Weather, Climate Change and Death in India

Robin Burgess Olivier Deschenes Dave Donaldson Michael Greenstone*

April 20, 2017

Abstract

This paper reveals a stark inequality in the effect of ambient temperatures on death in human populations. Using district-level daily weather and annual mortality data from 1957 to 2000, we find that hot days lead to substantial increases in mortality in rural but not urban India. Despite being far poorer, the mortality response in urban India is not dissimilar to that in the US over the same period. Looking into potential mechanisms we find that the rural death effects are driven by hot days in the growing season which reduce productivity and wages in agriculture. Consistent with a model of endogenous survival in the face of credit constraints, we also find that the expansion of bank branches into rural India helped to mitigate these effects. When coupled with a climatological model that predicts many more hot days in a typical year by the end of this century, these estimates imply considerable reductions in rural Indian, but not urban Indian or US, life expectancy *ceteris paribus*.

1 Introduction

Is humankind insulated—by and large—from the vagaries of nature? This question has been a central concern in the natural and social sciences for centuries, beginning with Montesquieu (1748), Malthus (1798) and Marsh (1865). A recent focus has been on developing countries, where

*Correspondence: ddonald@stanford.edu. Affiliations: Burgess: LSE and IGC; Deschenes: UCSB and NBER; Donaldson: Stanford and NBER; Greenstone: Chicago and NBER. We are grateful to Clare Balboni, Oriana Bandiera, Marianne Bertrand, Tim Besley, Bronwen Burgess, Klaus Desmet, Esther Duflo, Edward Glaeser, Matthew Gentzkow, Selim Gulesci, Matthew Kahn, Melanie Morten, Mushfiq Mobarak, Ben Olken, Torsten Persson, Imran Rasul, Jesse Shapiro, Nick Stern, Robert Townsend, Matthew Turner, Dean Yang and seminar participants at the Asian Development Review 2016 Conference, Asian Development Research Institute Patna, Berkeley, Boston University, Chicago Booth, Columbia, Harvard Kennedy School, IIES Stockholm, Indian Statistical Institute Development and Growth Conference, LSE, LSE/Grantham Adaptation, Green Growth and Urbanization Conference, Minnesota 2015 Climate Change Conference, MIT-Harvard, MOVE Barcelona Conference, NEUDC, Pakistan Institute of Development Economics Silver Jubilee Conference, Pompeu Fabra, Stanford, UCSB, Virginia Environment and Development Conference, the World Bank and Yale for helpful comments.

livelihoods are hypothesized to be especially fragile in the face of weather shocks (Sen, 1981; Dreze and Sen, 1989; Paxson, 1993; Townsend, 1994; Deaton, 1997 and Dasgupta, 2010). The realization that the earth's climate is warming, which will add to the weather risks that human populations face, only intensifies the contemporary interest in this topic.

An interesting question in this context concerns whether populations at different stages of development are affected differently by the same weather variation. This paper tries to make progress on this issue by using the best available district-level panel data from India, covering daily temperature and annual rural and urban mortality. To allow us to compare effects to those observed in a developed economy, we also gather panel data from the US over the same period. This analysis, in turn, provides insights into the future impacts of climate change—we combine estimates from our temperature-mortality analysis with a climatological model in order to gauge how the increase in hot days predicted to occur over this century may will affect life expectancy.

Our data therefore allow us to test whether economic development mediates the relationship between weather, climate change and human health as proxied by mortality. In India, urban areas are defined by having more than 75% of the working population engaged in non-agricultural activities (Indian Census, 2011). We are therefore testing whether the *same* daily temperature variation at the district level affects populations living in the countryside (who are mainly dependent on subsistence agriculture) differently from populations living in towns and cities (who are involved principally in services and manufacturing). Both rural and urban Indian populations contain large numbers of people living in extreme poverty and so we are comparing effects in two poor developing country populations to those in a rich developed country population.[1]

Panel A of Figure 1 sets the scene for the paper. It plots, for India and the US from 1957 to 2000, the estimated impact on annual mortality of having an extra day whose daily mean temperature lies in each of seven temperature bins. The reported estimates are relative to a day in a reference category bin, which we take to be 70°-74° F. A key contrast emerges. Equivalently hot days lead to much greater increases in the death rate in India than they do in the US—for example, a single day above 95° F (relative to a day in the reference category) elevates annual mortality by 0.74% in India but only by 0.03% in the US.

In panel B we split these results out for India's rural and urban populations. This makes it clear that the effects we see in panel A are driven by effects on the rural population. Hot days significantly increase mortality in rural but not urban India. Despite the remarkable extent to which within-village shocks appear to be shared and smoothed in rural India in seminal studies such as Townsend (1994), our results suggest that the smoothing of important cross-district shocks remains very much incomplete. This is the key result of our paper. As hot weather sweeps across an Indian district it affects those that live in the countryside very differently from those who live

[1] In 1957 62% of the rural population and 49% of the urban population was below the Indian poverty line. By 1999/2000 these numbers had dropped to 31% and 21% (Datt, 1998; Datt et al., 2003). In the US 0% of the population was below this extreme poverty line at these two points in time.

in towns and cities. Depite high rates of extreme poverty in both rural and urban India, living in urban areas helps to protect people from the deleterious effects of hot weather.

Panel C makes it clear that the urban Indian and US mortality responses are not dissimilar. Even though the income distributions of urban Indian and US populations are largely non-overlapping, living in an urban area even at low levels of development appears to confer a significant degree of protection against hot weather. Indeed this may be a key reason why people choose to live in urban areas in developing countries just as they did in developed countries in the past (Glaeser, 2014).

Worryingly, the hot days that are evidently so lethal among India's large rural population are predicted by climatologists to occur with increasing frequency throughout the rest of this century.[2] For example, in an average year between 1957 and 2000 there were five days with average daily temperatures above 95° F in India. But according to the Hadley (HadCM3) climate change model (under a business-as-usual scenario) there will be 75 such days in an average year between 2075 and 2099. The corresponding rise for the 90-94° F bin is from 21 days to 64. In the US, the number of days move from zero to 29 and from one to 26 in the $> 95°$ F and 90-94° F bins respectively.

This predicted warming of the climate combined with estimates from Figure 1, imply that *ceteris paribus,* life expectancy will fall by 2.37, 5.09 and 10.41 years for rural Indians born in 2015-2029, 2045-2059 and 2075-2099 respectively relative to life expectancy for those born in 2000.[3] The corresponding figures are 0.35, 1.05 and 2.82 years for urban Indians and 0.06, 0.15 and 0.28 years for US populations. Climate change could therefore have highly unequal effects across developed and developing countries but also across rural and urban populations within developing countries. Naturally, many limitations apply when using the impact of past interannual weather variation to predict the consequences of future climate change which is a gradual, anticipated and permanent change in the weather. Nonetheless, what is clear is that the pattern of predicted effects is alarming and requires our attention now.

The deleterious effects of hot weather on mortality seen in Figure 1 could be driven by two broad sets of factors: hot weather could affect health either directly (through human physiology or disease) or indirectly (through real incomes). In order to clarify the interplay between these channels, we develop an equilibrium model of intertemporal consumption choice in the presence of endogenous borrowing an savings constraints—an extension of canonical models of incomplete markets with heterogenous agents (Bewley (1986), Aiyagari (1994), and Acemoglu and Jensen (2015)) to incorporate endogenous health status as in Grossman (1972) and Becker, 2007. Households choose to spend the incomes they earn either on health goods (which improve health status but provide no utility directly) or on consumption goods (which provide utility). If health is threat-

[2]From 1960 to 2000 India moved from 82% to 72% rural so the at-risk population will continue to comprise a large share of the total population over the coming century.

[3]We do not try to model how life expectancy will evolve over the coming century. Our estimates are relative to a 2000 benchmark and are therefore capturing the degree to which the mortality impacts of rising temperatures will constrain the likely growth in life expectancy over the coming century.

ened or incomes fall as a result of high temperatures then households face a trade-off between health-enhancing and utility-enhancing expenditures. Our model makes clear that irrespective of whether citizens face direct shocks (that can be partially offset by health spending) or indirect shocks (that imply less resources available for health spending) any policy that that allows citizens to maintain expenditures on health goods can be effective in allowing them to avoid death.

Our model of financially-constrained intertemporal choice implies that an improved financial environment, such as one with better access to banks, should mitigate the impact of hot weather on ill-health as it would enable households to better smooth consumption (and hence their health status) in the face of weather shocks. Using the identification strategy of Burgess and Pande (2005) we find evidence that India's rural bank branch expansion program (which ran from 1977 to 1990) mitigated substantially the impact of hot weather on mortality. For example, according to our estimates, the median district in terms of bank branch penetration saw the impact of hot days on mortality fall by a factor of approximately ten. This result is important for two reasons. First, because it is consistent with one of the implications of our model—namely, that technologies which enable citizens to smooth spending on health-improving goods across periods in response to temporary shocks should assist them in avoiding premature mortality. And second, because it provides confirmation that the state can take actions which protect citizens from the mortality impacts of hot weather shocks.

To better understand the channels through which hot days increase mortality in rural but not urban India, we examine the effects of hot days on incomes in the agricultural (primarily rural) and registered manufacturing (primarily urban) sectors.[4] We find that hot weather sharply depresses agricultural yields and the wages of agricultural laborers in rural areas but exerts no impact on output and wages in the registered manufacturing sector. What is more, when we divide weather according to the growing season and non-growing seasons (based on the timing of monsoon onset) we find that it is only growing season weather that affects rural incomes. Both of these facts are consistent with the income effect of hot weather being more severe for rural versus urban populations. In line with this, the large effect of hot days on annual mortality in rural areas seen in Figure 1 is entirely driven by growing season days, despite the fact that the hottest period of the year is typically in the non-growing season. Urban areas, by contrast, see no effect of hot days during either the growing or the non-growing season.

By examining whether economic development can protect human populations from weather variation and climate change we build connections between three literatures. The first is a development economics literature that examines whether rural populations can smooth consumption over weather-induced income shocks (Rosenzweig and Stark, 1989; Rosenzweig and Wolpin, 1993; Paxson, 1993; Townsend, 1994; Deaton, 1997; Dercon and Krishnan (2000); Dercon (2005); Khandker (2012); Cole et al., 2013 and Chiappori et al. 2014). Here our main finding is that rural pop-

[4] In India the registered manufacturing sector is defined by the Factories Act of 1948 to include firms using electricity and 10 or more employees (or more than 19 employees without electricity)(Besley and Burgess, 2004).

ulations in India remain unprotected against aggregate, district-level weather variation. Neither informal risk sharing arrangements (Besley, 1995) nor public policy (Besley and Burgess, 2002) seem to be efficacious in protecting rural populations from the mortality effects of hot weather. The second is an urban economics literature that examines how moving from the countryside to towns and cities fosters development (Lewis, 1954; Harris and Todaro, 1970; Ciccone and Hall, 1996; Henderson, 2010; Kahn, 2013; Bryan et al., 2014; Glaeser, 2014 and Morten, 2016). Here the main insight of this paper is that living in towns and cities offers some insurance against inclement weather even at low levels of development when extreme poverty remains widespread.[5] Finally, the third is a literature that looks at the implications of climate change for the incomes and health of human populations (Nordhaus, 1994; Stern, 2006; Deschenes and Greenstone, 2007; Schlenker and Roberts, 2008; Deschenes and Greenstone, 2011; Deschenes, 2014; Burke et al., 2015; Barreca et al., 2016; Hsiang, 2016 and Carleton and Hsiang, 2016). Here our contribution is to reveal that climate change may have unequal effects human populations at different stages of development and to probe both theoretically and empirically the mechanisms that generate these effects.

The remainder of this paper proceeds as follows. The next section outlines a theoretical model that describes the mechanisms through which weather might be expected to lead to death, as well as how financial access could mitigate these effects. Section 3 describes the background features of India in our sample period from 1957-2000, as well as the data on weather, mortality, incomes and banking that we have collected in order to conduct our analysis. Section 4 outlines our empirical method. Section 5 presents the main results of the paper on weather and death. Section 6 presents the results on weather and income and Section 7 the results on climate change and death. Section 8 concludes.

2 A Model of Weather and Health with Financial Constraints

In this section we describe a theoretical framework within which to examine the potential for weather variation to pass through into mortality. In this setting, weather extremes can potentially harm human health both directly (through physiological stress, for example) and indirectly (because of lower incomes or those who work in weather-dependent occupations). We assume that individuals have access to a market for health-improving goods that enhance the probability of survival in the face of such shocks. However, individuals may face financial constraints in their ability to shift resources across time periods in the face of transitory shocks. Our model combines features of canonical models of endogenous health (for a summary, see, e.g. Becker (2007)) with canonical models of financial constraints among heterogeneous households in general equilibrium (due to Bewley (1986) and Aiyagari (1994)). We draw on the developments of Acemoglu and

[5]The economic history of developed countries suggests that migration from rural to urban areas provided a means to escape impoverishment in rural areas, especially following rural agricultural shocks such as famines (e.g. Lees and Modell, 1977 for the case of Ireland). The literature has also identified other channels through which urban residents are insured against the mortality impacts of environmental shocks, including greater resources devoted to zoning, building codes and early warning systems (Kahn, 2005; Kahn, 2013).

Jensen (2015) to derive comparative statics for such a model that hold despite the richness of all the heterogeneity and market incompleteness involved.

2.1 Model Setup

Consider an economy with a large number (formally, a continuum) of infinitely-lived agents indexed by i. Agents value both consumption c_{it} and health status h_{it}. In particular, agent i seeks to maximize

$$\mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t u_i(c_{it}, h_{it}) \right]. \tag{1}$$

We place no restrictions on $u_i(c_{it}, h_{it})$ other than that c_{it} and h_{it} are complements, as seems natural at this level of aggregation. In particular, every agent can have a different (sub)-utility function, $u_i(\cdot)$.

Health status h_{it} is determined by a mixture of the agent's choices and chance. That is, agent i can improve her health status by purchasing a health-improving good, q_{it}. But she is still subject to health shocks due to the vector of random variables z_{it}, which is drawn from some Markov process and invariant distribution that can be specific to agent i. All told, health status is given by

$$h_{it} = h_i(q_{it}, z_{it}), \tag{2}$$

where we define $h_i(\cdot)$ such that it is increasing in q_{it} (so that the heath-improving good always improves health status) and decreasing in z_{it} (such that z_{it} is a "bad" health shock). Otherwise, the function $h_i(\cdot)$ is unrestricted and, again, can differ across the heterogeneous agents.

The consumption good c_{it} is produced using an aggregate production function that requires capital and labor inputs. We denote this production function by $f(\cdot)$ and assume that it exhibits non-increasing returns to scale. We assume that the health good q_{it} can be bought and sold on world markets for price P_t, which the country under study is too small to affect. The price of the consumption good is taken as our numeraire.

Each agent is endowed with a_{it} units of efficiency-adjusted labor in period t and earns income by selling this labor at the going wage rate, w_t. Again, a_{it} is potentially random so we allow it to be an agent i-specific function of z_{it}, given by $a_{it} = a_i(z_{it})$. While the same shock z_{it} enters both the agent's productivity and health status, this is without loss of generality due to the fact that z_{it} is a vector and the functions $h_i(q_{it}, \cdot)$ and $a_i(\cdot)$ depend on z_{it} in unrestricted ways.

We assume that the agent has access to a means for (partial) saving and borrowing that allows her to smooth (partially) consumption and health status over time. Let x_{it} denote the agent's stock of savings at date t. The agent can save and earn an interest rate r_t, or borrow at this same rate. But x_{it} is restricted by a borrowing constraint which means that x_{it} cannot fall below $-\underline{b}_i(r_t)$, and similarly by a savings constraint which means it cannot exceed $\bar{b}_i(r_t)$. Both constraints are unrestricted (i.e. could be arbitrarily relaxed such that the agent is in fact unconstrained), free to vary across agents, and potentially an endogenous function of the interest rate r_t. Savings are transformed into capital which is used in production. These considerations imply that agent i faces a budget constraint each period that is given by

$$x_{i,t+1} \leq r_t x_{it} + w_t a_i(z_{it}) - P_t q_{it} - c_{it}. \tag{3}$$

We follow a timing convention in which agents' state variables (known at the start of period t when decisions about c_{it}, q_{it} and $x_{i,t+1}$ are made) are x_{it} and z_{it}. Agents take prices as given but have rational expectations about how these prices are formed in equilibrium. The state variables in the first period (x_{i0}, z_{i0}) are exogenous and given.

Finally, in equilibrium all markets must clear each period, which then determines endogenously the factor prices (r_t, w_t). Non-increasing returns to scale production implies that the returns to capital and labor will be functions of the aggregate amount of capital in the economy, which we denote $K_t \equiv \int x_{it} di$. These factor prices will also be a function of the aggregate amount of labor endowment but that will be constant, as we discuss below, so we suppress this dependency for now. We therefore denote these dependencies on aggregate capital by $r_t = r(K_t)$ and $w_t = w(K_t)$.

To summarize, this is an Aiyagari-Bewley model (with extensions due to Acemoglu and Jensen (2015)), but with two additional features that are relevant for our setting: (i) the endogenous health status h_{it} driven both by a vector of shocks z_{it} and choices q_{it}; and (ii), savings constraints in the form of $x_{it} \leq \bar{b}_i(r_t)$. Like in canonical Aiyagari-Bewley models, this setting features no *aggregate* uncertainty. While it is certainly the case that each agent faces a rich set of idiosyncratic shocks (the vector of random variables z_{it} affects both the agent's health status $h_{it} = h_i(q_{it}, z_{it})$ and her productivity $a_{it} = a_i(z_{it})$), and so there is a great deal of potential for *individual-level* uncertainty, on average these shocks cancel out when it comes to determining aggregates such as K_t or (r_t, w_t). As is well known, this greatly simplifies the study of heterogenous-agent general equilibrium models.

So far this model has not explicitly featured a role for weather shocks in determining either health or production. But we can easily think of a subset of the elements of the vector z_{it} as determined by the weather faced by agent i in year t. This subset could include many different components of weather variation: hot weather (and different forms of hot weather), extremely cold weather, precipitation, and humidity. One important caveat is that the model focuses on the idiosyncratic components of weather variation in the modeled country. But for a large country such as India, one would expect agents' weather shocks to become increasingly distinct, leading to only vanishing aggregate components of weather variation in the limit, and it is this limiting case that we study here.

2.2 Mechanisms Relating Weather to Health

There are two fundamental reasons for the weather to potentially harm an agent's health status in this model. The first is what we refer to as a *direct* health effect: random shocks z_{it}, potentially including weather shocks, enter the agent's health status directly, as in equation (2). That is, holding constant the agent's income and consumption decisions, we expect a negative weather shock (such as hot temperatures) to affect this agent's health adversely. An extensive public health literature discusses the potential for such direct effects of high temperatures on human health (see for example, Basu and Samet (2002) for a comprehensive review).[6] Periods of excess temperature place additional stress on cardiovascular and respiratory systems due to the demands of body temperature regulation. This stress is known to impact on the elderly and the very young with particular severity, and can, in extreme cases, lead to death (Klineberg (2002); Huynen et al. (2001)). An alternative direct effect of extreme weather on health in India could include the possibility that disease pathogens (for example, diarrhoeal diseases) thrive in hot and wet conditions, or that some vectors of disease transmission (such as mosquitoes in the case of malaria) thrive in hot and wet environments. All of these possibilities can be represented by the vector of random shocks z_{it} and can affect agent i in her own heterogeneous manner (which could capture her region, her climate, or her own physiological individuality) through the agent i-specific health function, $h_i(\cdot)$.

The second, more *indirect,* mechanism through which weather can affect health in this model is through the agent's productivity, $a_i(z_{it})$. This term also depends on the vector of random shocks z_{it} that could again depend on various aspects of the weather that matter for productivity (whether they are the same weather attributes that matter for health or not simply depends on how the arguments that enter $h_i(\cdot, z_{it})$ compare to those that enter $a_i(z_{it})$, and these are unrestricted). Such a dependency is extremely likely in rural areas (as we document empirically below) where incomes depend largely on agricultural activities or their complements. But it is also possible that some agents' incomes in non-agricultural settings, such as in urban areas, are similarly exposed to weather variation (though below we find no evidence for this in the Indian context).

In the face of these two sorts of potential weather shocks, direct or indirect, any agent i in our model would seek to minimize the damage that such negative shocks do to her utility, $u_i(c_{it}, h_{it})$. Mitigation can be achieved through two broad mechanisms. First, the agent can purchase some of the health-improving good, q_{it}, to help buttress her health status to the extent that her function $h_i(q_{it}, z_{it})$ allows. This health good could include traditional health products such as medicine or services such as a health center visit. Equally, it could include a subsistence component of food consumption (that which increases health status but is not valued in $u_i(c_{it}, \cdot)$ directly). Or, given

[6]Extremely cold temperatures can affect human health adversely through cardiovascular stress due to vasoconstriction and increased blood viscosity. Deschenes and Moretti (2009) find evidence of a moderate effect of extreme cold days on mortality (especially among the elderly) in the United States, though this effect is concentrated among days below 10° F. Days in this temperature range are extremely rare in India."

our focus below on temperature shocks, an important health good might be the use of cooling technologies like air conditioning. More broadly, this health good could also represent any leisure or rest (i.e. foregone earning opportunities) that the agent might decide to 'purchase' so as to improve her health. This could include the decision to work indoors rather than outdoors when it is hot, or to accept a job at lower pay so as to avoid working outside on a hot day. Of course, there are limits to the impact that such vehicles can have, both because of potential curvature in $\frac{\partial h_i(q_{it},z_{it})}{\partial q_{it}}$ and because any funds spent on the health-improving good in period t reduce what is left to be spent on other components of utility—the consumption good c_{it} or either good in the next period ($c_{i,t+1}$ or $q_{i,t+1}$).

A second form of mitigation potentially available to the weather-shocked agent would be to reallocate resources between periods. This would involve either borrowing or reducing savings in the decision period (when income is temporarily low or health status receives a temporary insult) in order to attempt to smooth marginal utility as much as possible over time. However, borrowing and savings constraints, to the extent that they are close to binding for agent i, may limit this smoothing: a binding borrowing constraint places an upper bound on $u_i(c_{it}, h_{it})$ irrespective of what is expected in future periods, and a binding savings constraint limits the extent to which the agent can accumulate precautionary savings (as she would desire to do in the presence of a borrowing constraint). In such cases of limited smoothing, the resulting impact of a weather shock on health (whether it occurs because of direct health effects or because of indirect effects on income) could be substantial.

It is important to stress that, while the above discussion introduced a conceptual distinction between direct and indirect effects of weather on health, for many of the most important policy questions the distinction is moot. The presence of health-improving goods (which we believe deserves a wide interpretation, as discussed above) means that there is no fundamental difference between a weather shock that harms the agent's health indirectly through lower income or directly through the health status function. In the former case, the weather shock affects the agent's budget constraint, meaning that health suffers as less of the health-improving good can be purchased. And in the latter case, the agent simply uses her income to buy health goods so as to offset the direct effect of weather on health, which effectively reduces the budget remaining for consumption goods. Indeed, it is straightforward to see that for any given observed effect of weather on health, any policy that distributes funds to agents directly, or that distributes or subsidizes health goods, could be designed optimally regardless of the underlying reason (that is, direct or indirect) that weather matters for human health.

2.3 Comparative statics: equilibrium effects of weather on health

We now describe the implications of the model in Section 2.1 for the equilibrium relationships between weather and health one might expect to see in a population. Our focus here is on qualitative comparative statics that guide intuition for the empirical work to follow. For this reason, the comparative statics tools developed in Acemoglu and Jensen (2015) are particularly useful.

Somewhat surprisingly, even though the agents in our model feature unrestricted heterogeneity in virtually every sense—their health status functions, their utility functions, their earnings functions, their initial assets, the stochastic processes from which their individual shocks are drawn, and the savings and borrowing constraints that they are potentially subject to—and even though there are general equilibrium feedback interactions among agents (via endogenous wages, endogenous interest rates, and endogenous borrowing and savings constraints), Acemoglu and Jensen (2015) offer methods for characterizing how aggregate variables of the sort that interest us here will respond to exogenous changes in the economic environment. Further, even in situations like ours where the economic primitives are not sufficiently restrictive to ensure that the economy features a unique equilibrium, we can derive "robust" comparative statics results (in the spirit of Milgrom and Roberts (1994)) that ask whether the range of values that a given variable of interest can take (across all equilibria) will move in a monotone way. [7]

Throughout, we focus on the following aggregate variable, which aims to summarize the weather-health nexus:

$$H_t \equiv \left| \frac{\int h_{it} z_{it}^{(T)} di}{\int h_{it} di} \right|. \tag{4}$$

Here, we let the scalar $z_{it}^{(T)}$ denote one particular component of interest from the shock vector z_{it}. This component stands in for any particular weather shock (though, naturally, temperature shocks are our main focus, hence the "T" superscript). Then the expression in equation (4) represents the aggregate correlation between health status h_{it} and this weather shock $z_{it}^{(T)}$, scaled by the aggregate amount of h_{it}. Recall that we have normalized the shock vector z_{it} such that these shocks harm agents' health status. This means that H_t increases when there is a greater (absolute value of) correlation between health status among individuals i and time periods t for which the weather shock $z_{it}^{(T)}$ was positive. While we focus on one element $z_{it}^{(T)}$, our results extend trivially to any other element as well.

This correlation (or, in some domains, the lack thereof) is the focus of our empirical work in Section 5 and so it is rightly our focus here. But there is an obvious distinction between the focus here, health status, and that in our empirical work below, where the main proxy for health status at our disposal is the mortality rate. In order to make this leap from the study of health to the study of death we simply note that the two are likely to be highly correlated, especially in a low-income environment. This is also an assumption of convenience because of the many complications

[7] The results in Acemoglu and Jensen (2015) imply that when certain mild regularity conditions (continuity of agents' utility, health status, and productivity functions, as well as of the aggregate production function; compactness of agents' choice sets; and that the stochastic processes satisfy the Feller property) then a stationary equilibrium exists.

an Acemoglu and Jensen (2015) model with endogenous mortality would need to overcome (such as the non-stationarity of the population distribution, the need to introduce an annuity market to prevent agents from dying with debt or savings, and the non-recursive nature of the optimization problem involved).

To the extent that health status h_{it} and mortality risk are strongly correlated, H_t from equation (4) will be correlated with the object of empirical study in Section 5 below, namely the extent to which a district sees more mortality than usual when its weather is hotter than usual. While there is clearly not a continuum of districts in India, the fact that there are approximately 400 implies that the cross-district correlation under study in Section 5 will be well approximated by the cross-individual correlation in equation (4).

Comparing urban to rural India:

We begin with a comparison between a purely urban economy and a rural economy. As discussed above, one key distinction between rural and urban regions in India is that in the rural economy there is potentially a greater exposure of residents' incomes—because they are so likely to come from the agricultural sector—to weather variation. Another distinction could be that the availability of health goods is considerably lower. A third distinction is that incomes are typically lower. In order to capture the model's predictions about a comparison between rural and urban areas, we consider that the above economy actually comprises two subsets of agents, those who live in rural ($i \in \mathcal{R}$) and urban ($i \in \mathcal{U}$) areas of the country respectively.

The comparative static exercise that compares rural to urban agents takes all rural agents and changes their economic attributes, discretely, from values that correspond to urban areas to those that correspond to rural areas. Specifically, we consider changes to the subset of rural agents $i \in \mathcal{R}$ that simultaneously: (i) increase the responsiveness of $a_i(z_{it})$ to the weather component $z_{it}^{(T)}$; (ii) raise the price that rural residents pay for health goods (denoted P_t^R) above the price for the same goods in urban areas (P_t); and (iii) lower the incomes of rural residents via a strict reduction in $a_i(z_{it})$ at any level of z_{it}. In the Acemoglu and Jensen (2015) framework, all of these changes are "positive" shocks to the group of rural residents, meaning that any one of these changes (or all three at the same time) would cause an increase in H_t. Further, if we were to create the value of H_t that corresponds to integrating only over rural residents ($i \in \mathcal{R}$) in equation (4) and define this as H_t^R, then this comparative static would cause H_t^R to increase as well. This implies that we should expect to see a greater correlation between weather shocks and health (and health proxies such as the mortality rate) in rural areas than in urban areas. Naturally the same is true if we interpret the urban areas of this model as similar to the US—simply an economy with relatively little dependency of productivity on weather, relatively cheap access to health-improving goods, and relatively high average incomes.

Proposition 1. *Relative to urban areas, rural areas (those with any combination of relatively* *higher dependence of productivity $a_i(z_{it})$ on weather shocks $z_{it}^{(T)}$, higher prices P_t for health-improving goods, or lower productivity a_{it} among any subset of agents in rural areas) will feature a stronger correlation between health and weather shocks (greater lowest- and highest-equilibrium values of H_t^R).*

The effect of improved bank access:

We now consider an economy in which a subset of agents enjoy an improvement in their ability to access the formal banking sector. Recall that each agent faces (potentially non-binding) financial constraints such that her savings x_{it} when the interest rate is r_t cannot fall below -$\underline{b}_i(r_t)$ and cannot exceed $\overline{b}_i(r_t)$. An improvement in the banking sector could be thought of as increasing the values of either or both of $(\underline{b}_i(r_t), \overline{b}_i(r_t))$, and hence relaxing the constraints faced by some agent i. The intuition is that the way households manage life in the presence of saving and borrowing constraints is to build up a buffer stock of savings. Empirically we model the arrival of banks as relaxing the savings constraint (an increase in $\overline{b}_i(r_t)$)[8] as rural bank branches enabled rural consumers to deposit (and earn positive interest) on thier savings.[9] In contrast, loans from rural bank branches were intended to be for production as opposed to consumption purposes.[10] By making it easier for many agents to deposit saving funds, greater access to formal sector banks would enable rural citizens to build buffer stocks of savings that could be drawn down if they experienced a weather shock. Our next comparative static result studies this change.

Proposition 2. *A relaxation in savings constraints (an increase in $\overline{b}_i(r_t)$ among any subset of agents) will decrease the aggregate correlation between weather and health status (smaller lowest- and highest equilibrium values of H_t).*

Summary of model implications:

The discussion above illustrates how, under extremely general conditions, there are clear predictions about what one might expect to see in an economic environment such as the one we study empirically below. We have allowed for a model in which agents' health status is endogenous but can potentially depend on weather shocks. These shocks work through either a "direct" mechanism (harming human physiology or worsening the disease environment, to the extent that these effects cannot be offset by the choice to purchase health-improving goods) or an "indirect" mechanism (lower incomes, and hence less resources for the purchase of health-improving goods). In this sense it is natural to expect that these effects—be they direct or indirect in origin—will be stronger in areas characteristic of rural India (greater dependency of income on weather, weaker

[8]By an increase in $\overline{b}_i(r_t)$ we mean, following Acemoglu and Jensen (2015), an increase in the value that $\overline{b}_i(r_t)$ takes at any fixed level of r_t

[9]These deposit accounts were considered attractive, on the margin, relative to other forms of saving.

[10]Which would imply a muted of bank access on -$\underline{b}_i(r_t)$.

access to health-improving goods, and lower overall income levels) than in urban India. This is the prediction behind Proposition 1.

Further, the overall impact of weather shocks on health, again whether they work through direct or indirect mechanisms, will depend on the ability of agents to smooth out shocks by reallocating consumption of the health-improving good across time periods. This smoothing requires the ability to borrow and save, and so we expect that any improvement in the ability to save (such as heightened access to formal banks that accept deposits) will reduce the effect of weather on health status. This is what Proposition 2 predicts.

We now turn to an empirical exploration of these possibilities.

3 Background and Data

To examine the relationship between weather variation and death outlined in our theoretical framework (Section 2) we put together a comprehensive Indian district-level data set for the 1957-2000 period.[11] We have gathered daily weather (temperature and precipitation) records for 334 districts covering the whole of India for this period. We combine this data with data on annual all-age mortality for rural and urban populations in these districts to look at the overall effect of the *same* variation in ambient temperatures on these two populations.[12] Data on annual output and wages in agriculture and registered manufacturing allow us to ascertain whether or not weather-income effects line up with weather-death effects. Data on rural bank branch expansion over 1977-1990 allows us to check whether, in line with our model, an intervention intended to relax saving constraints for rural citizens mitigated the impact of weather on death. Finally we use a climate change model to predict daily weather realizations in our 334 districts out to 2099. This enables us to look at the projected impact of climate change on life expectancy across this century. To allow us to compare our estimated effects with those in a developed country, we put together a data set covering daily weather (historical and predicted), mortality and life expectancy for the US over the same period.

The rural-urban distinction is critical to our analysis. In India it rests primarily on employment which is predominately agricultural in the case of rural citizens and non-agricultural in the case of urban citizens.[13] India's 1999/2000 National Sample Survey reveals that 76% of rural citizens belonged to households that draw their primary incomes from employment in the agricultural sector

[11] Please see Table 1 for summary statistics and the Data Appendix for a detailed description of how the different variables covered in this section were constructed.

[12] Beacuse we map the weather data to the district centroid we are looking at whether identical district-level weather affects resident rural and urban populations differently.

[13] The rural/urban assignment is based on the following criteria, used throughout official Indian statistics. Urban areas comprise 1. All places with a municipality, corporation, cantonment board or notified town area committee, 2. All other places which satisfied the following criteria: i) A minimum population of 5,000; ii) At least 75 per cent of the male main working population engaged in non-agricultural pursuits; and iii) A density of population of at least 400 persons per sq. km. (Indian Census, 2011).

compared to only 9% in urban areas. So we are comparing the effects of weather on a population that derives its income primarily from agriculture to one that derives its income primarily from services and manufacturing.

Despite the dramatic extent to which the world has urbanized in the last 60 years, India remains primarily a rural country—in 2000 72% of Indians lived in rural areas, only modestly down from 82% in 1960. In contrast, the US moves from 30% to 21% rural between these two years (World Bank Group, 2016). So we are comparing a developing country which is largely rural with a developed country which is largely urban.

The overriding distinction between India and the US is that the former is much poorer.[14] In contrast with the US both rural and urban Indian populations contain large numbers of people in extreme poverty. At the start of period (1957) the majority of the rural population (62%) and half (49%) of the urban population are poor and by the end of our period (2000) this has dropped to one third (31%) and one fifth (21%).[15] As regards undernutrition in 1999-2000, the proportion of the population living in households with per capita calorie consumption below adequate levels (2,400 per day rural, 2,100 per day urban) was 74% in rural areas and 58% in urban areas (Deaton and Dreze, 2009).[16] In 1998/1999 50% of rural children and 38% of urban children under the age of three are classified as wasted and 48% and 36% as stunted.[17]

So overall we are comparing two poor populations, one of whose incomes are relatively delinked from agriculture, with a rich population where incomes are derived almost entirely from services and manufacturing.

3.1 Weather

3.1.1 Historical Weather

A key finding from Deschenes and Greenstone (2011) is that a careful analysis of the relationship between temperature and mortality requires *daily* temperature data. This is because the relationship between mortality and temperature is highly nonlinear and the nonlinearities would be missed with annual or even monthly temperature averages. This message is echoed in the agronomic and agricultural economics literatures where plant survival falls off precipitously at high temperatures

[14] In 2000 India was home to one third of the world's population living in extreme poverty (Chen and Ravallion, 2008).

[15] These numbers are based on the Indian poverty line which corresponds to a $1 PPP a day international poverty line. This is below the $1.25 PPP a day poverty line used to define extreme poverty across the world in our period (Chen and Ravallion, 2008). The fractions living in poverty in rural and urban India would therefore be higher if we used the latter line.

[16] The importance of food in consumption, which is another marker of poverty, shows a similar pattern—for example, Deaton and Dreze (2009) report that in 2001, 58% and 45% of rural and urban residents' budgets were spent on food.

[17] These figures are from the 1998/1999 round of the National Family Health Survey where wasting is defined as being two standard deviations below typical weight for age and stunting two standard deviations below typical height for age (IIPS, 2000).

(Deschenes and Greenstone, 2007; Schlenker and Roberts, 2008).

Although India has a system of thousands of weather stations with daily readings dating back to the 19th century, the geographic coverage of stations that report publicly available temperature readings is poor (and surprisingly the public availability of data from these stations drops sharply after 1970). Further, there are many missing values in the publicly available series so the application of a selection rule that requires observations from 365 days of the year would yield a database with very few observations.

As a solution, we use data from gridded daily datasets. These data sets use more complete, non-public weather station data and sophisticated climate models to construct daily temperature and precipitation records for 1° (latitude) × 1° (longitude) grid points (excluding ocean sites). We use gridded data produced by the India Meteorological Department (IMD) (Srivastava et al., 2009; Pai et al., 2014) to construct a complete record for daily average temperatures and total precipitation for the period 1950-2000. We match these grid points to each of the districts in our sample by taking weighted averages of the daily mean temperature and total precipitation variables for all grid points within 100 kilometers of each district's geographic center. The weights are the inverse of the squared distance from the district center.[18]

We also use an additional and independent gridded data set to construct instrumental variables for the temperature variables derived from the IMD data. This additional data set, called NCC (NCEP/NCAR Corrected by CRU) provides a complete record for daily average temperatures and total precipitation for the period 1950-2000 across the whole world.[19] The instrumental variables approach is motivated by the fact that the IMD (and NCC) data necessarily assigns weather variables to grid points that are mis-measured representations of the "true" weather variables. In that case, instrumental variables estimates can correct for measurement error bias under the assumption that the errors contained in the IMD and NCC variables are uncorrelated.

In order to capture the distribution of daily temperature variation within a year, we use two different approaches. The first of these assigns each district's daily mean temperature realization to one of eight temperature categories (see Figure 1). These categories are defined to include daily mean temperatures $< 65°$ F, $> 95°$ F, and the six 5° F-wide bins in between. The 365 daily weather realizations within a year are then distributed over these eight bins. This binning of the data preserves the daily variation in temperatures, which is an improvement over much of the previous research on the relationship between weather and health and economic economic outcomes that had obscured much of the variation in temperature. The second way we capture

[18]On average, there are 1.9 grid points within a 100 kilometer radius circle. The results are insensitive to taking weighted averages across grid points across distances longer than 100 kilometers and using alternative weights (e.g., the distance, rather than the squared distance). After the inverse distance weighting procedure, 334 out of a possible 339 districts have a complete weather data record. The five districts that are dropped in this procedure are Alleppey (Kerala); Simla and Mahasu (Punjab); Sikkim (Sikkim); Laccadive, Minicoy, and Amindivi Islands; and the Nicobar and Andaman Islands.

[19]Here NCEP/NCAR stands for National Center for Environmental Prediction/National Center for Atmospheric Research and CRU is the Climate Research Unit at the University of East Anglia.

temperature variation is to bin the data our data into larger $< 75°$ F, $75 - 89°$ F and $> 90°$ F bins which we use when looking at the heterogeneous effects of temperature shocks (see Table 1 for summary statistics).

While the primary focus of our study is the effect of high temperatures on mortality, we use data on rainfall to control for this potential confounding variable (to the extent that temperature and rainfall are correlated). Table 1 reports annual precipitation totals. The striking feature of rainfall in India is its intra-annual distribution: in an average location, over 95% of annual rainfall arrives after the arrival of the southwest monsoon, a stark arrival of rain on the southern tip of the subcontinent around June 1^{st} which then moves slowly northwards such that the northern-most region of India experiences the arrival of the monsoon by the start of July—see, for example, Wang (2006). Naturally this sudden arrival of rainfall after a period of dryness triggers the start of the agricultural season in India. We exploit this feature of the timing in our analysis.

An important component of the analysis in this paper is the contrast of estimated temperature effects between India and the US. We obtain daily average temperature and precipitation data for the US from the National Atmospheric and Oceanic Administration (NOAA) "Global Historical Climatology Network-Daily" (GHCN-Daily) database of weather station readings that are subjected to a common set of quality assurance checks. According to the National Climatic Data Center, GHCN-Daily contains the most complete collection of US daily climate summaries available.

Since the US mortality data is available at the monthly level, we construct monthly measures of weather from the daily records. Specifically, we construct temperature bin variables that count the number of days per month with temperatures $< 65°$ F, $> 95°$ F and the six $5°$ F-wide bins in between. To this end, we select weather stations that have no missing records in any given year over the period 1957-2000. The station-level data is then aggregated to the county level by taking an inverse-distance weighted average of all the measurements from the selected stations that are located within a fixed 200km radius of each county's centroid. Finally, the county-level variables are aggregated to the state-year-month level by taking a population-weighted average over all counties in a state, where the weight is the county-year population. More details about the processing of the India and US historical weather data are presented in the Data Appendix.

3.1.2 Climate Change

In order to determine the predicted impact of future climate change we need predicted data on the future climate. To this end, we use daily data obtained from the Hadley Centre Coupled Model, version 3 (HadCM3), based on the A1FI scenario. A1FI is a "business-as-usual" scenario that assumes the world does not implement significant greenhouse gas mitigation policies. The model produces daily minimum and maximum temperatures and we only consider grid points over land in India and the US. Similar to the weather data, we assign measures to each of the districts in our sample with an inverse-distance weighted average of the daily mean temperature and total precipitation variables for all HadCM3 grid points within 300 km of each district's centroid.[20] This approach produces a complete data record on the annual realization of the eight daily average temperature bins and annual precipitation for each district in India for the period 1990-2099. For the US, we aggregate data to the state level by taking a simple average of all counties within a state. Following Deschenes and Greenstone (2011), we correct the Had3CM predictions for systematic model error by comparing the model's predictions for the 1990-2002 period with the historical data for the same time period.

Figure 2 illustrates the average variation in daily temperature readings across the eight temperature bins over the 1957-2000 period for India (panel A) and the US (panel B). The height of each bar corresponds to the mean number of days that the average person in India or the US experiences in each bin. Figure 2 also displays the average temperature distribution for the 2075-2099 period predicted by the error-corrected HadCM3 data described above. As can be seen from this plot the temperature distribution in India and the US will move sharply to the right with there being many more hot days by the end of this century.

3.2 Mortality

The cornerstone of the analysis in this paper is district-level mortality data taken from the *Vital Statistics of India* (VSI) publications for 1957-2000, which were digitized for this project. The VSI data represent the universe of registered deaths in each year and registration was compulsory in India throughout our sample period. This source contains the most detailed possible panel of district-level mortality for all Indian citizens broken out by rural and urban areas separately. Death tallies in the VSI are only presented for infants (deaths under the age of one) and for all others (deaths over the age of one). From this information we construct our primary measure of rural, urban and total mortality—an all-age mortality rate, defined as the total number of deaths normalized by the population in thousands.

Table 1 summarizes the mortality rates derived from the VSI data for the 1957-2000 period. The data covers 315 districts spanning 15 of India's largest states (which account for over 85% of India's population).[21] The table reveals a remarkable reduction in mortality over the sample period. In both rural and urban areas, the all age mortality rate declined from 10-11 per 1000 to 4-6 per 1000 between the early part of the sample (1957-1971) and the most recent (1987-2000). In Section 4 below, we describe our strategy to avoid confounding these trends in mortality rates with any time trends in temperatures. It is important to stress that these mortality rates are

[20] The Hadley grid is 2.5° × 3.75°, so not as "dense" as the 1° × 1° historical data for India and so we need a larger radius 300 km to make sure we have complete data assignment.

[21] These states are (in 1961 borders and names): Andhra Pradesh, Bihar, Gujarat, Himachal Pradesh, Jammu and Kashmir, Kerala, Madhya Pradesh, Madras, Maharashtra, Mysore, Orissa, Punjab, Rajasthan, Uttar Pradesh, and West Bengal. These are the states with a consistent time series of observations in the VSI data. The results in this paper are largely insensitive to the inclusion of all observations in the VSI data.

almost surely underestimates of the extent of mortality in India. Despite compulsory registration of births and deaths, many areas of the country suffer from significant under-reporting.[22]

In order to compare our estimates for India with corresponding estimates for the US, we use all-age mortality rates per 1,000 population, constructed by combining mortality counts digitized from Mortality Statistics of the US annual volumes and machine-readable Multiple Cause of Death Files (see Barreca et al., 2016). We obtain all-age mortality at the state-year-month level for the period 1957-2000, allowing us to estimate annualized temperature-mortality curves for the US that can be directly compared to those we estimate for India.

In order to quantify the impact of climate change on health, we first calculate the predicted effect of climate change on mortality rates. This is obtained by applying the predicted future temperature and rainfall distribution associated with the error-corrected Hadley 3 A1FI model (as reported in Figure 2) to the regression estimates of the temperature-mortality relationship for India and the US based on historical 1957-2000 data (as reported in Figure 1). By combining the counterfactual mortality rates associated with climate change with current life tables (base year 2000), we can then calculate the predicted effect impact of climate change on life expectancy for rural India, urban India and the US.

3.3 Income

It is natural to expect that the weather plays an important role in the agricultural economy in India. In turn, the agricultural economy may play an important role in the health of rural citizens who draw their incomes from agriculture. To shed light on these relationships we draw on the best available district-level agricultural data in India. The data on agricultural outputs, prices, wages and employment come first from the India Agriculture and Climate Data Set, which was prepared by the World Bank (Sanghi et al., 1998). This file contains detailed district-level data from the Indian Ministry of Agriculture and other official sources from 1956 to 1987. We then obtain similar data from the Center for Monitoring of the Indian Economy (CMIE) India Harvest database for the period 1987-2000. From these combined sources we construct two distinct district-level variables on the agricultural economy: yields and real wages.[23] Our final sample contains data from 294 districts for the years 1956-2000.

We construct a measure of annual, district-level yields by aggregating over the output of each

[22]According to the National Commission on Population of India, only 55% of births and 46% of deaths were being registered in 2000. These estimates were obtained from India's Sample Registration System (SRS), which administers an annual survey of vital events to a nationally representative sample of households. The data published by the SRS, however, are only available at the state level.

[23]One difference relative to the mortality data concerns our adjustment of the timing of the weather data when relating it to agricultural outcomes. The agricultural yield data are based on measures of total output produced during the agricultural year (from June 1st to May 31st). So weather variables in year t used in our agricultural income specifications are calculated from the 365 days beginning on June 1st of calendar year t. In practice this distinction matters little because the bulk of output is produced between June 1st and December 31st, and there are typically relatively few hot days early in the calendar year.

of the 15 main crops available in the combined World Bank and CMIE dataset. To do this we first create a measure of real agricultural output for each year using a consumer price index[24] and then divide this by the total cultivated area in the district-year. Table 1 reports on the resulting yield measure. All of the major agricultural states are included in the database, with the exceptions of Kerala and Assam.

A second important metric of rural incomes is the real daily wage rate earned by agricultural laborers.[25] The Agricultural Wage in India datasets contain information on daily wages, as collected by government surveys of randomly chosen villages in each district and year. All figures are given in nominal wages per day, and are converted into equivalent daily rates to reflect the (low) degree of variation in the number of hours worked per day across the sample villages. We divide the reported, nominal wage rate by the agricultural price index described in the Data Appendix to construct an estimate of the real rural agricultural wage in each district-year.[26] As can be seen in Table 1, the level of real wages is low throughout the period—never rising above 13.24 rupees, or approximately 2 US dollars (base year 2000) per day in PPP terms.

We also investigate the extent to which economic conditions in urban areas react to temperature fluctuations. To this end we have collected the best available data on urban economic conditions. It is important to stress at the outset that, perhaps because of the over-riding current and historical importance of agriculture for economic welfare in India, the statistics on India's urban economy are not as detailed as those on India's rural, agricultural economy. All of the sources listed below report data on urban outcomes at the *state* level, whereas all of the rural equivalents introduced above were available at the *district* level.

India's manufacturing sector (especially its registered or formal manufacturing sector) is almost entirely located in urban areas. For this reason we use a measure of state-level registered manufacturing productivity (real output per worker) as one measure of the productivity of the urban area of each state in each year. We draw this data from Besley and Burgess (2004), who collected the data from publications produced by India's Annual Survey of Industries.

The second measure of incomes in urban areas that we exploit comes from manufacturing wage data. To construct this variable we first use data on nominal (registered) manufacturing wages, as surveyed by the Annual Survey of Industries and published in the annual *Indian Labour Yearbook*, which was collected by Besley and Burgess (2004). We then divide nominal manufacturing wages by an urban consumer price index to create a measure of real manufacturing wages.

[24]We construct an agricultural price index for each district and year which provides a simple proxy for the real cost of purchasing food in each district-year relative to a base year. Our simple price index weights each crop's price (across the 15 crops in the World Bank / CMIE sample) by the average value of district output of that crop over the period. Annual, district-level consumption data, which would be required to construct a more appropriate *consumption*-based price index, are not available in India.

[25]This is often treated as a welfare measure in its own right as it constitutes the main source of income of Indian society who own little or no land and have limited non-farm employment opportunities (Dreze and Sen, 1989).

[26]A better real wage measure would of course also incorporate price information on non-agricultural items in the rural consumption basket. Unfortunately, the price and quantity information that would be required to do this are unavailable annually at the district level in India.

3.4 Banks

We make use of the data from Burgess and Pande (2005) on rural bank branch expansions in India to test whether better access to credit markets reduces the impact of extreme temperature on mortality. The data set includes 330 districts over the period 1961-2000 and reports the number of commercial bank branches opened in rural, unbanked areas in a given year and district. For ease of interpretation we normalize this variable by population. Table 1 shows the remarkable increase in credit market access (as defined by the presence of a bank branch) in rural India over the period 1957-2000. Over this period, bank branches per 1000 population increased from 0.02 (1957-1971) to 0.23 (1972-1986) and 0.46 (1987-2000). This reflects the rapid expansion of bank branches into rural areas of India between 1977 and 1990 which lowered the rates of interest that rural Indians could borrow at and gave them access to interest bearing savings accounts.

4 Empirical Method

This section describes the econometric methods used to establish relationships between a series of outcomes and temperature realizations. Several of the relationships are expected to be non-linear, so we have chosen an approach that allows for flexibility in the response function while still retaining some parsimony. Our *first approach* to estimating these relationships was introduced briefly in the Introduction, and results based on it were presented in Figure 1. We now discuss the details.

Estimates for this approach are obtained from the following flexible model of the effects of daily temperature on outcomes:

$$\begin{aligned} Y_{it} &= \sum_{j=1}^{8} \theta_j \mathrm{TMEAN}_{itj} + \sum_{k} \delta_k \mathbf{1}\left\{\mathrm{RAIN}_{it} \text{ in tercile } k\right\} \\ &+ \alpha_i + \gamma_t + \lambda_r^1 t + \lambda_r^2 t^2 + \varepsilon_{it}, \end{aligned} \tag{5}$$

where Y_{it} is the outcome variable (e.g., the log of the mortality rate) in district (India) or state (US) i in year t.[27] The key explanatory variables of interest are those that capture the distribution of daily temperatures in district i within year t. The variable $TMEAN_{itj}$ denotes the number of days in district i and year t on which the daily mean temperature fell in the jth of the eight temperature bins. We estimate separate coefficients θ_j for each of these temperature bin regressors. Since the number of days in a standardized year always sums to 365, we use the bin for temperatures between 70° F and 74° F as a reference category whose coefficient is consequently normalized to zero. Importantly, across our outcome variables for India we cannot reject the null hypothesis that the number of days with a mean temperature <65° F has zero effect on the outcome. Further, there

[27] Given that the bulk of analysis is on India for brevity we refer to i as district except where it refers specifically to the US.

are not enough high temperature days to obtain meaningful estimates for separate temperature bins $\geq 95°$ F. We therefore bin into eight temperature categories $< 65°$ F, $> 95°$ F, and the six 5° F-wide bins in between.

This approach makes three assumptions about the effect of daily temperature on the outcome variable. First, this approach assumes that the impact is governed by the daily mean alone. Since daily data on the intra-day variation of temperatures in India over this time period is unavailable, this assumption is unavoidable. Second, the approach assumes that the impact of a day's mean temperature on the annual mortality rate is constant within 5° F degree intervals. Our decision to estimate separate coefficients θ_j on each of seven temperature bin coefficients represents an effort to allow the data, rather than parametric assumptions, to determine the various temperature response functions, while obtaining estimates that are precise enough to have empirical content. This degree of flexibility and freedom from parametric assumptions is only feasible because of the use of district (India) and state (US) data spanning 44 years. Third, by using as a regressor the total count of days in each bin in year t, we are assuming that the sequence of relatively hot and cold days is irrelevant for how hot days affect the annual outcome variable. This is a testable assumption, that we do not reject at standard levels.[28]

The second set of variables on the right-hand side of equation (5) aims to capture variation in precipitation. Our primary focus is on the effects of temperature on death, so the coefficients on rainfall regressors are of secondary importance. However, because it is possible that temperature variation is correlated with rainfall variation, the inclusion of these rainfall variables is important. We model rainfall in a manner that is fundamentally different from our approach to modeling temperature because of one key difference between temperature and rainfall: rainfall is far more readily stored (in the soil, in tanks and irrigation systems, and in stagnant water that might breed disease) than is temperature. Given this distinction, we model the effect of rainfall as the impact of sums over daily accumulations. Specifically, we calculate whether the total amount of rainfall RAIN_{it} in year t in district i was in the upper, middle or lower tercile of annual rainfall amounts in district i over all years in our sample. These are the regressors $\mathbf{1}\{\text{RAIN}_{it} \text{ in tercile } k\}$. We estimate a separate coefficient on each of the three tercile regressors, though we treat the middle tercile regressor as the omitted reference category.

The specification in equation (5) also includes a full set of district fixed effects, α_i, which absorb all unobserved region-specific time invariant determinants of the outcomes. So, for example, permanent differences in the supply of medical facilities will not confound the weather variables in equations for the log mortality rate. The equation also includes unrestricted year effects, γ_t. These fixed effects control for time-varying differences in the dependent variable that are common across regions (e.g., changes in health expenditures related to the 1991 economic reforms).

[28]For example, if we augment the regression in equation (6) below with an additional regressor that counts the number of spells with three or more days in a row on which the mean temperature exceeds 90° F, this regressor is insignificant and our two main temperature coefficients are virtually unchanged.

Since shocks or time-varying factors that affect health may not be common across regions, we emphasize specifications that include separate quadratic time trends for each region. For India we also control for climatic region subscripted by r in equation (5) where India's Meteorological Department divides the country into five such regions that are deemed to have similar climates. Because a key component of our analysis contrasts urban and rural regions of Indian districts, whenever we present such results they come from separate regressions run on urban and rural subsamples. This effectively allows for separate district fixed-effects and climatic region trend terms for urban and rural regions. Specification checks below confirm the robustness of our estimates to alternative approaches to these time-varying controls.

For the US we use the same definition of the daily temperature bins and rainfall terciles as for India. However, because US mortality rate data are recorded at the state-year-month level, we fit the regression including state fixed effects, year-by-month fixed effects, and state-by-month quadratic trends. Monthly regression estimates for the US are then converted to annual levels by multiplying the estimated coefficients by 12.

We use the approach described above for graphical analyses, like Figure 1, throughout the paper. We also employ a more parsimonious approach when attempting to tease out heterogeneous effects of temperature shocks, such as how these effects interact with bank penetration. Consequently, our *second approach* to modeling the temperature response functions estimates fewer parameters while still doing some justice to the non-linear nature of the relationships. This second approach focusses on two critical temperature ranges in the estimated response functions: the number of days with average temperature between 75° F and 89° F and the number of days with average temperature exceeding 90° F. The estimated regression model corresponds to:

$$Y_{it} = \beta_1 TMEAN_{it}^{75-89} + \beta_2 TMEAN_{it}^{>90} + \sum_k \delta_k \mathbf{1}\left\{\text{RAIN}_{it}\text{in tercile } k\right\} + \alpha_i + \gamma_t + \lambda_r^1 t + \lambda_r^2 t^2 + \varepsilon_{dt}, \quad (6)$$

where the variables $TMEAN_{it}^{75-89}$ and $TMEAN_{it}^{>90}$ represent the number of days in district i and year t where the daily average temperature ranges between between 75° F and 89° F, and in excess of 90° F, respectively.[29] This approach to modeling temperature has the advantage of estimating only two temperature coefficients rather than seven, while still capturing the non-linearity evident from the seven coefficient estimates in Figure 1.

The validity of this paper's empirical exercise rests crucially on the assumption that the estimation of equations (5) and (6) will produce unbiased estimates of the θ_j, β and δ_k parameters. These parameters are identified from district-specific deviations in weather about the district averages that remain after adjustment for the year fixed effects and region-specific quadratic time trends. Due to the randomness and unpredictability of weather fluctuations, it seems reasonable

[29] This is a particular restriction on the flexible approach in equation (5)—where all of the temperature bin coefficients θ_j below 75° F are restricted to be zero, the three coefficients between 75° F and 89° F are restricted to be equal, and the two coefficients above 90° F are restricted to be equal.

to presume that this variation is orthogonal to unobserved determinants of the outcomes.

There are two further points about estimating equations (5) and (6) that bear noting. First, it is likely that the error terms are correlated within districts over time. Consequently, the paper reports standard errors that allow for heteroskedasticity of an unspecified form and that are clustered at the district level. We also consider an alternative approach to clustering in a robustness check below. Second, we fit weighted versions of equations (5) and (6), where the weight is the square root of the population in the district, for two complementary reasons. First, the estimates of mortality rates from large population districts are more precise, so this weighting corrects for heteroskedasticity associated with these differences in precision. Second, the results reveal the impact on the average person rather than on the average district, which we believe to be more meaningful.[30]

5 Weather and Death

This section presents our empirical results on the relationship between weather and mortality in India and the US. We begin with a simple baseline comparison of the two countries and dig into the Indian results by exploring heterogeneity across rural/urban regions, the timing of the agricultural cycle, and whether access to bank branches in rural India can mitigate the effect of hot weather on mortality.

5.1 Baseline Analysis

Figure 1(a)—as discussed in the Introduction—sets the scene for our empirical analysis. The Figure plots, for the Indian and US populations from 1957-2000, the impact of having an extra day whose *daily mean* temperature lies in each of seven temperature bins relative to a day in the reference category bin of 70°-74° F. These response functions are based on rich statistical models that control for location, time and climate region-year fixed effects as well as for rainfall—as per equation (5). All of the coefficient estimates underlying Figure 1 can be found in Appendix Table A2.

Figure 1(a) shows that the same inter-annual variation in temperature exerts a much stronger impact on mortality in India than in the US (though the estimated effects are statistically significant in both cases). Notably, the non-parametric relationship between daily temperatures and mortality that we estimate for India shows particularly large effects at higher temperature ranges, increasing steeply when there are more days at or above 90° F. These effects are large—for example, a *single* additional day with a mean temperature above 95° F, relative to a day with a mean temperature in the 70-74° F range, increases the annual mortality rate by roughly 0.7%. The corresponding figure for the US is 0.03%, an order of magnitude smaller. Clearly, the effects

[30]For analogous reasons, when estimating relationships in which the outcome variable concerns agricultural income we weight by cultivated area.

of hot weather are highly unequal across these two populations.

Figure 1(b) displays an analogous cut of the data, but within India—comparing the effects of the same temperature shocks on mortality among the rural and urban populations within each Indian district. Figure 1(b) demonstrates that the all-India results from Figure 1(a) appear to be entirely driven by India's rural areas. The estimated impacts in urban India are small and never statistically significant at the 5% level. Conversely, all of the estimated rural coefficients above the 70-74° F reference category range are statistically significant (and for all those above 79° F we can reject the equality of urban and rural impacts). This stark inequality is the central result of our paper.

Finally, Figure 1(c) plots the US and urban Indian coefficients alone. Up to the (non-negligible) Indian sampling variation, the two sets of results are surprisingly similar. In Appendix Table A2 we see that though the US and urban India coefficients on temperature bins above 70-74° F are close to each other but typically ten times smaller than the rural India coefficients. Living in urban areas, even when large sections of the population remain poor, seems to confer protection against the mortality impacts of hot weather. Our paper thus points to one key benefit of living in towns and cities.

Appendix Table A2 also reports the coefficients on the two rainfall variables in equation (5) for all of the sample regressions above. While the coefficient estimates are in line with what one might expect from the literature on famines in India (see, e.g., Dreze and Sen, 1989), with positive estimates on mortality from relatively little rainfall in rural (but not urban) areas, we do not have the power to estimate these effects precisely. Throughout the remainder of our analysis we continue to control for these rainfall variables in all regressions but report these coefficients only in certain cases.

5.2 Growing versus Non-Growing Seasons

As discussed above, one important distinction between India's rural and urban economies is the extent to which households derive income from agricultural sources. This difference could offer one explanation for the rural-urban heterogeneity in the health impacts of hot temperatures seen in Figure 1(b). To probe this possibility further we now exploit the timing of India's agricultural cycle.

The starkest change in the Indian agricultural calendar comes with the arrival of the southwest monsoon rains, after which the vast majority of an average district's annual rainfall arrives. This monsoon onset is the beginning of the *kharif* agricultural season, the first and most important growing season of each year. On approximately June 1^{st} of every year the southwest monsoon begins to arrive on the subcontinent at its southern tip (roughly the state of Kerala). The monsoon moves slowly northwards, typically reaching India's northern limits by the start of July. Because of this slow progression, the arrival of the monsoon, and therefore the start of the main agricultural season, varies throughout the country.

We utilize spatial variation in this dramatic change in growing conditions in our analysis as follows. We have obtained data on each district's "typical" (i.e. long-run average) date of monsoon arrival from the Indian Meteorological Department. Within a calendar year, we define all dates after a given district's typical date of monsoon arrival as the growing season. To define the non-growing season we take all dates that are within the three-month window prior to each district's typical date of monsoon arrival.[31] This is clearly a simple way of segmenting the year based on agricultural activity but, as we discuss in Section 6, it is successful at predicting the temperature shocks that actually affect agricultural output in rural India.

Figure 3(a) repeats the rural mortality analysis in Figure 1(b), but with the temperature data partitioned into two segments: that in the growing season and that in the non-growing season. Across our 1957-2000 period the average temperature during the GS is 78° F and during the NGS is 86° F. This is because the NGS is just before the onset of the monsoon which is the hottest and driest period in India. Despite this in Figure 3(a) we see that an equivalently hot day has a substantially greater impact on rural mortality when that day occurs inside the growing season (GS) rather than in the non-growing season (NGS). All of the GS warm weather (i.e. above 74° F) coefficients are statistically significantly different from zero, yet none of NGS coefficients are. And all of the warm weather GS coefficient estimates are greater than their corresponding NGS coefficients.[32]

Rural residents are therefore sheltered from the mortality risk of hot days during NGS when agricultural activity is low depite these bieng the three hottest months of year. In urban areas, neither GS nor NGS hot days elevate mortality (see Figure 3(b)). Consonant with this, we find in Section 6 that warm GS weather affects agricultural (but not manufacturing) productivity and wages. This evidence is consistent with the unequal effects of temperature on death across rural and urban India and across India and the US being driven, in part, by hot weather lowering incomes of those who work in weather-dependent agricultural occupations.

Development, in the form of a move to greater dependence on non-agricultural employment, therefore seems to carry the benefit of protecting citizens from deleterious effects of hot weather. In terms of coefficient magnitudes, Figure 3(a) tells us that a year whose growing season features just one more day in the hottest ($\geq$ 95° F) bin rather than in the reference (70-74° F) bin will have approximately 1% more rural death within the year. This compares to 0.7% across the whole year (Figure 1(b)).

[31]The use of three months rather than the entire year (up to the monsoon onset date) matters little because there are so few hot days in the first months of the calendar year. We pursue this approach because it is more appropriate for use with the agricultural data (which provide measures of output by agricultural year rather than by calendar year) employed in Section 6. In many regions the entire growing season, typically the combination of the *kharif* and then the *rabi* seasons, can be as long as nine months, so the first few months of a calendar year are typically the final months of the previous year's agricultural season.

[32]Although this difference is never statistically significant, we do reject equality of the GS and NGS coefficients overall (the F-statistic on this test is 4.49, $p \leq 0.001$).

As discussed in Section 4, in order to streamline the presentation of the remaining mortality results in this section we now shift to a more parsimonious presentation of the relationship between temperature and mortality. That is, we report the estimates of the two coefficients (β_1 and β_2) from equation (6) rather than the seven coefficients (θ_1 through θ_7) from equation (5).

These results are presented in Table 2. For the sake of comparison we begin in panel A with results from the analog of Figure 1(b); as expected, in both the moderate (75-89° F) and high (> 90° F) temperature categories, hot days appear to kill rural residents but not urban ones. Panel B then reports the corresponding results disaggregated by growing and non-growing seasons. As in Figure 3, hot temperatures on growing season days increase rural mortality, but the same is not true in the non-growing season.[33] Indeed, the reported test of equality across agricultural seasons demonstrates that we reject such equality (within each of the two temperature categories) at the 5% level.

Columns 2(a) and 2(b) of Table 2 (panel B) then report the equivalent GS/NGS disaggregation for urban areas. As one might expect, given the lack of an overall effect in urban areas, there is no apparent distinction between the impact of hot temperature GS and NGS shocks in urban areas. It is again striking that it is the GS weather rather than the hottest NGS weather that elevates mortality in rural areas. Similarly that the hottest NGS days do not significantly elevate mortality in urban areas.

Our results so far make it clear that the fatal weather-death relationship in India follows a pattern that is highly correlated with agricultural activity—it is present only in rural areas and only in the agricultural growing season. There are multiple possible explanations for this finding. A first class of possibilities has nothing to do with agriculture—it could simply be the case that the timing of the agricultural season is correlated with other mechanisms that relate temperature to death. The only such mechanism of which we are aware is the possibility of diseases that favor hot and wet conditions and are more deadly in rural areas. While we cannot definitively rule out this interpretation of our results, it faces three difficulties. First, infectious diseases associated with hot and wet weather might be expected to manifest both in urban and rural areas (e.g. diarrhoea, dysentery, malaria). Second, as we show in Appendix Table A3, there is no evidence of an (economically or statistically) important interaction effect between rainfall and temperature on mortality in either rural or urban India. And further, our main temperature coefficients are extremely stable despite the introduction of controls for temperature-rainfall interaction terms. Third, for the case of malaria—one particular disease that seems a leading candidate here and for which we have data—the results in Appendix Figure A1 demonstrate that malarial incidence does not appear to respond to temperature shocks in the same way that overall mortality does.[34]

[33] If anything, the impact of moderately hot days, relative to the reference category in the non-growing season (reported in column 1(a)) appears to be negative. This can be reconciled with Figure 3(a) by noting that the reference temperature category in Table 2 is ≤ 74° F, whereas that in Figure 3(a) is 70-74° F.

[34] Hot days appear to have no statistically signifiant effect on reported malarial deaths (panel A), on the share of the population testing positive once per year for any malarial parasite (except for the 90-94° F bin where the effect

A second class of possible explanations for our findings does relate to agriculture, but we stress that there are several ways in which agriculture could matter. First, the growing season corresponds to the period in which agricultural workers would need to be outside working the most (and hence exposed to hot ambient temperatures). Second, most agricultural crops will also suffer during hot temperatures, and this will reduce agricultural incomes and hence the possibilities for households to use income to respond to health shocks. Third, workers may respond to dangerous ambient temperatures by working outside less intensively, and this may also reduce agricultural output and hence incomes. Consistent with income playing a role, our results in Section 6 confirm that hot weather in the growing season is associated with falls in agricultural output and wages. The fact that the growing season corresponds to the "hungry season" in India (since it occurs in the run-up to the year's major harvest) only makes these income-related effects all the more plausible. If households can expend income on improving health then, as discussed in Section 2, the household's linear budget constraint means that they will view these three effects as perfect substitutes on the margin. So while we cannot separately identify these three possible agriculture-related mechanisms, any policies that enable households to maintain expenditures on survival-enhancing health goods should mitigate the impact of hot temperatures on death. We examine one such policy related to rural bank branch expansion in Section 5.4 and discuss a wider range of policies in the conclusion.

5.3 Robustness and Heterogeneity

We now probe further the relationship between hot temperatures and mortality in India using the more parsimonious empirical specification of equation (6).

Table 3 begins with an investigation of the effects of lagged temperatures on Indian mortality. This is important because of the possibility that hot days kill people not only in the current year (which would be within a time frame of 4-9 months, given that most hot days in India take place in March through September and our mortality data track calendar years) but also in subsequent years. Such a lagged impact would be natural if there is any sense in which households draw down assets (such as financial assets, physical capital or human or health capital) in the face of a shock and then are more exposed to even relatively minor shocks in subsequent years.

Row 2 of panel A repeats the simple specification of equation (6), which has no lagged terms, on the subsample of data for which a lag analysis is feasible.[35] We see similar results to those in our baseline sample (row 1), with the exception of the moderate temperature coefficient in rural areas which is found to be lower than (but still within the 95% confidence interval of) our baseline estimate.

is negative—panel B), or on the share of blood slides examined by the National Malaria Eradication Program that test positive for *Plasmodium falciparum* (the species of malaria that is known to be most fatal in India—panel C).

[35] Because the vital statistics data that we use to measure mortality have missing values for some observations (within each district's rural or urban time series), our analysis here of three lagged effects can only be estimated with non-missing observations whose preceding three observations are also non-missing.

Panel B of Table 3 extends this analysis by estimating equation (6), but now with additional controls for the two temperature regressors in the previous three years.[36] Rather than reporting all four of these coefficients, we report the contemporaneous coefficient (to confirm that it does not change appreciably when lags are controlled for) and the sum of the three lagged coefficients. We also report the total impact of temperature shocks, summing across both the contemporaneous and lagged impacts. The key message of this table is that, in rural India, hot days kill people both immediately and with a lag of up to three years. By contrast, these effects never appear to arise (either immediately or with a lag) in urban segments of each Indian district.

Table 4 (panel A) reports on three additional specification checks concerning our rural Indian mortality results. First, because the vital statistics data underpinning our dependent variable are collected by Indian states, we compute here standard errors that are clustered by state rather than by district.[37] While the standard errors rise, the estimated effects of temperature here are still statistically signifiant at standard levels. Second, we drop the controls for climatological region-specific quadratic polynomial trends (from equation 6). And third, we instead replace these trend terms with a more flexible set of fixed effects for the interaction between climatological regions and years. The results from these last two checks demonstrate that our baseline results are, if anything, conservative with respect to how we control for unobserved, time-varying climatological features.

In panel B of Table 4 we probe whether the effect of hot temperatures on rural deaths has been falling throughout our sample period (1956-2000) as rural areas have grown richer (for instance, because of developments like the Green Revolution in Indian farming after the 1960s). There is surprisingly weak evidence for such a change—the effect of moderate temperatures (75-89°) has barely budged over the nearly half century we study and the effect of hot temperatures ($> 90°$ F) has fallen but only in the most recent third of our sample.

In Appendix Figure A2 we contrast the effects between hot and cold districts in rural and urban areas. The coefficients for hot districts in rural areas (those whose long-run average number of days $> 90°$ F are larger than the national average) tend to lie below those for cold districts though both sets of coefficients are statistically significant. This is to be expected if adaptations to hot temperatures (such as crop choice, growing techniques and building structures) are more feasible in the long-run. For urban populations coefficients for hot and cold district lies on top of one another.

Our final set of mortality specification checks (reported in Table 5) is motivated by the possibility of measurement error in our independent variable, the temperature indicators. As discussed in Section 3, we are helped here by the existence of two very different attempts to measure daily temperatures at high (district-scale) spatial resolution throughout our sample period: IMD data

[36] We use three lagged terms and no lead terms in this analysis because among the set of all specifications with up to three lags and up to three leads that reported here had the highest value of the Akaike information criterion.

[37] This also offers a simple check on whether spatial correlation of the error term is likely to mean that our baseline standard errors are substantially understated.

constructed for India alone; and NCC data, which uses a different methodology because of its global coverage. Since the potential sources of measurement error in these two datasets are likely to have different origins, and hence be uncorrelated with one another, we report in Table 5 instrumental variable (IV) regressions where any variable constructed from the IMD data (used in all regressions reported thus far) is instrumented for by its analog in the NCC data. Consistent with the measurement error interpretation, the IV estimates in Table 5 are indeed higher than their OLS equivalents in Table 2. In the interests of proceeding conservatively we continue to stress the OLS results throughout the rest of our analysis but Table 5 suggests that the true impact of temperature on death in India is even larger than Figure 1 implies.

5.4 Bank Branch Expansion

The model in Section 2 demonstrated that if rural citizens had access to improved means of smoothing their consumption (and hence their survival) then this should mitigate the impact of hot weather on mortality. One such method of consumption smoothing, which has been the subject of an extensive theoretical and empirical literature (Besley, 1995; Townsend, 1994), is access to borrowing and saving opportunities.

In the Indian context, an important vehicle through which many households borrow and save is via subsidized loans in the formal banking sector. But access to such banks has long proven difficult—for example, as discussed in Burgess and Pande (2005), the 1961 census documents that commercial banks had an extremely limited presence in rural areas. A 1977 policy, later repealed in 1990, was designed to improve this state of affairs. It stipulated that for each bank branch opened in a (typically urban) banked location, a commercial bank would have to open four branches in (typically rural) unbanked locations.

This "1:4" policy effectively forced banks to open branches in rural, unbanked locations. As a result, bank branch expansion occurred differentially in financially underdeveloped regions after 1976, reversing pre-existing (1961-1976) trends that saw branch openings concentrated in financially developed regions. The opposite then occurred after the policy was revoked in 1990, as banks were free to return to their laissez-faire expansion strategies.

In order to evaluate how this bank expansion program mediated the weather-death relationship in rural India, we follow Burgess and Pande (2005) and construct two instruments for rural bank access: (i) a trend-break, between the pre-program period 1961-1976 and the program period 1977-1989, in the relationship between a district's 1961 financial development and its subsequent rural branch expansion; and (ii) the equivalent trend break after the program, 1990 onwards.[38]

[38]This is a district-level extension of the state-level 2SLS analysis in Burgess and Pande (2005)—while our focus is on how banks change the temperature-mortality relationship, which is estimable at the district-level, Burgess and Pande (2005) estimated the impact of rural branch expansion on poverty, for which district-level data is sparse. Specifically, we use 2SLS to estimate the following augmented version of equation (6):

$$Y_{it} = \beta_1 TMEAN_{it}^{75-89} + \beta_2 TMEAN_{it}^{>90} + \rho_1 B_{it}^R \times TMEAN_{it}^{75-89} + \rho_2 B_{it}^R \times TMEAN_{it}^{>90} + \gamma' X_{it} + \varepsilon_{it}, \quad (7)$$

Table 6 presents the results of this 2SLS procedure. Column 1 repeats our baseline OLS temperature-death regression (as in Table 2) but for the subset of observations that enters the 2SLS analysis, confirming that the results are largely unchanged by this restriction.[39] Column 2 then reports 2SLS results that include interactions between our two temperature variables and the density of rural bank branches (as well as the un-interacted banks variable), but where these three bank variables are instrumented with the policy trend-break expressions as described above.[40] We see that the effects of both moderate temperatures (75-89° F) and hot temperatures (> 90° F) are mediated by bank access in a statistically significant manner. The median district began in 1961 with no rural bank branches, and ended our sample period (in 2000) with 0.44 bank branches in previously unbanked rural areas (per 10,000 persons). The results in column 2 of Table 6 therefore imply that such a district saw especially large rural impacts of hot days in 1961 when no banks were present (e.g. 1.2% higher death rate for every day above 90° F) and yet a much weaker impact of equally hot days by 2000 when rural banks were more prevalent (e.g. 0.3% higher death rate for every day above 90° F, since $0.012 + (0.44)\times(-0.021) = 0.3$).

Finally, in column 3 we report the results from an analogous exercise to that in column 2, but where the (log) urban mortality rate is now used as the dependent variable. The Burgess and Pande (2005) identification strategy is only applicable for predicting rural bank expansion so the spirit of this exercise is a "placebo" check: we are asking whether the rural bank expansion driven by our instrumental variables appears to have had a spurious impact (or potentially one working through urban-rural spillovers within districts) on urban mortality or its responsiveness to hot days. Reassuringly, there is no evidence for such an impact.

These results show the power of bank access in enabling rural residents to smooth the health and

where: B^R_{it} represents the cumulative number of rural bank branch openings per 10,000 rural residents in district i and year t; and X_{dt} represents a vector of control variables comprising rainfall controls, district fixed-effects, year fixed-effects, the level effect of B^R_{it}, plus a trend term interacted with the number of rural bank branches per capita in 1961 (which we denote $B^R_{i,1961}$, with 1961 chosen because it is the first year in our sample period for which the number of banks per capita can be calculated). We omit here the quadratic polynomial trend terms for each climatic region of equation (6) because the Burgess and Pande (2005) instruments, based on trend breaks, no longer pass weak instrument diagnostics when these terms are included; but the point estimates (and standard errors) reported in Table 6 change very little upon controlling for these region-specific trends. Our instruments are based on the following trend-break model first-stage:

$$B^R_{it} = \alpha_1 B^R_{i,1961} \times t \times \mathbb{I}(t > 1977) + \alpha_2 B^R_{i,1961} \times t \times \mathbb{I}(t > 1990) + \gamma' X_{it} + \varepsilon_{it}, \tag{8}$$

where $\mathbb{I}(x)$ is an indicator variable that equals one if x is true and equals zero otherwise. In addition, we control for the level effects of $B^R_{i,1961} \times \mathbb{I}(t > 1977)$ and $B^R_{i,1961} \times \mathbb{I}(t > 1990)$ throughout. In practice, we instrument for the the level effect of B^R_{it} with the instruments in equation (8) and for the bank-temperature interaction terms with the instruments from equation (8) interacted with the temperature variables and estimate these regressions on data from 1961 onwards (because the instruments are parameterized on the basis of 1961 penetration).

[39]The two samples differ both because our banks analysis starts in 1961 (rather than 1957) and contains the smaller subset of districts (270 rather than XXX) for which bank branch data are available.

[40]Column 2 reports the F-statistics corresponding to each of the three first-stage equations, following the procedure devised by Sanderson and Windmeijer (2016) to check for weak instruments in settings (like ours) with more than one endogenous variable. Importantly, the two first-stages corresponding to the bank-weather interaction terms have an F-statistic above the usual rule-of-thumb level of 10.

income shocks induced by hot weather shocks. Given that these banks offered subsidized loans at lower interest rates than were formerly available from informal moneylenders (as discussed in Burgess and Pande, 2005), a natural interpretation of these findings is that this enhanced smoothing was due to enhanced household borrowing and saving itself. Consistent with this interpretation, in Appendix Figure A3 we plot estimates of the effect of hot days, using the bins method of equation (5), on some measures of banking behavior from 1972-2000. Panel A does this for the (log of the) stock of bank deposits held per capita, and it is clear that hot weather does cause rural deposits to fall, and statistically significantly so. By contrast, the effect in urban areas is weaker and not statistically significant.[41] So rural residents seem to make greater use of dissaving to smooth out weather shocks than do urban residents. This finding casts doubt on the hypothesis that there is no impact of weather on death in urban areas because households there are richer or otherwise better able to adapt to shocks than their rural counterparts, since it seems natural to expect that one form of such adaptation would occur via urban bank deposits.

While the findings of Figure A3 provide one interpretation for how banks intervened in rural India's weather-death connection, it is also possible that improved banking had additional effects on rural households. This could include an improved ability to borrow as entrepreneurs, or general equilibrium effects working through wages (as banks may have facilitated business expansions through improved credit access).[42] While we are unable to disentangle these possibilities, the simple fact that banks matter—and that a feasible government policy like rural bank expansion can eliminate the fatal effects of weather on death in rural India—is undeniable. This offers evidence for a potential adaptation strategy in the face of hotter temperatures that we return to when discussing climate change in Section 7.

6 Weather and Income

This paper's main results, described in the previous section, concern the impact of weather on death. In this section we complement these mortality results with additional results on income. We view the results in this section as interesting in their own right, given their welfare implications for the survivors of weather shocks. But they also shed light on the hypothesis that weather-induced income shocks can explain why we see so much death from hot temperatures in rural India. A necessary condition for the hypothesis that temperature-driven income shocks can explain why we see so much death from hot days during the growing season (but not during the non-growing season) in rural (but not urban) India is evidence for falling incomes due to hot growing season (but

[41]Panel B of Figure A3 goes on to report a similar set of results for the (log of the) stock of outstanding credit in rural and urban areas. As with the case of deposits, we see a reduction in bank credit. Given that commercial bank loans were typically intended to finance costs associated with farming and other rural businesses (Burgess and Pande, 2005) these results suggests that demand for business loans falls during periods of hot weather.

[42]Indeed, Burgess and Pande (2005) find support (at the 10% level of statistical significance) for an effect of bank expansion on agricultural wages.

not non-growing season) temperatures in rural (but not urban) India. We present such evidence here.

6.1 Weather and Income in Rural India

We begin with estimates of the weather-income relationship in rural India. As discussed above, the vast majority of households in rural India derive income from the agricultural sector. Measures of income in that sector, namely agricultural yields and wages, therefore comprise the best measures of rural income throughout our sample period.

We focus on the overall effect of temperature on income, which can be displayed graphically (as in Figure 1, but with the point estimates reported in Appendix Table A4). Figure 4(a) plots the seven temperature bin coefficients (following the specification in equation 5) when the log of total (all crops) agricultural yields are used as the dependent variable. This variable is intended to proxy for total district-level revenue per acre, deflated by an agricultural price index. It may therefore be a useful proxy for the incomes of owner-cultivators. Echoing the findings in a large agronomic and economic literature, we estimate large negative effects of high temperature days on agricultural yields (Guiteras, 2009; Dell et al., 2014; Deschenes and Greenstone, 2007; Deschenes, 2014). The precise shape of the relationship in Figure 4(a) is difficult to ascertain due to the width of the 95% confidence interval. But if we take the point estimates literally they imply a response function that is somewhat different at the highest temperatures from that seen in the existing literature. One potential explanation for this is our focus on total (all-crop) agricultural output rather than on any single crop.

In Appendix Table A5 we also report the estimated impacts of rainfall on agricultural output. These results—that scanty rainfall reduces agricultural yields—confirm relationships estimated in prior work on Indian agriculture, such as Duflo and Pande (2007), Rosenzweig and Binswanger (1992), Jayachandran (2006) and Townsend (1994).

Figure 4(b) then extends this result—along the lines of our mortality analysis in Section 5.2—by disaggregating the daily temperature data into days that take place during the growing season and those that do not. As would be expected, hot days during the growing season correlate strongly and statistically significantly with agricultural yields, but hot days in the non-growing season do not (an exception being the 80-84° F bin). These results imply that a single day in (for example) the 85-89° F range, relative to the 70-74° F range, during the growing season reduces yields by fully 0.7%.

Finally, we turn to a different measure of agricultural income: the real wage earned by agricultural workers. As described in Section 3, this variable is calculated as the average district-level wage of agricultural day laborers (which is calculated only among those working) deflated by our district-specific agricultural price index. Figure 4(c) illustrates how this proxy for rural incomes also responds adversely to hot temperatures. This finding is particularly significant as agricultural laborers constitute one of the largest and poorest groups in India (Deaton and Dreze, 2009). While the agricultural yield and agricultural wage response functions are not of precisely the same shape (though as before this comparison is not precise due to the size of the standard errors on each coefficient), the distinction is of the expected form if workers suffer health consequences from working outside on hot days and need to be compensated for this.

Overall, we conclude that, during the 1956-2000 period, high temperature days depress agricultural incomes whether they are measured with yields (as would matter to owner-cultivators) or real wages (as would matter to agriculturally-engaged workers).

6.2 Weather and Income in Urban India

We turn now to the urban Indian analogs of the results in the previous subsection. As discussed in Section 3, state-level data output and wages in the registered manufacturing sector (a sector that is only rarely active in rural regions) are our best available proxies for urban incomes.

Figure 4(d) displays the seven temperature bin coefficient estimates we obtain when we estimate a state-level analog of equation (5) and use the log of registered manufacturing GDP as the dependent variable.[43] There is no statistically significant support for an effect of temperature on urban output—though as might be expected, the confidence intervals on these estimates from state-level data are wider than those from our district-level rural outcomes (in Figures 4(a) and 4(c)).[44] Similarly, there is as expected no evidence for an effect of rainfall on manufacturing output, as shown in Appendix Table A4, which reports all of the coefficients from equation (5).

Figure 4(e) confirms that the weak evidence for an effect of temperature on registered manufacturing output applies during both the growing and non-growing seasons, in contrast to the rural equivalent illustrated in Figure 4(b). Finally, in Figure 4(f) we check that, as expected, the registered manufacturing wage is largely unaffected by temperature.

In summary, from the best available data, we conclude that there is no strong temperature-income relationship in urban India, which stands in stark contrast to the rural-agricultural results discussed in Section 6.1.

7 Climate Change and Death

The results above suggest that extreme temperatures have strong effects on mortality (Section 5) and income (Section 6) in rural areas of India. Both sets of results are important in their own

[43]Specifically, these specifications involve: temperature variables that are based on the state-wide mean daily temperature (aggregating using population-weighted averages of district-level weather within each state); we control for state-level aggregate rainfall tercile indicators, state fixed-effects, year fixed-effects, climatic region-specific quadratic polynomials in time; and clustered standard errors that are clustered at the state level.

[44]A nascent literature estimates the effect of temperature on manufacturing output—see, Dell et al. (2014); Burke et al. (2015).

right as they suggest that weather fluctuations matter a great deal for the welfare of poor, rural citizens in developing countries today. But these estimates may also provide some insight into the potential consequences of future climate change both in a developing country (India) and a developed country (the US). While challenges in formulating such predictions result in limitations to our analysis (as we discuss below), we believe that the exercise here forms a useful first step in understanding the pattern of predicted effects in populations at different stages of development.

Our calculations begin by drawing on the predicted change in India and the US's climate that emerges from a leading global circulation model (HadCM3, under the business as usual A1F1 scenario), as described in Section 3. To align the predicted climatic variation with the inter-annual climatic variation we use throughout the paper, we obtain predictions from the HadCM3 model about the average number of days, in each year from 2015-2099, on which the mean temperature will fall into each of the eight temperature bins from equation (5). This amounts to constructing a projected value for $TMEAN_{ijt}$—the number of days in year t when the mean temperature in district (India) or state (US) i falls into temperature bin j—for a year t in the future. We do the same for the rainfall variables in equation (5).[45] In practice, we employ the average of this variable over 15-year periods to smooth out noise in the annual climate projections. We split 2015-2099 into five 15-year periods and refer to, for example, $TMEAN_{ij,2015-2029}$ as the average temperature variable (for district (India) or state (US) i and bin j) in the 2015-2029 period.

Figure 2 illustrates how climate change is expected, according to the Hadley 3 model, to change the full distribution of daily mean temperatures in India (panel A) and the US (panel B). According to this model, India will exchange a large number of days that we have estimated, in Figure 1, to cause relatively low annual mortality for days that we have estimated to cause high mortality.

We now calculate what our estimates imply for the predicted effect of climate change on life expectancy in rural India, urban India and the US. We do this in three steps:

1. Calculate the predicted change—according to our regression estimates in Figure 1(b) (Rural and Urban India) and Figure 1(a) (US) —in the mortality rate due to climate change for each future year t. This is simply: $\Delta\widehat{Y_t} = \sum_j \widehat{\theta}_j \Delta TMEAN_{jt}$, where $\Delta\widehat{Y_t}$ is the predicted change in the log mortality rate, $\widehat{\theta}_j$ is the estimated coefficient on temperature bin j obtained from Figures 1(a) and 1(b), and $\Delta TMEAN_{jt}$ is the predicted (according to the Had3CM A1FI model) change in the number of days on which the mean temperature will fall into temperature bin j in period t averaged across all districts (India) and states (US).
2. Apply $\Delta\widehat{Y_t}$ to 2000 life tables separately for rural India, urban and the US.
3. Calculate the change in life expectancy at birth due to the projected change in the

[45] There is considerably less agreement in the climatological literature about how climate change will affect precipitation patterns (particularly in India, where the complex dynamics of the monsoon are not well understood). This uncertainty is unlikely to matter a great deal for our mortality predictions, however, because our estimates of the effect of rainfall variation on mortality are relatively small.

life tables for each year, relative to baseline 2000 life expectancy, for every year from 2015-2099. Group these predicted changes in life expectancy into 15-year periods.

Before turning to a discussion of the results, it is important to be clear about the meaning of each of these calculations. For each of the 15-year periods, the above exercise produces an estimate of the life expectancy effect for a person hypothetically born in a future period with a 2000 Indian's or American's life expectancy, but who will continue to experience the climate (and the associated additional mortality risks) of the future period in which they were born in all periods throughout his/her life. For example, the reported number corresponding to 2030 refers to the reduction in life expectancy (relative to a year 2000 benchmark) for a hypothetical person born with the climate of 2030 and for whom the climate is constant (at 2030 levels) throughout his/her life. Consequently, this calculation does not capture the full reduction in life expectancy that would be predicted from a continually worsening climate as this hypothetical person ages. In this respect, it is appropriate to interpret the calculation as the life expectancy effect of alternative climate change scenarios. For example, the 2030 result corresponds to the predicted reduction in life expectancy in a hypothetical scenario in which the climate worsens up to 2030 but then stabilizes at that 2030 level.

Figure 5 reports on the results from this exercise. The rural India results are striking as they reveal a substantial predicted reduction in life expectancy (relative to 2000 life expectancy) that begins imminently and grows dramatically throughout the remainder of the 21st century.[46] Specifically, the average Indian born in a rural area during the 2015-2029 period is expected to lose 2.4 years (with a standard error of 0.75 years) of life expectancy over the course of his or her life (relative to year 2000 life expectancy), even under the conservative assumption that the climate does not worsen throughout this individual's lifetime after he or she is born in 2015-2029.[47] This increases to losses of 5.1 years and 10.4 years for those born in 2045-2059 and 2075-2099, respectively.

The impact on life expectancies in urban areas (relative to 2000) is projected to be substantially smaller (and with a 95% confidence band that includes zero impact until the estimate for 2075-2099) throughout the 21st century. For example, the 2075-2099 prediction is a loss of 2.8 years, comparable to the predicted life expectancy effect in rural areas for the relatively moderate period 2015-2029. Effects for the US in Figure 5 show that climate change could have important consequences for US mortality, but those effects pale in comparison to the future that our results predict for rural Indians. Relative to 2000 life expectancy we see a reduction of 0.06 years in

[46] Our estimates are relative to a 2000 life expectancy benchmark and are therefore capturing the degree to which the mortality impacts of rising temperatures will constrain the likely growth in life expectancy (due to improved medical tenchnologies, better nutrition and rising incomes) over the coming century and suggest that these constraints will be felt most keenly by rural Indians. Our estimates therefore should not be interpreted as effects on actual life expectancy in the periods being studied.

[47] Standard errors are calculated using the "delta-method", which takes the uncertainty in the estimates of the betas (i.e. temperature coefficients) into account (clustering at the district (India) or state (US) level), and treats the climate change predictions and life table entries as non-stochastic.

2015-2029, 0.15 years in 2045-2059 and 0.28 years in 2075-2099.[48]

Table 7 extends the above analysis in a number of respects. In each case, we present the effects for rural areas (as computed from the estimates in Figure 1(b)) in columns 1-3, for urban areas (also from Figure 1(b)) in columns 4-6, and for India as a whole (computed from the population-weighted average of the rural and urban results in Figure 1(a)) in columns 7-9. First, in row 2, we take seriously the changing impact of temperature over time found in Table 4, and utilize only the latest of our estimates (that pertain to 1987-2000) to predict climate change impacts. These impacts are lower, but still startling: 6.17 years lower rural Indian life expectancy by 2075-2099 though urban life expectancy effects (0.34 years) do move closer to US estimates. Second, in row 3, we use only the coefficients that are estimated from hot districts and continue to observe large impacts. Third, in rows 4 and 5, we alter our approach to computing the all-India impact in columns 7-9 to account for different rates of urbanization. Row 4 uses the 2000 population-weighted average (with 32% weight on the urban estimates, rather than the 24% historical average weight used in previous rows) of the estimated rural and urban impacts, and row 5 uses equal weights to account for the future urbanization that is likely to occur in India. Under both scenarios our estimates remain striking.

Finally, the last two rows of Table 7 use the same method to forecast the impacts of climate change on the income outcomes—measures of real output and wages—from Section 6. As above, because climatological models predict that the coming 85 years will bring so many more hot days to India in any given year, our estimates imply substantially reduced incomes in rural areas: 14.3% (0.155 log points) lower real agricultural output per acre and 26.5% (0.308 log points) lower real agricultural wages. These magnitudes imply that there is likely to be substantial rural economic suffering, even among survivors, due to the many hot days expected in India's future.

These calculations suggest that the health costs of predicted climate change in India could be severe. When standard models of climate change are used in combination with our estimates of the weather-death relationship, these models predict large reductions in life expectancy in rural India over the remainder of this century. Further, it is important to recall that the estimates applied here are deliberately on the conservative end of those presented in Section 5. Both IV models (to correct for measurement error in the weather data) and dynamic models (to sum over contemporaneous and lagged effects of weather on death) suggested effects that were about twice as large as the baseline estimates that we emphasize in our climate change calculations.

These estimates underline how a warming climate may have unequal effects on human populations at different stages of development. They suggest that a changing climate will amplify the death effects we observed across the 1957-2000 period leading to significant reductions in life expectancy of rural Indians across this century. These results are important in terms of pointing to

[48]The much smaller magnitudes for the US are the result both of hot days having a much smaller impact on mortality (Figure 1), the increase in the number of hot days in the US being more modest (see Figure 2) and because the reduction in cold days that climate change brings actually helps reduce mortality (see Appendix Table A.1).

how climate change may have dramatically different effects on populations both within developing countries and across developing and developed countries. In common with the 1957-2000 results they suggest that the structure of production and employment is key to determining how human health is affected by the weather.

As forecasts our estimates are subject to a number of limitations. The first obvious limitation is that the validity of our estimates of the impacts of climate change depends on the accuracy of the climate change predictions that feed into our analysis. We have used a business as usual scenario from a well-known climatological model but this should be conceived of as a single realization from a super-population of models and scenarios. The sources of uncertainty in these models and scenarios are unclear and cannot readily be incorporated into our estimates of the impacts of climate change (Burke et al., 2015). Also if attempts to mitigate climate change by reducing greenhouse gas emissions were even partially successful then this would bring down our estimates as the earth would warm more slowly.

Another key limitation is that our procedure draws on estimates of the impact of year-to-year weather fluctuations to gauge the potential impact of climate change, a slower, more permanent and more forecastable weather shock. Individuals can adapt to worsening climatic conditions for example by shifting away from climate exposed occupations and regions, by using more heat resistant agricultural technologies or by adopting protective technologies such as fans and air conditioning. These actions and the actions taken by government to protect its citizens may help to constrain the impact of climate change but modeling this adaptation is outside the scope of this paper.

The inter-annual weather shocks that we use are also different from a longer-run change in climate in a spatial sense. Whereas our backward looking estimates isolate spatially idiosyncratic variation, climate change will affect the whole of India and the US. Because implicit cross-regional insurance against differential regional-level shocks (such as from government policies or inter-regional integration of financial, labor, or product markets) is likely to be substantial, this means that our estimates could be understating the impact of a spatially correlated shock like climate change.

Finally there is a sense in which our study is incomplete as our estimates cannot capture all of the potential negative health impacts of climate change. For example, in India, coastal inundation, changes in the timing of the monsoon, the melting of the Himalayan glaciers or desertification of soil could greatly increase income losses. The climatological models whose climate change predictions we have used here do not incorporate any possibility of catastrophic changes in India's or the US's climate as a result of a rise in greenhouse gas emissions. While some climatological models predict that modest rises in temperatures may have catastrophic knock-on effects, we have deliberately obtained our climate predictions from climatological models in which these catastrophic, but highly uncertain and controversial, effects are not in operation.

The bottom line is that we cannot capture in our analysis the full set of adpative responses that individuals and governments will make to climate change and the extent to which these responses will be constrained by the fact that the citizens of India and the US will be living in world where the climate is permanently and significantly hotter. However, the fact that rural Indians, despite significant actions by individuals and governments, have not been fully protected against the health damaging impacts of high temperatures during a period (1957-2000) when the climate was much less inclement than it is predicted to be for the remainder of this century should be a significant concern.

8 Conclusion

If we look across the world it is striking how much of humanity still lives in the rural areas of developing countries. This can be seen in night images of the earth—whereas the developed nations are lit up vast tracts of the developing nations remain black with the darkness punctuated only by the bright spots of towns and cities. Out of a global population of 7.3 billion, 0.6 billion live in low income countries, 2.9 billion in lower middle income countries (which includes India), 2.5 billion in upper middle income countries (which includes China) and the remaining 1.2 billion in high income or developed countries (which includes the US). The proportions living in rural areas in these four groupings are 70%, 61%, 37% and 19% (World Bank Group, 2014). So 43% of humanity (3.1 billion) live in the rural areas of developing countries.[49] The key contribution of this paper has been to show that the health impacts of hot weather are highly unequal and inflict the greatest damage on these populations.

A great deal of work has been done on how informal institutions in poor, rural settings have evolved to enable households to deal with the income volatility induced by weather variation (Townsend, 1994; Paxson, 1993; Deaton, 1997; Dercon and Krishnan, 2000; Dercon, 2005; Khandker, 2012; Chiappori et al., 2014; Cole et al., 2013). There is also a large literature on the state's central role in protecting rural citizens from the deleterious effects of inclement weather (Dreze and Sen, 1989; Besley and Burgess, 2002; Imbert and Papp, 2015). India, for example, is spending a significant fraction of its GDP on public distribution and public works systems. The sobering conclusion from this paper is that, despite all these efforts, rural Indians are not fully protected against the death effects of hot weather. This accords well with the intuition that—though citizens can organize themselves to protect themselves against idiosyncratic shocks—they are less able to deal with aggregate, district-level shocks.[50] Therefore though famines have been eliminated from

[49] Of the billion people who live in extreme poverty, around 70% live in rural areas and over 80% of these rural households rely to some extent on farming (IFAD, 2010).

[50] In Townsend (1994) the village budget constraint is the aggregate budget constraint and so if the incomes of all village citizens fall as a result of, for example, a district-wide temperature shock then smoothing consumption (and survival) within the village becomes problematic. Munshi and Rosenzweig (2016) point out that insurance networks often extend beyond the village. What our results suggest is that, even with these wider networks and the plethora of policies that enable resources to flow into villages in response to shocks, rural Indians are not fully insured against the mortality elevating impacts of aggregate hot weather shocks.

modern day India (Sen, 1981), inclement weather continues to kill people in large numbers in the Indian countryside. This is a striking finding given that India is not anywhere near the bottom of the world income distribution.

An equally striking result is that, despite widespread extreme poverty, urban populations in India are protected from the mortality impacts of hot weather. It is well-known that people are attracted to living in towns and cities by higher wages and better public amenities (Ciccone and Hall, 1996; Glaeser and Gottlieb, 2009; Glaeser, 2014). What this paper points to is the insurance value of living in urban areas—relative to the countryside, towns and cities serve as havens where citizens are better insulated against the deleterious effects of hot weather. This has been little emphasized in the existing literature. The fact the economy-wide urban India and US temperature-mortality response functions resemble one another more than the rural and urban India functions underlines how considerable this insurance value is.

At the time of writing in 2016 India was experiencing one of the hottest years on record. According to the Indian Meterological Department, India has seen 12 of its 15 warmest years ever since 2001 and the past decade was the warmest decade on record (IMD, 2016). This warming pattern is common across the world—the National Centers for Environmental Information report that 15 of the 16 warmest years between 1880 and 2015 occurred in the 21st century (NOAA, 2016). Against this background of a warming climate, the question of how ambient temperatures affect human populations gains added salience. A key contribution of the paper is to point to the vulnerability of rural, developing country populations to the health damages caused by a warming climate. We find that, *ceteris paribus*, the large increase in hot days by the end of this century will result in a large loss in life expectancy for rural Indian, but not urban Indian or US, populations.

This is a paper that throws some big new facts out into the world. It, however, begs the question of what mechanisms underlie these facts. Our model provides a framework for thinking about this. It makes clear that high temperatures may affect mortality both via direct health and indirect income effects. Empirically we cannot disentangle these direct and indirect channels but we do find evidence consistent with the income effect being more prominent in rural areas. This interesting as it suggests that even if rural and urban citizens are equally poor rural citizens are more *vulnerable* to their health being damaged by weather shocks. The importance of these volatile environments where weather-dependent agriculture is the main employer is also what distinguishes developed from developing countries (Koren and Tenreyro, 2007). Protecting citizens in developing countries from weather risk is therefore more difficult not just because these countries are (by definition) poorer and have fewer resources at their disposal but also because they face significantly greater weather-induced income volatility.

In the face of a warming planet much new thinking will need to go into designing effective social protection policies for these at-risk populations. A key insight from our model is that, irrespective of the channel, any policy that relaxes an individuals budget constraint and enables her/him to maintain expenditure on health goods should help them to avert death. Our finding that the arrival of bank branches the rural areas mitigated the impact of hot weather on death confirms this intuition and illustrates how public policy can play a key role in protecting poor, rural populations from weather risk. There is a multitude of options here from subsidizing health goods through to encouraging migration, diversifying of economic activities and targeting transfers to households hit by inclement weather. Much work in coming years will have to focus on evaluating such policies.

By applying our methodology to data sets for other countries it would also be useful to find out how widespread the vulnerability of human health to hot weather is. We would want to know whether rural populations across the developing world are affected but also whether these effects extend to urban populations, particularly those in the poorest nations. This would allow us to build up a more general picture of whether and how development protects human populations from hot weather. On mechanisms it would be useful to use experimental or other estimates of the elasticity of mortality with respect to income so that we can break-down how much of the overall effect we observe is going though the indirect income channel and how much through the direct health channel. Large income transfer programs like NREGA in India seem like promising candidates in this regard. In all this work it would be useful to know more about occupation, cause of death, age at death, saving, borrowing, lending and transfers so that we can better chart out the channels of influence.

Viewed from space the earth consists of a patchwork of areas at different stages of development. As hot weather passes over these areas we find that it has very different effects on human health. In a world that is warming and where 43% of humanity still lives in the rural areas of developing countries, this presents some key challenges for humankind which will need to be taken up in the coming decades.

References

Acemoglu, Daron and Martin Kaae Jensen, "Robust Comparative Statics in Large Dynamic Economies," *Journal of Political Economy*, 2015, *123* (3), 587–640.

Aiyagari, S. Rao, "Uninsured Idiosyncratic Risk and Aggregate Saving," *Quarterly Journal of Economics*, 1994, *109*, 659–684.

Barreca, Alan, Karen Clay, Olivier Deschenes, Michael Greenstone, and Joseph S Shapiro, "Adapting to Climate Change: The Remarkable Decline in the US Temperature-Mortality Relationship over the Twentieth Century," *Journal of Political Economy*, 2016, *124* (1), 105–159.

Basu, Rupa and Jonathan M. Samet, "Relation Between Elevated Ambient Temperature and Mortality: A Review of the Epidemiologic Evidence," *Epidemiologic Review*, 2002, *24 (December)*, 190–202.

Becker, Gary, "Health as human capital: synthesis and extensions," *Oxford Economic Papers*, 2007, *59*, 379–410.

Besley, T. and R. Burgess, "The political economy of Government responsiveness: thery and evidence from India," *Quarterly Journal of Economics*, 2002, *117* (4).

Besley, Timothy, "Savings, credit and insurance," in Jere Behrman and T.N. Srinivasan, eds., *Handbook of development economics (volume 3A)*, Elsevier, 1995, pp. 2123–2207.

— **and Robin Burgess**, "Can Labor Regulation Hinder Economic Performance? Evidence From India," *Quarterly Journal of Economics*, 2004, *119(1)*, 91–134.

Bewley, Truman, "Stationary Monetary Equilibrium with a Continuum of Independently Fluctuating Consumers," in Werner Hildenbrand and Andreu Mas-Colell, eds., *Contributions to Mathematical Economics in Honor of Gerard Debreu*, Amesterdam: North Holland, 1986.

Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak, "Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh," *Econometrica*, 2014, *82* (5), 1671–1748.

Burgess, Robin and Rohini Pande, "Can Rural Banks Reduce Poverty? Evidence from the Indian Social Banking Experiment," *American Economic Review*, 2005, *95(3)*, 780–795.

Burke, Marshall, Solomon M Hsiang, and Edward Miguel, "Global non-linear effect of temperature on economic production," *Nature*, 2015.

Carleton, Tamma A and Solomon M Hsiang, "Social and economic impacts of climate," *Science*, 2016, *353* (6304), aad9837.

Chen, Shaohua and Martin Ravallion, "The developing world is poorer than we thought, but no less successful in the fight against poverty," *World Bank Policy Research Working Paper Series, Vol*, 2008.

Chiappori, Pierre-André, Krislert Samphantharak, Sam Schulhofer-Wohl, and Robert M Townsend, "Heterogeneity and risk sharing in village economies," *Quantitative economics*, 2014, *5* (1), 1–27.

Ciccone, Antonio and Robert Hall, "Productivity and the Density of Economic Activity," *American Economic Review*, 1996, *86* (1), 54–70.

Cole, Shawn, Xavier Giné, Jeremy Tobacman, Petia Topalova, Robert Townsend, and James Vickery, "Barriers to household risk management: Evidence from India," *American Economic Journal: Applied Economics*, 2013, *5* (1), 104–135.

Dasgupta, Partha, "The place of nature in economic development," *Handbook of Development Economics*, 2010, *5.*

Datt, Gaurav, "Poverty in India and Indian states: an update," *The Indian Journal of Labour Economics*, 1998, *41* (2), 191–211.

_ **, Valerie Kozel, and Martin Ravallion**, "A Model-based Assessment of India's Progress in Reducing Poverty in the 1990s," *Economic and Political Weekly*, 2003, pp. 355–361.

de Montesquieu, Baron, *The Spirit of Laws*, Hayes Barton Press, 1748.

Deaton, Angus, *The analysis of household surveys: a microeconometric approach to development policy*, World Bank Publications, 1997.

_ **and Jean Dreze**, "Food and Nutrition in India: Facts and Interpretations," *Economic and Political Weekly*, 2009, *XLIV (7)*, 42–65.

Dell, Melissa, Benjamin F Jones, and Benjamin A Olken, "What do we learn from the weather? The new climate–economy literature," *Journal of Economic Literature*, 2014, *52* (3), 740–798.

Dercon, Sefan and Pramila Krishnan, "In Sickness and in Health: Risk Sharing within Households in Rural Ethiopia," *Journal of Political Economy*, 2000, *108* (4), 688–727.

Dercon, Stefan, *Insurance against poverty*, Oxford University Press, 2005.

Deschenes, Olivier, "Temperature, human health, and adaptation: A review of the empirical literature," *Energy Economics*, 2014, *46*, 606–619.

_ **and Enrico Moretti**, "Extreme Weather Events, Mortality and Migration," *Review of Economics and Statistics*, 2009, *91* (4), 659–681.

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

_ **and** _ , "Climate Change, Mortality, and Adaptation: Evidence from Annual Fluctuations in Weather in the US," *American Economic Journal: Applied Economics*, 2011, *3* (4), 142–85.

Dreze, Jean and Amartya Sen, *Hunger and Public Action*, Oxford University Press, 1989.

Duflo, Esther and Rohini Pande, "Dams," *Quarterly Journal of Economics*, 2007, *122* (2), 601–646.

Glaeser, Edward L, "A world of cities: the causes and consequences of urbanization in poorer countries," *Journal of the European Economic Association*, 2014, *12* (5), 1154–1199.

_ **and Joshua D Gottlieb**, "The wealth of cities: Agglomeration economies and spatial equilibrium in the United States," *Journal of Economic Literature*, 2009, *47* (4), 983–1028.

Grossman, M., "On the concept of health capital and the demand for health," *Journal of Political Economy*, 1972, *80*, 223–55.

Guiteras, Raymond, "The Impact of Climate Change on Indian Agriculture," 2009. University of Maryland mimeo.

Harris, John R and Michael P Todaro, "Migration, unemployment and development: a two-sector analysis," *The American economic review*, 1970, pp. 126–142.

Henderson, J Vernon, "Cities and development," *Journal of Regional Science*, 2010, *50* (1), 515–540.

Hsiang, Solomon M, "Climate Econometrics," Technical Report, National Bureau of Economic Research 2016.

Huynen, Maud, Pim Martents, Dieneke Schram, Matty P. Weijenberg, and Anton E. Kunst, "The Impact of Heat Waves and Cold Spells on Mortality Rates in the Dutch Population," *Environmental Health Perspectives*, 2001, *109 (May)*, 463–70.

IFAD, "Rural Poverty Report 2011," Technical Report, International Fund for Agricultural Development 2010.

IIPS, "India National Family Health Survey (NFHS-2), 1998-99," Technical Report, International Institute for Population Sciences (IIPS) and ORC Macro 2000.

Imbert, Clement and John Papp, "Labor market effects of social programs: Evidence from india's employment guarantee," *American Economic Journal: Applied Economics*, 2015, *7* (2), 233–263.

IMD, "Statement on Climate of India during 2015," Technical Report, Earth System Science Organization, Ministry of Earth Sciences, India Meteorological Department 2016.

Indian Census, "Census of India 2011," *Office of the Registrar General Census Commissioner, India*, 2011.

Jayachandran, Seema, "Selling Labor Low: Wage Responses to Productivity Shocks in Developing Countries," *Journal of Political Economy*, 2006, *114(3).*

Kahn, Matthew E, "The death toll from natural disasters: the role of income, geography, and institutions," *Review of economics and statistics*, 2005, *87* (2), 271–284.

_ , *Climatopolis: how our cities will thrive in the hotter future*, Basic Books, 2013.

Khandker, Shahidur, "Seasonality of income and poverty in Bangladesh," *Journal of Development Economics*, 2012, *97* (2), 244–256.

Klineberg, Eric, *Heat Wave: A Social Autopsy of Disaster in Chicago*, Chicago: Chicago University Press, 2002.

Koren, Miklós and Silvana Tenreyro, "Volatility and development," *The Quarterly Journal of Economics*, 2007, pp. 243–287.

Lees, Lynn H and John Modell, "The Irish Countryman Urbanized:" A Comparative Perspective on the Famine Migration"," *Journal of Urban History*, 1977, *3* (4), 391.

Lewis, W Arthur, "Economic development with unlimited supplies of labour," *The manchester school*, 1954, *22* (2), 139–191.

Malthus, Thomas Robert, *An essay on the principle of population* 1798.

Marsh, George Perkins, *Man and nature, physical geography as modified by human action*, Charles Scribner, 1865.

Milgrom, Paul and John Roberts, "Comparing Equilibria," *American Economic Review*, 1994, *84*, 441–459.

Morten, Melanie, "Temporary migration and endogenous risk sharing in village india," Technical Report, National Bureau of Economic Research 2016.

Munshi, Kaivan and Mark Rosenzweig, "Networks and misallocation: Insurance, migration, and the rural-urban wage gap," *The American Economic Review*, 2016, *106* (1), 46–98.

NOAA, "State of the Climate: Global Analysis for Annual 2015," Technical Report, NOAA National Centers for Environmental Information 2016.

Nordhaus, William D, *Managing the global commons: the economics of climate change*, Vol. 31, MIT press Cambridge, MA, 1994.

Pai, DS, Latha Sridhar, M Rajeevan, OP Sreejith, NS Satbhai, and B Mukhopadhyay, "Development of a new high spatial resolution (0.25× 0.25) long period (1901–2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region," *Mausam*, 2014, *65* (1), 1–18.

Paxson, Christina, "Consumption and Income Seasonality in Thailand," *Journal of Political Economy*, 1993, *101* (1), 39–72.

Rosenzweig, Mark R and Hans P Binswanger, *Wealth, weather risk, and the composition and profitability of agricultural investments*, Vol. 1055, World Bank Publications, 1992.

_ **and Kenneth I Wolpin**, "Credit market constraints, consumption smoothing, and the accumulation of durable production assets in low-income countries: Investments in bullocks in India," *Journal of political economy*, 1993, pp. 223–244.

_ **and Oded Stark**, "Consumption smoothing, migration, and marriage: Evidence from rural India," *The Journal of Political Economy*, 1989, pp. 905–926.

Sanderson, Eleanor and Frank Windmeijer, "A weak instrument F-test in linear IV models with multiple endogenous variables," *Journal of Econometrics*, 2016, *190* (2), 212–221.

Sanghi, Apurva, KS Kavi Kumar, and James W McKinsey Jr, "India agriculture and climate data set," *World Bank*, 1998.

Schlenker, Wolfram and Michael Roberts, "Estimating the Impact of Climate Change on Crop Yields: The Importance of Nonlinear Temperature Effects," 2008. NBER Working Paper 13799.

Sen, Amartya, *Famines: An Essay on Entitlement and Deprivation*, Oxford University Press, 1981.

Srivastava, AK, M Rajeevan, and SR Kshirsagar, "Development of a high resolution daily gridded temperature data set (1969–2005) for the Indian region," *Atmospheric Science Letters*, 2009, *10* (4), 249–254.

Stern, Nicholas, "Stern review report on the economics of climate change," 2006.

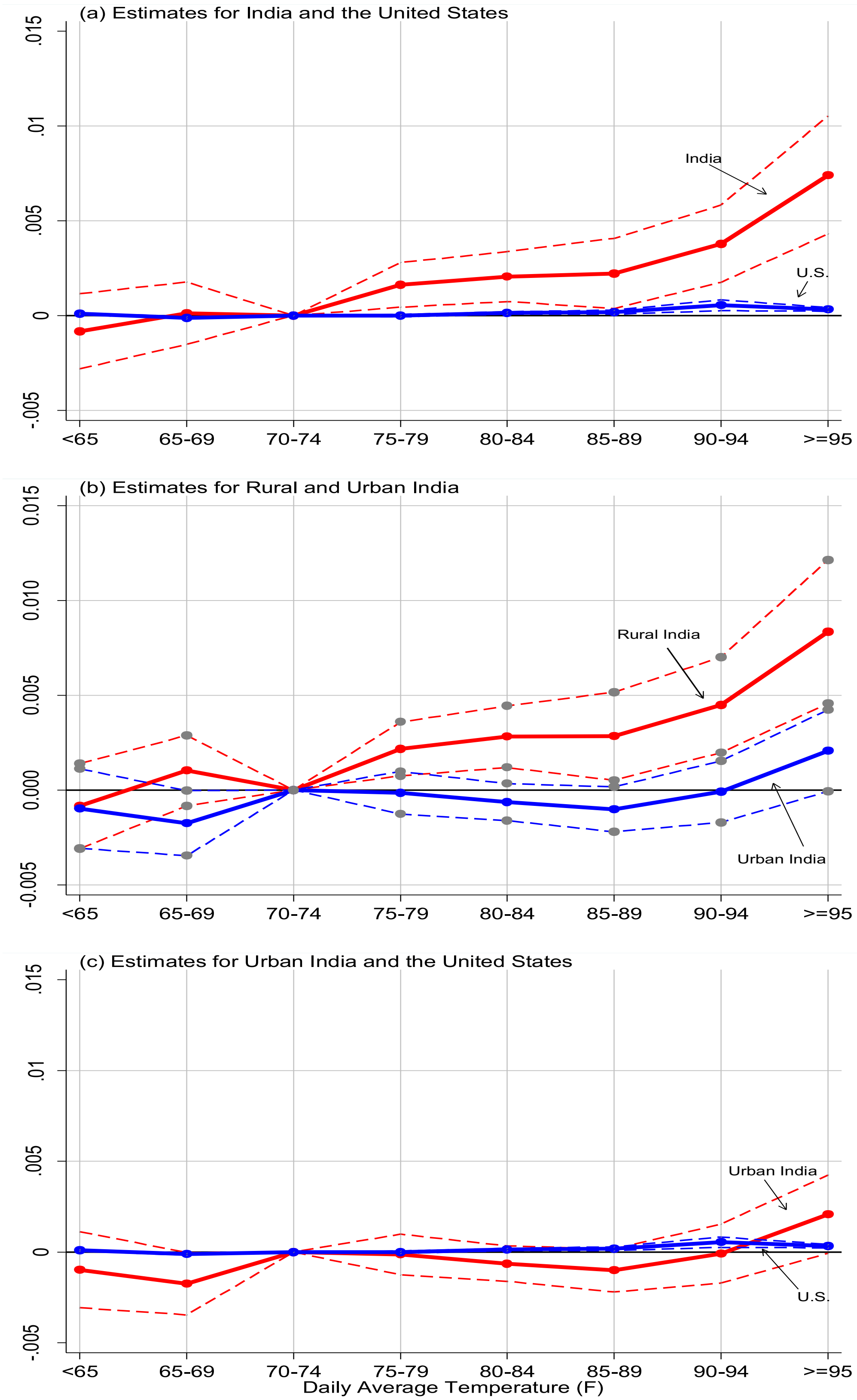
Townsend, Robert, "Risk and Insurance in Village India," *Econometrica*, 1994, *62(3)*, 539–591.

Wang, Bin, *The Asian Monsoon*, Berlin: Springer-Verlag, 2006.

World Bank Group, *World Development Indicators 2014*, World Bank Publications, 2014.

_ , *World Development Indicators 2015*, World Bank Publications, 2016.

Figure 1: Impact of Daily Temperature on Log All-Age Mortality Rates in India and the United States.

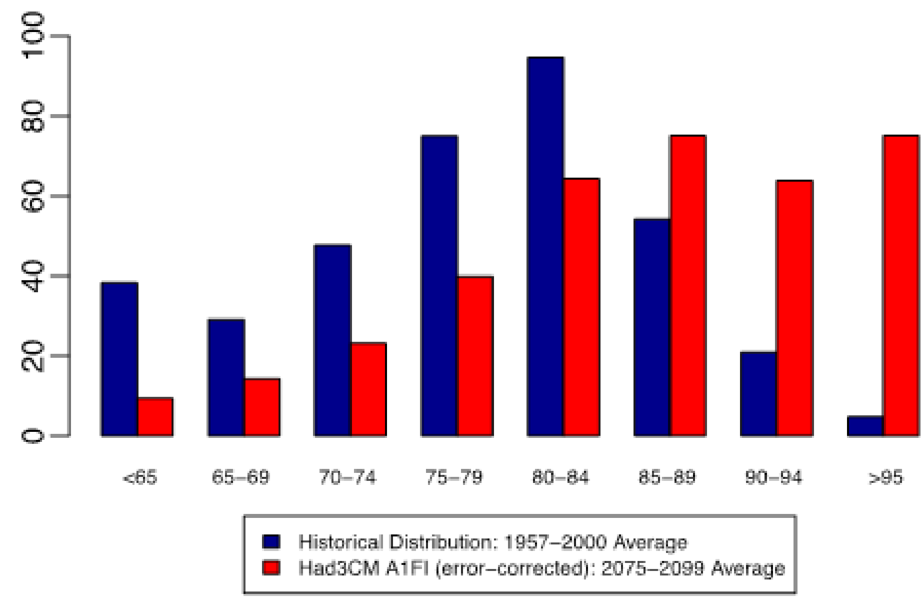


Notes: The plotted lines report 7 coefficient estimates (circle markers), representing the effect of a single day in each of the corresponding 7 temperature bins, relative to the effect of a day in the 70-

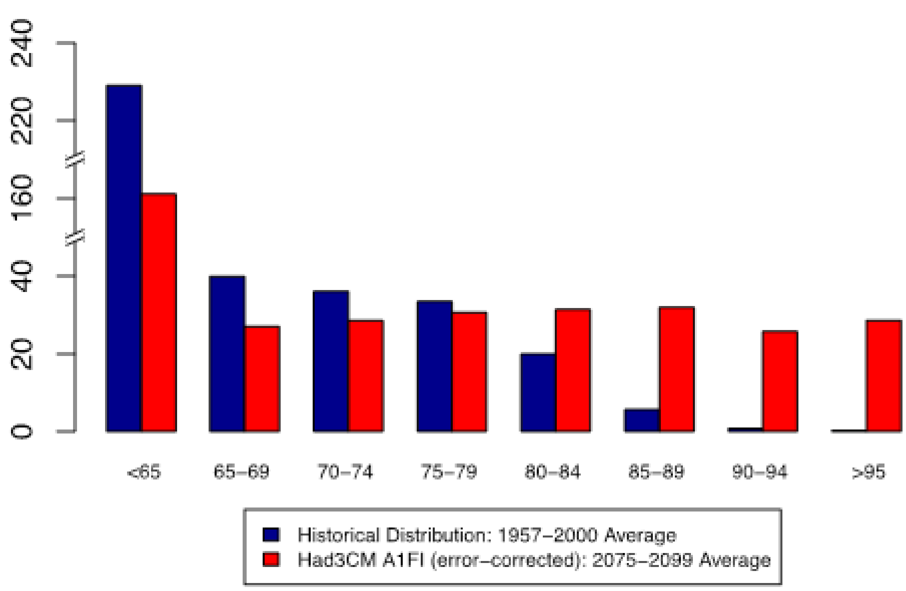
74 °F reference bin, on annual all-age log mortality rate. Dashed lines represent the 95% confidence interval of the estimates. The mortality rate data for the US is taken from Barreca, Clay, Deschenes, Greenstone, and Shapiro (2016). Daily average temperature defined as the simple average of daily minimum and maximum temperature. Standard errors clustered on district (India) and on state (US). See the text for more details.

Figure 2: Historical and Predicted Distribution of Daily Average Temperatures

(a) India



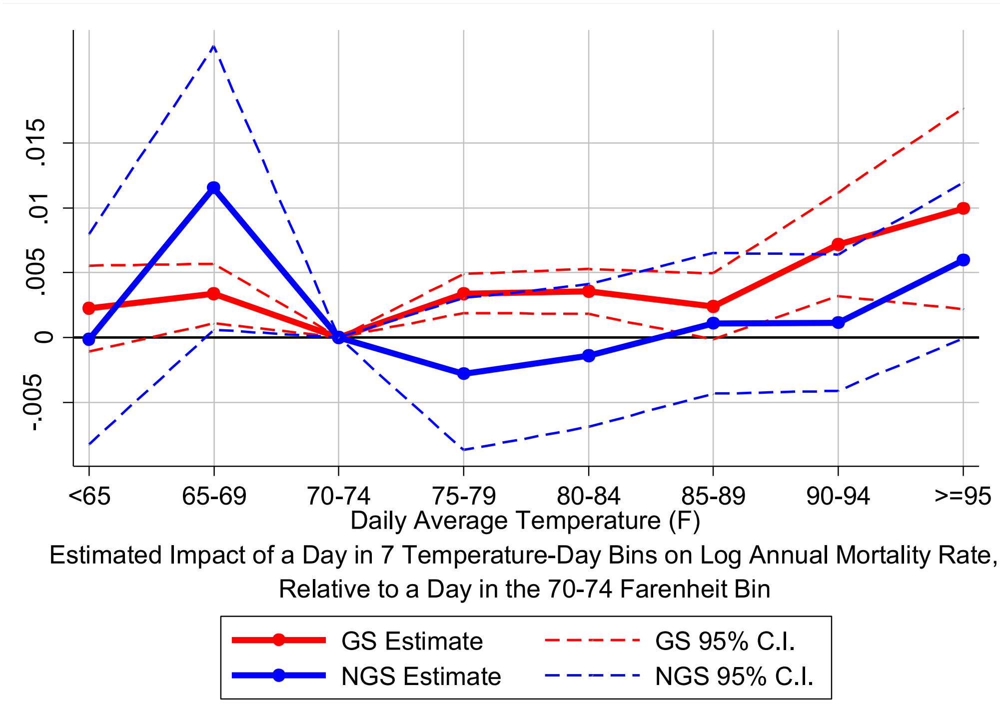
(b) United States



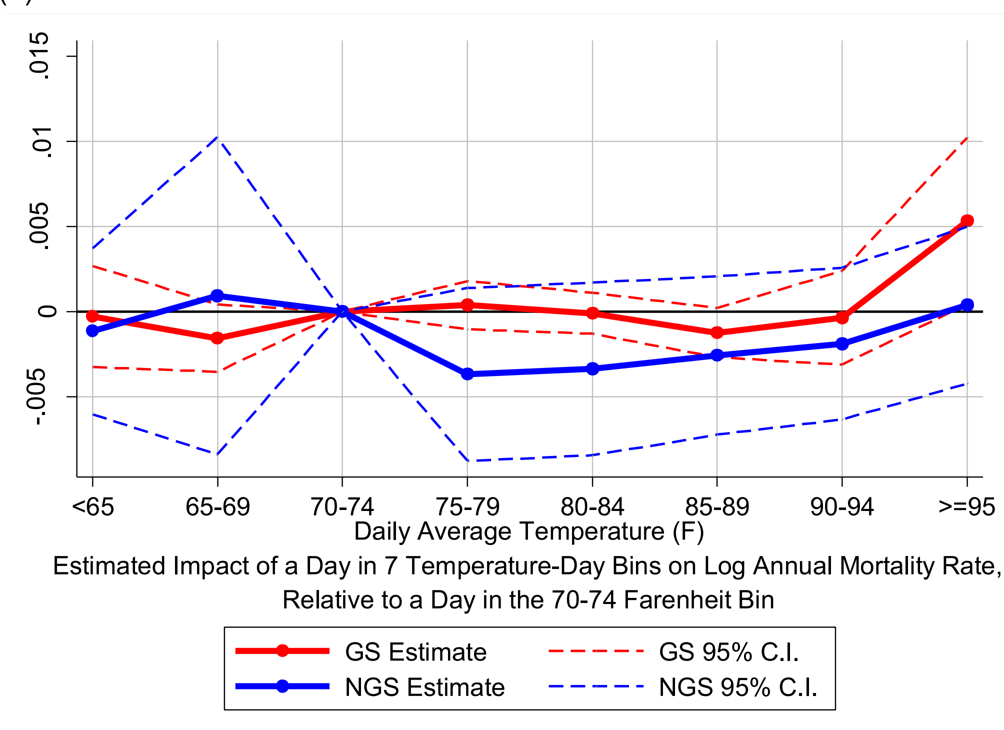
Notes: Figure 2 shows the historical average distribution of daily mean temperatures and predicted future distribution across 8 temperature bins. Blue bars represent the average number of days per year in each temperature bin over 1957-2000. Red bars show the corresponding predicted distribution derived using daily data from error-corrected Had3CM A1FI model data for the period 2075-2099. Both averages weighted by 1957-2000 district population in India and state population in the United States. See the text for more details.

Figure 3: Impact of Daily Temperature on Log All-Age Mortality Rate in India, by Agricultural Growing Season

(a) Rural



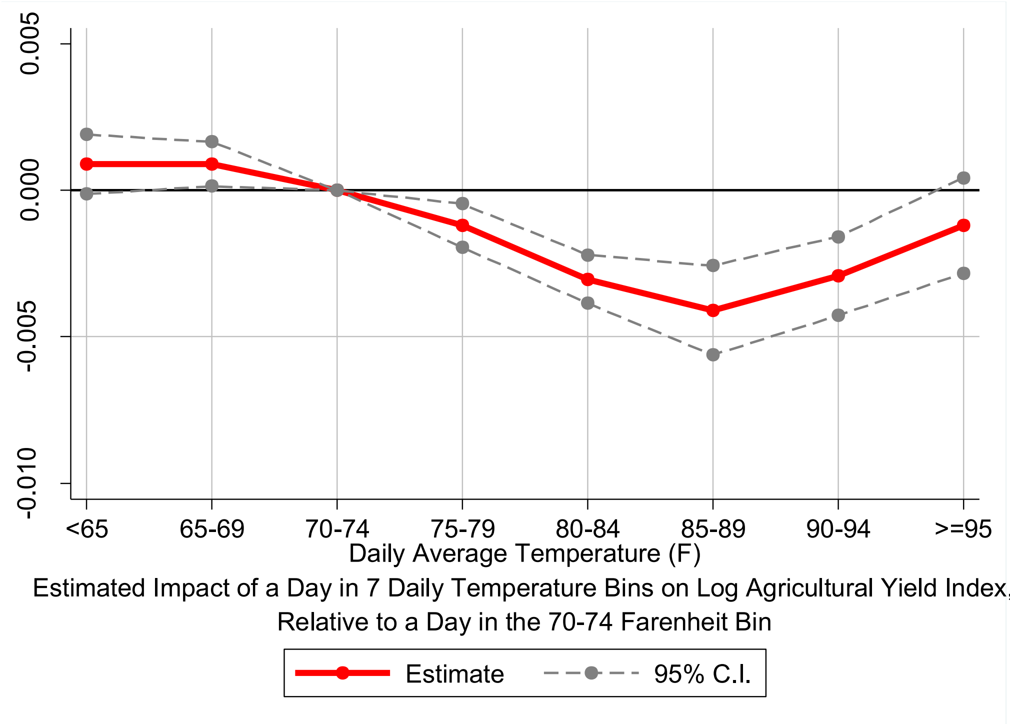
(b) Urban



Note: The red (growing season) and blue (non-growing season) lines report 7 coefficient estimates (circle markers), representing the effect of a single day in each of the corresponding 7 temperature bins, relative to the effect of a day in the 70-74 °F bin, on annual log all-age mortality rate by agricultural growing season in rural and urban India. The regressions also control for rainfall (upper/lower terciles indicators), district fixed effects, year fixed effects, and climatic region specific time trends. Growing season defined using district-specific historical average monsoon arrival date from the Indian Meteorological Department. Growing and non-growing season coefficients are estimated in the same regression that is weighted by district population. Dashed lines represent the 95% confidence interval of the estimates. The F-statistic testing the hypothesis of no difference in the temperature-all age mortality relationships by agricultural growing season in rural India is 4.49 (p-value <= 0.001). The corresponding F-statistic for urban India is 1.37 (p-value = 0.218). See the text for more details.

Figure 4: Impact of Daily Temperature on Proxies of Income

(a) Impact of Daily Temperature on Log Agricultural Yield Index



(b) Impact of Daily Temperature on Log Agricultural Yield Index, by Agricultural Growing Season

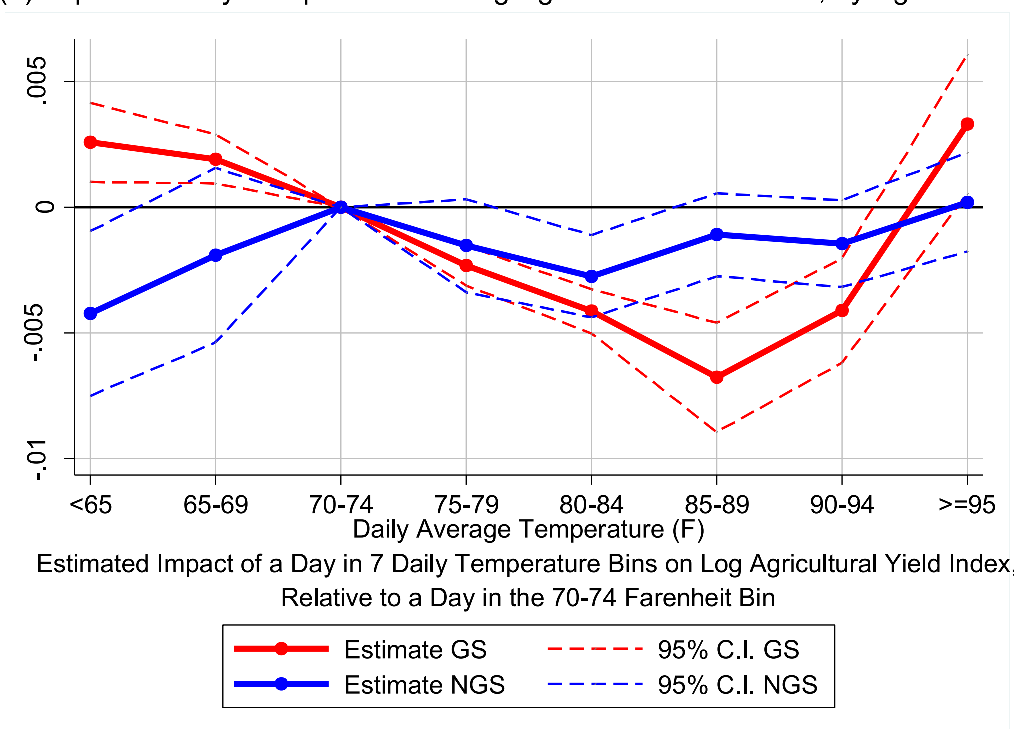
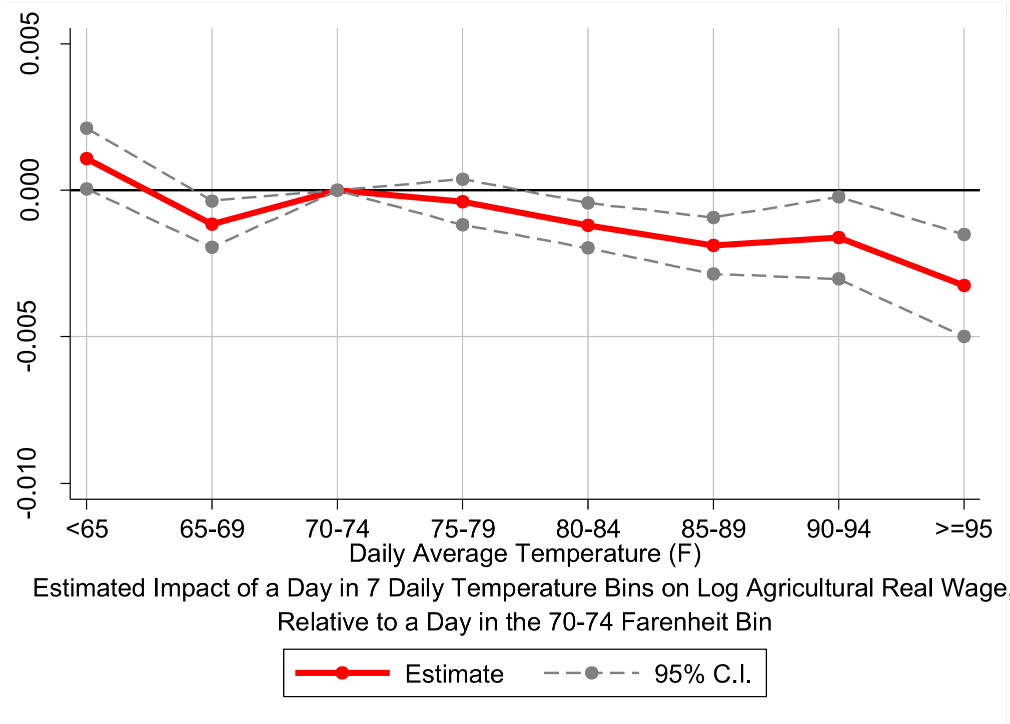


Figure 4: Impact of Daily Temperature on Proxies of Income (ctd)

(c) Impact of Daily Temperature on Log Agricultural Real Wage



(d) Impact of Daily Temperature on Log Manufacturing Output

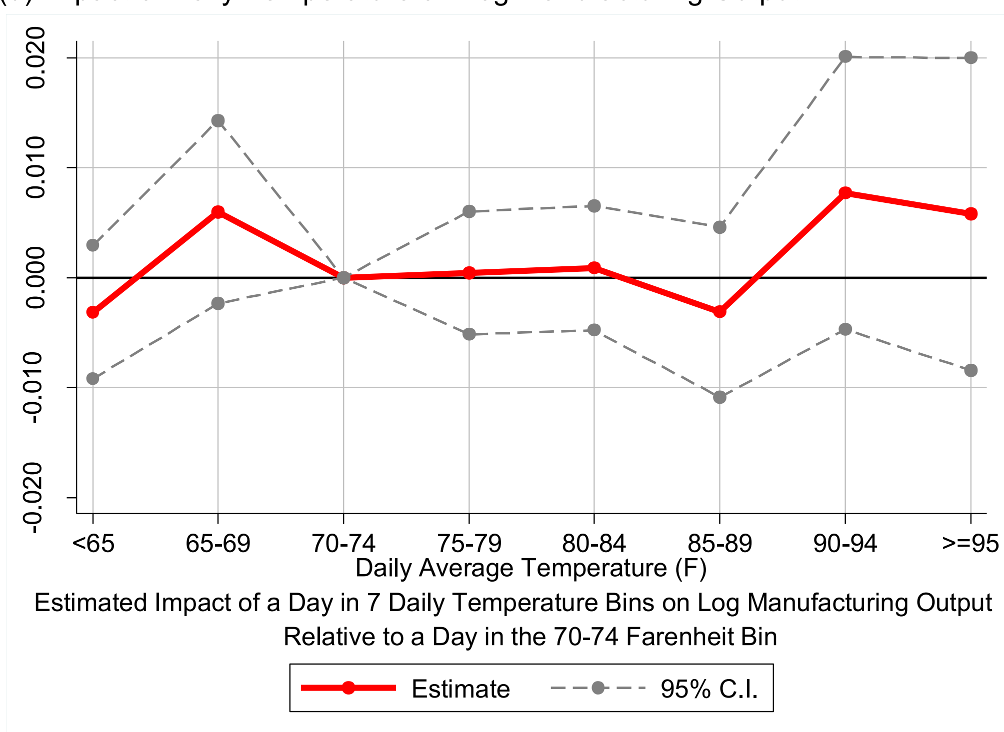
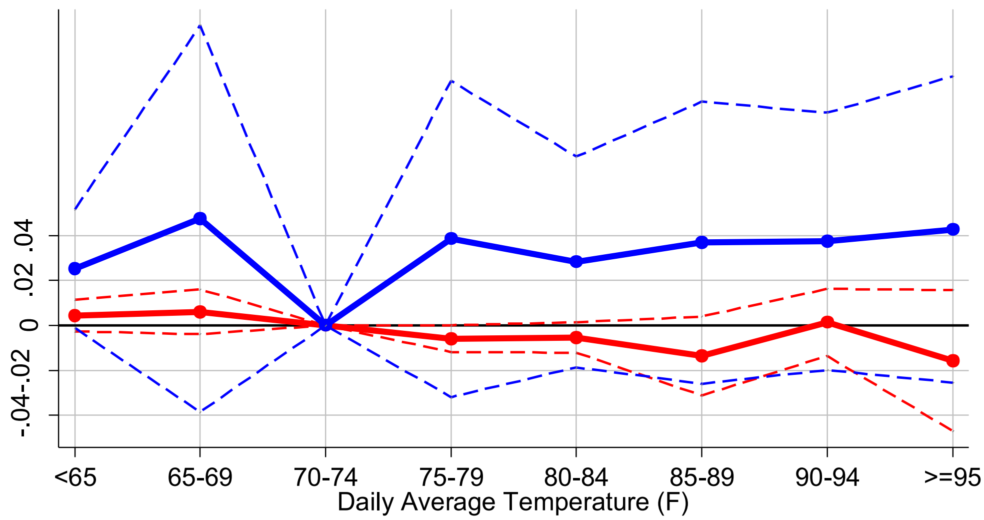


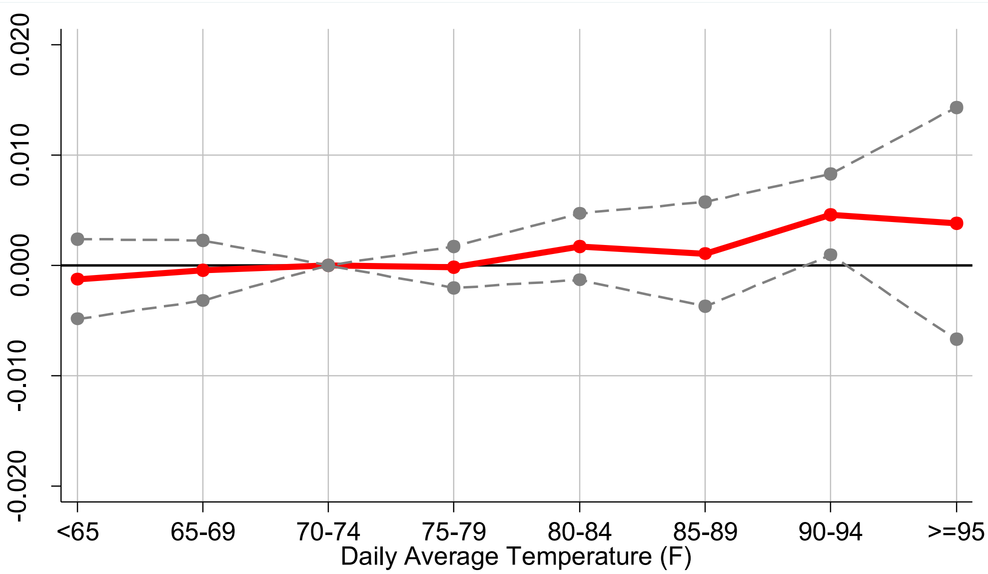
Figure 4: Impact of Daily Temperature on Proxies of Income (ctd)

(e) Impact of Daily Temperature on Log Manufacturing Output, by Agricultural Growing Season



Estimated Impact of a Day in 7 Daily Temperature Bins on Log Manufacturing Output Relative to a Day in the 70-74 Farenheit Bin

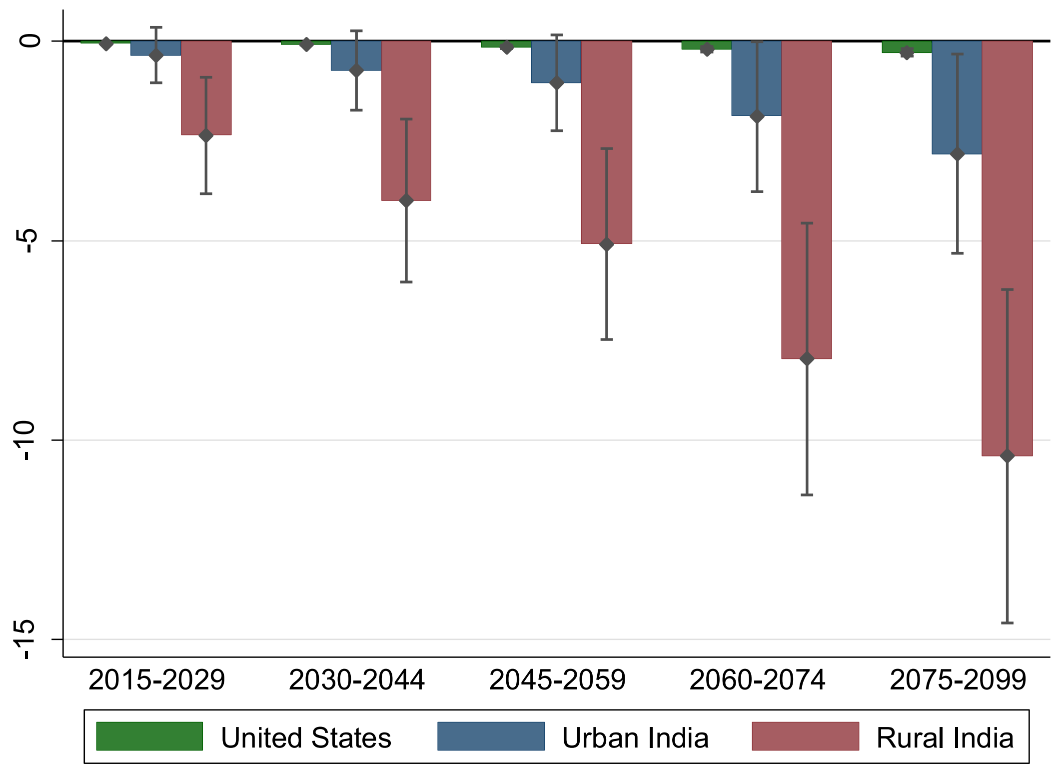
(f) Impact of Daily Temperature on Log Manufacturing Real Wage



Estimated Impact of a Day in 7 Daily Temperature Bins on Log Manufacturing Real Wage, Relative to a Day in the 70-74 Farenheit Bin

Note: The figures report 7 coefficient estimates (circle markers), representing the effect of a single day in each of the corresponding 7 temperature bins, relative to the effect of a day in the 70-74 °F bin, on district-level log agricultural yield (Figures 4(a), and 4(b)), log agricultural real wage (Figure 4(c)), state-level log manufacturing productivity (Figure 4(d) and 4(e)), and the state-level log manufacturing real wage (Figure 4(f)). Daily average temperature defined as the simple average of daily minimum and maximum temperature. Manufacturing productivity defined as state level GDP in registered manufacturing sector. The regressions also control for rainfall (upper/lower terciles indicators), district (or state) fixed effects, year fixed effects, and climatic region specific time trends. Growing and non-growing season coefficients are estimated in the same regression that is weighted by district population. Dashed lines represent the 95% confidence interval of the estimates. See the text for more detail.

Figure 5: Predicted Impact of Climate Change on Indian and US Life Expectancy at Birth, Based on Error-Corrected Hadley 3 A1FI Model: 2015-2099



Notes: The entries in Figure 5 correspond to the predicted impact of climate change on life expectancy in India and the U.S. The estimates are obtained by combining climate change prediction from the error-corrected Hadley 3 A1FI model (as reported in Figure 2) with regression estimates of the temperature-mortality relationship for India and the U.S. based on historical 1957-2000 data (as reported in Figure 1). Baseline life expectancy is taken from year 2000 life tables for rural India, urban India, and for the U.S. Counterfactual life expectancy associated with climate change is obtained by predicting the change in mortality rate due to changes in temperature and precipitation using to the estimated temperature-mortality relationships based on the historical data. For example, the entries for rural India for the period 2015-2029 represent the difference in life expectancy for an individual born in 2000 and exposed to the 1957-2000 climate in rural India compared to the same individual born in 2000 but exposed to the 2015-2029 predicted climate in rural India. Standard errors are clustered at the district level (India) and state level (U.S.). The gray lines represent the 95% confidence intervals. See the text for more detail.

TABLES

Table 1: Summary Statistics for the India and United States Samples

	Rural India			Urban India			United States		
	1957-1971	1972-1986	1987-2000	1957-1971	1972-1986	1987-2000	1957-1971	1972-1986	1987-2000
Total Mortality Rate Per 1,000 Population	10.26	6.42	4.17	10.96	7.62	6.07	9.55	8.81	8.64
	(6.51)	(4.36)	(2.54)	(5.10)	(3.78)	(2.35)	(0.95)	(1.00)	(1.15)
Agricultural Yield Index (kg/hectare)	40.72	54.57	81.23	---	---	---	---	---	---
	(19.60)	(28.83)	(41.78)						
Agricultural Real Wage (Rs/day)	7.17	8.77	13.24	---	---	---	---	---	---
	(3.75)	(4.50)	(5.56)						
Bank Branches Per 1,000 Population	0.02	0.23	0.46	---	---	---	---	---	---
	(0.04)	(0.18)	(0.16)						
State-Level Registered Manufacturing Output (100000 Rs / annun)	---	---	---	19,957	107,071	753,295	---	---	---
				(18,992)	(108,724)	(774,267)			
				1,958.5	6,286.7	12,970.2	---	---	---
State-Level Reg. Manuf. Earnings Per Worker (Rs / annun)	---	---	---	(595.2)	(2,782.4)	(3,156.8)			
Number of Days with Temperature in 75-89 °F	225.34	226.30	228.12	240.35	235.10	239.37	55.00	57.88	63.00
	(57.22)	(56.36)	(57.39)	(59.40)	(58.44)	(59.98)	(41.34)	(45.41)	(47.48)
Number of Days with Temperature > 90 °F	23.34	26.24	24.52	20.79	24.76	23.25	0.54	0.92	1.50
	(19.72)	(21.46)	(21.57)	(19.96)	(21.88)	(22.40)	(2.77)	(4.17)	(6.34)
Annual Total Precipitation (in)	47.21	45.04	45.27	47.19	45.69	46.79	36.69	39.93	38.70
	(25.81)	(24.66)	(25.96)	(27.93)	(26.48)	(28.08)	(11.82)	(13.21)	(13.81)

Notes: The entries are sample averages and standard deviations (parenthesis). Summary statistics from India are computed from district-level data (except state-level manufacturing data) and are weighted by the relevant population or geographical area. Summary statistics for the United States are from state-level data and weighted by the relevant population.

Table 2: Impact of Daily Temperature on All-Age Mortality in Rural and Urban India.

	Rural		Urban	
	Temperature		Temperature	
	75-89°F	**>90°F**	**75-89°F**	**>90°F**
	(1a)	(1b)	(2a)	(2b)
A. Exposure over Calendar Year				
Impact of Temperature	0.0021**	0.0047***	0.0001	0.0012
(std error)	(0.0007)	(0.0010)	(0.0005)	(0.0008)
B. Exposure over Agricultural Calendar				
Impact of Growing Season Temperature	0.0028***	0.0067***	0.0005	0.0016
(std error)	(0.0008)	(0.0016)	(0.0006)	(0.0013)
Impact of Non-Growing Season Temperature	-0.0047*	-0.0024	-0.0032	-0.0020
(std error)	(0.0023)	(0.0023)	(0.0017)	(0.0016)
Test of equality across agricultural seasons	0.002	0.001	0.058	0.060

Notes: The coefficient estimates correspond to the effect of a single day with daily temperatures in the 75-89 °F and >90 °F ranges on annual all-age log mortality rate, relative to days with daily temperatures <75 °F. The regressions also control for rainfall (upper/lower tercile indicators), district fixed effects, year fixed effects, and climatic region specific time trends. Regressions are estimated separately by rural and urban sectors, and weighted by district population. Number of observations: 11,399 (Rural) and 11,450 (Urban), corresponding to an unbalanced sample of 330 districts over 1957-2000. Asterisks denote p-value < 0.05 (*), <0.01 (**), <0.001 (***).

Table 3: Dynamic Analysis of the Impact of Daily Temperature on All-Age Mortality in Rural and Urban India.

	Rural		Urban	
	Temperature		Temperature	
	75-89°F	**>90°F**	**75-89°F**	**>90°F**
	(1a)	(1b)	(2a)	(2b)
A. Contemporaneous Model (no lags)				
Impact	0.0021***	0.0047***	0.0001	0.0012
(std error)	(0.0007)	(0.0010)	(0.0005)	(0.0008)
[N]	[11,399]	[11,399]	[11,450]	[11,450]
Impact (sample = 3 lags model)	0.0006	0.0043***	0.0003	0.0013
(std error)	(0.0008)	(0.0011)	(0.0005)	(0.0009)
[N]	[8,671]	[8,671]	[8,706]	[8,706]
B. Lag Model (3 lags)				
Contemporaneous impact	0.0002	0.0037***	0.0004	0.0015
(std error)	(0.0006)	(0.0009)	(0.0004)	(0.0008)
Lagged impact	0.0036*	0.0042	-0.0004	0.0012
(std error)	(0.0017)	(0.0026)	(0.0009)	(0.0016)
Total impact	0.0039	0.0079*	0.0000	0.0027
(std error)	(0.0021)	(0.0031)	(0.0011)	(0.0021)
[N]	[8,671]	[8,671]	[8,706]	[8,706]

Notes: The coefficient estimates correspond to the effect of a single day with daily temperatures in the 75-89 °F and >90 °F ranges on annual all-age log mortality rate, relative to days with daily temperatures <75 °F. The regressions also control for rainfall (upper/lower tercile indicators), district fixed effects, year fixed effects, and climatic region specific time trends. Regressions are estimated separately by rural and urban sectors, and weighted by district population. The Contemporaneous model refers to the baseline regression that only includes temperature variables for the current year. The Lag model coefficients are cumulative (summed) effects from regressions that includes temperature variables for the current year and from the 3 prior years. The Lagged model Asterisks denote p-value < 0.05 (*), <0.01 (**), <0.001 (***). See the text for more details.

Table 4: Robustness and Heterogeneity of the Impact of Daily Temperature on All-Age Mortality in Rural India.

	Temperature	
	75-89°F	**>90°F**
	(1a)	(1b)
A. Robustness		
Baseline estimate	0.0021***	0.0047***
(std error)	(0.0007)	(0.0010)
[std error clustered on state]	[0.0010]	[0.0020]
Without region-specific time trends	0.0023**	0.0059***
	(0.0008)	(0.0012)
With region-year fixed effects	0.0033***	0.0061***
	(0.0007)	(0.0012)
B. Heterogeneity		
Estimate for years 1957-1971	0.0020**	0.0060***
	(0.0008)	(0.0016)
Estimate for years 1972-1986	0.0020**	0.0058***
	(0.0007)	(0.0012)
Estimate for years 1987-2000	0.0019*	0.0030***
	(0.0007)	(0.0012)

Notes: The coefficient estimates correspond to the effect of a single day with daily temperatures in the 75-89 °F and >90 °F ranges on annual all-age log mortality rate, relative to days with daily temperatures <75 °F. The regressions also control for rainfall (upper/lower tercile indicators), district fixed effects, year fixed effects, and climatic region specific time trends. Regressions are weighted by district population. Number of observations: 11,399, corresponding an unbalanced sample of 330 districts over 1957-2000. Asterisks denote p-value < 0.05 (*), <0.01 (**), <0.001 (***). See the text for more details.

Table 5: Instrumental Variables Estimates of Impact of Daily Temperature on Mortality in Rural India.

	Temperature	
	75-89°F	**>90°F**
	(1a)	(1b)
A. Exposure over Calendar Year		
Impact of Temperature	0.0039**	0.0095***
(std error)	(0.0012)	(0.0021)
B. Exposure over Agricultural Calendar		
Impact of Growing Season Temperature	0.0040*	0.0143***
(std error)	(0.0016)	(0.0035)
Impact of Non-Growing Season Temperature	0.0031	0.0076
(std error)	(0.0044)	(0.0047)

Notes: The coefficient estimates correspond to the effect of a single day with daily temperatures in the 75-89 °F and >90 °F ranges on the listed outcome, relative to days with daily temperatures <75 °F. The regressions also control for rainfall (upper/lower tercile indicators), district fixed effects, year fixed effects, and climatic region specific time trends. The coefficients are from instrumental variables regressions where the IMD temperature variables are instrumented with the NCC temperature variables. Number of observations: 11,399, corresponding to an unbalanced sample of 330 districts over 1957-2000. Asterisks denote p-value < 0.05 (*), <0.01 (**), <0.001 (***). See the text for more details.

Table 6: Effect of Rural Banking Sector Expansion on the Temperature-Mortality Relationship in Rural India

	Panel Estimates, Rural		2SLS, Rural		2SLS, Urban	
	Temperature		Temperature			
	75-89 °F	>90 °F	75-89 °F	>90 °F	75-89 °F	>90 °F
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)
Dependent Variable is Log Annual Mortality Rate						
Main Effect of Temperature	0.0024***	0.0067***	0.0026***	0.0121***	0.0000	0.0017
(std error)	(0.0009)	(0.0012)	(0.0010)	(0.0019)	(0.0007)	(0.0021)
Main Effect of Number of Bank Branches			1.0633		0.1262	
(std error)			(0.7407)		(0.1143)	
Temperature x Bank Branches			-0.0041*	-0.0216***	-0.0010	-0.0013
(std error)			(0.0021)	(0.0055)	(0.0006)	(0.0014)
First-stage F-statistics						
Bank branches			6.28		7.71	
Bank branches x (75-89 °F)			12.28		11.70	
Bank branches x (>90 F)			15.13		9.26	

Notes: The coefficient estimates are from the baseline regression for log mortality rate specification, augmented with a main effect in the number of rural bank branches, and its interaction with the number of days with temperatures in the 75-89 °F and >90 °F ranges. Following the approach in Burgess and Pande (2005), we instrument for the number of bank branches with the number of bank branches in 1961 interacted with separate post 1976 and post 1989 time trends. The temperature × interactions are instrumented in a similar way, that is, with interactions between the temperature variables and the instruments for the number of bank branches. The regressions also control for rainfall (upper/lower tercile indicators), district fixed effects, and year fixed effects. Regressions are weighted by district population. Number of observations: 9,054, corresponding to an unbalanced sample of 330 districts over 1961-2000. Asterisks denote p-value < 0.05 (*), <0.01 (**), <0.001 (***). First-stage F-statistics are the Sanderson-Windmeijer (2015) F statistics for multiple endogenous regressors. The corresponding Stock and Yogo (2005) critical value for 10% maximal IV relative bias is 7.77. See the text for more details.

Table 7: Climate Change Impact on Life Expectancy in India and the United States

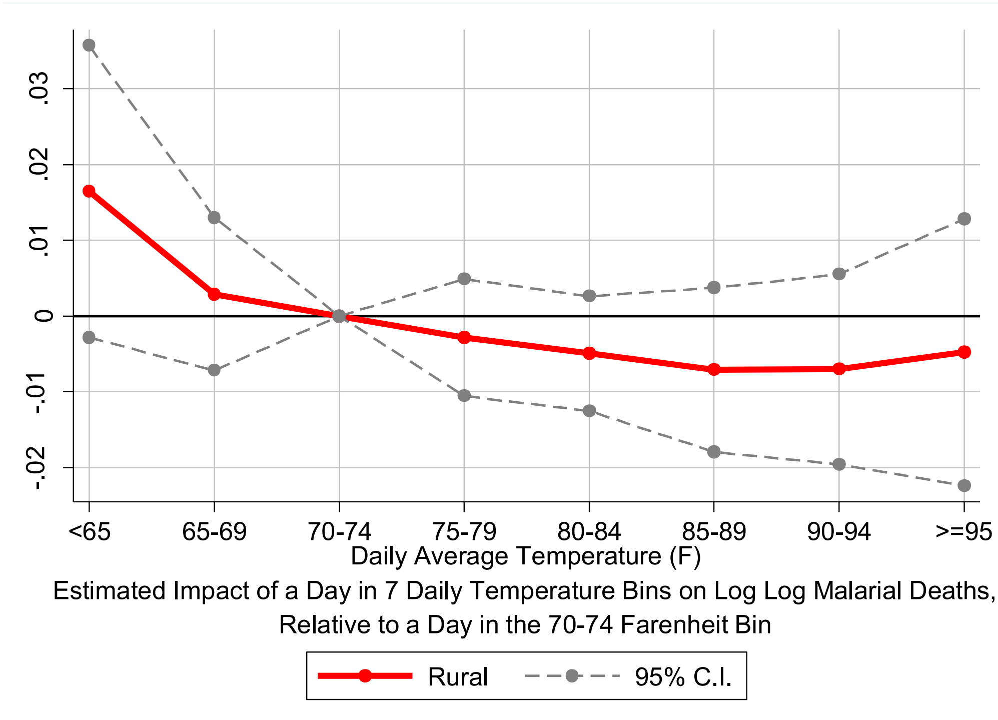
	Rural India			Urban India			All India		
	2015-29	2045-59	2075-99	2015-29	2045-59	2075-99	2015-29	2045-59	2075-99
Predicted Impact on Life Expectancy in India									
Baseline (Coefficients for 1957-2000)	**-2.37**	**-5.09**	**-10.41**	**-0.35**	**-1.05**	**-2.82**	**-1.88**	**-4.12**	**-8.59**
	(0.75)	**(1.22)**	**(2.14)**	**(0.36)**	**(0.61)**	**(1.27)**	**(0.65)**	**(1.08)**	**(1.93)**
Based on 1987-2000 coefficients	-1.14	-2.75	-6.17	0.18	0.17	-0.34	-0.82	-2.05	-4.77
	(0.68)	(1.35)	(3.02)	(0.38)	(0.76)	(1.91)	(0.61)	(1.21)	(2.76)
Based on "hotter" district coefficients	-2.78	-5.71	-10.87	-0.03	-0.67	-2.25	-2.12	-4.50	-8.80
	(0.70)	(1.13)	(2.00)	(0.41)	(0.68)	(1.34)	(0.63)	(1.02)	(1.84)
Based on 2000 urbanization (32%)	-2.37	-5.09	-10.41	-0.35	-1.05	-2.82	-1.72	-3.79	-7.98
	(0.75)	(1.22)	(2.14)	(0.36)	(0.61)	(1.27)	(0.62)	(1.03)	(1.86)
Asssuming 50% urbanization	-2.37	-5.09	-10.41	-0.35	-1.05	-2.82	-1.36	-3.07	-6.61
	(0.75)	(1.22)	(2.14)	(0.36)	(0.61)	(1.27)	(0.55)	(0.92)	(1.70)
Predicted Impact on Log Output	**-0.047**	**-0.072**	**-0.155**	**0.112**	**0.239**	**0.428**	**---**	**---**	**---**
	(0.020)	**(0.033)**	**(0.072)**	**(0.097)**	**(0.196)**	**(0.387)**			
Predicted Impact on Log Real Wage	**-0.068**	**-0.145**	**-0.308**	**0.050**	**0.090**	**0.175**	**---**	**---**	**---**
	(0.019)	**(0.036)**	**(0.081)**	**(0.059)**	**(0.134)**	**(0.341)**			

Predicted Impact on Life Expectancy in the U.S.	**2015-29**	**2045-59**	**2075-99**
Baseline (Coefficients for 1957-2000)	**-0.06**	**-0.15**	**-0.28**
	(0.01)	**(0.02)**	**(0.05)**

Notes: The entries in Table 7 correspond to the predicted impact of climate change on life expectancy, output, and real wages in India and life expectancy in the U.S. The life expectancy estimates are obtained by combining climate change prediction from the error-corrected Hadley 3 A1FI model (as reported in Figure 2) with regression estimates of the temperature-mortality relationship for India and the U.S. based on historical 1957-2000 data (as reported in Figure 1). Baseline life expectancy is taken from year 2000 life tables for rural India, urban India, and for the U.S. Counterfactual life expectancy associated with climate change is obtained by predicting the change in mortality rate due to changes in temperature and precipitation using to the estimated temperature-mortality relationships based on the historical data. For example, the entries for rural India for the period 2015-2029 represent the difference in life expectancy for an individual born in 2000 and exposed to the 1957-2000 climate in rural India compared to the same individual born in 2000 but exposed to the 2015-2029 predicted climate in rural India. The output and real wage estimates are obtained by applying the predicted temperature and precipitation distributions for each future time period (2015-29, 2045-59, 2075-99) to the regression estimates from Figure 4(a), 4c, 4d, and 4(f). Standard errors are clustered at the district level (India) and state level (U.S.). See the text for more detail.

Figure A.1: Impact of Daily Temperature on Malaria Incidence in Rural India

(a) Log Malarial Deaths



(b) Log Share of Positive Tests for Plasmodium Falciparum

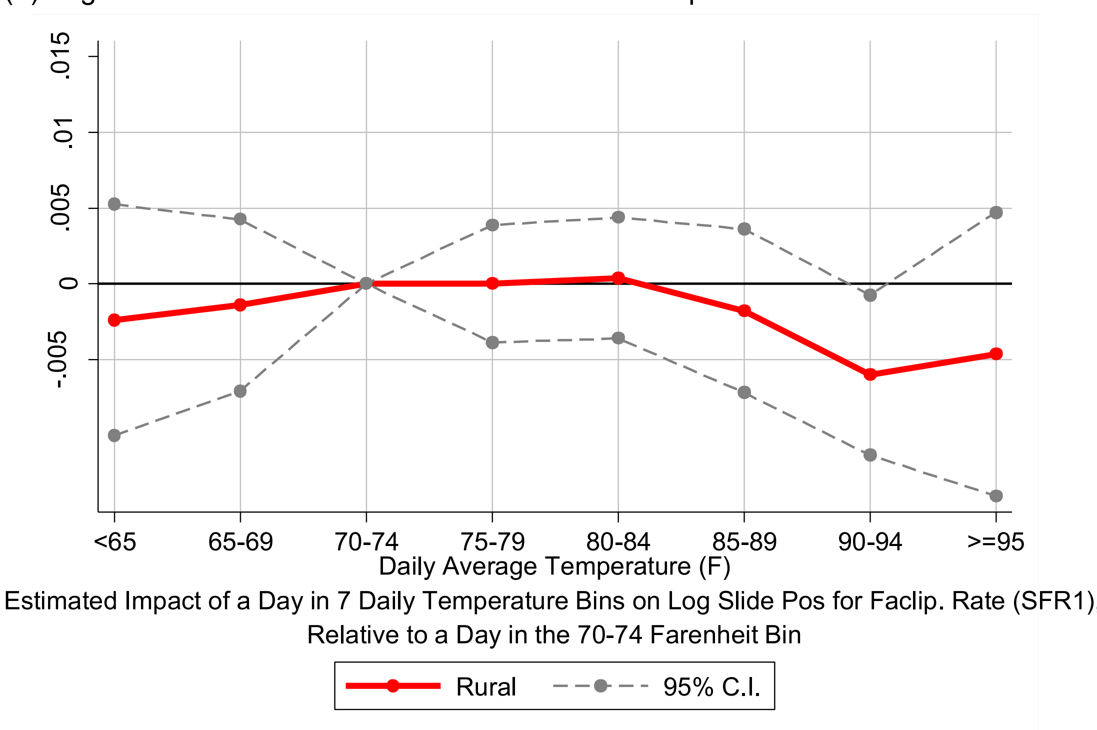
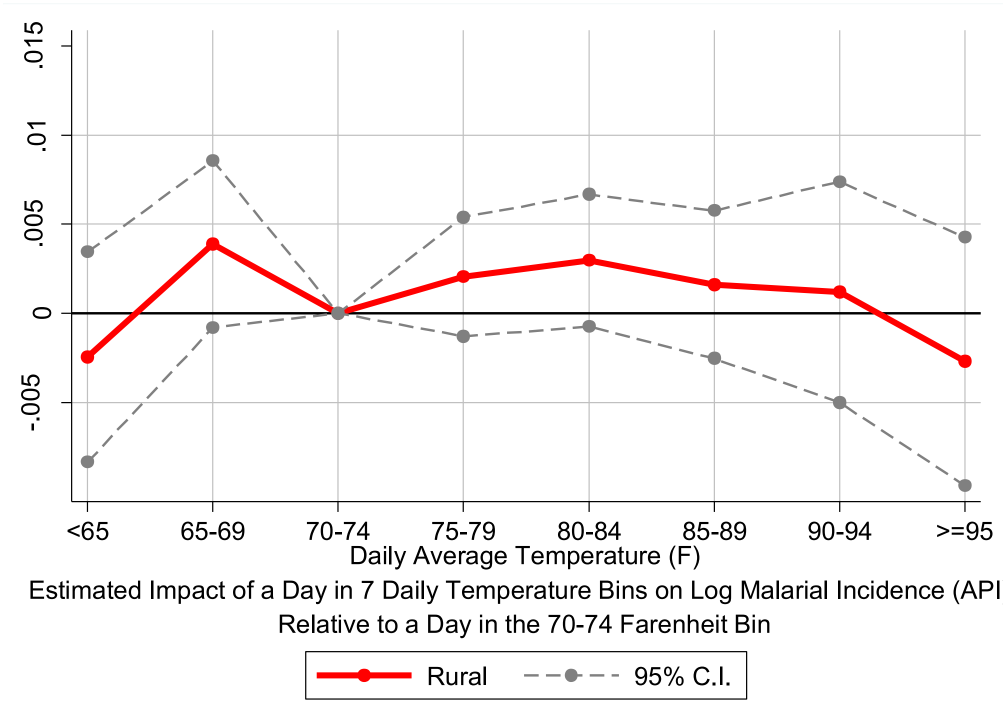


Figure A.1: Impact of Daily Temperature on Malaria Incidence in Rural India (ctd)

(c) Log Malarial Incidence (API)

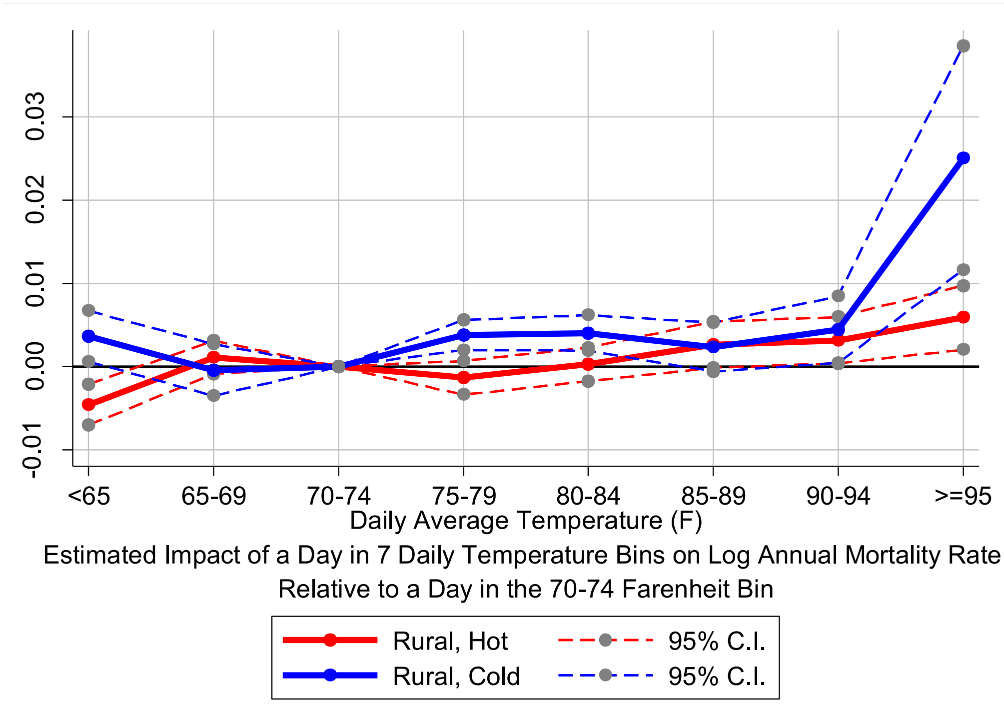


Estimated Impact of a Day in 7 Daily Temperature Bins on Log Malarial Incidence (API), Relative to a Day in the 70-74 Farenheit Bin

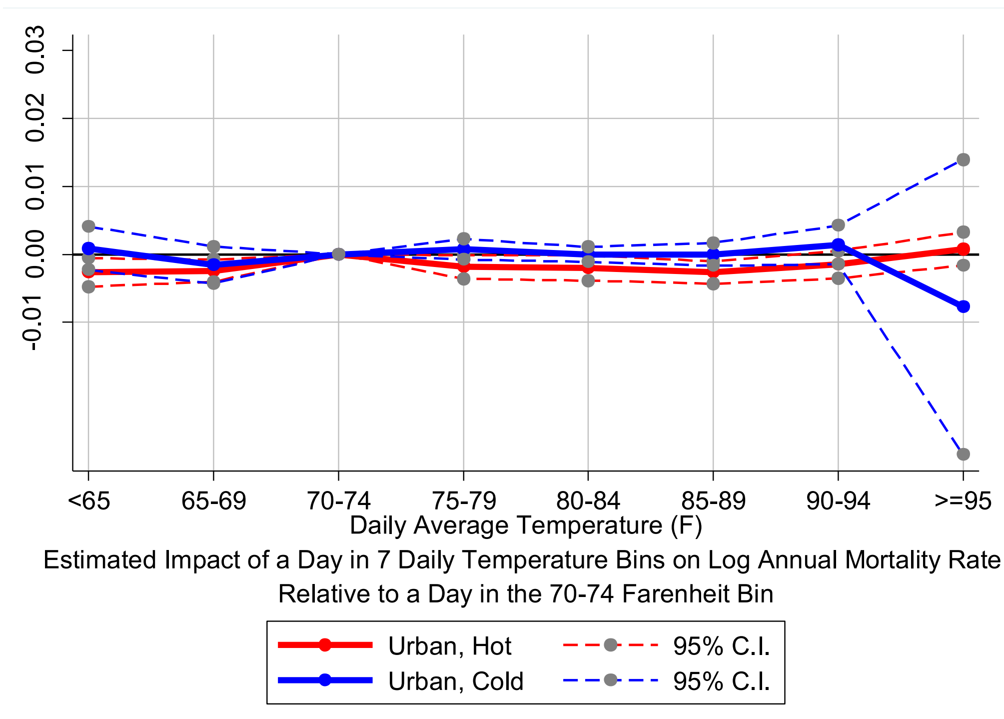
Note: The plotted lines report 7 coefficient estimates (circle markers), representing the effect of a single day in each of the corresponding 7 temperature bins, relative to the effect of a day in the 70-74 °F reference bin, on the listed indicators of malaria incidence in rural India. The regressions also control for rainfall (upper/lower terciles indicators), district fixed effects, year fixed effects, and climatic region specific time trends. Regressions are weighted by district population. Dashed lines represent the 95% confidence interval of the estimates. See the text for more details.

Figure A.2: Impact of Daily Temperature on Log All-Age Mortality Rate in India, by Hotter/Colder District

(a) Rural



(b) Urban



Note: The plotted lines report 7 coefficient estimates (circle markers), representing the effect of a single day in each of the corresponding 7 temperature bins, relative to the effect of a day in the 70-74 °F reference bin, on log annual mortality rate. The regressions are estimated separately by rural and urban districts. “Cold” and “hot” districts defined as districts with long-term average number of days with temperature > 90 °F above (hot) or below (cold) the national average among rural districts (25). The regressions also control for rainfall (upper/lower terciles indicators), district fixed effects, year fixed effects, and climatic region specific time trends. Regressions are weighted by district population. Dashed lines represent the 95% confidence interval of the estimates. See the text for more details.

Table A.1 Coefficient Estimates from Temperature-Mortality Regressions

	All India	United States	Rural India	Urban India
	(1)	(2)	(3)	(4)
Dependent Variable is Log Mortality Rate				
A. Coefficients on Temperature				
<65°F	-0.0008	0.0001***	-0.0008	-0.0010
	(0.0010)	(0.0000)	(0.0011)	(0.0011)
65-69°F	0.0001	-0.0001***	0.0010	-0.0017*
	(0.0008)	(0.0000)	(0.0009)	(0.0009)
70-74°F	0	0	0	0
	---	---	---	---
75-79°F	0.0016**	0.0000	0.0022**	-0.0001
	(0.0006)	(0.0000)	(0.0007)	(0.0006)
80-84°F	0.0021**	0.0001***	0.0028***	-0.0006
	(0.0007)	(0.0000)	(0.0008)	(0.0005)
85-89°F	0.0022*	0.0002**	0.0028*	-0.0010
	(0.0009)	(0.0001)	(0.0012)	(0.0006)
90-94°F	0.0038***	0.0005***	0.0045***	-0.0001
	(0.0010)	(0.0001)	(0.0013)	(0.0008)
>95°F	0.0074***	0.0003***	0.0084***	0.0021
	(0.0016)	(0.0000)	(0.0019)	(0.0011)
B. Coefficients on Rainfall Terciles				
Lowest Tercile	0.0157	0.0002**	0.0164	0.0005
	(0.0127)	(0.0001)	(0.0156)	(0.0094)
Highest Tercile	0.0045	-0.0001	0.0014	0.0015
	(0.0125)	(0.0001)	(0.0153)	(0.0104)
Observations	22,849	25,872	11,399	11,450

Notes: The table reports the full set of coefficient estimates underlying Figure 1a and 1b correspond to the effect of a single day in a given temperature bin on annual all-age log mortality rate, relative to a day with daily temperatures in 70-74 °F. The Indian regressions control for rainfall (upper/lower tercile indicators), district fixed effects, year fixed effects, climatic region specific time trends, and are estimated separately by rural and urban sectors. The U.S. regressions control for rainfall (upper/lower tercile indicators), state fixed effects, year-month fixed effects, state-month quadratic time trends. Regressions are weighted by district population (weighted by state population for the U.S. regressions). Asterisks denote p-value < 0.05 (*), <0.01 (**), <0.001 (***).

Table A.2 Estimates from Model with Temperature × Rainfall Tercile Interactions, Rural and Urban India

	Rural India		Urban India	
	(1)	(2)	(3)	(4)
Dependent Variable is Log Mortality Rate				
A. Coefficients on Temperature				
75-89°F	0.0021**	0.0021**	0.0001	0.0002
	(0.0007)	(0.0007)	(0.0005)	(0.0005)
>90°F	0.0047***	0.0045***	0.0012	0.0013
	(0.0010)	(0.0011)	(0.0008)	(0.0008)
B. Coefficients on Rainfall Terciles				
Lowest Tercile	0.0176	-0.0227	-0.0018	-0.0469
	(0.0148)	(0.0651)	(0.0089)	(0.0430)
Highest Tercile	-0.0007	0.0087	0.0041	0.0628
	(0.0151)	(0.0594)	(0.0105)	(0.0430)
C. Interaction Coefficients				
75-89°F × Lowest Tercile	---	0.0001	---	0.0002
		0.0002		0.0001
>90°F × Lowest Tercile	---	0.0006	---	0.0001
		0.0007		0.0004
75-89°F × Highest Tercile	---	-0.0001	---	-0.0003
		0.0002		0.0001
>90°F × Highest Tercile	---	0.0002	---	0.0002
		0.0007		0.0005

Notes: The coefficient estimates correspond to the effect of a single day with daily temperatures in the 75-89 °F and >90 °F ranges on annual all-age log mortality rate, relative to days with daily temperatures <75 °F. The regressions also control for rainfall (upper/lower tercile indicators), rainfall indicators interacted with temperature, district fixed effects, year fixed effects, and climatic region specific time trends. Regressions are estimated separately by rural and urban sectors, and weighted by district population. Number of observations: 11,399 (Rural) and 11,450 (Urban), corresponding to an unbalanced sample of 330 districts. Asterisks denote p-value < 0.05 (*), <0.01 (**), <0.001 (***).

Table A.3 Coefficient Estimates from Temperature-Output Regressions

	Log Agricultural Yield Index	Log Manufacturing Output
	(1)	(2)
A. Coefficients on Temperature		
<65°F	0.0009	-0.0032
	(0.0005)	(0.0028)
65-69°F	0.0009*	0.0060
	(0.0004)	(0.0039)
70-74°F	0	0
	---	---
75-79°F	-0.0012**	0.0004
	(0.0004)	(0.0026)
80-84°F	-0.0030***	0.0009
	(0.0004)	(0.0026)
85-89°F	-0.0041***	-0.0031
	(0.0008)	(0.0036)
90-94°F	-0.0029***	0.0077
	(0.0007)	(0.0058)
>95°F	-0.0012	0.0058
	(0.0008)	(0.0066)
B. Coefficients on Rainfall Terciles		
Lowest Tercile	-0.0814***	0.1486
	(0.0086)	(0.1044)
Highest Tercile	-0.0012	0.0373
	(0.0060)	(0.0859)
Observations	12,619	584

Notes: The table reports the full set of coefficient estimates underlying Figure 3a and 3d correspond to the effect of a single day in a given temperature bin on log agricultural yield index (measured at the district level) and log manufacturing output (measured at the state level) temperatures in 70-74 °F. The regressions control for rainfall (upper/lower tercile indicators), district (column 1) or state (column 2) fixed effects, year fixed effects, climatic region specific time trends, and are estimated separately by rural and urban sectors. Asterisks denote p-value < 0.05 (*), <0.01 (**), <0.001 (***).