise-linear	-22.95	(6.84)	-23.59	(6.22)	-35.37	(7.27)	-72.26	(14.71)
Piecewise-linear (Time Series)	-18.51	(2.24)	-19.18	(2.06)	-29.37	(2.32)	-65.67	(4.17)
Piecewise-linear (Cross Section)	-24.39	(1.81)	-27.79	(1.84)	-40.25	(2.01)	-71.75	(2.05)
Piecewise-linear (Cross Section + Soil)	-25.29	(1.82)	-28.71	(1.86)	-41.43	(2.00)	-72.90	(2.04)

Notes: Table lists predicted climate change impacts under various sensitivity checks for the slowest warming scenario (B1) as well as the most rapid warming scenario (A1FI). The dummy variables regression are the baseline results reported in Table 3. The last four rows of each crop report the predicted climate change impacts of a more parsimonious piecewise-linear function with only two temperature variables. Rows labeled time series use 56 area-weighted average yields (as well as area-weighted average climatic variables) in the regression. Rows labeled cross section use the average yield (and climatic variables per county) in the regression, both with and without controls for soil quality.

Table 5: Sensitivity of Predicted Climate Impacts Under Various Model Assumptions (Percent)

	Medium-term (2020-2049)				Long-term (2070-2099)			
	B1	**(t-val)**	**A1FI**	**(t-val)**	**B1**	**(t-val)**	**A1FI**	**(t-val)**
Corn								
Dummy Variables (Baseline)	-22.34	(21.03)	-28.54	(21.14)	-43.16	(19.50)	-78.59	(14.75)
Dummy Variables (Northern Counties)	-20.63	(16.16)	-19.92	(17.69)	-50.23	(16.45)	-92.89	(15.74)
Dummy Variables (Interior Counties)	-37.86	(9.98)	-44.54	(9.03)	-60.41	(8.10)	-90.75	(5.94)
Dummy Variables (Southern Counties)	-35.60	(17.65)	-36.91	(17.59)	-49.66	(18.17)	-80.12	(15.32)
Dummy Variables (1950-1977)	-22.33	(14.95)	-29.29	(15.57)	-46.99	(16.80)	-87.62	(23.75)
Dummy Variables (1978-2005)	-22.52	(18.36)	-28.36	(18.86)	-40.85	(17.32)	-70.77	(10.48)
Dummy Variables (Linear Interpolation)	-22.87	(20.90)	-29.33	(20.74)	-45.18	(18.07)	-82.85	(10.51)
Dummy Variables (Area Planted)	-27.82	(18.56)	-35.31	(19.70)	-52.15	(20.57)	-87.78	(21.73)
Soybeans								
Dummy Variables (Baseline)	-18.62	(21.10)	-23.04	(21.76)	-36.10	(22.94)	-73.64	(19.53)
Dummy Variables (Linear Interpolation)	-19.70	(22.43)	-24.42	(23.00)	-38.42	(23.92)	-75.21	(14.42)
Dummy Variables (Area Planted)	-22.47	(19.22)	-27.87	(19.43)	-43.13	(19.74)	-82.53	(19.67)
Cotton								
Dummy Variables (Baseline)	-21.71	(6.58)	-21.59	(5.53)	-31.08	(5.59)	-67.18	(7.97)
Dummy Variables (Linear Interpolation)	-22.74	(6.67)	-23.11	(5.79)	-35.21	(6.56)	-76.71	(12.53)
Dummy Variables (Area Planted)	-29.26	(7.75)	-29.79	(6.80)	-42.36	(7.30)	-80.42	(12.28)

Notes: Table lists predicted climate change impacts under various sensitivity checks for the slowest warming scenario (B1) as well as the most rapid warming scenario (A1FI). The dummy variables models are the baseline results reported in Table 3. To show stability of the estimates we report results when using various geographic and temporal subsets of corn. The last two rows for each crop report results when (i) we use a linear rather than sinusoidal interpolation between minimum and maximum temperature in each day to derive the time a crop is exposed to each temperature and (ii) yields are calculated as total production divided by the land area planted instead of the land area harvested.

The stability of predicted impacts across models and subsets of data shows the robustness of the general findings to specification and sources of climate variation used to identify the model. Predictions are similar whether just warmer southern counties, cooler northern counties, or all counties are used for estimation. Predictions are similar whether earlier or more recent half of the sample is used for estimation. Predictions are also similar whether derived from aggregate time-series variation in aggregate weather or derived from the cross-section of average county-level yields and climate outcomes. This stability is also replicated across all three crops.

Stability of the predictions across models and sources of identification lends strong support to the idea that the underlying weather-yield relationships are in fact causal. It also suggests that scope for adaptation within a crop species, at least using current and historical seed varieties and management strategies, is limited. Where identification using just time series variation uses arguably random year-to-year weather variation and thus reflects a causal link, it accounts for little adaptation. In contrast, identification using the cross-section of average climate outcomes compares warmer and cooler areas, and much like a hedonic model, implicitly accounts for farmers' managerial adjustments in response to differing climates. And predictions based on the cross-section are robust to controls for soils, which suggest omitted variables biases are less likely.

6 Conclusions

This paper examines links between US corn, soybeans, and cotton yields to a new fine-scale data set of daily weather records and that considers the entire distribution of temperatures within each day and each county. We find a robust and significant nonlinear relationship between temperature and yields that shows yields increasing in temperature up to a critical threshold of 29°C for corn, 30°C for soybeans, and 32°C for cotton, above which higher temperatures significantly harm yields. Our model is significantly better at predicting yields than existing statistical models in the literature. Moreover, the sharp nonlinearity has strong implications for a warming climate.

The same basic yield-temperature relationship is observed for various subsets of the data, such as warmer southern states, cooler northern states, and in both earlier and later years of the sample. It is also observed if statistically identified using only time-series (year-to-year variation in aggregate temperature and yield outcomes) or cross-sectional (variation in county-average yields in relation to average temperature distributions) sources of variation.

Furthermore, corn futures respond sharply to extremely warm temperatures suggesting that financial markets expect extreme temperatures to drive down yields. The basic relationship is also robust to the definition of the growing season and to various controls for soils and unobserved factors, accounted for using county fixed effects. Taken collectively, these findings suggest the statistical link is both causal and robust to many kinds of adaptation. They also imply that relative heat tolerance has not increased over time, which may be surprising given technological change has led to a three-fold increase in corn yields over the sample period examined.

We use the estimated link between weather and yields to derive yield predictions under the latest climate change scenarios. Predicted damages are large, highly significant, and robust to various model specifications. Using our preferred model, nationwide average yields for corn, soybeans, and cotton yields for the years 2070-2099 are predicted to decline by 43%, 36%, and 31%, respectively, under the slow-warming scenario (B1), and by 79%, 74%, and 67% under the rapid warming scenario (A1FI).

These rather dire predictions hold growing areas and planting dates fixed. A sensitivity check shows that moving planting dates one month forward would mitigate damages somewhat but would still result in sizeable impacts. While more elaborate structural models are needed to estimate shifts in growing regions, the results and data presented here will provide valuable inputs into such an analysis.[28]

In future work, it may be fruitful to apply some of the basic techniques developed here for predicting future crop choices and/or for modeling weather and climate effects more generally. Such efforts may facilitate estimates of more comprehensive equilibrium price and welfare effects likely to occur. While the analysis presented here cannot speak to these broader implications, it sets forth a valuable new data set, statistical approach, and pattern of yield responses to temperatures. We see the estimated yield responses as a critical first step in such an analysis. Moreover, the findings do suggest that large shifts in the supply of food and fiber are a distinct possibility.

[28]Yield declines may also be offset by CO_2-fertilization. Plants use CO_2 as an input in the photosynthesis process and increasing CO_2 levels might spur plant growth. While higher CO_2 concentrations may boost yields, the magnitude of the effect is still debated. Long et al. (2005) and Long et al. (2006) recently stressed that existing laboratory studies and field experiments might overestimate this effect.

References

Adams, Richard M., Ronald A. Flemming, Ching-Chang Chang, Bruce A. McCarl, and Cynthia Rosenzweig, "A Reassessment of the Economic Effects of Global Climate Change on U.S. Agriculture," *Climatic Change*, June 1995, *30* (2), 147–167.

Ashenfelter, Orley and Karl Storchmann, "Using a Hedonic Model of Solar Radiation to Assess the Economic Effect of Climate Change: The Case of Mosel Valley Vineyards," *NBER Working Paper 12380*, July 2006.

Black, J. Roy and Stanley R. Thompson, "Some Evidence on Weather-Crop-Yield Interaction," *American Journal of Agricultural Economics*, August 1978, *60* (3), 540–543.

Brown, Robert. A. and Norman J. Rosenberg, "Climate Change Impacts on the Potential Productivity of Corn and Winter Wheat in Their Primary United States Growing Regions.," *Climatic Change*, January 1999, *41* (1), 73–107.

Campbell, Sean D. and Francis X. Diebold, "Weather Forecasting for Weather Derivatives," *Journal of the American Statistical Association*, March 2005, *100* (469), 6–16.

Conley, Timothy G., "GMM Estimation with Cross Sectional Dependence," *Journal of Econometrics*, September 1999, *92* (1), 1–45.

Darwin, Roy, "A Farmers view of the ricardian approach to measuring agricultural effects of climate change," *Climatic Change*, March-April 1999, *41*, 371–411.

Deschenes, Olivier and Michael Greenstone, "The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather," *American Economic Review*, March 2007, *97* (1), 354–385.

Diebold, Francis X. and Roberto S. Mariano, "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, July 1995, *13* (3), 253–263.

Fisher, Anthony C., W. Michael Hanemann, Michael J. Roberts, and Wolfram Schlenker, "Potential Impacts of Climate Change on Crop Yields and Land Values in U.S. Agriculture: Negative, Significant, and Robust," *Working Paper*, 2007.

Kelly, David L., Charles D. Kolstad, and Glenn T. Mitchell, "Adjustment costs from environmental change," *Journal of Environmental Economics and Management*, November 2005, *50* (3), 468–495.

Long, Stephen P., Elizabeth A. Ainsworth, Andrew D.B. Leakey, and Patrick B. Morgan, "Global food insecurity. Treatment of major food crops with elevated carbon dioxide or ozone under large-scale fully open-airconditins suggests recent models may have overestimated future yields.," *Philosophical Transaction of the Royal Society B*, 2005, *360*, 2011–2020.

———, ———, ———, **Josef Nosberger, and Donald R. Ort**, "Food for Thought: Lower-Than-Expected Crop Yield Simulation with Rising CO2 Concentrations," *Science*, 30 June 2006, *312* (5782), 1918–1921.

Mearns, L. O., W. Easterling, C. Hays, and D. Marx, "Comparison of Agricultural Impacts of Climate Change Calculated from High and Low Resolution Climate Change Scenarios: Part I. The Uncertainty Due to Spatial Scale," *Climatic Change*, November 2001, *51* (2), 131–172.

Mendelsohn, Robert, William D. Nordhaus, and Daigee Shaw, "The Impact of Global Warming on Agriculture: A Ricardian Analysis," *American Economic Review*, September 1994, *84* (4), 753–771.

Nakicenovic, Nebojsa, ed., *Special Report on Emissions Scenarios*, Cambridge University Press, 2000.

Ritchie, J. T. and D. S. NeSmith, "Temperature and Crop Development," in John Hanks and J. T. Richie, eds., *Modeling Plant and Soil Systems. Agronomy 31*, American Society of Agronomy Madison, WI 1991, pp. 5–29.

Schlenker, Wolfram and Michael Roberts, "Non-linear Effects of Weather on Corn Yields," *Review of Agricultural Economics*, 2006, *28* (3), 391–398.

———, **W. Michael Hanemann, and Anthony C. Fisher**, "Will U.S. Agriculture Really Benefit From Global Warming? Accounting for Irrigation in the Hedonic Approach.," *American Economic Review*, March 2005, *95* (1), 395–406.

———, ———, **and** ———, "The Impact of Global Warming on U.S. Agriculture: An Econometric Analysis of Optimal Growing Conditions," *Review of Economics and Statistics*, February 2006, *88* (1), 113–125.

Sinclair, Thomas R. and Noam G. Seligman, "Crop modeling: from infancy to maturity," *Agronomy Journal*, September-October 1996, *88* (5), 698–704.

——— **and Noam Seligman**, "Criteria for publishing papers on crop modeling," *Field Crops Research*, November 2000, *68* (3), 165–172.

Snyder, R. L., "Hand Calculating Degree Days," *Agricultural and Forest Meterology*, October 1985, *35* (1-4), 353–358.

Stockle, Claudio O., Marcello Donatelli, and Roger Nelson, "CropSyst, a cropping systems simulation model," *European Journal of Agronomy*, January 2003, *18* (3-4), 289–307.

Thom, H. C. S., "Normal Degree Days Above Any Base by the Universal Truncation Coefficient," *Monthly Weather Review*, July 1966, *94* (7), 461–465.

Timmins, Christopher, "Endogenous Land Use and the Ricardian Valuation of Climate Change," *Environmental and Resource Economics*, January 2006, *33* (1), 119–142.

Tubiello, Francesco N., Jean-Francois Soussana, and S. Mark Howden, "Crop and pasture response to climate change," *Proceedings of the National Academy of Sciences of the United States*, December 11 2007, *104* (50), 19686–19690.