The Impact of Climate Change on Indian Agriculture

Raymond Guiteras*,†
University of Maryland

September 2009

Abstract

This paper estimates the impact of climate change on Indian agriculture. I use a 40-year district-level panel data set covering over 200 Indian districts to estimate the effect of random year-to-year variation in weather on agricultural output. These panel estimates incorporate farmers' within-year adaptations to annual weather shocks. These estimates, derived from short-run weather effects, are relevant for predicting the medium-run economic impact of climate change if farmers are unable to adapt quickly. I find that projected climate change over the period 2010-2039 reduces major crop yields by 4.5 to 9 percent. The long-run (2070-2099) impact is dramatic, reducing yields by 25 percent or more in the absence of long-run adaptation. These results suggest that climate change is likely to impose significant costs on the Indian economy unless farmers can quickly recognize and adapt to increasing temperatures. Such rapid adaptation may be less plausible in a developing country, where access to information and capital is limited.

Keywords: climate change, India, agriculture, panel data

JEL Classification Codes: Q54, Q56, O13, C23

*I thank my advisors, Esther Duflo and Michael Greenstone, for their encouragment and guidance. I thank C. Adam Schlosser of MIT's Center for Energy and Environmental Policy Research for guidance with the NCC data set and Kavi Kumar of the Madras School of Economics for guidance with the World Bank India Agriculture and Climate Data Set. Nivedhitha Subramanian and Henry Swift provided excellent research assistance. I thank Daron Acemoglu, Joseph Aldy, Miriam Bruhn, Jessica Cohen, Greg Fischer, Douglas Gollin, Rick Hornbeck, Jeanne Lafortune, Doug Miller, Ben Olken, José Tessada, Robert Townsend and Maisy Wong for their comments and suggestions. The paper also benefited from comments by seminar participants at University of British Columbia, University of California Merced, Colgate University, Harvard University Kennedy School of Government, MIT, Resources for the Future, Smith College, University of Maryland, University of North Carolina and Williams College. All errors are my own.

†Department of Economics; University of Maryland; College Park, MD 20742; email: guiteras@econ.umd.edu.

1 Introduction

As the scientific consensus grows that significant climate change, in particular increased temperatures and precipitation, is very likely to occur over the 21st century (Christensen and Hewitson, 2007), economic research has attempted to quantify the possible impacts of climate change on society. Since climate is a direct input into the agricultural production process, the agricultural sector has been a natural focus for research. The focus of most previous empirical studies has been on the US, but vulnerability to climate change may be greater in the developing world, where agriculture typically plays a larger economic role. Credible estimates of the impact of climate change on developing countries, then, are valuable in understanding the distributional effects of climate change as well as the potential benefits of policies to reduce its magnitude or promote adaptation. This paper provides evidence on the impact of climate change on agriculture in India, where poverty and agriculture are both salient. I find that climate change is likely to reduce agricultural yields significantly, and that this damage could be severe unless adaptation to higher temperatures is rapid and complete.

Most previous studies of the economic effects of climate change have followed one of two methodologies, commonly known as the *production function approach* and the *Ricardian approach*. The production function approach (also known as *crop modeling*) is based on controlled agricultural experiments, where specific crops are exposed to varying climates in laboratory-type settings such as greenhouses, and yields are then compared across climates. This approach has the advantage of careful control and randomized application of environmental conditions. However, these laboratory-style outcomes may not reflect the adaptive behavior of optimizing farmers. Some adaptation is modeled, but how well this will correspond to actual farmer behavior is unclear. If farmers' actual practices are more adaptive, the production function approach is likely to produce estimates with a negative bias. On the other hand, if the presumed adaptation overlooks constraints on farmers' adaptations or does not take adjustment costs into account, these estimates could be overoptimistic.

The Ricardian approach, pioneered by Mendelsohn et al. (1994), attempts to allow for the full range of compensatory or mitigating behaviors by performing cross-sectional regressions of land prices on county-level climate variables, plus other controls. If markets are functioning well, land prices will reflect the expected present discounted value of profits from all, fully adapted uses of land, so, in principle, this approach can account for both the direct impact of climate on specific crops as well farmers' adjustment of production techniques, substitutions of different crops and even exit from agriculture. However, the success of the Ricardian approach depends on being able to account fully for all factors correlated with climate and influencing agricultural productivity. Omitted variables, such as unobservable farmer or soil quality, could lead to bias of unknown sign and magnitude.[1]

More recently, economists studying the US have turned to a panel data approach, using presumably random year-to-year fluctuations in realized weather across US counties to estimate the effect of weather on agricultural output and profits (Deschênes and Greenstone, 2007; Schlenker and Roberts, 2009). This fixed-effects approach has the advantage of controlling for time-invariant district-level unobservables such as farmer quality or unobservable aspects of soil quality. Furthermore, unlike the production function approach, the use of data on actual field outcomes, rather than outcomes in a laboratory environment, means that estimates from panel data will reflect intra-year adjustments by farmers, such as changes in inputs or cultivation techniques. However, by measuring effects of annual fluctuations, the panel data approach does not reflect the possibility of longer-term adaptations, such as crop switching or exit from farming.

Agriculture typically plays a larger role in developing economies than in the developed world. For example, agriculture in India makes up roughly 20% of GDP and provides nearly 52% of employment (as compared to 1% of GDP and 2% of employment for the US), with

[1] Deschênes and Greenstone (2007) and Fisher et al. (2009) discuss the reliability of hedonic methods in the U.S. context.

the majority of agricultural workers drawn from poorer segments of the population (FAO, 2006). Furthermore, it is reasonable to expect that farmers in developing countries may be less able to adapt to climate change due to credit constraints or less access to adaptation technology. However, the majority of the economics literature on the impact of climate change has focused on developed countries, in particular the US, presumably for reasons of data availability. Most research in developing countries has followed the production function approach, finding alarmingly large possible impacts (Cruz et al., 2007). A true Ricardian study would be difficult to carry out in a developing country context, because land markets are less likely to be well-functioning and data on land prices are not generally available. Instead, a *semi-Ricardian* approach has used data on average profits instead – the idea is that the land price, if it were available, would just be the present discounted value of profits. The major developing country semi-Ricardian studies, of India and Brazil, found significant negative effects, with a moderate long-run climate change scenario (an increase of $2.0°C$ in mean temperature and seven percent increase in precipitation by the end of the 21st century) leading to losses on the order of 10% of agricultural profits (Sanghi et al., 1998b, 1997).

This paper applies the panel data approach to agriculture in India, using a panel of over 200 districts covering 1960-1999.[2] The basic estimation strategy, following Deschênes and Greenstone (2007), is to regress yearly district-level agricultural outcomes (in this case, yields) on yearly climate measures (temperature and precipitation) and district fixed effects. The resulting weather parameter estimates, then, are identified from district-specific deviations in yearly weather from the district mean climate. Since year-to-year fluctuations in the weather are essentially random and therefore independent of other, unobserved determinants

[2] Auffhammer et al. (2006) also employ the panel data methodology to study Indian agriculture, rice in particular. Their study uses state-level data on rice output and examines the impact of climate as well as atmospheric brown clouds, the byproduct of emissions of black carbon and other aerosols. They find a negative impact of increased temperature, as does this paper. Dell et al. (2008) examine a broad panel of countries over 50 years, and find that higher temperatures reduce growth of aggregate output in poor countries. Felkner et al. (2009) provide a detailed study of the impact of weather fluctuations on 137 rice-cultivating households in Sisaket province, Thailand, over 5 years.

of agricultural outcomes, these panel estimates should be free of the omitted variables problems associated with the hedonic approach. The use of district-level data is important to obtain adequate within-year climate variation, thereby distinguishing climate impacts from other national-level yearly shocks. I also include smooth regional time trends so that the effect of a slowly warming climate over the second half of the twentieth century is not confounded with improvements in agricultural productivity over the same period. The predicted mean impact of climate change is then calculated as a linear combination of the estimated weather parameters and the predicted changes in climate.

The paper finds significant negative impacts, with medium-term (2010-2039) climate change predicted to reduce yields by 4.5 to nine percent, depending on the magnitude and distribution of warming. Long-run climate change (2070-2099) is even more detrimental, with predicted yields falling by 25 percent or more. Because these large changes in long-run temperatures will develop over many decades, farmers will have time to adapt their practices to the new climate, likely lessening the negative impact. However, estimates from this panel data approach may be more relevant for the medium-run scenario, since, as the paper's theoretical section argues, developing country farmers face significant barriers to adaptation, which may prevent rapid and complete adaptation.

This negative impact of climate change on agriculture is likely to have a serious impact on poverty: recent estimates from across developing countries suggest that one percentage point of agricultural GDP growth increases the consumption of the three poorest deciles by four to six percentage points (Ligon and Sadoulet, 2007). The implication is that climate change could significantly slow the pace of poverty reduction in India.

2 Theoretical Framework

Because this paper will attempt to estimate the impacts of climate change based on the effects of annual fluctuations in the weather, it is worthwhile to consider the relationship between the two.

2.1 Short-run weather fluctuations versus long-term climate change

Consider the following simple model of farmer output. A representative farmer's production function is $f(T, L, K)$, where T represents temperature, L represents an input that can be varied in the short run, which we shall call labor for concreteness, and K represents an input that can only be varied in the long run, which we call capital. Labor and capital should not be thought of literally, nor are the distinctions between inputs that are flexible in the short and long run so sharp in reality. The point is that some inputs, such as fertilizer application or labor effort are relatively easy to adjust, while other inputs, such as crop choice or irrigation infrastructure, may be more difficult to adjust or may be effectively fixed at the start of the growing season. The farmer, taking price and temperature as given, solves the following program:

$$\max\left\{p \cdot f(T, L, K) - wL - rK\right\} \tag{1}$$

where for simplicity we assume linear input costs. For a given temperature T, with all inputs fully flexible, the farmer will choose profit-maximizing $L(T)$ and $K(T)$ and obtain a maximized profit of $\pi(T, L(T), K(T))$. Now consider a small change in temperature to $T' > T$. First, consider the case where the farmer is not allowed to make any changes, i.e. L and K are held fixed at $L(T)$ and $K(T)$, respectively. In this case, the farmer obtains profit $\pi(T', L(T), K(T))$. To the extent that the production function approach discussed in the introduction understates or ignores the possibility of adaptation, that approach estimates the effect of climate change on profits as $\widehat{\Delta\pi}_{PF} = \pi(T', L(T), K(T)) - \pi(T, L(T), K(T))$.

Next, consider the case where the farmer can carry out short-run adjustments, which in this model we capture as reoptimizing L, but is constrained from long-run adjustments of

K. In this case, the farmer obtains $\pi(T', L(T'), K(T))$.[3] The panel data approach followed in this paper, where farmers are free to make all intra-season adjustments but not longer-run adjustments, estimates the effect of climate change as $\widehat{\Delta\pi}_{FE} = \pi(T', L(T'), K(T)) - \pi(T, L(T), K(T))$.

Finally, consider the case where the farmer is allowed to reoptimize all factors. In this case, the farmer obtains $\pi(T', L(T'), K(T'))$ and the true effect of climate change is $\Delta\pi = \pi(T', L(T'), K(T')) - \pi(T, L(T), K(T))$. Since greater choice can only help the farmer, we have

$$\Delta\pi \geq \widehat{\Delta\pi}_{FE} \geq \widehat{\Delta\pi}_{PF} \tag{2}$$

This framework is illustrated in Figure 1. The first point to note is that the panel data approach should better approximate the true effect of climate change than a production function approach that does not allow for adaptation. The second point is that, for small changes in climate, the panel data approach may provide a reasonable approximation to the true effect of climate change. However, for large changes in climate, the panel data approach will overstate the costs of climate change relative to the true long-run cost, when farmers have re-optimized.

Furthermore, the panel data approach may also provide a reasonable approximation if farmers are unable to reoptimize along some margins or do so only slowly. If long-term reoptimization is slow or incomplete, it is plausible that the panel data approach will provide a good estimate of the costs incurred over the medium run, while not all adjustments have been carried out. There are several reasons to expect that agricultural practice may adapt slowly to climate change. First, the signal of a changing mean climate will be difficult to extract from the year-to-year weather record. The IPCC calculates that a discernible signal of a warmer mean climate for the South Asian growing season will take 10-15 years to emerge

[3]For simplicity of notation, we suppress the fact that the optimal labor effort will in general reflect the fact that capital is fixed, i.e. the actual optimal labor input will be $L(T'|K = K(T))$ and in general this is different than $L(T')$.

from the annual noise (Christensen and Hewitson, 2007). This is for South Asia as a whole, greater noise in particular locations will slow the signal's emergence further. If farmers' practices are based on perceived weather distributions based on historical experience, then this difficulty in discerning climate change could lead to farming practices significantly out of phase with the true optimum. Second, many of the investments associated with long-term reoptimization – new irrigation, new crop varieties, or migration – involve both fixed costs and irreversibilities, both of which can delay investment in the presence of uncertainty (Bertola and Caballero, 1994; Dixit and Pindyck, 1994).

Additionally, it is reasonable to expect developing country agriculture to face even greater difficulties adjusting. Incomplete capital markets, poor transmission of information, and low levels of human capital are all pervasive and likely to slow adaptation. Topalova (2005) provides evidence that factors, especially labor, are relatively immobile in India. Furthermore, slow adaptation of profitable agricultural practices is a long-standing puzzle in development economics (Foster and Rosenzweig, 1995; Duflo et al., 2005).

Empirically, the impacts of climate shocks appear to persist strongly. In a study of the U.S. Dust Bowl, Hornbeck (2008) finds that long-term adjustments recovered only 14%−28% of short-run costs. Dell et al. (2009) compare climate-income associations from municipal-level data on labor income for 11 Latin American countries and the U.S. to results from their panel study of 136 countries over 50 years (Dell et al., 2008) and find that adaptation offsets only half of the negative impact of higher temperatures.

2.2 Caveats

Several important caveats may limit the applicability of the above model. First, data on annual agricultural profits are not available.[4] This paper will use data on annual yields

[4] Sanghi et al. (1998b) use average profits over a 20-year period. Their imputed labor inputs are based on agricultural labor quantities measured by decadal censuses, with linear interpolations for non-census years. This is appropriate for their purpose, which is to asses the relationship between average climate and average profits, but not appropriate for this paper, where the emphasis is on annual fluctuations.

(output per hectare) as a proxy instead and explore the impact on those inputs for which annual data are available. This may overstate the impact on welfare if farmers reduce their use of inputs in response to a negative weather shock. The empirical analysis explores the impacts on those inputs for which yearly, district-level data are available. Second, it is not possible for a panel study to assess the impact of weather on output through its effects on stock inputs. For example, if climate change hurts agriculture by depleting aquifers but one year's drought does not appreciably deplete an aquifer, the panel data approach will not capture this effect. Finally, the panel approach cannot assess the impact of variables that vary only slowly over time. For example, it is believed that the same increased levels of carbon dioxide (CO_2) that are causing global warming may be beneficial to agriculture, since carbon dioxide is important to plant development.[5] Since the level of CO_2 changes only slowly, it is not possible to separate its effect from that of, for example, smooth technological progress over time. However, since CO_2 levels are roughly constant across space, the Ricardian approach is not able to capture this effect either.

3 Data Sources and Summary Statistics

The analysis is performed on a detailed 40-year panel of agricultural outcomes and weather realizations covering over 200 districts. Although Indian districts are generally somewhat larger than US counties, the district is the finest administrative unit for which reliable data are available. This section describes the data and provides some summary statistics.

3.1 Agricultural outcomes

Detailed district-level data from the Indian Ministry of Agriculture and other official sources on yearly agricultural production, output prices and acreage planted and cultivated for 271

[5] Recent research in the crop modelling school has cast doubt on the magnitude of beneficial effects from CO_2 fertilization (Long et al., 2006).

districts over the period 1956-1986 have been collected into the "India Agriculture and Climate Data Set" by a World Bank research group, allowing computation of yield (revenues per acre) and total output (Sanghi et al., 1998a). This dataset covers the major agricultural states with the exceptions of Kerala and Assam. Also absent, but less important agriculturally, are the minor states and Union Territories in northeastern India, and the northern states of Himachal Pradesh and Jammu-Kashmir. These 271 districts are shown in Figure 2.A. The production, acreage and price data for major crops were extended through 1999 by Duflo and Pande (2007), allowing computation of yields (output per acre) for these major crops.[6] 218 districts have data for all years 1960-1999; these are the districts that will be included in the regressions. These districts are mapped in Figure 2.B. The bulk of the districts lost are in the East, in particular Bihar and West Bengal.

Because markets are not well-integrated, local climate shocks could affect local prices. These price effects make estimating effects on revenue undesirable. While the price response to a negative climate shock will reduce the impact on farmers, calculating the effect of climate on revenues will ignore the effect on consumer surplus. In this context, the impact on *yields* better approximates the overall welfare effects, as pointed out by Cline (1992). To avoid these potential pitfalls from endogenous prices, then, I hold prices fixed at their 1960-1965 averages.

The World Bank dataset also includes input measures, such as tractors, plough animals and labor inputs, as well as prices for these inputs. However, many of these inputs, in particular the number of agricultural workers, are only measured at each 10-year census, with annual measures estimated by linear interpolation. This precludes construction of annual profits data, a theoretically preferable measure. This paper will use data on fertilizer inputs, the agricultural wage and the extent of double-cropping, each of which is measured annually at the district level, to estimate the extent of within-year adaptation to negative

[6] The six major crops are rice, wheat, jowar (sorghum), bajra (millet), maize and sugar. These comprise roughly 75% of total revenues.

climate shocks.

3.2 Weather data

Recent research in economics and agricultural science has pointed to the importance of daily fluctuations in temperature for plant growth (Schlenker and Roberts, 2009). Commonly available data, such as mean monthly temperature, will mask these daily fluctuations, so it is important to obtain daily temperature records. Recent economics research in the US has used daily records from weather stations to construct daily temperature histories for US counties. However, the publicly available daily temperature data for India are both sparse and erratic. The main clearinghouse for daily data, the Global Summary of the Day (GSotD, compiled by the US National Climatic Data Center on behalf of the World Meteorological Organization) has at most 90 weather stations reporting on any one day and contains major gaps in the record – for example, there are no records at all from 1963–1972. Furthermore, these individual stations' reports come in only erratically – applying a reasonable sample selection rule such as using stations that report at least 360 days out of the year or 120 days out of the 122 day growing season would yield a database with close to zero observations.

To circumvent this problem, I use data from a gridded daily dataset that use non-public data and sophisticated climate models to construct daily temperature and precipitation records for $1^\circ \times 1^\circ$ grid points (excluding ocean sites). This data set, called NCC (NCEP/NCAR Corrected by CRU), is a product of the Climactic Research Unit, the National Center for Environmental Prediction / National Center for Atmospheric Research and the Laboratoire de Météorologie Dynamique, CNRS. NCC is a global dataset from which Indian and nearby grid points were extracted, providing a continuous record of daily weather data for the period 1950-2000 (Ngo-Duc et al., 2005). To my knowledge, this paper is the first use of these data in the economics literature. To create district-level weather records from the grid, I use a weighted average of grid points within 100 KM of the district's geographic center.[7] The weights are the inverse square root of the distance from the district center. Both the actual correlation of weather patterns across space and this method of mapping weather to districts will lead to spatial correlation in the econometric analysis. In the econometric methods section below, I detail my approach to accounting for this correlation.

I employ two methods to convert these daily records to yearly weather metrics for analysis. The first, *degree-days*, reflects the importance of cumulative heat over the growing season, but may fail to capture important nonlinear effects. The second, less parametric approach, counts the number of growing-season days in each one-degree C temperature bin. This approach is more flexible, but imposes a perhaps-unrealistic additive separability assumption. However, the results are similar between the two approaches. Details of the methods follow.

3.2.1 Temperature: Degree-days

Agricultural experiments suggest that most major crop plants cannot absorb heat below a temperature threshold of $8°C$, then heat absorption increases roughly linearly up to a threshold of $32°C$, and then plants cannot absorb additional heat above this threshold. I follow the standard practice in agronomics, then, by converting daily mean temperatures to *degree-days* by the formula

$$D(T) = \begin{cases} 0 & \text{if } T \leq 8°C \\ T - 8 & \text{if } 8°C < T \leq 32°C \\ 24 & T \geq 32°C \end{cases}$$

Degree-days are then summed over the summer growing season, which for India is defined as the months of June through September, following Kumar et al. (2004). Fixing the growing season avoids endogeneity problems with farmers' planting and harvesting decisions. It should be noted that the degree-day thresholds were developed in the context of US agri-

[7] Alternative radii did not appreciably affect the district-level records.

culture. Crops cultivated in a warmer climate may have different thresholds, in particular a higher upper threshold. For comparability with other research, I use the standard $8°C$ and $32°C$ thresholds in the empirical results that follow, but the results are not sensitive to the use of alternative upper thresholds $(33°C, 34°C)$. I also allow for the possibility that heat in excess of a threshold may be damaging by including a separate category of *harmful degree-days*. Each day with mean temperatures above $34°C$ is assigned difference between that day's mean temperature and $34°C$; these harmful degree-days are then summed over the growing season. Again, the results are not sensitive to alternate thresholds $(33°C,\ 35°C)$.

3.2.2 Temperature: One-degree bins

Schlenker and Roberts (2009) emphasize the importance of using daily records in the context of nonlinear temperature effects. Consider the following simple example: imagine that increased temperature is initially beneficial for plants, but then drastically damaging above $30°C$. Consider two pairs of days, the first pair with temperatures of $(30°C, 30°C)$ and the second pair with temperatures of $(29°C, 31°C)$. Although both pairs have the same mean temperature, their effects on yields will be very different, with the second much less beneficial. To capture such potential nonlinearities, I employ a nonparametric approach, counting the total number of growing season days in each one-degree C interval and including these totals as separate regressors. That is, for each grid point g, I construct

$$T_{c,g,y} = \{\#\text{ of growing season days with mean temperature in the interval } \left((c-1)\,°C, c°C\right]\}$$

for year y and for each of $c = 1, \ldots, 50$. To obtain district-level measures from these measures at each grid point, I again take the weighted average of the number of days in that bin for each grid point within 100KM of the district center.

It is important to emphasize that the district-level bins are constructed by averaging over grid point temperature bins rather than constructing bins of district center temperatures.

Again, this is necessary to account for potential nonlinear temperature effects. To understand the reasoning, consider the following simplified example of a district center equidistant between two grid points. Suppose these are the only two grid points within 100 KM of the district center. As above, imagine that increased temperature is initially beneficial for plants, but then drastically damaging above $30°C$. Now suppose that one of the two grid points has a mean temperature of $29°C$ every day while the other grid point has a mean temperature of $31°C$. The mean temperature calculated at the district center will be $30°C$ each day, but the bin-by-bin experience of the district as a whole would be better captured by assigning half a day to each of the bins corresponding to $29°C$ and $31°C$. This methodology does lead to districts having fractional number of days in bins, but the total over all bins still sums to 122, the number of days in the growing season, for each district.

The mean number of days in each bin across all districts is plotted in Figure 3. Because of the scarcity of observations above $38°C$ and below $22°C$, each of these will be collected into single bins. The tradeoff here is between precision of estimation (aided by grouping these observations) and estimation of nonlinearities at extreme temperatures.

3.2.3 Precipitation

Precipitation data are summed by month to form total monthly precipitation for each month of the growing season, during which the vast majority of annual precipitation occurs. Including separate monthly measures, rather than merely summing over the growing season, allows the timing of precipitation, in particular the arrival of the annual monsoon, to affect output. To test robustness, I also run regressions with total growing season precipitation.

3.3 Climate change predictions

I compute estimated impacts for three climate change scenarios. First, I examine the short-term (2010-2039) South Asia scenario of the Intergovernmental Panel on Climate Change's latest climate model (Cruz et al., 2007), which is an increase of $0.5°C$ in mean temperature and four percent precipitation for the growing season months of June–September. This scenario corresponds to the "business-as-usual" or highest emissions trajectory, denoted A1F1 in the IPCC literature. However, because most of the short-run component of climate change is believed to be "locked-in", i.e. already determined by past emissions, these short run projections are not very sensitive to the emissions trajectory. For example, the short-term South Asia scenario associated with the lowest future emissions trajectory, denoted B1 in the IPCC literature, differs by less than $0.05°C$ for the growing season months. The impact of this scenario on the distribution of growing season temperatures is plotted in Figure 4.

The IPCC does not report higher moments of predictions for this consensus scenario. However, considering just a mean shift in temperatures would overlook the potentially important effects of the distribution of temperatures, in particular nonlinearities at temperature extremes. Furthermore, this consensus scenario is given as a uniform change across all regions, whereas it is likely that climate change will develop differently across different regions of India. To assess the effects of changing distributions of temperatures and to account for regional differences, I use daily predictions from the Hadley Climate Model 3 (HadCM3) data produced by the British Atmospheric Data Centre for the A1F1 business-as-usual scenario. These predictions are given for points on a 2.5° latitude by 3.5° longitude grid. I calculate the average number of days in each one-degree interval, by region, for the years 1990-1999, 2010-2039 and 2070-2099. The changes in the distribution of temperatures are then applied to the district-level temperature distributions derived from the historical NCC data to obtain district-level changes in temperature distributions.[8] These changes are plotted in Figures 5 and 6. The contrast with the mean-shift scenario of the IPCC is apparent

[8]The Hadley data for 1990-1999 display both a higher mean and variance than the NCC data for the same period. Since the estimation of temperature effects is performed with the NCC data, calculating projected impacts using temperature changes based on the Hadley data would not be properly scaled. In the projections, I rescale the level of the Hadley data so that the 1990s means by region match the 1990s NCC data. I also rescale the spread so that the root mean squared errors around each gridpoint's monthly mean match for the 1990s.

in the greater relative mass in the right tails. Significantly, the *increase* in the mean number of growing-season days with temperatures above $38°C$ is greater than the mean *number* of such days in the historical data: while the average district experienced just 0.4 such days observed per year in the historical data, the mean number of days is expected to increase by nearly 2 for the period 2010-2039 and nearly 10 for the period 2070-2099. Because the effect of these extreme temperatures is only imprecisely estimated, this will add uncertainty to the estimated impacts.

3.4 Summary statistics

Summary statistics of the key variables of interest are presented in Table 1.A. This table presents the sample used in the analysis, covering 1960-1999 and including only the 218 districts with full records of output and yields. Noteworthy points in this table include the high productivity, irrigation and use of high-yield varieties (HYV) of the Northern states (Haryana, Punjab and Uttar Pradesh). Significant poverty reduction, defined relative to state- and sector-specific thresholds for minimum adequate calorie consumption, is also visible, although poverty remains high, especially in the Eastern states. Panel I of Table 1.B compares the 218 districts analyzed in this paper to the sample of 271 districts for the period 1966–1986 studied in Sanghi et al. (1998b) (referred to as SMD98 hereafter). The two samples are very similar. Panel II of Table 1.B looks at the 218 districts over time. Noteworthy trends include the increase in agricultural productivity revealed by increasing yields, the large increase in irrigation an high-yield varieties, and warming (observed in mean temperatures and degree-days).

3.5 Residual variation

Because this paper uses district fixed-effects to strip out time-invariant unobservables that could be confounded with mean climate, it is important to consider how much variation in climate will be left over after these fixed effects and other controls have been removed. This section assess the extent of this residual variation.

3.5.1 Mean temperatures, degree-days and precipitation

Table 2.A reports the results of an exercise designed to assess the extent of residual variation in mean temperatures, degree-days and precipitation. I regress each weather measure on various levels of fixed effects – none, district, district and year, district and region-by-year, district and state-by-year. The residual from this regression is a measure of remaining variation. For example, the residual from the regression with no fixed effects is simply the deviation of that district-by-year observation from the grand mean of the sample, the residual from the regression with district fixed effects is the deviation of that district-by-year observation from the district mean, etc. I then count how many observations have residuals of absolute value greater than certain cutoffs – for mean growing season temperature, for example, steps of $0.5°C$ up to $2.5°C$. Ideally, there should be a substantial number of observations with deviations greater than the predicted change in climate. If this is the case, then the effect of weather variation of similar magnitude to the predicted climate change would be identified from the data, rather than from functional form extrapolations.

Unfortunately, the fixed effects do wipe out a great deal of variation. Consider the sixth row of Panel 2, which examines the results for district and year fixed effects for the sample that will be the focus of the regression analysis: the 218 districts with output data for all years 1960-1999. Here, we see that just 15 district-by-year observations differ from the predicted value – which, in this case, is the district mean plus the deviation of the national mean for that year from the national mean for the sample period – by more than 120 degree-days (which corresponds roughly to a $1.0°C$ mean temperature increase), while no observations differ from the predicted value by more than 180 degree-days.

These findings are less than ideal, since they mean that only a few observations are available to identify even small weather fluctuations. To recapture some of this variation, I retreat from year fixed effects and add smooth time trends (linear, quadratic, cubic) to district fixed effects. This way, I remove possible confounding from correlated trends in temperature and technological progress. If yearly weather fluctuations are indeed random, then in expectation they will be uncorrelated with other economic shocks and therefore the consistency of the estimates will not be affected. Looking at the fifth row of Panel 2, we see that we now have 161 observations differing from the predicted value by more than 120 degree-days. Although this is an improvement relative to the year fixed-effects, there is still not an overwhelming amount of variation: we still have no observations differing from the predicted value by more than 240 degree-days. However, not much variation is lost relative to the district fixed-effects alone (the first row of each panel). In Appendix Table 2, I experiment with alternative upper bounds for the degree-day measure, but this does not revive much variation. These results should lead to caution in interpreting predicted impacts for large changes in temperature, since these will depend on functional form assumptions.

In the case of precipitation, there is no lack of underlying variation, as is made clear by Panel 3. Estimates of precipitation effects will be well-identified from the data.

3.5.2 Temperature bins

To assess the extent of the residual variation within temperature bins, I calculate the sum of the absolute value of the residuals from a regression of the value of the bin variable on different levels of fixed effects. That is, for each bin $c =< 20, 21, \ldots, 40, > 40$, I estimate

$$T_{c,d,t} = \sum_f FE_f + \varepsilon_{c,d,t}$$

where $\{FE_f\}$ is some set of fixed effects (e.g. none, district, district and year, district and region-by-year, state-by-year) and calculate the average value of the absolute residuals,

$$\overline{AVR_c} = \frac{1}{D \times T} \sum_{d,t} |\hat{\varepsilon}_{c,d,t}|$$

I also perform similar calculations for regression models incorporating smooth functions of time rather than year fixed effects, e.g.

$$T_{c,d,t} = \sum_d FE_d + \gamma_1 Y + \gamma_2 Y^2 + \gamma_3 Y^3 + \varepsilon_{c,d,t}$$

The results of these calculations of mean sums of absolute residuals are reported in Table 2.B. Each entry represents the mean across districts and years, so the mean times the number of district-by-year observations (here $218 \times 40 = 8270$) yields the number of observations available to identify the effect of that interval. For example, looking at the fifth row, corresponding to the regression model with district fixed effects and a cubic time trend, there are roughly $0.05 \times 8720 \approx 435$ observations available to identify the extremal bin collecting all days with mean temperatures above $40°C$. Because of the scarcity of observations above $38°C$ and below $22°C$, each of these will be collected into single bins. The tradeoff here is between precision of estimation (aided by grouping these observations) and estimation of nonlinearities at extreme temperatures. The results for the specification of district fixed effects and a cubic year trend are plotted in Figure 7.

4 Econometric Strategy

4.1 Semi-Ricardian method

This section describes the econometric framework used in the *semi-Ricardian approach* of Sanghi et al. (1998b) in order to make clear the difference between that approach and the panel approach considered here. The cross-sectional model is

$$\bar{y}_d = \mathbf{X}_d'\beta + \sum \theta_i f_i\left(\bar{W}_{id}\right) + \varepsilon_d \tag{3}$$

where $\bar{y}_d$ is the mean agricultural outcome of interest for district d, $\mathbf{X}_d$ is a vector of observable district characteristics (such as urbanization, soil quality, etc.), $\bar{W}_{id}$ is a climate variable of interest (temperature, precipitation) and ε_d is the error term. In SMD98, the climate variables are monthly mean temperature and precipitation for the months of January, April, July and October, as well as their squares and within-month interactions. As noted above, SMD98 diverge from the traditional Ricardian or hedonic approach by using an average of profits, output and other flow variables rather than land values in a year, for reasons of data availability.

For the coefficients of interest θ_i to be estimated consistently, it is necessary that

$$E\left[f_i\left(\bar{W}_{id}\right)\varepsilon_d \,|\, \mathbf{X}_d\right] = 0$$

for all i. Intuitively, climate must be uncorrelated with unobserved determinants of agricultural productivity, after controlling for observed determinants of agricultural productivity. Note that this requires that the influence of the $\mathbf{X}_d$ has been correctly specified. SMD98 include measures of soil quality, population density and other plausible determinants of agricultural productivity. However, the possibility remains that unobserved determinants of output, such as unobserved soil quality, farmer ability, or even government institutions are correlated with the error term ε_d, which would bias the estimated coefficients $\hat{\theta}_i$ and therefore the imputed impact of climate change.

In the language of the model in Section 2.1, the semi-Ricardian method estimates the impact of a shift in climate from T to T' by comparing observed $\pi\left(T', L\left(T'\right), K\left(T'\right)\right)$ with observed $\pi\left(T, L\left(T\right), K\left(T\right)\right)$, with the observations taking place in two different districts. However, there may be other unobserved components of the profits function, so in truth the semi-Ricardian method would be comparing $\pi\left(T', L\left(T'\right), K\left(T'\right), \tilde{\varepsilon}\right)$ with $\pi\left(T, L\left(T\right), K\left(T\right), \varepsilon\right)$, while the true long-run impact of climate change for the district currently at climate T would be $\pi\left(T', L\left(T'\right), K\left(T'\right), \varepsilon\right) - \pi\left(T, L\left(T\right), K\left(T\right), \varepsilon\right)$.

4.2 Panel approach

This paper follows the panel data approach, estimating

$$y_{dt} = \alpha_d + g_r\left(t\right) + \mathbf{X}'_{dt}\beta + \sum \tilde{\theta}_i f_i\left(W_{idt}\right) + \varepsilon_{dt} \tag{4}$$

There are a number of important differences between equation (4) and equation (3). First, note that the dependent variable, y_{dt}, is a yearly measure rather than an average. In the models estimated below, this is annual yields (output per hectare). Second, the regressors of interest are functions of yearly realized weather W_{idt}, rather than climate averages. Third, as discussed in the theory section, the coefficients on short run fluctuations need not be the same as those on long run shifts, i.e. $\tilde{\theta}_i \neq \theta_i$. Fourth, $g_r\left(t\right)$ controls for smooth changes in productivity over time. Finally, the district fixed effects α_d will absorb any district-specific time-invariant determinants of y_{dt}.

The consistency of fixed-effects estimates of $\tilde{\theta}_i$ rests on the following assumption:

$$E\left[f_i\left(W_{idt}\right)\varepsilon_{dt} \,|\, \mathbf{X}_{dt}, \alpha_d, g_r\left(t\right)\right] = 0$$

Intuitively, $\tilde{\theta}_i$ is identified from district-specific deviations in weather about the district averages after controlling for a smooth time trend. This variation is presumed to be orthogonal to unobserved determinants of agricultural outcomes, so it provides a potential solution to the omitted variables bias problems that could prevent consistent estimation of equation (3).

Because outcomes are likely autocorrelated between years for a given district, feasible generalized least squares (FGLS) estimation of the fixed-effects model can improve efficiency. Examining the residuals from the fixed effects regression reveals that an AR(2) process best fits the data. However, as Hansen (2007) emphasizes, conventional estimation of the parameters of the autocorrelation model are biased in a fixed effects framework, so I compute these parameters using Hansen's bias-corrected method.

While an AR(2) process describes the observed data best, it is unlikely that the true underlying error-generating process is literally AR(2). Therefore, I construct standard errors for the FGLS estimates that are robust to clustering at the district level.[9] Because of the likelihood of spatial correlation, I also compute standard errors that are additionally robust to arbitrary contemporaneous correlation within state, using the multi-way clustering procedure of Cameron et al. (2006).[10] In the empirical results that follow, standard errors robust to clustering at the district level are reported in parentheses, while standard errors robust to clustering at the district and state-by-year level are reported in brackets.

[9]That is, rather than computing the standard error as $\hat{\sigma}^2\left(\widetilde{X}'\hat{\Omega}^{-1}\widetilde{X}\right)^{-1}$ (where $\widetilde{X}$ denotes the regressors with fixed effects removed), which would be appropriate if the data truly were governed by an AR(2) process, I compute $\left(\widetilde{X}'\hat{\Omega}^{-1}\widetilde{X}\right)^{-1}\hat{W}\left(\widetilde{X}'\hat{\Omega}^{-1}\widetilde{X}\right)^{-1}$, where $\hat{W}$ is the robust sum of squared residuals matrix. $\hat{W}=\sum_{j=1}^{N}\hat{u}_j'\hat{u}_j$, where $\hat{u}_j=\sum_{t=1}^{T}\hat{e}_{jt}\tilde{x}_{jt}^*$. $\hat{e}_{jt}$ is the residual from the FE regression and and $\tilde{x}_{jt}^*$ is the j,t row of the matrix $\hat{\Omega}^{-1/2}\tilde{X}$. Wooldridge (2003) and Wooldridge (2006) discuss cluster-robust standard errors for FGLS estimators.

[10]Cameron et al. (2006) show that an estimated variance-covariance matrix robust to clustering in two dimensions can be constructed as $\hat{V}_{1\cup 2}=\hat{V}_1+\hat{V}_2-\hat{V}_{1\cap 2}$, where $\hat{V}_1$ is robust to clustering in dimension 1, $\hat{V}_2$ is robust to clustering in dimension 2, and $\hat{V}_{1\cap 2}$ is robust to clustering at the intersection of dimensions 1 and 2. In this paper, dimension 1 is the district, i.e. observations over time of the same district, dimension 2 is state-by-year, i.e. all districts in a state in a given year, and $1\cap 2$ is district-by-state-by-year, which is just the district-by-year observation. In this case, where observations in the $1\cap 2$ dimension are single observations, $\hat{V}_{1\cap 2}$ is the standard HEW heteroscedasticity-robust variance-covariance matrix.

5 Results

5.1 Regression results

5.1.1 Modelling temperatures with degree-days

The first set of regressions models temperature using growing season degree-days (and its square) and harmful growing season degree-days. As noted above, this reflects the agronomic emphasis on cumulative heat over the growing season, but may be overly restrictive in its functional form. In particular, I estimate

$$\begin{aligned} y_{dt} &= \alpha_d + g_r(t) + \tilde{\theta}_{DD} GSDD_{dt} + \tilde{\theta}_{DD^2} GSDD_{dt}^2 + \tilde{\theta}_{HDD} HDD_{dt} \\ &+ \tilde{\theta}_P P_{dt} + \tilde{\theta}_{P^2} P_{dt}^2 \\ &+ \tilde{\theta}_{I_{DD,P}} \left(GSDD_{dt} \cdot P_{dt} \right) + \tilde{\theta}_{I_{HDD,P}} \left(HDD_{dt} \cdot P_{dt} \right) + \varepsilon_{dt} \end{aligned} \tag{5}$$

where the dependent variable is the major crop yield (output per hectare in 2005 US dollars). I include district fixed effects and region-specific cubic time trends. The time trends allow productivity to improve as the climate slowly warms over the latter half of the 20th century, avoiding confounding of temperature warming with technological progress. Table 3.A, column (1) reports the results of this FGLS regression. The results are overall as expected: yields are increasing in the linear temperature and precipitation terms, but decreasing in the squares. Harmful degree-days are indeed very harmful, although when considering the magnitude of the coefficient it should be kept in mind that an increase of 100 harmful degree-days would be quite a radical increase in temperatures. Interestingly, the interaction term between degree-days and precipitation is negative, which runs counter to the received agronomic wisdom that extra moisture helps shield plants from extra heat. However, precipitation does appear to shield plants from extreme heat, at least if we take seriously the positive point estimate of positive interaction between harmful degree-days and precipitation.

In column (2) of Table 3.A, I report results for a regression that includes monthly precipitation (and squares), in an attempt to capture the importance of the timing of precipitation. The estimates on precipitation are generally sensible, with yields increasing in linear precipitation terms and decreasing in their squares. The exception is August precipitation, where the signs are reversed, presumably reflecting the negative impact of a late-arriving monsoon.

5.1.2 Temperatures in nonparametric bins

Because of concerns that the degree-day specification may be overly restrictive, I also estimate models where the regressors are the number of days in each of 20 temperature bins. In particular, I estimate

$$y_{dt} = \alpha_d + g\left(t\right) + \sum_{c} \tilde{\theta}_{T_c} T_{c,dt} + \tilde{\theta}_P P_{dt} + \tilde{\theta}_{P^2} P_{dt}^2 + \varepsilon_{dt} \tag{6}$$

where the temperature bins are $c =< 22, 21, \ldots, 38, > 38$. The $(29, 30]$ bin is omitted as the reference category, so each coefficient $\tilde{\theta}_{T_c}$ represents the impact of an additional day in the bin $(c-1, c]$ compared to a day in the $(29, 30]$ bin. The main functional form restriction this framework imposes on the temperature effects is that the effect is constant within each bin. This seems a reasonable approximation for the interior bins but of course cannot be true for the extremal bins (< 22 and > 38). As above, I run models with aggregate growing season precipitation (and its squares) and with monthly precipitation (and squares). Again, I include district fixed effects and regional cubic time trends.

The coefficient estimates are given in Table 3.B, but are perhaps more easily assessed in graphical form, presented in Figure 8. The clear pattern is that cooler temperatures increase agricultural productivity and warmer temperatures are harmful, relative to the $(29, 30]$ bin. For example, imagine that climate change shifts one day from the P29 bin (temperatures in $(28, 29]$ C) to the P31 bin. Since the estimated coefficients on these bins are 0.013 and -0.082, respectively, the total estimated impact of such a shift would be -0.095. Notably, the effects of the highest temperature bins, while negative, are imprecisely estimated.

This additional flexibility does come with a cost: as is readily visible in equation (6), there is a strong assumption of additive separability. That is, I am implicitly assuming that the marginal effect of, for example, a day in the $(34, 35]$ bin is the same in a relatively warm year as in a relatively cool year. This is unlikely to be true. However, in the U.S. context, Schlenker and Roberts (2006) find that a similar assumption of additive separability performs well. Furthermore, as will be apparent in the predicted impacts, the results from this nonparametric approach are reasonably close to those obtained from the degree-day approach, which does take into account cumulative heat over the growing season.

5.2 Predicted Impact of Climate Change

To incorporate the estimated coefficients into a climate change prediction, I calculate the discrete difference in predicted yields at the projected temperature and precipitation scenario from the predicted yield at the historical mean. That is, in the case of the nonparametric bins, I calculate

$$
\begin{aligned}
\widehat{\Delta y} &= \widehat{y_1} - \widehat{y_0} \\
&= \sum_c \left\{ \hat{\theta}_{T_c} \left(\overline{T_{c,1}} - \overline{T_{c,0}} \right) \right\} \\
&+ \sum_m \left\{ \hat{\theta}_{P_m} \left(\overline{P_{m,1}} - \overline{P_{m,0}} \right) + \hat{\theta}_{P_m^2} \left(\overline{P_{m,1}^2} - \overline{P_{m,0}^2} \right) \right\}
\end{aligned}
$$

where $\hat{\theta}_{T_c}$ is the estimated coefficient on temperature bin c, $\hat{\theta}_{P_m}$ on precipitation in month m, and $\hat{\theta}_{P_m}$ on squared month-m precipitation. $\overline{T_{c,1}}$ represents the mean number of days in bin c in the projected climate, $\overline{T_{c,0}}$ the mean in the historical climate, and similarly for the precipitation variables. The calculation is similar for the degree-days approach.

Table 4.A presents results for the degree-days approach, using both total precipitation and monthly precipitation. The underlying regression coefficients are taken from the corre-

sponding columns in Table 3.A. In each case, impacts are estimated for each of three climate change scenarios: the IPCC 2010-2039 consensus A1F1 (business-as-usual) scenario ($+0.5°C$ uniform temperature increase, +4% precipitation increase), the Hadley 2010-2039 A1F1 temperature predictions with +4% precipitation, and the Hadley 2070-2099 A1F1 temperature predictions with +10% precipitation. The aggregate impact is negative for all three scenarios, with mildly positive precipitation effects outweighed by negative temperature effects. Even the moderate IPCC 2010-2099 scenario reduces yields by roughly 4.5%. The Hadley scenarios are even more detrimental. In the medium run (2010-2039), yields are predicted to fall by approximately nine percent, while the long run effect is over 40% of yields. However, this latter estimate is in the absence of long-run adaptation, and therefore likely represents an upper bound on damages.

Table 4.B presents results for the temperature bins approach. The national results, reported in column (1) are broadly similar to those from the degree-days approach: the mild (IPCC) medium-run scenario reduces yields by roughly 4.5 percent, and damages increase for the Hadley scenarios, emphasizing the importance of the shift into the highest temperature bins. Notably, the long-run Hadley scenario is not nearly as damaging as in the degree-days model, although yields are still predicted to fall by 25 percent. The difference can be attributed to the degree-days model's reliance on functional form: as temperatures increase, the negative coefficient on the quadratic degree-day term pushes yields far down.

To explore potentially heterogeneous impacts, columns (2)–(5) reports results from separate estimates by region. Effects are negative across all regions with the exception of the East, which is very imprecisely estimated due to the small sample size. The estimated long-run effect for the Northwest region is perhaps implausibly large (over 60% of yields), although this region also has a small sample size and this estimate is not very precise.

Table 4.C explores the possibility of heterogenous impacts over time. I split the sample into 1960-1979 and 1980-1999 and run the temperature bins regressions separately. The coef-

ficients on the temperature bins are plotted in Figures 9.A and 9.B. Inspection of these figures reveals the same pattern of beneficial lower temperatures and harmful higher temperatures. overall decline in temperature. As above, these coefficients (along with the coefficients on monthly precipitation) are combined with climate projections to obtain predicted impacts, reported in Table 4.C. For comparison, column (1) re-reports the results for the full sample, 1960-1999. Interestingly, the later period (reported in column (3)) shows greater sensitivity to climate than the earlier period (reported in column (2)), both absolutely and as a percentage of average yields. One possible explanation for this increased vulnerability is the higher prevalence of high-yield varieties (HYVs) in the latter period, as HYVs are believed to provide greater yields on average but are more sensitive to climate fluctuations. Another potential explanation is that temperatures in the second half of the period were somewhat higher. The important message is that technological progress need not reduce climate vulnerability.

5.3 Evidence on Adaptation

Three margins for adaptation can be explored with the available data. First, the application of fertilizer can be adjusted in the face of a harmful weather shock. The fertilizers reported in the data are nitrogen, phosphorus, and potassium, which are aggregated at mean 1960-1965 prices. Column (1) of Table 5 shows the estimated impact of a $1°C$ mean temperature increase on the quantities of fertilizers used per hectare. Fertilizer use falls by roughly 4.5 percent. This suggests that the true welfare impact of a climate shock may be slightly overstated by the effect on yields, since farmers can reduce their input use.

Second, farmers could respond to a harmful shock by planting in the second, winter season. However, column (2) shows that the extent of double-cropping[11] is not significantly affected by a one-degree temperature shock. This margin for adaptation may be limited, at

[11] As measured by the ratio of gross cropped area to net cropped area.

least in the short run.

Finally, although yearly data on labor inputs are not available, yearly wages are available at the district level. If we assume that the temperature affects the agricultural labor market mainly through the channel of labor demand rather than labor supply, the behavior of wages in response to temperature shocks can be informative about the response of labor demand to temperature. Column (3) shows that the wage falls by nearly two percent in response to a one-degree temperature increase.[12] In a full-employment context, we could interpret this as reducing the welfare effect of a climate shock, since farmers use fewer scarce resources, much as with fertilizer. However, given chronic unemployment in India, it is likely that these resources are left unemployed, so the welfare effect is ambiguous. The results of Topalova (2007), demonstrating the slow response of factor quantities in India to shocks, suggest that this effect could be persistent.

6 Conclusion

This paper employs a panel data methodology to show that the impact of climate change on Indian agriculture is likely to be negative over the short- to medium-term. The medium-term (2010-2039) impact on yields is estimated to be negative 4.5 to nine percent. Since agriculture makes up roughly 20 percent of India's GDP, this implies a cost of climate change of 1 to 1.8 percent of GDP per year over the medium run. Furthermore, agricultural productivity is particularly important for the well-being of the poor. A back-of-the envelope calculation using the estimate of Ligon and Sadoulet (2007) that each percentage point of agricultural GDP growth increases consumption of the lowest three deciles by four to six percent would imply that climate change could depress consumption among India's poor by at least 18 percent. In the absence of rapid and full adaptation, the consequences of long-run climate change could be even more severe, up to 25 percent of crop yields. The results of

[12] This is consistent with the effects for rainfall found by Jayachandran (2006).

this paper pose two important questions for future research. First, what are the factors explaining the difference between these negative consequences for a developing country and the mildly positive results for the U.S. found by Deschênes and Greenstone (2007)? Second, and crucial for the welfare of Indian agriculture, how quickly will developing country farmers be able to adjust their farming practices to adapt to the changing climate and what policies or technologies will enable rapid adaptation?

References

Maximilian Auffhammer, V. Ramanathan, and Jeffrey R. Vincent. Intergrated model shows that atmospheric brown clouds and greenhouse gases have reduced rice harvests in India. *Proceedings of the National Academy of Sciences*, 103(52):19668–19672, December 2006.

Guiseppe Bertola and Ricardo J. Caballero. Irreversibility and aggregate investment. *The Review of Economic Studies*, 61(2):223–246, April 1994. ISSN 0034-6527. URL `http://www.jstor.org/stable/2297979`.

A. Colin Cameron, Jonah B. Gelbach, and Douglas L. Miller. Robust inference with multi-way clustering. URL `http://papers.nber.org/papers/t0327`. NBER Technical Working Paper #T0327, September 2006.

Jens Hesselbjerg Christensen and Bruce Hewitson. Regional climate projections. In *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, chapter 11, pages 847–940. Cambridge University Press, 2007.

William R. Cline. *The Economics of Global Warming*. Peterson Institute for International Economics, 1992.

R.V. Cruz, H. Harasawa, M. Lal, S. Wu, Y. Anokhin, B. Punsalmaa, Y. Honda, M. Jafari, C. Li, and N. Huu Ninh. Asia. In M.L. Parry, O.F. Canziani, J.P. Palutikof, P.J. van der Linden, and C.E. Hanson, editors, *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovenmental Panel on Climate Change.*, chapter 10, pages 470–506. Cambridge University Press, Cambridge, UK, 2007.

Melissa Dell, Benjamin Olken, and Benjamin F. Jones. Climate shocks and economic growth:

Evidence from the last half century. URL http://www.nber.org/papers/w14132. NBER Working Paper 14132, June 2008.

Melissa Dell, Benjamin F. Jones, and Benjamin A. Olken. Temperature and income: Reconciling new cross-sectional and panel estimates. *American Economic Review*, 99(2):198–204, May 2009. URL http://www.aeaweb.org/articles.php?doi=10.1257/aer.99.2.198.

Olivier Deschênes and Michael Greenstone. The economic impacts of climate change: Evidence from agricultural profits and random fluctuations in the weather. *American Economic Review*, 97(1):354–385, March 2007.

Avinash Dixit and Robert S. Pindyck. *Investment Under Uncertainty*. Princeton University Press, 1994.

Esther Duflo and Rohini Pande. Dams. *Quarterly Journal of Economics*, pages 601–646, May 2007.

Esther Duflo, Michael Kremer, and Jonathan Robinson. Understanding fertiilzer adoption: Evidence from field experiments. MIT mimeo, 2005.

FAO. *FAO Statistical Yearbook 2005–2006*. Food and Agricultural Organization, 2006. URL http://www.fao.org/statistics/yearbook/.

John Felkner, Kamilya Tazhibayeva, and Robert Townsend. Impact of climate change on rice production in Thailand. *American Economic Review*, 99(2):205Ű–210, May 2009. URL http://www.atypon-link.com/doi/abs/10.1257/aer.99.2.205.

Anthony Fisher, Michael Hanemann, Michael J. Roberts, and Wolfram Schlenker. Climate change and agriculture reconsidered. URL http://www.columbia.edu/~ws2162/agClimateChange/agClimateChange.pdf. Manuscript, May 2009.

Andrew D. Foster and Mark R. Rosenzweig. Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of Political Economy*, 103(6): 1176–1209, December 1995. URL http://www.jstor.org/stable/2138708.

Christian Hansen. Generalized least squares inference in panel and multilevel models with serial correlation and fixed effects. *Journal of Econometrics*, 140:670–694, 2007. URL http://dx.doi.org/10.1016/j.jeconom.2006.07.011.

Richard Hornbeck. Quantifying long-term adjustment to environmental change: Evidence from the American Dust Bowl. URL http://www.economics.harvard.edu/faculty/hornbeck/files/Hornbeck_dustbowl_nov08.pdf. Mimeo, November 2008.

Seema Jayachandran. Selling labor low: Wage responses to productivity shocks in developing countries. *Journal of Political Economy*, 114(3):538–575, 2006.

K. Krishna Kumar, K. Rupa Kumar, R.G. Ashrit, N.R. Deshpande, and J.W. Hansen. Climate impacts on Indian agriculture. *International Journal of Climatology*, 24:1375–1393, 2004. doi: 10.1002/joc.1081.

Ethan Ligon and Elizabeth Sadoulet. Estimating the effects of aggregate agricultural growth in the distribution of expenditures. Background paper for the World Development Report 2008, 2007.

Stephen P. Long, Elizabeth A. Ainsworth, Andrew D. B. Leakey, Josef Nosberger, and Donald R. Ort. Food for thought: Lower-than-expected crop yield stimulation with rising CO2 concentrations. *Science*, 312:1918–1921, June 2006. doi: 10.1126/science.1114722.

Robert Mendelsohn, William D. Nordhaus, and Daigee Shaw. The impact of global warming on agriculture: A Ricardian analysis. *American Economic Review*, 84(4):753–771, 1994.

T. Ngo-Duc, J. Polcher, and K. Laval. A 53-year forcing data set for land surface models. *Journal of Geophysical Research*, 110(D06116):1–16, March 2005. doi: 10.1029/2004JD005434.

Apruva Sanghi, D. Alves, R. Evenson, and R. Mendelsohn. Global warming impacts on Brazilian agriculture: Estimates of the Ricardian model. *Economia Aplicada*, 1997.

Apurva Sanghi, K.S. Kavi Kumar, and James W. McKinsey Jr. India agriculture and climate dataset. Technical report, World Bank, 1998a. Compiled and used in the study "Measuring the Impact of Climate Change on Indian Agriculture," World Bank Technical Paper No. 402. Funded jointly by the World Bank Research Budget and by the Electric Power Research Institute in Palo Alto, California. Available online at the UCLA / BREAD website.

Apurva Sanghi, Robert Mendelsohn, and Ariel Dinar. The climate sensitivity of Indian agriculture. In *Measuring the Impact of Climate Change on Indian Agriculture*, chapter 4, pages 69–139. World Bank, 1998b.

Wolfram Schlenker and Michael J. Roberts. Estimating the impact of climate change on crop yields: The importance of non-linear temperature effects. URL `http://www.columbia.edu/~ws2162/SchlenkerRoberts.pdf`. Working paper, September 2006.

Wolfram Schlenker and Michael J. Roberts. Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. *Proceedings of the National Academy of Sciences*, 106(37):15594–15598, 2009. doi: 10.1073/pnas.0906865106. URL `http://www.pnas.org/content/106/37/15594.abstract`.

Petia Topalova. Factor immobility and regional effects of trade liberalization: Evidence from India. IMF mimeo, 2005.

Petia Topalova. Trade liberalization, poverty and inequality: evidence from Indian districts. In Ann Harrison, editor, *Globalization and Poverty*, chapter 7, pages 291–336. National Bureau of Economic Research, March 2007. URL `http://www.nber.org/chapters/c0110.pdf`.

Jeffrey Wooldridge. Cluster-sample methods in applied econometrics. *American Economic Review*, 93(2):133–138, 2003.

Jeffrey M. Wooldridge. Cluster-sample methods in applied econometrics: An extended analysis. URL `https://www.msu.edu/~ec/faculty/wooldridge/current%20research/clus1aea.pdf`. Mimeograph, 2006.

Table 1.A: Descriptive Statistics

Variable	Units	All	North	NorthWest	East	South
Output	2005 USD (000)	4,564.2	7,580.2	2,289.0	4,098.6	3,682.8
		(4,813.8)	(6,860.9)	(2,567.6)	(2,396.3)	(3,092.3)
Yield (output per hectare)	2005 USD	15.2	19.6	10.4	10.9	14.8
		(11.0)	(11.5)	(9.0)	(4.8)	(10.9)
Mean temperature (growing season)	Deg. C	28.5	29.9	29.3	28.0	27.5
		(1.8)	(1.4)	(1.3)	(0.7)	(1.6)
Degree-days (growing season)	Deg. C	2,464.6	2,610.5	2,556.4	2,426.4	2,357.5
		(194.9)	(149.6)	(125.1)	(85.2)	(175.7)
Precipitation (growing season)	mm	775.0	807.8	691.6	1,011.2	760.4
		(302.5)	(267.5)	(271.3)	(167.8)	(325.7)
Share of cropland irrigated		0.28	0.48	0.21	0.21	0.20
		(0.23)	(0.22)	(0.12)	(0.16)	(0.19)
Share of cropland HYV		0.23	0.31	0.16	0.16	0.21
		(0.19)	(0.22)	(0.13)	(0.12)	(0.19)
Share below poverty line (1973)		0.45	0.36	0.42	0.70	0.49
		(0.14)	(0.11)	(0.16)	(0.14)	(0.10)
Share below poverty line (1999)		0.23	0.15	0.17	0.42	0.27
		(0.12)	(0.10)	(0.08)	(0.15)	(0.10)
Number of districts		218	61	36	11	110
Number of observations		8,720	2,440	1,440	440	4,400

Notes: Regression sample: 1960-1999, 218 districts with output data for all years 1960-1999. Regions defined as: North (Haryana, Punjab and Uttar Pradesh); Northwest (Gujarat, Rajasthan); East (Bihar, Orissa, West Bengal); South (Andhra Pradesh, Karnakata, Madhya Pradesh, Maharastra, Tamil Nadu). Standard deviations in parentheses.

Table 1.B: Supplemental Descriptive Statistics

Panel I: Comparison of Regression Sample with World Bank Sample

Variable	Units	Regression Sample, 1960-1999	World Bank Sample, 1966-1986	Regression Sample, 1966-1986
Output	2005 USD (000)	4,564.2	4,296.8	4,245.7
		(4,813.8)	(3,806.1)	(3,912.0)
Yield (output per hectare)	2005 USD	15.2	13.3	13.7
		(11.0)	(8.3)	(8.6)
Mean temperature (growing season)	Deg. C	28.5	28.4	28.4
		(1.8)	(1.8)	(1.8)
Degree-days (growing season)	Deg. C	2,464.6	2,458.6	2,454.8
		(194.9)	(195.0)	(193.7)
Precipitation (growing season)	mm	775.0	829.6	781.2
		(302.5)	(339.7)	(303.3)
Share of cropland irrigated		0.28	0.26	0.28
		(0.23)	(0.23)	(0.23)
Share of cropland HYV		0.23	0.22	0.22
		(0.19)	(0.19)	(0.19)
Share below poverty line (1973)		0.45	0.47	0.45
		(0.14)	(0.15)	(0.14)
Share below poverty line (1999)		0.23		
		(0.12)		
Number of districts		218	271	218
Number of observations		8,720	5,670	4578

Panel II: Regression Sample By Decade

Variable	Units	1960-1969	1970-1979	1980-1989	1990-1999
Output	2005 USD (000)	2,777.3	3,956.2	5,365.1	6,158.1
		(2,110.4)	(3,346.6)	(5,025.6)	(6,713.1)
Yield (output per hectare)	2005 USD	10.2	12.9	17.3	20.4
		(5.8)	(7.6)	(10.6)	(14.8)
Mean temperature (growing season)	Deg. C	28.5	28.3	28.6	28.6
		(1.9)	(1.8)	(1.8)	(1.8)
Degree-days (growing season)	Deg. C	2,461.3	2,443.5	2,476.5	2,477.3
		(199.3)	(196.0)	(192.6)	(190.0)
Precipitation (growing season)	mm	790.7	788.6	766.5	754.1
		(302.3)	(302.6)	(308.9)	(294.6)
Share of cropland irrigated		0.22	0.27	0.32	
		(0.19)	(0.21)	(0.25)	
Share of cropland HYV		0.05	0.20	0.35	
		(0.07)	(0.16)	(0.19)	
Share below poverty line			0.45	0.39	0.27
			(0.14)	(0.17)	(0.15)
Number of districts		218	218	218	218
Number of observations		2,180	2,180	2,180	2,180

Notes: Regression sample includes all 218 districts with output data for all years 1960-1999. Word Bank sample includes all 271 districts of the Sanghi, Mendelsohn and Dinar (1998) World Bank study. Standard deviations in parentheses.

Table 2.A: Residual Variation in District Weather Variables

Panel 1: Growing Season Mean Temperatures (C)

Mean: 28.5; N:8720		District*year observations differing from predicted value by more than									
		0.5 deg C		1.0 deg C		1.5 deg C		2.0 deg C		2.5 deg C	
Regressors	RMSE	Number	Share	Number	Share	Number	Share	Number	Share	Number	Share
Constant only	1.81	6853	0.786	5200	0.596	3619	0.415	2214	0.254	1230	0.141
District FEs	0.50	2624	0.301	358	0.041	69	0.008	6	0.001	0	0.000
District FEs, Linear Year	0.49	2551	0.293	330	0.038	46	0.005	4	0.000	0	0.000
District FEs, Quadratic Year	0.49	2557	0.293	303	0.035	53	0.006	5	0.001	0	0.000
District FEs, Cubic Year	0.49	2531	0.290	298	0.034	44	0.005	4	0.000	0	0.000
District and Year FEs	0.33	1011	0.116	32	0.004	0	0.000	0	0.000	0	0.000
District and Year*Region FEs	0.26	517	0.059	9	0.001	0	0.000	0	0.000	0	0.000
District and Year*State FEs	0.20	174	0.020	3	0.000	0	0.000	0	0.000	0	0.000

Panel 2: Growing Season Degree-Days (C)

Mean: 2,464.6; N:8720		District*year observations differing from predicted value by more than									
		60 deg-days (C)		120 deg-days (C)		180 deg-days (C)		240 deg-days (C)		300 deg-days (C)	
Regressors	RMSE	Number	Share	Number	Share	Number	Share	Number	Share	Number	Share
Constant only	194.95	6752	0.774	4914	0.564	3034	0.348	1441	0.165	867	0.099
District FEs	50.47	1926	0.221	189	0.022	25	0.003	0	0.000	0	0.000
District FEs, Linear Year	49.32	1783	0.204	168	0.019	20	0.002	0	0.000	0	0.000
District FEs, Quadratic Year	48.92	1735	0.199	168	0.019	22	0.003	0	0.000	0	0.000
District FEs, Cubic Year	48.89	1751	0.201	161	0.018	21	0.002	0	0.000	0	0.000
District and Year FEs	33.89	617	0.071	15	0.002	0	0.000	0	0.000	0	0.000
District and Year*Region FEs	27.27	323	0.037	3	0.000	0	0.000	0	0.000	0	0.000
District and Year*State FEs	20.49	85	0.010	2	0.000	0	0.000	0	0.000	0	0.000

Panel 3: Growing Season Precipitation (mm)

Mean: 775.0; N:8720		District*year observations differing from predicted value by more than									
		2 percent		4 percent		6 percent		8 percent		10 percent	
Regressors	RMSE	Number	Share	Number	Share	Number	Share	Number	Share	Number	Share
Constant only	302.47	8403	0.964	8058	0.924	7721	0.885	7382	0.847	7028	0.806
District FEs	183.38	8093	0.928	7439	0.853	6775	0.777	6220	0.713	5638	0.647
District FEs, Linear Year	182.15	8091	0.928	7473	0.857	6850	0.786	6219	0.713	5631	0.646
District FEs, Quadratic Year	182.15	8096	0.928	7462	0.856	6853	0.786	6222	0.714	5635	0.646
District FEs, Cubic Year	182.14	8091	0.928	7464	0.856	6851	0.786	6225	0.714	5637	0.646
District and Year FEs	149.57	8003	0.918	7228	0.829	6517	0.747	5802	0.665	5125	0.588
District and Year*Region FEs	133.05	7785	0.893	6912	0.793	6009	0.689	5229	0.600	4460	0.511
District and Year*State FEs	105.00	7382	0.847	6094	0.699	4952	0.568	3970	0.455	3232	0.371

Notes: Table counts residuals from regressions of district*year observations on regressors listed in row headings. Cell entries are number of residuals of absolute value greater than or equal to the cutoffs given in the column headings.Years: 1960-1999; Sample: 218 districts with output data for all years.

Table 2.B: Residual Variation in District Temperature Bins

	Bin										
Regressor(s)	<20	21	22	23	24	25	26	27	28	29	30
Constant	0.30	0.06	0.13	0.85	2.89	6.08	12.16	18.77	19.47	18.13	13.68
District FEs	0.04	0.03	0.09	0.42	0.92	2.24	4.08	4.85	5.65	4.92	4.06
District FEs, Linear Year	0.06	0.04	0.11	0.50	1.05	2.32	4.15	4.85	5.64	4.89	4.06
District FEs, Quadratic Year	0.06	0.04	0.11	0.52	1.05	2.32	4.15	4.84	5.63	4.90	4.05
District FEs, Cubic Year	0.06	0.04	0.11	0.52	1.05	2.31	4.16	4.84	5.64	4.90	4.05
District and Year FEs	0.07	0.04	0.14	0.61	1.12	2.43	4.19	4.68	5.48	4.79	4.01
District and Region*Year FEs	0.07	0.04	0.15	0.62	1.09	2.21	3.72	4.41	4.89	4.24	3.28
District and State*Year FEs	0.07	0.05	0.13	0.47	0.89	2.04	3.26	4.07	4.41	3.89	2.93

	Bin										
Regressor(s)	31	32	33	34	35	36	37	38	39	40	>40
Constant	7.48	5.20	4.33	3.91	3.35	2.51	1.54	0.76	0.30	0.09	0.03
District FEs	2.70	2.02	1.77	1.59	1.40	1.22	1.00	0.68	0.36	0.14	0.04
District FEs, Linear Year	2.70	2.02	1.77	1.59	1.40	1.22	1.01	0.69	0.36	0.15	0.05
District FEs, Quadratic Year	2.69	2.02	1.77	1.60	1.40	1.22	1.01	0.69	0.36	0.15	0.05
District FEs, Cubic Year	2.69	2.02	1.77	1.60	1.41	1.23	1.02	0.71	0.37	0.15	0.05
District and Year FEs	2.56	1.93	1.67	1.51	1.34	1.17	0.98	0.70	0.39	0.17	0.06
District and Region*Year FEs	2.22	1.75	1.54	1.41	1.24	1.08	0.90	0.64	0.37	0.16	0.06
District and State*Year FEs	1.99	1.54	1.35	1.25	1.09	0.93	0.74	0.53	0.32	0.14	0.05

	Alternative Extremal Bins			
Regressor(s)	<21	<22	>38	>39
Constant	0.36	0.49	0.42	0.12
District FEs	0.05	0.11	0.51	0.18
District FEs, Linear Year	0.05	0.12	0.52	0.19
District FEs, Quadratic Year	0.06	0.13	0.52	0.19
District FEs, Cubic Year	0.06	0.13	0.53	0.19
District and Year FEs	0.07	0.17	0.57	0.21
District and Region*Year FEs	0.07	0.18	0.53	0.20
District and State*Year FEs	0.08	0.17	0.45	0.18

Notes: This table assesses the extent of residual variation available after removing district fixed effects and other controls. For each bin, the number of days in that bin is regressed on the controls given in the row heading. The absolute value of the residual is then averaged over all district*year observations. The result can be interpreted as the mean number of days per district*year available to identify the effect of that bin. Years: 1960-1999; Sample: 218 districts with output and yield data for all years 1960-1999 (8720 total year*district observations)

Table 3.A: FGLS Estimates of Weather Variables' Effects on Major Crop Yields

	Total GS Precipitation (1)	Monthly GS Precipitation (2)
Growing Season Degree-days (100, C)	5.418	3.536
	(2.335)	(2.343)
	[5.579]	[4.354]
GSDD Squared	-0.125	-0.094
	(0.048)	(0.049)
	[0.114]	[0.094]
Harmful GSDD (100, C) with threshold 34	-3.508	-2.687
	(1.655)	(0.706)
	[4.326]	[2.381]
Total Growing Season Precipitation (100 mm)	1.620	
	(0.473)	
	[1.368]	
TotalGrowSeasonPrecip Squared	-0.021	
	(0.008)	
	[0.016]	
GrowSeasonDegreeDays*TotalGrowSeasonPrecip	-0.048	
	(0.017)	
	[0.058]	
HarmfulGSDD34*TotalGrowSeasonPrecip	0.068	
	(0.185)	
	[0.440]	
June Precipitation (100 mm)		0.520
		(0.225)
		[0.631]
June Precipitation Squared		-0.034
		(0.050)
		[0.096]
July Precipitation (100 mm)		0.272
		(0.214)
		[0.931]
July Precipitation Squared		-0.061
		(0.028)
		[0.103]
August Precipitation (100 mm)		-0.450
		(0.210)
		[0.540]
August Precipitation Squared		0.050
		(0.030)
		[0.067]
September Precipitation (100 mm)		0.533
		(0.230)
		[0.629]
September Precipitation Squared		-0.010
		(0.049)
		[0.143]
N	8,720	8,720

Notes: Dependent variable: major crop yields (2005 USD / HA). Regressions include district fixed effects and region*year cubic time trends (coefficients not reported). FGLS estimator uses bias-corrected AR(2) parameter estimates. Standard errors clustered by district in parentheses, standard errors twoway-clustered by district and state-by-year in brackets. Years: 1960-1999. Sample: 218 districts with output data for all years.

Table 3.B: FGLS Estimates of Effect of Days in One-Degree (C) Temperature Bins on Major Crop Yields

	Total GS Precipitation	Monthly GS Precipitation
	(1)	(2)
Days in <=22 bin	0.192	0.126
	(0.142)	(0.146)
	[0.266]	[0.250]
Days in P23 bin	0.081	0.094
	(0.059)	(0.059)
	[0.072]	[0.070]
Days in P24 bin	0.032	0.026
	(0.039)	(0.039)
	[0.064]	[0.063]
Days in P25 bin	0.040	0.038
	(0.019)	(0.019)
	[0.046]	[0.044]
Days in P26 bin	0.037	0.028
	(0.014)	(0.014)
	[0.044]	[0.042]
Days in P27 bin	0.031	0.025
	(0.015)	(0.015)
	[0.044]	[0.042]
Days in P28 bin	0.006	-0.001
	(0.015)	(0.015)
	[0.045]	[0.043]
Days in P29 bin	0.017	0.013
	(0.018)	(0.018)
	[0.049]	[0.047]
Days in P30 bin (omitted category)	-	-
	-	-
	-	-
Days in P31 bin	-0.073	-0.082
	(0.044)	(0.043)
	[0.063]	[0.062]
Days in P32 bin	-0.001	-0.013
	(0.046)	(0.047)
	[0.099]	[0.099]
Days in P33 bin	0.012	0.004
	(0.052)	(0.052)
	[0.084]	[0.086]
Days in P34 bin	-0.091	-0.082
	(0.040)	(0.041)
	[0.069]	[0.069]
Days in P35 bin	0.040	0.044
	(0.041)	(0.043)
	[0.063]	[0.068]
Days in P36 bin	-0.114	-0.123
	(0.048)	(0.049)
	[0.07]	[0.070]
Days in P37 bin	-0.021	-0.014
	(0.061)	(0.061)
	[0.083]	[0.087]
Days in P38 bin	-0.225	-0.209
	(0.124)	(0.125)
	[0.22]	[0.216]
Days in >38 bin	-0.092	-0.092
	(0.074)	(0.074)
	[0.137]	[0.138]

Table 3.B, continued

	Total GS Precipitation	Monthly GS Precipitation
	(1)	(2)
Total Growing Season Precipitation (100 mm)	0.387	
	(0.145)	
	[0.438]	
Growing Season Precipitation Squared	-0.016	
	(0.007)	
	[0.017]	
June Precipitation (100mm)		0.602
		(0.264)
		[0.681]
June Precipitation Squared		-0.050
		(0.056)
		[0.102]
July Precipitation (100mm)		0.361
		(0.213)
		[0.864]
July Precipitation Squared		-0.075
		(0.028)
		[0.097]
August Precipitation (100mm)		-0.441
		(0.221)
		[0.515]
August Precipitation Squared		0.050
		(0.032)
		[0.067]
September Precipitation (100mm)		0.614
		(0.232)
		[0.618]
September Precipitation Squared		-0.030
		(0.049)
		[0.139]
N	8720	8720

Notes: Dependent variable: major crop yields (2005 USD / HA). Each bin is identified as its upper limit (e.g. P35 includes temperatures in (34,35] C). FGLS estimator uses bias-corrected AR(2) parameter estimates; standard errors clustered by district in parentheses, standard errors twoway-clustered by district and state-by-year in brackets. Years: 1960-1999. Sample: 218 districts with output data for all years.

Table 4.A: Projected Impact of Climate Change on Major Crop Yields
from Aggregate Weather Regressions

	Total GS Precipitation (1)	Monthly GS Precipitation (2)
Mean of Dependent Variable	15.215	15.215
Number of Observations	8,720	8,720
Panel A: IPCC Medium-Run (2010-2039) S. Asia Scenario (Uniform +0.5 deg C, +4% precipitation)		
Temperature Effect	-0.566	-0.691
	(0.105)	(0.081)
	[0.275]	[0.316]
Precipitation Effect	0.373	0.026
	(0.120)	(0.011)
	[0.386]	[0.058]
Interaction Effect	-0.520	
	(0.176)	
	[0.639]	
Total Effect	-0.712	-0.665
	(0.082)	(0.086)
	[0.356]	[0.358]
Panel B: Hadley A1F1 Medium-Run (2010-2039) Scenario		
Temperature Effect	-1.358	-1.414
	(0.337)	(0.173)
	[0.889]	[0.745]
Precipitation Effect	0.373	0.026
	(0.120)	(0.011)
	[0.386]	[0.058]
Interaction Effect	-0.514	
	(0.315)	
	[0.961]	
Total Effect	-1.498	-1.387
	(0.176)	(0.179)
	[0.801]	[0.788]
Panel C: Hadley A1F1 Long-Run (2070-2099) Scenario		
Temperature Effect	-6.755	-6.506
	(1.722)	(0.894)
	[4.612]	[3.645]
Precipitation Effect	0.926	0.064
	(0.300)	(0.026)
	[0.964]	[0.140]
Interaction Effect	-1.329	
	(1.526)	
	[4.108]	
Total Effect	-7.158	-6.441
	(0.909)	(0.908)
	[4.062]	[3.745]

Notes: Projections are calculated as the discrete difference in yields (output per hectare) at the projected climate versus the historical climate. Coefficients are obtained from bias-corrected FGLS regressions of yields on growing season weather variables, regional cubic time trends and district fixed effects, weighted by area cropped. Weather variables in column (1) are growing-season degree-days, its square, harmful growing season degree days, total growing season precipitation, its square and the interaction of precipitation with growing-season degree-days and harmful degree-days. Column (2) substitutes monthly precipitation (and squares) for aggregate precipitation, and drops the interactions. Standard errors of the projection clustered by district are reported in parentheses, standard errors twoway clustered by district and state-by-year are reported in brackets. Sample: 218 districts with output data for all years 1960-1999

Table 4.B: Projected Impact of Climate Change on Major Crop Yields from Bins Regressions

	National (1)	North (2)	Northwest (3)	East (4)	South (5)
Mean of Dependent Variable	15.215	19.555	10.436	10.858	14.809
Number of Observations	8720	2440	1440	440	4400
Panel A: IPCC Medium-Run (2010-2039) S. Asia Scenario (Uniform +0.5 deg C, +4% precipitation)					
Temperature Effect	-0.727	-1.467	-0.888	-0.256	-0.452
	(0.069)	(0.205)	(0.185)	(0.218)	(0.085)
	[0.330]	[1.188]	[0.593]	.	[0.237]
Precipitation Effect	0.029	0.059	-0.015	-0.070	0.037
	(0.011)	(0.021)	(0.028)	(0.063)	(0.012)
	[0.055]	[0.144]	[0.064]	.	[0.032]
Total Effect	-0.699	-1.408	-0.903	-0.326	-0.415
	(0.075)	(0.215)	(0.209)	(0.258)	(0.093)
	[0.370]	[1.290]	[0.608]	.	[0.256]
Panel B: Hadley A1F1 Medium-Run (2010-2039) Scenario					
Temperature Effect	-1.225	-1.479	-0.871	0.255	-1.676
	(0.329)	(1.395)	(0.386)	(0.312)	(0.600)
	[0.711]	[2.722]	[1.367]	.	[1.113]
Precipitation Effect	0.029	0.059	-0.015	-0.070	0.037
	(0.011)	(0.021)	(0.028)	(0.063)	(0.012)
	[0.055]	[0.144]	[0.064]	.	[0.032]
Total Effect	-1.196	-1.420	-0.886	0.185	-1.639
	(0.332)	(1.400)	(0.398)	(0.321)	(0.604)
	[0.743]	[2.825]	[1.385]	.	[1.127]
Panel C: Hadley A1F1 Long-Run (2070-2099) Scenario					
Temperature Effect	-3.983	-4.514	-6.624	2.276	-2.615
	(0.976)	(4.593)	(1.780)	(3.321)	(1.195)
	[2.713]	[10.460]	[6.107]	.	[2.381]
Precipitation Effect	0.070	0.129	-0.046	-0.160	0.093
	(0.026)	(0.050)	(0.071)	(0.165)	(0.030)
	[0.135]	[0.343]	[0.160]	.	[0.078]
Total Effect	-3.913	-4.385	-6.670	2.116	-2.521
	(0.982)	(4.609)	(1.796)	(3.451)	(1.207)
	[2.803]	[10.727]	[6.114]	.	[2.420]

Notes: Projections are calculated as the discrete difference in yields (output per hectare) at the projected climate versus the historical climate. Coefficients are obtained from bias-corrected FGLS regressions of yields on growing season days in one-degree (C) temperature bins, monthly precipitation (and squares), regional cubic time trends and district fixed effects, weighted by area cropped. Standard errors of the projection clustered by district are reported in parentheses, standard errors twoway clustered by district and state-by-year are reported in brackets. Small-sample issues cause the twoway-clustered standard errors for region 3 to be nonpositive. Sample: 218 districts with output data for all years 1960-1999.

Table 5: Evidence on Within-Year Adaptation

Impact of Uniform One-Degree (C) Temperature Increase On:			
	Fertilizer Use (1)	Agricultural Wage (2)	Double-Cropping (3)
Effect	-5.556	-0.118	-0.002
	(0.828)	(0.008)	(0.001)
	[4.448]	[0.064]	[0.003]
Mean of Dependent Variable	125.3	7.0	1.2
N	7588	7588	7570

Notes: Projections are calculated as the predicted change from a one-degree C increase in temperature relative to the base period of 1960-1987. Coefficients are obtained from bias-corrected FGLS regressions of yields on growing season days in one-degree (C) temperature bins, monthly precipitation (and squares), regional cubic time trends and district fixed effects. Standard errors of the projection clustered by district are reported in parentheses, standard errors twoway clustered by district and state-by-year are reported in brackets. Sample: all 271 districts, 1960-1987.

Figure 1: Impact of Climate Change With Various Degrees of Adaptation

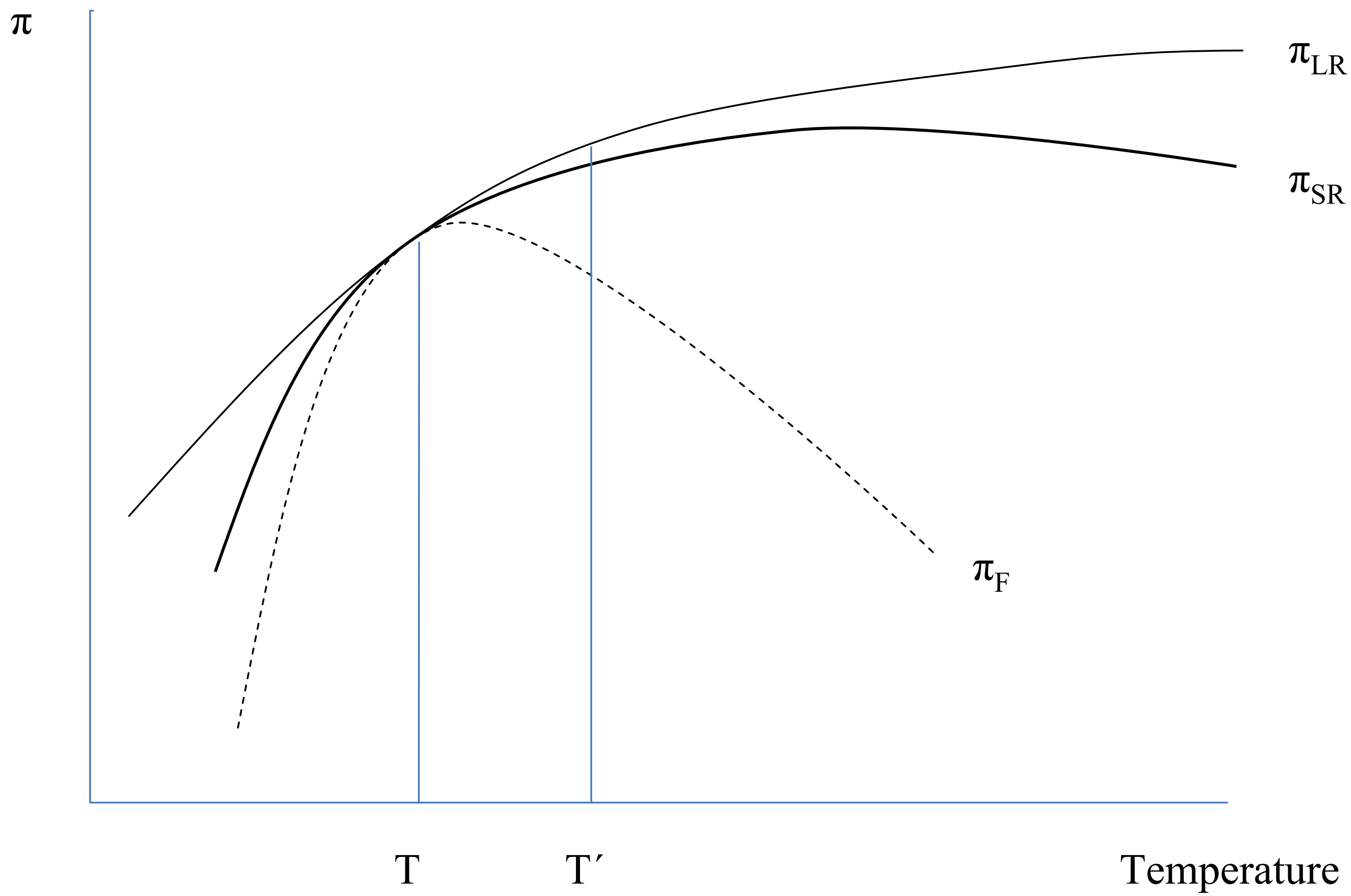


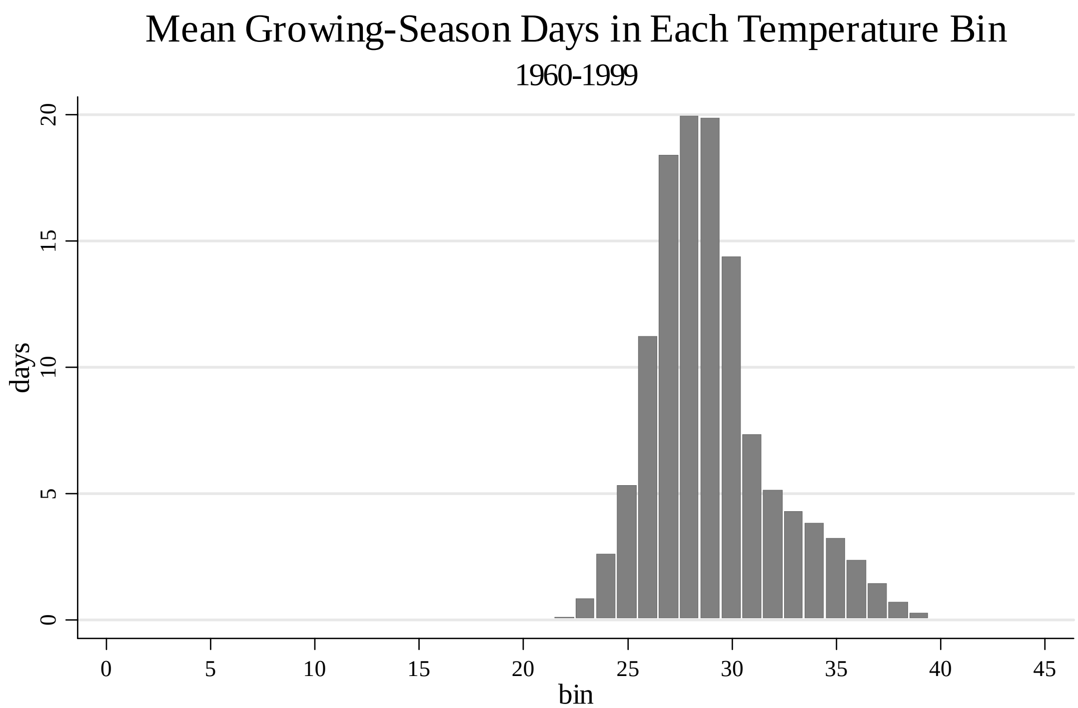
Figure 2.A: Districts Included in SMD98 Study

Notes: This map shows the 271 districts included in the Sanghi, Mendelsohn and Dinar 1998 study. The states included are: Haryana, Punjab and Uttar Pradesh (North); Gujarat, Rajasthan (Northwest); Bihar, Orissa, West Bengal (East); Andhra Pradesh, Karnakata, Madhya Pradesh, Maharastra, Tamil Nadu (South). The major agricultural state excluded is Kerala.

Figure 2.B: Districts Included in Regressions

Notes: This map shows the 218 districts with output data for all years 1960-1999. The bulk of the lost districts (relative to the SMD98 dataset) are from the East, especially Bihar and West Bengal.

Figure 3:



Notes: Each one-degree C bin is indicated by the upper limit, e.g. 20 indicates a day with mean temperature in (19,20]. Growing season is June-September (122 days).

Figure 4:

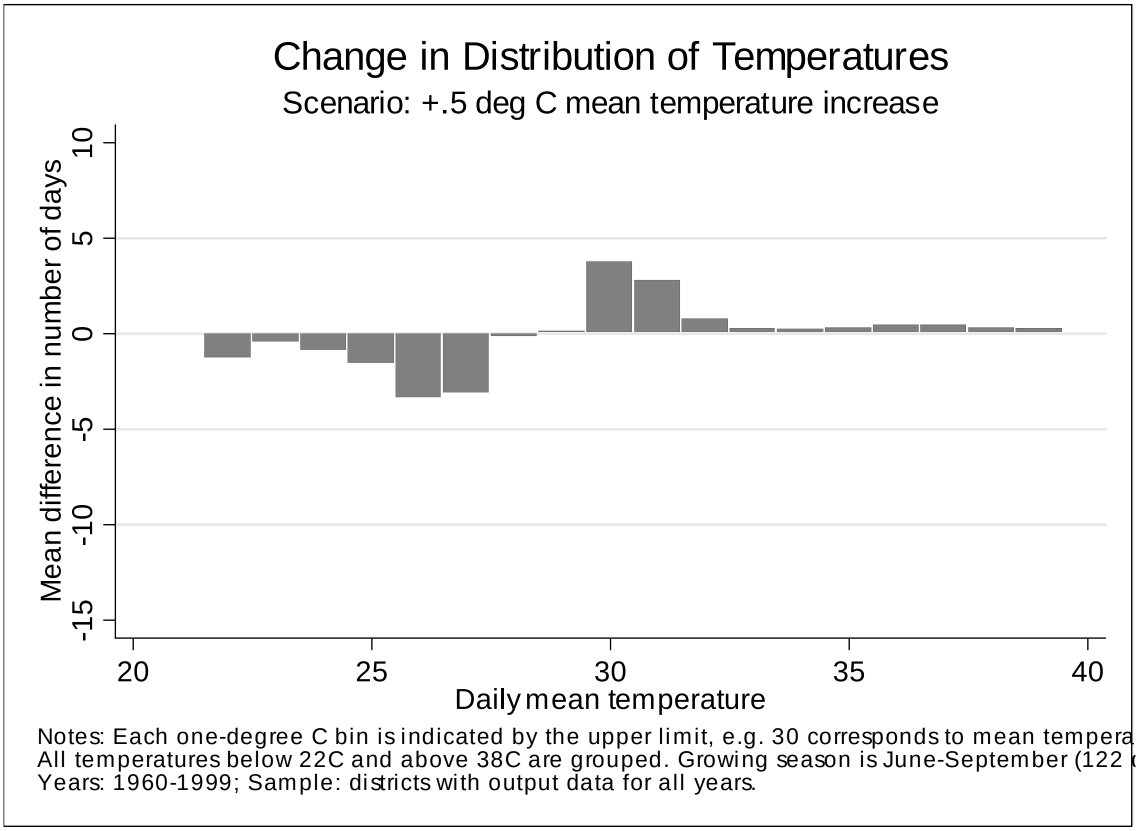
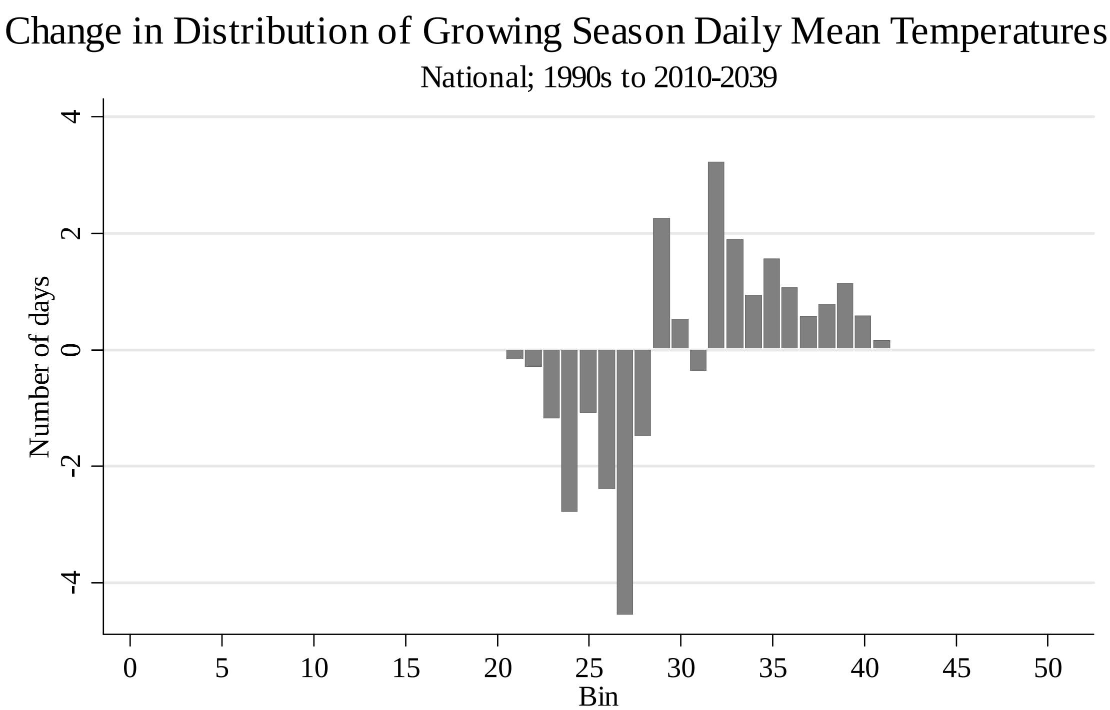
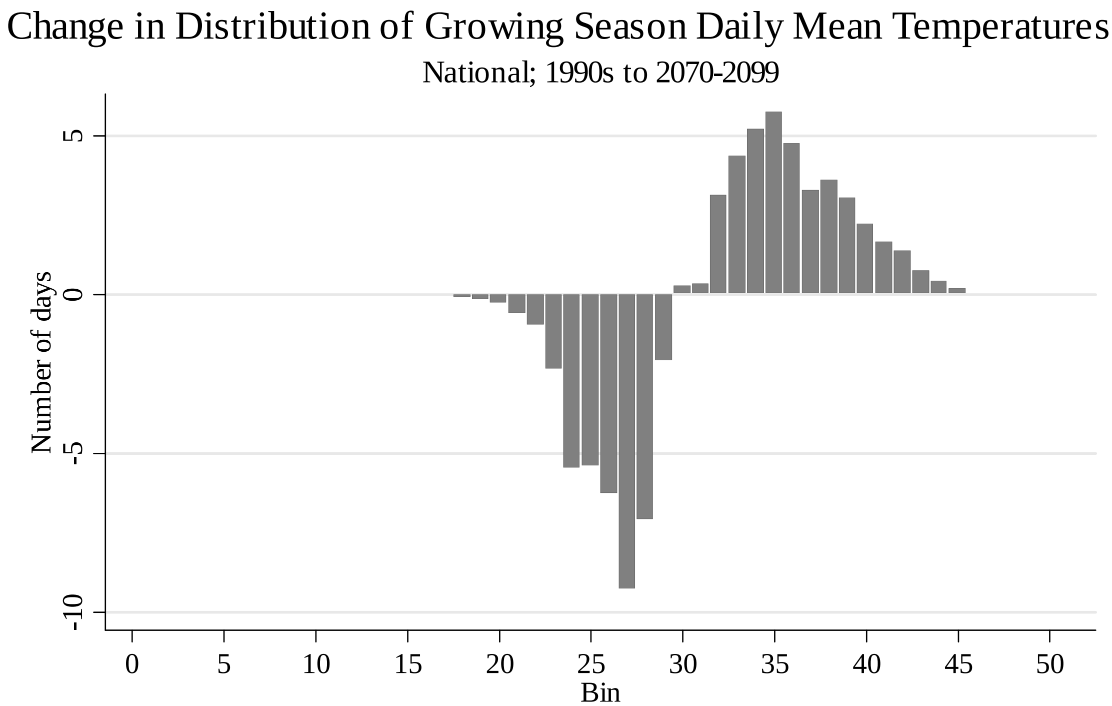

Change in Distribution of Temperatures
Scenario: +.5 deg C mean temperature increase
Mean difference in number of days
10
5
0
-5
-10
-15
20
25
30
35
40
Daily mean temperature
Notes: Each one-degree C bin is indicated by the upper limit, e.g. 30 corresponds to mean tempera
All temperatures below 22C and above 38C are grouped. Growing season is June-September (122
Years: 1960-1999; Sample: districts with output data for all years.

Figure 5:



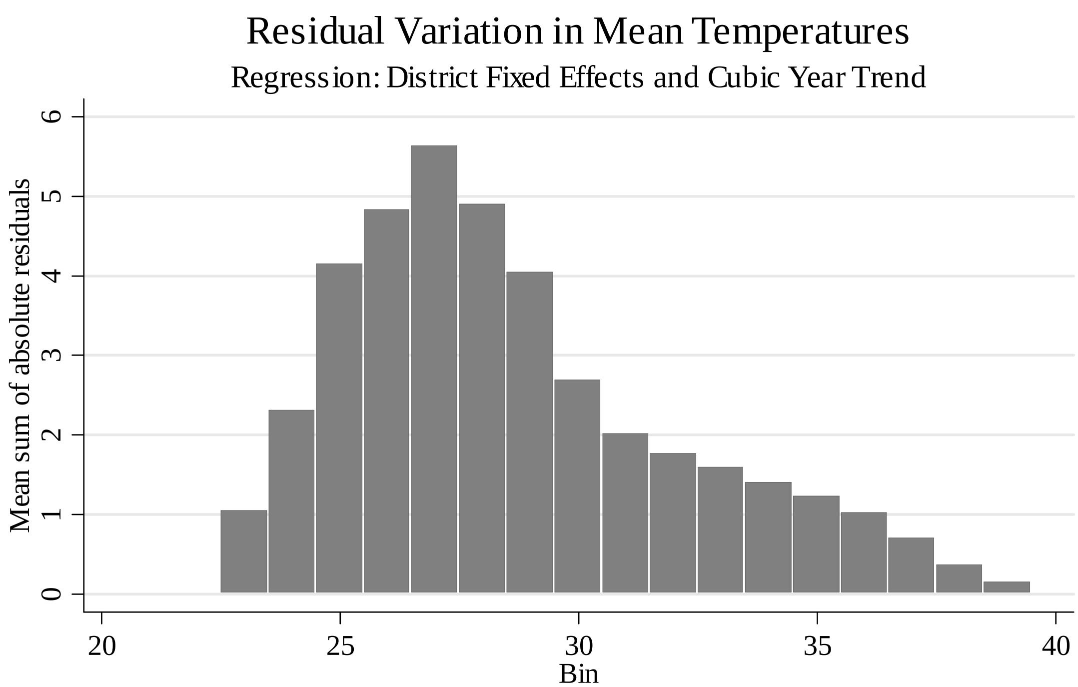
Notes: Each one-degree C bin is indicated by its upper limit, e.g. 30 corresponds to mean temperature (29,30].
Source: Hadley A1F1 model, rescaled. Growing season is June-September (120 days).

Figure 6:



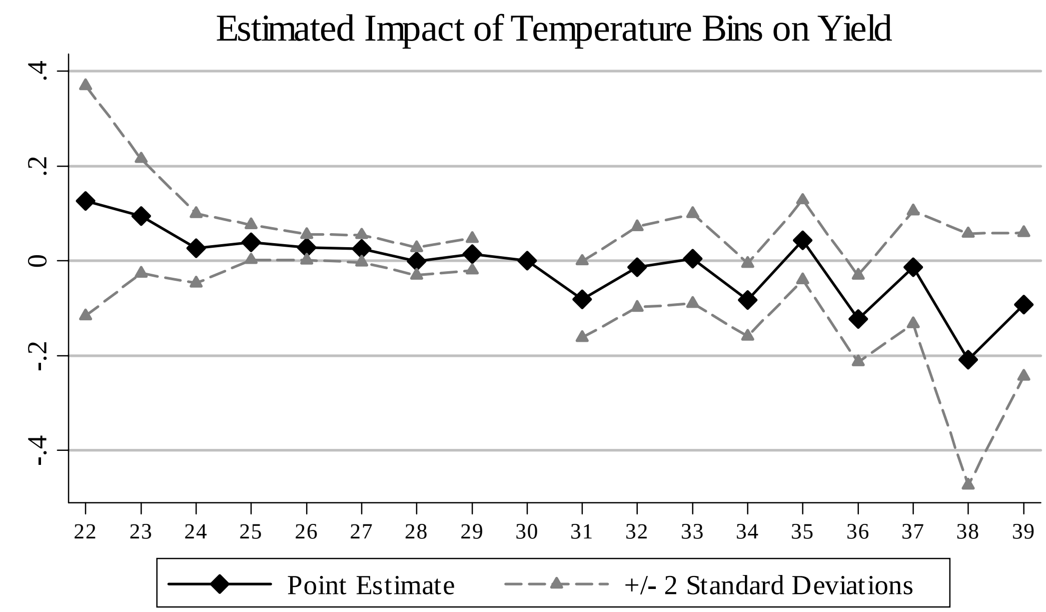
Notes: Each one-degree C bin is indicated by its upper limit, e.g. 30 corresponds to mean temperature (29,30].
Source: Hadley A1F1 model, rescaled. Growing season is June-September (120 days).

Figure 7:



Notes: Each one-degree C bin is indicated by the upper limit, e.g. 30 corresponds to mean temperature (29,30]. All temperatures below 22C and above 38C are grouped. Growing season is June-September (122 days). Years: 1960-1999; Sample: districts with output data for all years.

Figure 8:



Notes: Plots coefficients from FGLS regression of yield on number of growing season days in each bin, plus district fixed effects, monthly precipitation (and squares) and regional cubic time trends. Each 1-degree C bin is indicated by upper limit, e.g. 24 indicates a day in with mean temperatures in (23,24] C. The (29,30] bin is the omitted category, so estimates are relative to the effect of a day in (29,30]. Temperatures below 22 C and above 38C are grouped. Sample: 218 districts with output data for all years 1960-1999.

Appendix Table 2.A: Residual Variation in Growing Season Degree-Days with Alternative Upper Bounds

Panel 1: Degree-Day Upper Bound of 33 C

		District*year observations differing from predicted value by more than									
Mean: 2,479.1; N:8720		60 deg-days (C)		120 deg-days (C)		180 deg-days (C)		240 deg-days (C)		300 deg-days (C)	
Regressors	RMSE	Number	Share	Number	Share	Number	Share	Number	Share	Number	Share
Constant only	204.38	6826	0.783	5069	0.581	3304	0.379	1730	0.198	994	0.114
District FEs	53.72	2168	0.249	235	0.027	40	0.005	2	0.000	0	0.000
District FEs, Linear Year	52.58	2058	0.236	206	0.024	33	0.004	0	0.000	0	0.000
District FEs, Quadratic Year	52.15	2031	0.233	212	0.024	36	0.004	1	0.000	0	0.000
District FEs, Cubic Year	52.10	2025	0.232	208	0.024	33	0.004	0	0.000	0	0.000
District and Year FEs	35.52	753	0.086	18	0.002	0	0.000	0	0.000	0	0.000
District and Year*Region FEs	28.51	374	0.043	3	0.000	0	0.000	0	0.000	0	0.000
District and Year*State FEs	21.29	104	0.012	2	0.000	0	0.000	0	0.000	0	0.000

Panel 2: Degree-Day Upper Bound of 34 C

		District*year observations differing from predicted value by more than									
Mean: 2,489.3; N:8720		60 deg-days (C)		120 deg-days (C)		180 deg-days (C)		240 deg-days (C)		300 deg-days (C)	
Regressors	RMSE	Number	Share	Number	Share	Number	Share	Number	Share	Number	Share
Constant only	211.53	6855	0.786	5146	0.590	3489	0.400	1984	0.228	1093	0.125
District FEs	56.43	2379	0.273	295	0.034	47	0.005	3	0.000	0	0.000
District FEs, Linear Year	55.28	2287	0.262	247	0.028	41	0.005	2	0.000	0	0.000
District FEs, Quadratic Year	54.83	2271	0.260	245	0.028	44	0.005	3	0.000	0	0.000
District FEs, Cubic Year	54.76	2264	0.260	244	0.028	42	0.005	2	0.000	0	0.000
District and Year FEs	37.00	853	0.098	22	0.003	0	0.000	0	0.000	0	0.000
District and Year*Region FEs	29.62	418	0.048	5	0.001	0	0.000	0	0.000	0	0.000
District and Year*State FEs	22.03	130	0.015	2	0.000	0	0.000	0	0.000	0	0.000

Panel 3: Degree-Day Upper Bound of 35 C

		District*year observations differing from predicted value by more than									
Mean: 2,495.9; N:8720		60 deg-days (C)		120 deg-days (C)		180 deg-days (C)		240 deg-days (C)		300 deg-days (C)	
Regressors	RMSE	Number	Share	Number	Share	Number	Share	Number	Share	Number	Share
Constant only	216.36	6876	0.789	5193	0.596	3602	0.413	2144	0.246	1165	0.134
District FEs	58.51	2517	0.289	325	0.037	55	0.006	5	0.001	0	0.000
District FEs, Linear Year	57.33	2419	0.277	282	0.032	48	0.006	3	0.000	0	0.000
District FEs, Quadratic Year	56.88	2424	0.278	278	0.032	52	0.006	5	0.001	0	0.000
District FEs, Cubic Year	56.78	2427	0.278	270	0.031	47	0.005	3	0.000	0	0.000
District and Year FEs	38.24	948	0.109	30	0.003	1	0.000	0	0.000	0	0.000
District and Year*Region FEs	30.55	477	0.055	8	0.001	0	0.000	0	0.000	0	0.000
District and Year*State FEs	22.71	143	0.016	3	0.000	0	0.000	0	0.000	0	0.000

Panel 4: Degree-Day Upper Bound of 36 C

		District*year observations differing from predicted value by more than									
Mean: 2,499.5; N:8720		60 deg-days (C)		120 deg-days (C)		180 deg-days (C)		240 deg-days (C)		300 deg-days (C)	
Regressors	RMSE	Number	Share	Number	Share	Number	Share	Number	Share	Number	Share
Constant only	219.12	6881	0.789	5224	0.599	3660	0.420	2246	0.258	1233	0.141
District FEs	59.91	2601	0.298	356	0.041	64	0.007	7	0.001	0	0.000
District FEs, Linear Year	58.72	2521	0.289	318	0.036	50	0.006	6	0.001	0	0.000
District FEs, Quadratic Year	58.24	2517	0.289	300	0.034	53	0.006	5	0.001	0	0.000
District FEs, Cubic Year	58.13	2519	0.289	300	0.034	48	0.006	5	0.001	0	0.000
District and Year FEs	39.13	1011	0.116	32	0.004	1	0.000	0	0.000	0	0.000
District and Year*Region FEs	31.23	519	0.060	9	0.001	0	0.000	0	0.000	0	0.000
District and Year*State FEs	23.24	170	0.019	3	0.000	0	0.000	0	0.000	0	0.000

Notes: Table counts residuals from regressions of district*year observations on regressors listed in row headings. Cell entries are number of residuals of absolute value greater than or equal to the cutoffs given in the column headings.Years: 1960-1999; Sample: 218 districts with output data for all years.