The Impact of Temperature on Productivity and Labor Supply: Evidence from Indian Manufacturing*

E. Somanathan, Rohini Somanathan, Anant Sudarshan[†] and Meenu Tewari

Hotter years are associated with lower economic output in country-level data. We show that the effect of temperature on labor is an important part of the explanation. Using high-frequency micro data from selected firms in India, we find that worker productivity falls on hot days. Workers exposed to high temperatures are also more likely to be absent. Using a nationally representative panel of manufacturing plants, we find that annual output falls by about 2 percent per degree Celsius. This response appears to be driven by a reduction in the output elasticity of labor. Estimated effect sizes for both the high-frequency case studies and the national panel are consistent with prior cross-country panel studies and large enough to account for the whole of the country-level response of GDP to high temperature.

Keywords: temperature, heat stress, worker productivity, climate change.

JEL: Q54, Q56, J22, J24

*__Acknowledgements:__ This work was made possible by funding from the Rockerfeller Foundation, ICRIER, and PPRU of the ISI. We are grateful to Mehul Patel for field assistance and to Sheekha Verma for research assistance. This paper has benefited from helpful comments from four anonymous referees, Michael Greenstone, and participants at several conferences and seminars.

†Corresponding Author.

1 Introduction

Recent research has uncovered a systematic negative correlation between temperature and aggregate national output, especially in tropical developing countries (Dell, Jones, and Olken, 2012; Burke, Hsiang, and Miguel, 2015). High temperatures are associated with reduced crop yields as well as lower output in non-agricultural sectors.[1] Explanations for this relationship include heat stress on workers and temperature-related increases in mortality, conflict, and natural disasters.[2] Establishing and quantifying the relative importance of these mechanisms is crucial for identifying possibilities of adapting to a hotter world. In this paper we focus on understanding the role of heat stress.

The physiological effects of temperature on human beings can be seen very quickly, unlike other mechanisms examined in the literature which operate over longer time scales. We assemble high-frequency, worker-level output data from three different manufacturing settings; cloth weaving, garment sewing, and steel products. We use this dataset to test whether we can detect heat-induced declines in output at a granular level.

There are two channels through which high temperatures might affect these factory workers. They may produce less while at work and also be absent more often. We separately identify both these effects. We find that the output of individual workers and worker teams declines on hot days as well as in weeks with more hot days. We also find that exposure to high temperatures increases absenteeism. Absenteeism is influenced both by contemporaneous temperatures as well as those experienced over the preceding week. Stronger effects are visible for paid leave, with a weaker temperature-absenteeism relationship for unpaid leave. Climate control in the workplace eliminates productivity declines but not absenteeism.

[1]For evidence on yields, see Mendelsohn and Dinar (1999), Auffhammer, Ramanathan, and Vincent (2006), Schlenker and Roberts (2009), Lobell, Schlenker, and Costa-Roberts (2011), and Gupta, Somanathan, and Dey (2017).

[2]Hsiang (2010) discusses heat stress, Hsiang, Burke, and Miguel (2013) identify a temperature-conflict relationship and Burgess et al. (2017) study effects on mortality.

To examine whether the temperature effects for workers in these firms are more generally reflected in India's factory sector, we use a 15-year nationally representative panel of manufacturing plants. We find that the value of plant output declines in years with more hot days. Annual output is predicted to fall by 2.1 (± 0.8) percent if every day warms by 1^o C. We use a Cobb-Douglas specification to show that temperature-induced reductions in the output elasticity of labor appear to drive this response. We also use manufacturing-sector GDP for Indian districts for the period between 1998 and 2009 to directly estimate the impact of temperature on district output. We find that district annual output declines by about 3 percent per degree Celsius, consistent with the magnitude of the plant response.

To situate these findings within the context of the country-level relationships that motivate this paper, it is helpful to compare the temperature-output relationship estimated at several different levels of aggregation. Putting together our results from worker, plant, and district data, we find that effect sizes in all three cases are similar. Strikingly, these effects are large enough to explain most of the country-level response to temperature observed in the literature, suggesting that heat stress on labor may be a much more important mechanism than previously believed.

Notwithstanding the importance of these temperature effects, adaptation through climate control does not always occur. For example, the cloth-weaving firms we study are labor-intensive but do not use climate control. Given the costs of electricity, value added per worker may be too low to justify these investments. In the garment firms, value addition by workers is greater and we see more climate control. In our national plant panel, we find that temperature effects on output fall over time, possibly the result of investments in adaptation.

If heat stress plays an important role in reducing output, then firms that do make costly climate control investments should strategically allocate these resources towards protecting those activities that are labor-intensive and add significant value. We surveyed the management of 150 plants in the diamond processing industry to test these hypotheses. We find that air-conditioning is selectively used in rooms containing activities that are both labor-intensive and critical in determining diamond quality.

Taken together, this evidence suggests that the effect of heat on human beings is an economically important phenomenon. After presenting our main results, we consider some important alternatives to the heat stress channel, including natural disasters, power outages, and conflict. For the years covered by our plant panel, we collect data on instances of flooding, power shortages, and workdays lost in all recorded industrial disputes. We find that these variables have no additional explanatory power when incorporated in our empirical models.

The paper is organized as follows. Section 2 summarizes the physiological evidence on heat stress. Section 3 describes our data sources. Our main results are in Section 4. In Section 5 we compare effect sizes from our worker, plant, and district level data and show that these are of similar magnitude and are also consistent with country-level estimates in the literature. Section 6 examines the adoption of climate control investments within firms. Section 7 discusses alternative explanations and the robustness of our main results. Section 8 concludes.

2 Prior Literature

The science of how temperature affects human beings is straightforward. Heat generated while working must be dissipated to maintain body temperatures and avoid heat stress. If body temperatures cannot be maintained at a given activity level, it becomes necessary to reduce the intensity of work (Kjellstrom, Holmer, and Lemke, 2009; Iso, 1989). The efficiency of this process depends primarily on ambient temperature but is also influenced by humidity and wind speed (Parsons, 1993; Iso, 1989). Laboratory studies often use an adjusted measure of heat that accounts for these factors - the wet bulb temperature (WBT) (Lemke and

Kjellstrom, 2012). Unfortunately, outside the lab, data on humidity is often unavailable. For this reason, and to assist comparisons with prior work, we use daily maximum temperatures as our measure of heat throughout this paper.

Exposure to high ambient temperatures can reduce physical productivity and also affect our willingness and ability to go to work. There have been a number of studies on temperature and productivity. Mackworth (1946) conducted an early artefactual field experiment with wireless telegraph operators and found that they made more mistakes at high temperatures. Parsons (1993) and Seppanen, Fisk, and Faulkner (2003) summarize important findings in this area. Hsiang (2010) presents a meta analysis of recent laboratory evidence which shows that once wet bulb temperatures rise above 25 degrees celsius, task efficiency appears to fall by approximately 1 to 2 percent per degree. A WBT of 25 degrees at 65 percent relative humidity is roughly equivalent to a temperature of 31 degrees celsius in dry conditions.[3] These temperatures are not considered unsafe from the point of view of occupational safety and commonly occur in many developing countries.[4]

Controlled experiments in the laboratory or workplace provide a useful benchmark but do not fully capture real manufacturing environments. Workers and management generally operate well within physical limits and have room to increase effort in response to incentives. The output-temperature relationship therefore depends on the physical as well as behavioral aspects of employment such as the wage contract, particularities of production, management techniques, and mechanization. This makes data from non-experimental settings particularly valuable. As early as 1915, Huntington exploited daily variations in temperatures experienced by workers and students performing various tasks and found that high temperatures appeared to reduce output (Huntington, 1915).[5] More recently, Adhvaryu, Kala, and

[3] The WBT scale is also compressed relative to temperature, so a one degree change in WBT corresponds to a higher than one degree change in temperature.

[4] Temperature exposure in sectors such as mining can be high enough to create serious health hazards. These settings have long been used for research on heat stress and occupational safety (Wyndham, 1969).

[5] We are grateful to an anonymous reviewer for pointing us to some of this literature.

Nyshadham (2019) exploit variation in workplace temperatures induced by low-heat LED lighting and conclude that worker productivity increases when temperatures are reduced.

On absenteeism, there is much less evidence. Zivin and Neidell (2014) study time allocated between outdoor and indoor activities in response to extreme temperatures in the United States. Their unit of analysis is the individual rather than the plant, so they do not estimate the effect of these changes on labor supply within firms. Our data allows us to go further as we are able to directly estimate temperature-related absences within firms.

3 Data Sources

Our labor and output data are at three levels of aggregation: the worker or worker-team, the plant, and the district. For each data set, described below, we match output to measures of temperature. In addition, we use a survey of diamond firms to study the selective use of climate control. Official data in India is typically available for financial years, which run from April 01 through March 31. When referring to a financial year, we use the initial calendar year.

3.1 Worker Data

We collected worker output and attendance data from selected firms in three industries: cloth weaving, garment sewing, and the production of large infrastructural steel products. Figure A.1 in the Appendix has photographs from a typical production line in each of these industries. Our three cloth-weaving factories are all located in the industrial city of Surat in the state of Gujarat, in western India. Our garment factories are managed by a single firm, with six plants located in the National Capital Region (NCR) in North India, and two others in the cities of Hyderabad and Chhindwara in Central India. Our steel production data are from the rail and structural mill of a large public sector steel plant in the town of Bhilai in central India. Each of these micro-data sites is part of an important manufacturing sector in the Indian and global economy. The textile sector (which includes spinning, weaving, and dyeing) employs about 12 percent of factory workers in India. The garment sector employs about 7 percent of factory workers, and the Bhilai steel mill is the largest producer of steel rails in the world.[6]

For the three cloth-weaving factories, we gathered daily data on meters of cloth woven and attendance of 147 workers employed during the financial year starting April 2012. A worker in each of these factories operates about 6 mechanized looms producing woven cloth. Workers are engaged in monitoring looms, adjusting alignment, restarting feeds when interrupted, and making other necessary corrections. The cloth produced is sold in wholesale markets or to dyeing and printing firms. Workers are paid based on the meters of cloth woven by these looms and no payments are made for days absent. Protection from heat is limited to the use of windows and some fans. We obtained payment slips for each day and digitized these to generate a worker-level dataset of daily output and attendance. For most types of cloth, workers were paid 2 rupees per meter.[7]

For garment sewing, we have data from eight factories owned by a single firm producing garments for foreign apparel brands. Unlike in the cloth-weaving firms described above, these workers are paid monthly wages that do not directly penalize workers for small variations in productivity or occasional absences. In each plant, production is organized in sewing lines of 10-20 workers, with each line creating part or all of a clothing item. Lines are usually stable in their composition of workers, while the garment manufactured by a given line changes based on production orders. Our productivity measure relates to the entire sewing line. The

[6]For employment shares, see Annual Survey of Industries, 2009-10, Volume 1. A description of the steel plant at Bhilai is available from the Steel Authority of India Ltd. The steel rails from Bhilai are used for the entire network of public railroads in the country.

[7]Since payments are strictly based on production, incentive effects on output arising from non-linearities caused by minimum wages can be ignored (Zivin and Neidell, 2012).

garment sector is highly competitive and firms track worker output in sophisticated ways. In our case, the firm used an hourly production target for each line, based on the time taken to complete the desired garment by an experienced line of 'master craftsmen'. The actual hourly output, controlling for the target, provides a measure of the line productivity. The target is not revised each day so it is not sensitive to daily temperatures. The firm management provided us with daily production from 103 sewing lines for all 730 days over two calendar years, 2012 and 2013. They also gave us attendance records over the same time period for plants in the National Capital Region, allowing us to construct a daily count of absences within sewing lines in these factories.

These garment factories also provide us an opportunity to study the effects of climate control investments on productivity. During the period for which we have data, the firm was in the process of installing cooling equipment on its shopfloors. This installation of climate control had been completed in five of the manufacturing units in the capital region (NCR) before 2012 but the sixth unit did not get this until 2014. Two factories in Hyderabad and Chhindwara were also without climate control, but average temperatures in these areas are lower than in the NCR. This phased roll-out allows us to compare temperature effects in co-located factories with and without climate control.

The rail and structural mill in Bhilai is the primary supplier of rails to the Indian Railways and also produces steel products used for large infrastructural projects. Rectangular blocks of steel called *blooms* form the basic input for all these products. They enter a furnace and are then shaped into rails or structurals to meet ordered specifications.[8] When a bloom is successfully shaped, it is said to have been *rolled*. The number of blooms rolled in an eight-hour shift is our measure of output.

There are three shifts on most days, starting at 6 a.m., and workers are assigned to one of

[8]Structurals refer to a miscellaneous set of steel products used mostly in construction projects such as roads and bridges.

three teams which rotate across these shifts. The median number of workers on the factory floor is 66. Our production data records the team and the number of blooms rolled for each working shift during the period 1999-2008. We observe a total of 9172 shifts over 3339 working days. In addition to the team output in each shift, we also have team-level absences over a shorter period of 857 working days between February 2000 and March 2003.[9]

Unlike the weaving and garment units, the production of rails is highly mechanized and the mill runs continuously with breaks only for repair, maintenance, and adjustment for different products. Workers who manipulate the machinery used to shape rails sit in air-conditioned cabins. Others perform operations on the factory floor. This is the most capital-intensive of our case study sites with both automation and climate control.

3.2 Panel of Manufacturing Plants

We purchased secondary data from the Annual Survey of Industry (ASI) covering the financial years 1998-99 to 2012-13. The ASI is a Government of India census of large plants and a random sample of about one-fifth of smaller plants registered under the Indian Factories Act. Large plants are defined as those employing over 100 workers.[10] The ASI provides annual data on output, the value of fixed assets, debt, cash on hand, inventories, input expenditures, and the employment of workers and management. The format is similar to census data on manufacturing in many other countries.[11]

The ASI provides plant identifiers for the period 2000-2010 but not in other years. To create a longer panel requires matching observations across different years using time-invariant plant characteristics. Following a procedure similar to Allcott, Collard-Wexler, and O'Connell

[9]These data were first used by Das et al. (2013), who provide a detailed account of the production process in the mill.

[10]For regions with very little manufacturing, the ASI covers all plants irrespective of their size.

[11]See Berman, Somanathan, and Tan (2005) for a discussion on the measurement of variables in the ASI and its comparability with manufacturing data in other countries.

(2016), we create an unbalanced panel of 58377 plants over 1998 to 2012.[12] We match plants to temperature and rainfall at the level of the district.[13]

3.3 District Panel of Manufacturing GDP

The Planning Commission of India has published data on district-level manufacturing sector GDP over a 12-year period from 1998 to 2009. These figures include ASI plants as well as estimates from unregistered manufacturing and smaller factories not covered by the ASI. We use these statistics to directly estimate the effect of temperature on economic output, aggregated at the level of districts. Unfortunately, after 2009 this information has not been systematically compiled. Data for some districts was either not available in this dataset, or not reliable because of changes in boundaries over this period. Kumar and Somanathan (2009) provide a review of these boundary modifications. Therefore our estimates are based on a sub-sample of 438 districts with static boundaries and at least 2 non-missing observations over this period.

[12] Appendix Section A.4 provides details on panel construction.

[13] There are 529 districts with at least one plant in the data set. Figure A.4 in the Appendix shows the geographic distribution of ASI plants and locations of our micro-data sites.

Table 1: Summary of Worker and Firm Datasets

Data Source	Location	Unit (# of obs)	Dependent Variables	Time Period	Climate Control
Cloth Weaving Firms	Surat	Worker (147)	Meters of cloth woven, Worker Attendance	365 days	No
Garment Sewing Plants	NCR, Hyderabad, Chhindwara	Sewing Line (103)	Operations completed, Worker Attendance	730 days (varying by line)	Partial (74 lines)
Steel Mill	Bhilai	Shift-Team (9)	Blooms rolled, Team Absences	3337 days (Production)	Yes
Association of Diamond Firms	Surat	Plants × Operations (150×5)	Climate Control Indicator	Cross-section	Partial
Annual Survey of Industry	National	Plant (58,377)	Value of output	15 years	Not Observed
Planning Commission of India	National	District (438)	Manufacturing GDP	12 years	Not Observed

3.4 Weather Data

Our weather data come from two sources. We use recordings from public weather stations within the cities where our cloth-weaving and garment-sewing factories are located. We also use a $1^\circ \times 1^\circ$ gridded data product sold by the Indian Meteorological Department (IMD), which provides daily historical temperature and rainfall measurements interpolated over the IMD's network of monitoring stations across the country. The first of these provides a more precise measure for locations near a weather station. The second is best suited to averaging over larger areas.[14]

In the case of our worker data, we know the precise factory locations and can use data from nearby public weather stations wherever available. We characterize the temperature of a day using the daily maximum temperature, which occurs during working hours and is therefore a useful proxy for heat exposure at the workplace.

There were no public weather stations in the proximity of the Bhilai Steel Plant over the period for which we have data. For this plant, we instead rely on the IMD gridded dataset and use an inverse distance weighted average of grid points within 50 km of the plant to assign daily maximum temperature values.

For our annual panel of manufacturing plants we use daily maximum temperatures from the IMD gridded datasets as well as daily precipitation. Since we do not have precise location coordinates from the ASI, we assign to each plant the temperature and rainfall corresponding to the district in which it is situated. These numbers are obtained by spatially averaging grid temperatures over the geographical boundaries of each district. Additional details are in Appendix Section A.4.

[14] The physiology literature often uses wet bulb temperatures (WBT) to study heat stress. This measure combines temperature and humidity. We are not aware of a good source of time varying measures of wet bulb temperatures for the whole country. For this reason, and to ease comparison with previous work, we use maximum temperatures throughout the main paper.

When using the ASI data, in our main specification we aggregate daily temperatures up to the annual level using counts of the number of days in the year falling within different temperature bins. We use temperature bins defined as $\{(0, 20], (20, 25], (25, 30], (30, 35], (35, 50]\}$. To summarize the temperature distribution over the year, we construct a vector $\mathbf{T} = (T^1, T^2, T^3, T^4, T^5)$ with counts of the number of days in each of these bins. This is calculated for every district and each year. Taken together, these bins are non-overlapping and span the observed range of temperatures in the data, so that any given day is assigned to exactly one bin. We also estimate additional specifications using polynomials of of daily maximum temperatures over the year as well as a degree day measure. These are described in more detail in Section 4.2.

When using worker-level data we also use similar binned specifications. The cut-offs and width of these bins vary from those described above, reflecting differences in the realized distribution of weather in the two cases. Bin definitions for workers are discussed in more detail in Section4.1 and shown in Figure 1.

3.5 Climate Control within Diamond Firms

In August 2014, we surveyed 150 diamond-cutting plants, randomly sampled from over 500 units formally registered with the industry association of the city of Surat (the same location as our cloth weaving units). Each plant carries out five operations: (i) sorting and grading (ii) planning and marking (iii) bruting (rounding a diamond) (iv) cutting (v) polishing. Although these factories are small and labor intensive, like the cloth-weaving plants, the value added in production is considerable and these units commonly deploy air-conditioning in at least some parts of the plant.

We asked the management of each firm about the number of workers and machines and the use of air-conditioning in each of the five operations. They were also asked to rate, on a scale of 1-5, the importance of each of these processes to the quality of final output. We use these responses to study the selective deployment of climate control.

4 Results

4.1 Temperature Effects on Worker Output

Temperature may influence worker output through different channels. People may be more likely to miss work on very hot days. They may also be less productive at the workplace because of heat stress. Both contemporaneous and lagged temperatures may affect output.

We begin by estimating the effects of temperature on the aggregate weekly output of workers. This reflects the combined effects of absenteeism and reduced productivity at work. We then use daily data to separately examine the non-linear effects of contemporaneous and lagged temperatures, on both productivity and attendance.

Weekly output is related to temperature using the following binned specification:

$$y_{iw} = \alpha_i + \gamma_M + \gamma_t + \sum_{j=2}^{J} \beta_j T_{iw}^{j} + \theta R_{iw} + \lambda X_{iw} + \epsilon_{id} \tag{1}$$

Our output measure is in physical units in each of the three types of firms that we study. For cloth weaving, y_{iw} is the inverse hyperbolic sin transformation of the daily meters of cloth produced by worker i averaged over the course of week w. If a worker is absent, we set output for that day at zero. We use this transformation instead of logarithms since our output indicator can take zero values. For the steel mill, y_{iw} is the logarithm of the average number of rectangular blooms rolled in shift i during week w. As described in Section 3, a bloom is an intermediate steel product that is used in the manufacture of railway tracks. There are three shifts in the workday each manned by a different worker team. For garment plants, y_{iw} is the logarithm of the 'efficiency' of each sewing line (a team of workers). 'Efficiency' is a performance metric used by the garment firm and it is based on the number of operations completed every hour by the sewing line. We also control for a line-specific target efficiency that is set by the firm, as described in Section 3. We do this because different lines carry out operations of differing complexity over time and the target helps to control for this. Note that the target itself is not updated daily and is therefore independent of temperature.

We include a range of fixed effects to control for idiosyncratic worker productivity and temporal and seasonal shocks. Fixed-effects for the i^{th} unit are denoted by α_i. A unit is an individual worker in the cloth-weaving firms, a sewing line in garment firms, and a team-shift for the steel mill. As mentioned in Section 3, for the steel mill there are 3 shifts a day, and three teams of workers rotating across shifts, producing a total of 9 indicator variables.

Output is likely to respond to (possibly seasonal) demand, so we also include month and year fixed-effects (γ_M, γ_t). R_{iw} is a control for weekly rainfall, and X_{iw} are other controls- the number of working days in the week, and additionally for garment workers, the target efficiency. T^j is a count of the number of days in the reference week that fall in a given temperature bin j. We define a set of temperature bins of width two degrees Celsius as follows: (0,19], (19,21], (21,23], (23,25], (25,27], (27,29], (29,31], (31,33], (33,35], (35,50]. Taken together, these capture the non-linear relationship between output and temperature.

The temperature range we observe for each unit depends on its location. For units in the National Capital Region (NCR) around Delhi we can use all 10 temperature bins. For each of the other factory locations, we combine some of the lower temperature bins since these sites cover a smaller temperature range. To facilitate comparison across sites, the highest bin is pegged at maximum temperatures above 35 degrees Celsius. The cloth weaving workers in Surat face warmer temperatures so our first bin ends at 29 degrees Celsius. This produces 5 bins: (0,29], (29,31], (31,33], (33,35], and (35,50]. For the steel plant and the garment sewing lines outside the NCR, we start at 27 degrees Celsius. Because the sum of all bin counts is a constant, we omit the lowest bin in our regressions. The estimate of the coefficient of T^j should therefore be interpreted as the effect of a single day in the week moving from the lowest (coldest) temperature bin, T^1, to a warmer temperature range corresponding to bin j.

Figure 1 presents coefficient estimates β_j for all worker sites, with 90 percent confidence intervals.

In the absence of climate control, output falls in weeks with more hot days.[15] In climate controlled garment plants in the NCR (Panel A), we see no negative effects of temperature on output. For the steel mill which is largely automated and has climate control, if anything, output rises slightly at higher temperatures (Panel D). This might occur if climate control is turned on only on hot days, making workplace conditions on those days actually *more* comfortable. It is also possible that foundry operations are negatively affected by cold weather because metal may set too quickly, causing faults in the final output (Fiorese et al., 2015). We return to the question of interactions of capital equipment with temperature in Section 4.2.[16]

Our estimates are heterogeneous across workplace settings. For garment plants in the NCR without climate control, the effect of an additional day in a week moving from the lowest to highest temperature bin is to reduce average daily efficiency by as much as 8 percent. On the other hand, the analogous point estimate for weaving workers is as low as 2 percent. A similar response is also observed for garment plants in Hyderabad and Chhindwara. This is not surprising because the omitted bin is not the same across sites - in warmer regions the omitted bin spans higher temperatures than sites such as the NCR that observe colder

[15] Large shop-floors are not cooled by typical air-conditioning units. Thus when we refer to climate control we mean a plant that has a centralized cooling system such as an air-washer installed.

[16] High temperatures could directly reduce productivity if they are associated with power outages. All the factories in our dataset have a power backup, so this is not a concern. Also, if outages were driving our results, we should expect to see this effect in plants with and without climate control.

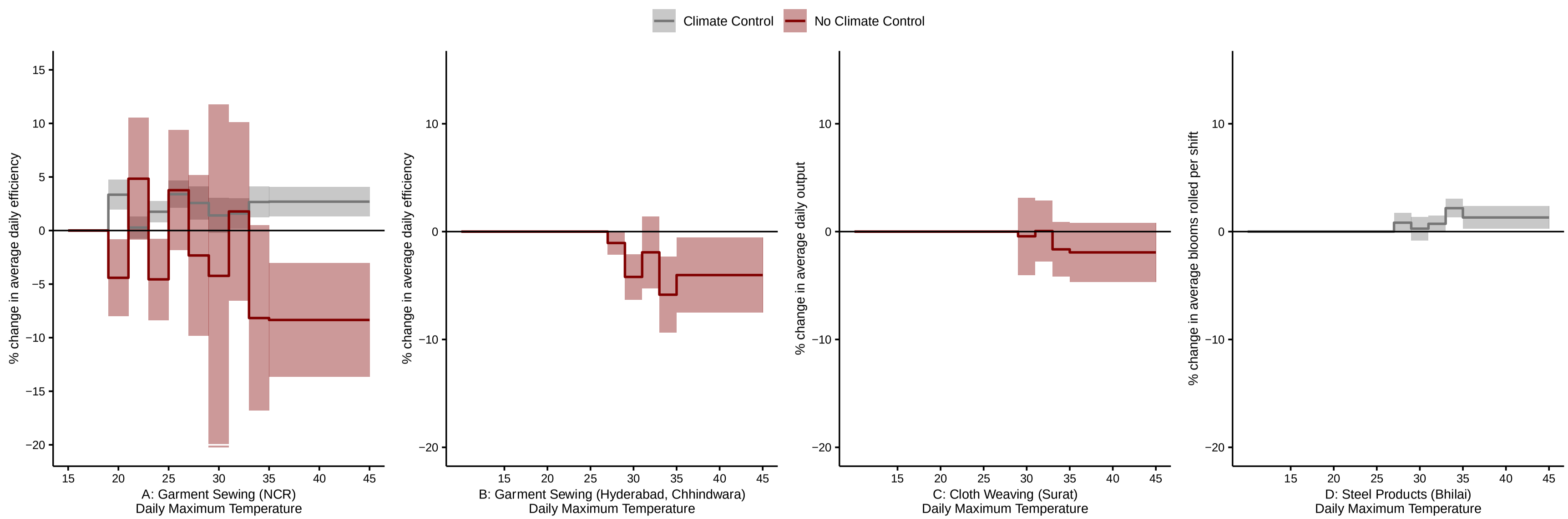

Figure 1: *The effect of temperature on worker output.* Estimates are percentage changes in daily output - averaged over a week - for a day in the week moving to a hotter temperature bin from the coolest (omitted) bin. Shaded areas represent 90 percent confidence intervals using robust standard errors clustered at the worker level. The number of bins varies across worker types, reflecting differences in observed temperatures. Panel A: Garment sewing lines in NCR, Panel B: Garment sewing lines in Hyderabad, Panel C: Cloth weaving in Surat, Panel D: Steel mill in Bhilai. The output variable for garment plants (A,B) is defined as the logarithm of the 'efficiency' measure of each sewing line. In Panel C, the output variable is defined as the inverse hyperbolic sin transformation of the meters of cloth woven by a worker. In Panel D, the output variable is the logarithm of blooms rolled by a team of workers.

temperatures. That said, different workplaces will vary in worker health and income, the nature of physical or cognitive tasks they perform, differences in the output measure, financial incentives and the nature of employment contracts etc. It is possible that these factors may lead to heterogeneous effects of heat even when observed temperature ranges are similar.

For worker sites, we are also able to obtain data on temperature and humidity and can estimate wet bulb temperatures (WBT), a quantity commonly used in the physiology literature to measure heat stress. Therefore in the Appendix (Figure A.3) we replicate the results in Figure 1 using bins in WBT instead of maximum temperatures. We find the same patterns of output response as we do when using maximum temperatures to proxy for heat. If anything, standard errors are smaller and effect sizes slightly larger.

Lagged Effects on Output and Absenteeism

To examine the effect of contemporaneous and lagged temperatures on workplace productivity and absenteeism, we turn to our disaggregated daily data. Exposure to very hot days may generate fatigue and illness, lowering output and increasing absenteeism. Strokes, fatigue, and even cases of organ damage have been directly linked to heat stress, and continued exposure may increase overall vulnerability (Kovats and Hajat, 2008). Other illnesses may be influenced by sustained warm weather through different mechanisms, for example, the increased breeding of pathogens and disease vectors.

We modify (1) to include lagged temperature bins. L^j_{id} is a count of the number of days falling in bin j in the six days *preceding* day d. Our output and other variables are as before, except now at the daily rather than weekly level:

$$Y_{id} = \alpha_i + \gamma_M + \gamma_t + \sum_j \beta_j T^j_{id} + \sum_j (\omega_j L^j_{id}) + \theta R_{id} + \lambda X_{id} + \epsilon_{id} \tag{2}$$

T^j is now an indicator for the day falling in temperature bin j. We use daily rainfall, R_{id}, and X_{id} now includes a day of the week fixed effect, and as before, for sewing lines we also include the target efficiency for the line. Our estimates from weekly data in Figure 1 suggest that most of the temperature effects occur in the two highest bins. We focus on these temperatures by aggregating over cooler bins. Therefore there are a total of three bins in both T and L.[17] Our results are in Table 2. Declines in daily output on hotter days are only seen in sites without climate control (Columns 1 and 3).[18] Lagged temperatures reduce output for some sites with the clearest effects on weaving workers where an additional day above 35^oC in the six preceding days causes a 2.7 percent decrease in contemporaneous daily output (Column 1, Row 4). Notice that lagged temperatures seem to matter even in climate controlled garment plants (Column 2). This may reflect exposure outside the workplace. This is related to our findings on absenteeism which we turn to next.

We have a daily indicator of absenteeism for our cloth-weaving workers. In the case of garment and steel plants, we have daily counts of the number of absences in the worker-team. Using these measures of absenteeism as the dependent variable, we estimate (2). From Table 3 we see that lagged high temperatures affect the likelihood of missed work in climate-controlled garment factories for paid leave, the steel plant, and the weaving plants. Standard errors for garment plants with no climate control are too high to draw any conclusions. Absenteeism effects are visible in settings with and without climate control.

The garment workers in our sample provide us with additional insight into worker responses to incentives. These workers are allocated a certain amount of paid leave. Our data on absenteeism distinguishes paid and unpaid absences for each worker. In climate-controlled garment plants in the NCR (columns 1-2) we find that the number of paid absences significantly increases with both contemporaneous and lagged temperatures but the probability

[17] Including lagged variables for all temperature bins increases the number of coefficients being estimated and results in imprecisely estimated coefficients.

[18] As before we see positive effects on output in the case of climate controlled sites. In addition, standard errors are high for the garment plants in central and south India and we are unable to draw clear conclusions here.

of *unpaid* leave does not change with temperature. Monetary disincentives may weaken the temperature-absenteeism link. The evidence here is only suggestive since in the case of non-climate-controlled garment plants (columns 5-6), our point estimates are too noisy to draw firm conclusions.

Absenteeism driven by contemporaneous high temperatures may be partially due to time-allocation decisions and labor-leisure trade-offs (Zivin and Neidell, 2014), while lagged effects may also reflect the effects of morbidity. Although workplace climate control may reduce the effects of temperature on worker productivity on the shop-floor, it may not remove negative output effects associated with greater absenteeism. Absenteeism might also result in costs we do not measure here, such as firms hiring redundant workers. The presence of redundant labor has been documented for the steel plant we study (Parry, 1999) and this might explain why we do not see output effects in climate controlled plants, even while capturing increased absenteeism.

For the garment and steel plants there is no straightforward way to translate increased absenteeism within worker teams into impacts on output. For weaving workers, an additional day above 35^oC in the six preceding days causes a 0.006 increase in the probability of missing work. The mean worker output, *on a day when the worker is present*, is 134.3 meters of cloth. Since absenteeism takes output to zero, this is equivalent to a reduction of 0.9 meters. Weaving workers come to work intermittently so their average daily output, *net* of absences, is about 51 meters of cloth per day. An additional hot day in the six preceding days therefore reduces output by about 1.8 percent through the absenteeism channel. This can be compared with a loss of 2.7 percent via the on-the-job productivity channel (Table 2).

Table 2: Effect of hot days on worker output

	Weaving	Garment Plants			Steel
	IHS Meters	Log Efficiency Measure			Log Blooms Rolled
	(1)	(2)	(3)	(4)	(5)
T (33-35 C)	-0.040^{**}	0.025^{**}	-0.129^{***}	-0.007	0.028^{*}
	(0.019)	(0.010)	(0.042)	(0.037)	(0.017)
T (above 35 C)	0.011	0.035^{***}	-0.154^{***}	0.008	0.020^{**}
	(0.022)	(0.014)	(0.041)	(0.046)	(0.009)
L (33-35 C)	-0.033^{***}	-0.004	-0.009	0.004	0.005
	(0.011)	(0.005)	(0.012)	(0.010)	(0.004)
L (above 35 C)	-0.027^{***}	-0.011^{**}	-0.019	0.015	-0.002
	(0.009)	(0.005)	(0.027)	(0.018)	(0.005)
Climate Control	No	Yes	No	No	Yes
Panel Width	147 workers	74 lines	10 lines	19 lines	9 shift-teams
Panel Length	365 days	578 days	316 days	628 days	3337 days

Notes: Data is at the worker and team level from our factory sites. *, **, *** denote estimates significant at 10, 5, 1 percent level. Robust standard errors clustered at worker level. T is an indicator for a day falling in the specified temperature bin. L is a count of the number of days falling in the specified temperature bins in the six preceding days. Models include individual (weaving workers) or team (garments and steel workers) fixed effects as well as month, year, and day-of-week fixed effects. Column 1 presents estimates from weaving workers in Surat with the output variable being an inverse hyperbolic sine transformation of meters of cloth woven. Output is set to zero when the worker is absent. Columns 2, 3 present estimates from garment sewing lines in the NCR with and without climate control respectively. The output variable is the logarithm of the efficiency measure. Column 4 presents estimates from garment sewing lines in South-Central India and (5) from the climate controlled steel mill. Coefficients on rainfall are omitted.

Table 3: Effect of hot days on worker absenteeism

	Climate Control			No Climate Control		
	Garments		Steel	Weaving	Garments	
	Paid	Unpaid	All	All	Paid	Unpaid
	(1)	(2)	(3)	(4)	(5)	(6)
T_33_35	0.082***	−0.083	−0.011	0.003	−0.001	0.796
	(0.022)	(0.065)	(0.048)	(0.004)	(0.128)	(0.678)
T_35_50	0.115***	0.031	0.051	−0.004	−0.034	1.001
	(0.027)	(0.049)	(0.068)	(0.004)	(0.117)	(0.862)
L_33_35	−0.018	−0.047	0.044***	0.006***	0.017	0.772
	(0.011)	(0.032)	(0.014)	(0.002)	(0.077)	(0.686)
L_35_50	0.021**	−0.001	0.045**	0.005***	0.078	0.567
	(0.010)	(0.022)	(0.020)	(0.002)	(0.083)	(0.426)
Panel Width	74 lines		9 teams	147 workers	42 lines	
Panel Length	587 days		3337 days	365 days	309-610 days	

Notes: Data is at the worker and team level from our factory sites.*, **, *** denote estimates significant at 10, 5, 1 percent level. Robust standard errors clustered at worker level. T is an indicator for a day falling in the specified temperature bin. L is a count of the number of days falling in the specified temperature bins in the six preceding days. Models include individual (weaving workers) or team (garments and steel workers) fixed effects as well as month, year, and day-of-week fixed effects. (1), (2) present the effect of temperature on the number of paid and unpaid leaves observed in sewing lines in climate-controlled garment plants. (3) reports the effect of temperature on absences in climate controlled steel worker teams. (4) reports the probability of a weaving worker being absent. (5), (6) present the effect of temperature on the number of paid and unpaid leaves observed in sewing lines in non climate-controlled garment plants.

4.2 Temperature Effects on Plant Output

Thus far we have used high-frequency data to show that worker productivity declines on hot days. We now turn to our nation-wide panel of manufacturing plants to examine whether there are similar temperature effects on the value of plant output and if so, whether they might be attributable to a decline in the productivity of labor.

We estimate a model analogous to (1):

$$y_{it} = \alpha_i + \gamma_t + \sum_{j=2}^{5} \beta_j T_{dt}^j + \theta R_{dt} + \epsilon_{it} \tag{3}$$

The dependent variable y is now the log of the value of annual plant output. Plant and year fixed effects are denoted by α_i and γ_t respectively while d denotes the district containing plant i. For every district d and year t, T_{dt}^j is the number of days in the year with maximum temperature falling in bin j, $j = 1, 2, \ldots, 5$: $\{(0, 20], (20, 25], (25, 30], (30, 35], (35, 50]\}$. All plants in a district are therefore mapped to the same temperature distribution. R_{dt} is rainfall in district d in year t.

We use wider bins here than with our worker data to preserve precision. We have a shorter panel with only 15 years of data, as opposed to the worker data where our shortest weekly panel is 52 weeks and our shortest daily panel covers 365 days. The topmost bin for both worker and plant models is identical.

Our coefficient estimates β_j are plotted in Figure 2 and indicate an inverse relationship between temperature and annual plant output, akin to the relationship we see between temperature and worker productivity.[19] Each β_j is the percentage change in annual plant output from a single day in the year moving from the coldest bin to bin j. Shaded areas represent 90

[19]Two recent studies from China have similar findings (Chen and Yang, 2017; Zhang et al., 2018).

percent confidence intervals with standard errors corrected for serial and spatial correlation following Conley (2008).[20] A day moving from the lowest to the highest temperature bin reduces annual output by 0.22 percent.

In Table 4 we present results from running a set of related specifications to probe the robustness of our results. Each row describes the effect of a different warming counterfactual on log output. Each column shows results from a different specification. Rows 1 through 4 provide the effect of one day in the year moving from the lowest bin to the second, third, fourth, and fifth bins respectively. More precisely we assume that days move from the top of bin k to the top of bin $k-1$, thus for example, row 1 represents a single day moving from 20^oC to 25^oC. Row 5 presents the effect of every day in the year warming by 1^oC so that the average annual maximum temperature increases by 1^oC. Across all models, the reduction in output ranges from 1.7 to 2.3 percent.

Row 6 provides a projection of the effects of long-term warming on levels of output. For every day in the year we first average over the 2075-2080 projections of daily temperature obtained from the RCP 8.5 scenario of the Hadley GEMS2 climate model. From this distribution, we subtract the average daily temperatures from the same model between 2005-2010. This yields daily estimates of warming that we add to a baseline distribution obtained by averaging over spatial and temporal values of daily temperatures in our data. Row 6 provides estimates of how log output changes under this warming scenario. Across models, the mean effect size ranges between roughly -4.5 to -7 percent. [21]

The columns of Table 4 present different specifications. Column 1 provides results under the main specification described in 3. Column 2 adds state-year polynomial time trends.

[20] Conley errors are presented assuming a 250km radius of spatial correlation. Standard errors do not significantly increase beyond 100km.

[21] Models that allow annual output shocks to influence year-on-year growth rates will predict larger changes. These models might be appropriate if plants permanently reduce productivity enhancing capital investment following low output years. Investigating this question lies outside the scope of our paper and we focus on the mechanism that might underpin the original output shock.

Columns 3 and 4 control for floods and industrial conflicts, and outages respectively. We discuss these further in Section 7.

Column 5 presents a model specified using *degree-days* instead of daily bin counts. Degree days are calculated in the following manner: a day with a temperature of 29 degrees contributes 20 degrees to the first bin (0-20], 5 degrees to the second bin (20-25], and 4 degrees to the third bin (25-30]. Under the scenario in Row 1, the effect of a single day moving from 20 degrees to 25 degrees is equal to an increase of 5 degrees in the second degree day bin and a change of zero degrees in all other bins. More formally, let D^k be the number of degree days in each bin k. Then we have:

$D^1 = max(0, T - 20)$

$D^2 = max(0, min(5, T - 20))$

$D^3 = max(0, min(5, T - 25))$

$D^4 = max(0, min(5, T - 30))$

$D^5 = max(0, T - 35)$

Columns 6 and 7 provide results from models where logged output depends on the sum of quadratic or linear polynomials of daily maximum temperature. This corresponds to specifications of the following type, where T_{dit} is the maximum temperature for plant i on day d of year t:

$$y_{it} = \alpha_i + \gamma_t + \sum_{d=1}^{365} \beta_1 T_{dit} + \sum_{d=1}^{365} \beta_2 T_{dit}^2 + \theta R_{dt} + \epsilon_{it}$$

Across all specifications we find an inverse relationship between measures of temperature and plant output, albeit with heterogeneity in effect sizes.[22]

[22]In Appendix Table A.2 we provide an additional set of specifications analogous to Table 4 but with a

Table 4: Effect of hot days on plant output

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Day in Bin 2	-0.00139**	-0.00093*	-0.00145**	-0.00173**	-0.00152***	-0.00017	-0.00023*
	(0.00043)	(0.00048)	(0.00044)	(0.00053)	(0.00060)	(0.00026)	(0.00014)
Day in Bin 3	-0.00164***	-0.0013**	-0.00163***	-0.00183***	-0.00116**	-0.00038	-0.00046*
	(0.00049)	(0.00056)	(0.00049)	(0.00055)	(0.00051)	(0.00040)	(0.00027)
Day in Bin 4	-0.00189***	-0.0016***	-0.00193***	-0.00225***	-0.00175***	-0.00064	-0.00069*
	(0.00051)	(0.00059)	(0.00051)	(0.00058)	(0.00058)	(0.00047)	(0.00041)
Day in Bin 5	-0.00217***	-0.00183***	-0.00223***	-0.00238***	-0.00133*	-0.00132*	-0.00114*
	(0.00056)	(0.00067)	(0.00056)	(0.00066)	(0.00079)	(0.00069)	(0.00069)
Mean shift by 1 degree	-0.02126***	-0.02085**	-0.02304***	-0.02267**	-0.01648*	-0.01887*	-0.01670*
	(0.00812)	(0.00848)	(0.00815)	(0.00965)	(0.00982)	(0.00979)	(0.01003)
Projected warming 2075-80	-0.06558**	-0.0629**	-0.07108**	-0.06006*	-0.04489	-0.07735*	-0.06680*
	0.02934	(0.03079)	(0.0295)	(0.03455)	(0.04348)	(0.04063)	(0.04012)
Plant FE	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y
State-Year Polynomial	N	Y	N	N	N	N	N
Conflict and Flood Controls	N	N	Y	N	N	N	N
Outage Controls	N	N	N	Y	N	N	N

Notes: Rows 1-4 present the effect on log output of an additional day in temperature bin T^j relative to T^1 for different specifications relating output to temperature. Bins are defined as $\{(0, 20], (20, 25], (25, 30], (30, 35], (35, 50]\}$. Row 5 presents the effect of a one degree increase in the temperature of every day in the year, starting from a baseline distribution obtained by averaging over daily temperatures observed in our panel of manufacturing plants. Row 6 presents the effect of adding daily deviations (increases) to this baseline temperature distribution, to simulate a warmer future temperature distribution in 2075-2080, as projected using the Hadley GEMS2 climate model under the RCP 8.5 scenario. Column 1 provides estimates from our preferred specification in Equation 3. Columns 2-4 add additional controls for state-year quadratic time trends, floods, conflict, and outages. Column 5 presents estimates from a model where temperature is described by the number of degree days falling in the temperature ranges described above. Column 6 presents estimates from a model where daily output is assumed to depend on a quadratic polynomial of daily temperature. Column 7 presents estimates from a model where daily output is assumed to depend linearly on temperature. Data on the value of output and inputs are at the plant level from the Annual Survey of Industry.*, **, *** denote estimates significant at 10, 5, 1 percent level. Standard errors are corrected for serial and spatial correlation following Conley (2008).

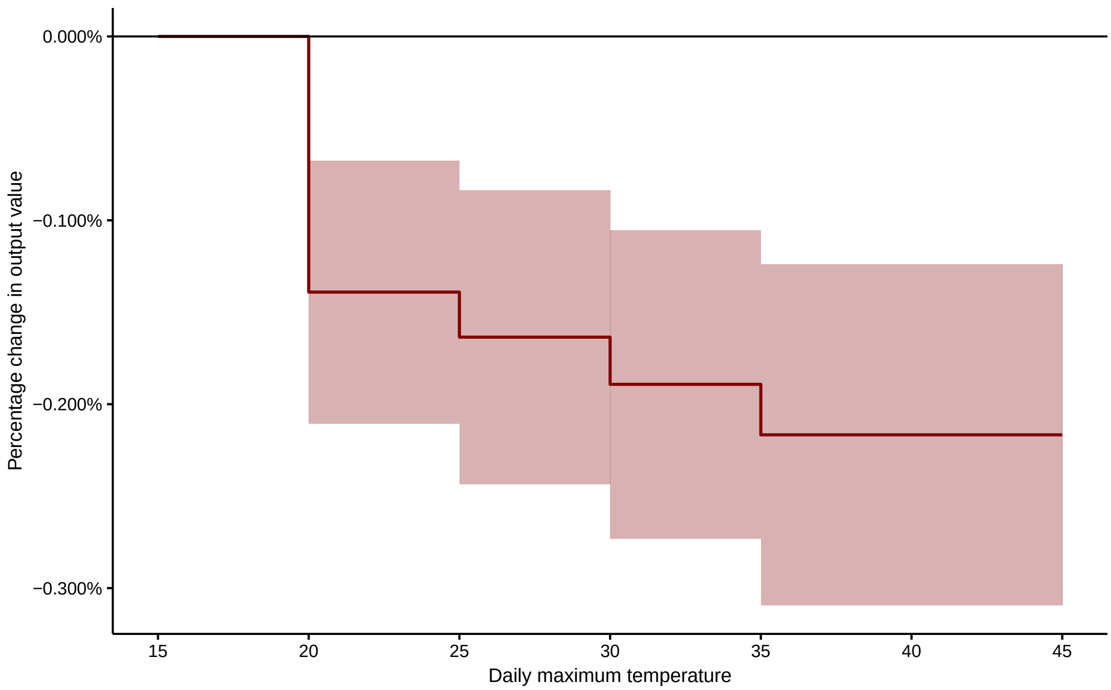

Figure 2: Effect of temperature on plant output, represented as the change in output on a warmer day relative to days with maximum temperature below 20 degrees Celsius. Shaded areas represent 90 percent confidence intervals with standard errors corrected for serial and spatial correlation following Conley (2008). Output data comes from the Annual Survey of Industry.

Temperature Response and the Labor Channel Table 4 and Figure 2 suggest a clear inverse relationship between temperature and plant output. Evidence from worker-level output additionally suggests that the productivity of labor might play a role in explaining the aggregate response. Using our nation-wide industry panel, we now examine whether there is evidence suggesting that our aggregate effects might be explained by reductions in the productivity of labor.

To do so we begin by specifying a Cobb-Douglas production function in log form:

$$y = \alpha(\mathbf{T}) + \omega(\mathbf{T})k + \beta(\mathbf{T})l \tag{4}$$

Logged values of output, capital and labor respectively are denoted by y, k, and l respectively.

different set of bin cutoffs.

$\mathbf{T}$ is a vector with counts of the number of days in each temperature bin. The total factor productivity α, output elasticity of labor β, and the output elasticity of capital ω, are all allowed to depend on the vector $\mathbf{T} = (T^1, T^2, \ldots, T^5)$. In this Cobb-Douglas formulation, the output elasticities will equal cost shares on average, but they may not do so in any given year because year-to-year variations in the temperature distribution are not predictable. Therefore, the quantities of labor and capital will not depend on the realization of the vector $\mathbf{T}$.

We assume that α, ω and β are all linear in temperature bins indexed by j.

Thus we have,

$$\alpha(\mathbf{T}) = \alpha_o + \sum_{j=2}^{5} \alpha_j T^j$$

$$\omega(\mathbf{T}) = \omega_o + \sum_{j=2}^{5} \omega_j T^j$$

$$\beta(\mathbf{T}) = \beta_o + \sum_{j=2}^{5} \beta_j T^j$$

Making these substitutions in (4) we obtain

$$y = \alpha_o + \sum_{j=2}^{5} \alpha_j T^j + \omega_o \cdot k + \sum_{j=2}^{5} \omega_j T^j k + \beta_o \cdot l + \sum_{j=2}^{5} \beta_j T^j l \tag{5}$$

For the plants in our panel, we define capital as the net value of equipment and machinery at the start of each year, and we use the number of full-time workers as our measure of labor. We add controls for plant and year fixed effects as well as rainfall to (5) and then estimate the resulting model. We can use these estimates of ω_j, β_j, and α_j to examine the effect of temperature on output, acting through changes in the elasticity of capital, labor, or the residual.

Table 5: Temperature interactions with factor inputs

	Log Plant Output				
	(1)	(2)	(3)	(4)	(5)
T^2	0.00256**	0.00008	-0.00008		
	(0.00097)	(0.00211)	(0.00164)		
T^3	0.00147	-0.00205	-0.00009		
	(0.00103)	(0.00229)	(0.00172)		
T^4	0.00081	-0.00094	-0.00028		
	(0.00108)	(0.00237)	(0.00178)		
T^5	0.00003	-0.00499*	-0.00171		
	(0.00118)	(0.00259)	(0.00194)		
l	0.8612***		0.91426***	0.36520***	
	(0.0957)		(0.10322)	(0.07038)	
k		0.20433**	0.06629		
		(0.05674)	(0.04215)		
$l \times T^2$	-0.00098***		-0.00134***	-0.00056**	
	(0.00027)		(0.00035)	(0.00026)	
$l \times T^3$	-0.00067**		-0.00104***	-0.00038*	
	(0.00027)		(0.00029)	(0.00020)	
$l \times T^4$	-0.00052*		-0.00077***	-0.00030	
	(0.00027)		(0.00029)	(0.00020)	
$l \times T^5$	-0.00036		-0.00075***	-0.00039*	
	(0.00029)		(0.00031)	(0.00021)	
$k \times T^2$		-0.00009	0.00028*		
		(0.00015)	(0.00015)		
$k \times T^3$		0.00005	0.00022*		
		(0.00016)	(0.00012)		
$k \times T^4$		-0.00003	0.00016		
		(0.00016)	(0.00012)		
$k \times T^5$		0.00022	0.00024**		
		(0.00018)	(0.00013)		
$T^a \times Q^l_2$					-0.0375**
					(0.0150)
$T^a \times Q^l_3$					-0.0628***
					(0.0161)
$T^a \times Q^l_4$					-0.1116***
					(0.0207)
$T^a \times Q^k_2$					0.03545**
					(0.0141)
$T^a \times Q^k_3$					0.03361**
					(0.0147)
$T^a \times Q^k_4$					0.00032
					(0.0161)
Observations	181442		181442	181442	181442

Notes: Data are at the plant level from the Annual Survey of Industry. Standard errors are corrected for serial and spatial correlation following Conley (2008). Models include plant and year fixed effects. Coefficients on rainfall are omitted. Temperature bins are $\{(0, 20], (20, 25], (25, 30], (30, 35], (35, 50]\}$. T^j is the number of days in the jth bin. T^1 is the omitted bin. (1) and (2) add interactions with labor and capital to our base model. (3) presents OLS estimates of the production function. (4) presents the first stage of a Levinsohn-Petrin estimate of the production function. Capital-temperature interactions and residual temperature effects are subsumed in a non-linear control function and not separately reported. (5) interacts annual temperatures with quartiles of labor and capital intensities.

Coefficient estimates from this model are presented in Column 3 of Table 5. We see that the temperature-labor interaction terms, β_j are all negative and significant, while temperature effects on the output elasticity of capital are positive. Controlling for temperature interactions with labor and capital, the residual effect of temperature is also insignificant, as seen in the first four rows. These results suggest that it is temperature-induced declines in labor productivity that drive the negative effects of temperature on output.

One concern in estimating production functions of this type is the potential endogeneity of labor (Ackerberg, Caves, and Frazer, 2006; Levinsohn and Petrin, 2003). This may not be a significant concern in our setting, given India's notoriously inflexible labor market. In 2017, the World Bank ranked India as low as 130 on its global *Ease of Doing Business* index, citing rigid labor laws as a primary reason for the country's poor performance. Among several other weaknesses, the report draws attention to India's Industrial Dispute Resolution Act (IDA) of 1947, which requires that firms with more than 100 employees obtain explicit government approval before dismissing workers. Since our measure of capital is the value of plant and machinery at the start of the year, this too is relatively inflexible and cannot be influenced by temperature shocks during the year.

Nevertheless, as a robustness check, we also estimate the production function in (5) using the Levinsohn-Petrin (LP) estimator that allows for endogenous labor (Levinsohn and Petrin, 2003). This approach assumes that labor is highly flexible and chosen by the firm in each period, after the realization of any shocks. Section A.5 of the Appendix describes the way in which we apply this method to our data and Column 4 of Table 5 reports the relevant coefficient estimates. The point estimates for the labor-temperature interactions are smaller but remain negative and are statistically indistinguishable from those in Column 3.[23]

Lastly we investigate how temperature effects vary by labor and capital intensity. We mea-

[23] Capital-temperature interactions and residual temperature effects are subsumed in a non-linear control function and not separately estimated here. See Section A.5 for details.

sure labor intensity by the ratio of the total annual wage bill to total annual output for all plants in our sample. We measure capital intensity by the ratio of the value of capital to annual output. Based on their mean value of labor and capital intensity across all years, we classify plants into quartiles, Q^{lj} and Q^{kj} and estimate the model below:

$$y_{it} = \alpha_i + \gamma_t + \beta_0 T^a_{it} + \sum_{j=2}^{4} \beta^l_j T^a_{it} Q^{lj} + \sum_{j=2}^{4} \beta^k_j T^a_{it} Q^{kj} + \theta R_{it} + \epsilon_{id} \tag{6}$$

Column 5 of Table 5 reports coefficients β^l_j and β^k_j from this model. The negative effects of the annual average of daily maximum temperature (T^a) are greatest in plants with high wage-share output ratios. On the other hand, capital intensity is positively associated with temperature. Importantly, these models include plant fixed effects and are therefore not simply driven by plant size. [24]

As a robustness check, in Columns 1 and 2 we report estimates from reduced form specifications that include interactions with either labor or capital respectively. As before labor-temperature interactions in (1) are negative and mostly statistically significant, while capital-temperature interactions in (2) are insignificant and close to zero. These are not our preferred specifications, since they include only one factor input.

Taken together, the evidence in this section not only suggests that temperature may negatively effect the output of manufacturing plants, but also that this response may run through the labor channel.

[24] For parsimony, this model interacts only the average daily maximum temperature with quartile dummies. We obtain similar results using days in the highest temperature bin rather than average maximum temperature. We could also interact all temperature bins with quartile dummies but this produces a large number of imprecisely estimated coefficients.

5 Comparison with Macro-level Estimates

In this section, we show that our estimated temperature effects at worker and plant levels are consistent with each other, and with estimates based on district-level manufacturing output. We also compare our results with prior country-level studies. These comparisons suggest that temperature effects on labor are large enough to account for the whole of the country-level response of manufacturing GDP to temperature.

Prior country-level studies have estimated the effect of a one-degree increase in annual temperature on country GDP. To compare our estimates with these, we must report our worker and plant results in similar terms. This requires specifying how the distribution of daily temperatures across the year changes such that the average annual temperature increases by one degree. There is of course, no unique way to map changes in temperature distribution to changes in annual average temperatures. We simply assume that every day in the year warms by one degree. Under this assumption, the change in plant output for our primary specification is -2.1 ($\pm$ 0.8) percent (Row 5, Column 1 of Table 4). This is plotted in Bar 2 of Figure 3.

For worker-level data our results in Figure 1 suggest significant heterogeneity across the type of work, as well as protection from heat when cooling technologies are used. Noting that no single setting is representative of all workers, we estimate the effect of a one-degree uniform increase in the daily temperature distribution for garment workers in the NCR who are not working in cooled environments. We use this site because it has the widest temperature range that corresponds most closely to that in the nationally representative plant data. The estimated percentage reduction in output is 3 ± 1.35 (Bar 1 of Figure 3).[25]

If the output from manufacturing plants drops in hot years, we should see correspond-

[25] The relationship between temperature and output for workers is described using a 'days in temperature-bins' specification so calculating this effect requires translating a one-degree increase in the daily temperature into corresponding changes in temperature bins.

ing changes in manufacturing GDP at the sub-national level. We use panel data on the manufacturing GDP for Indian districts described in Section 3. We regress district-level manufacturing GDP on average annual maximum temperature, T^a, controlling for rainfall as well as district and year fixed effects. The coefficient on T^a, gives us the effect of a one degree increase in temperature on district output. The estimated percentage reduction in manufacturing GDP is -3.5±2.6 (Bar 3 of Figure 3).[26]

The last two bars in Figure 3 depict estimates from two recent country-level studies; Dell, Jones, and Olken (2012) and Burke, Hsiang, and Miguel (2015). These numbers are derived using annual average temperatures for many countries across the world, observed over long periods of time. The earliest temperature-country observations used in both papers are from 1950. Thus while these estimates are not directly comparable with our results, they provide a useful and independent benchmark.

The fourth bar, with the label DJO, provides the contemporaneous effect of temperature on industrial sector output from Dell, Jones, and Olken (2012). The effect size reported in that paper is statistically indistinguishable from those we obtain at lower levels of aggregation. The last bar, with the label BHM, provides the temperature effect on *country GDP*, encompassing all sectors reported in Burke, Hsiang, and Miguel (2015). The authors do not separately provide estimates for the manufacturing sector alone but it is interesting to observe that the percentage reduction of output for the economy as a whole is similar to that observed for manufacturing. One possible explanation could be that reductions in labor productivity will affect all sectors of the economy.

[26]We are unable to study temperature effects using Reserve Bank of India GDP figures for Indian states because these data are interpolated in several years and therefore unreliable. In our district panel we observe missing data in some years, but not imputed estimates.

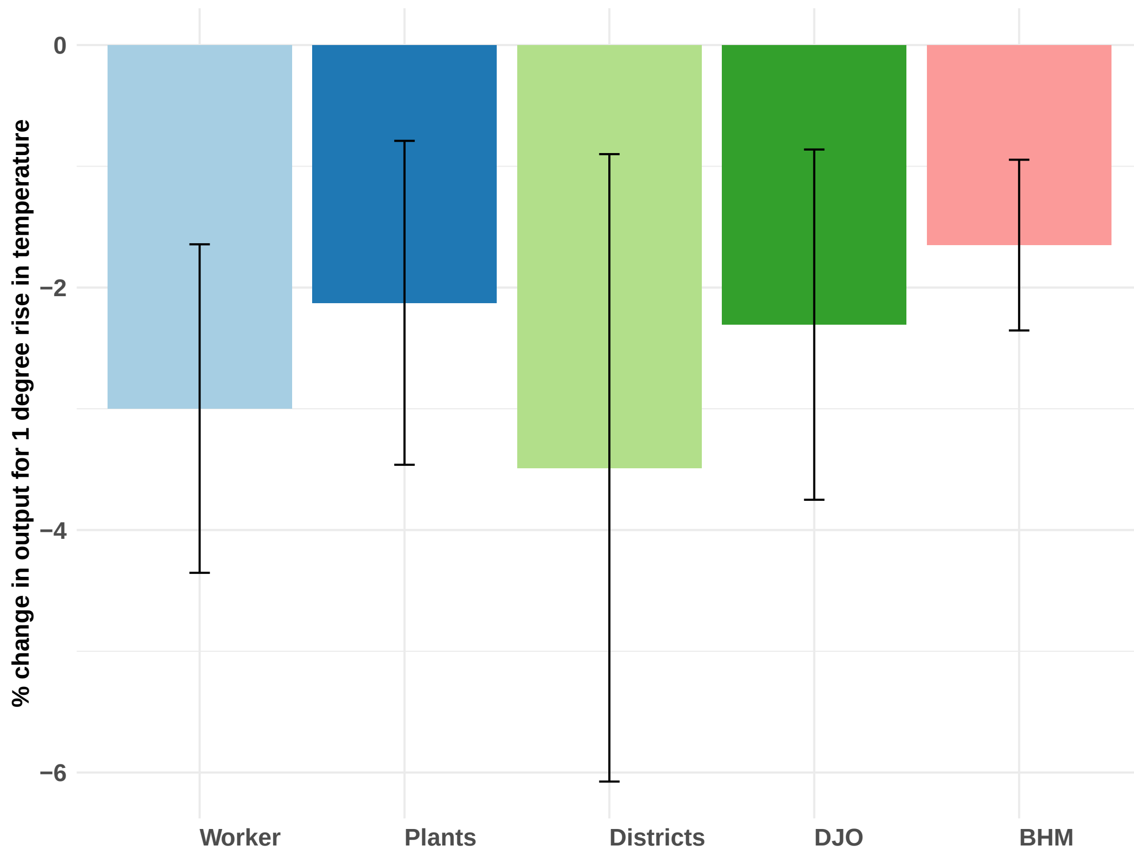

Figure 3: Marginal effect of temperature on log output at different levels of production with 90 percent confidence intervals. Legend entry **DJO Industry** refers to the contemporaneous effect of temperature on industrial sector output reported in Dell, Jones, and Olken (2012). Legend entry **BHM Country** refers to the contemporaneous effect of temperature on country-level GDP from Burke, Hsiang, and Miguel (2015). All other estimates come from data in this paper.

6 Adaptation

The loss in output caused by high temperatures encourages adaptive responses by firms. In the medium-run, they may introduce climate control. Over longer time periods, firms may increase automation, relocate plants, or change the composition of output.

Decisions to invest in climate control depend on the costs of cooling, relative to the expected output losses resulting from heat stress. In Appendix Section A.10.2, we carry out a back-of-the envelope cost benefit analysis of climate control for weaving plants and show that electricity costs may be too high relative to output losses to justify these investments.

Firms may also selectively invest in climate control. If labor productivity plays an important role in output losses associated with hot days, we would expect that processes which are labor-intensive and add high value would be preferentially protected. To study this we conducted a survey of 150 diamond cutting factories located in the same city of Surat as these cloth weaving units. These are drawn randomly from all factories registered with the local diamond industry association.

Diamond processing plants involve a set of distinct processes, some of which are largely mechanized (such as cutting stones), while others are carried out by human beings (such as sorting uncut diamonds by quality). Our survey allowed us to study the selective adoption of air-conditioning *within* plants. We find that climate control is indeed more likely to be used for processes that are labor-intensive and contribute most to diamond quality. We describe our data and results in Appendix Section A.10.3.

In our national plant panel we find that the effects of a degree-rise in temperature seem to be falling over a 15-year period. We modify (3) to include a full set of interactions of temperature bin counts with a continuous time variable. The negative effect on output from an additional day in the fourth and fifth temperature bins reduces by about 5 percent per year. Column 1 of Table A.5 in the Appendix provides these coefficient estimates.[27] As countries grow richer, it is possible that their manufacturing sector becomes less vulnerable to output losses associated with heat.

7 Alternative Explanations

We have argued that high temperatures result in a fall in manufacturing output through declines in labor productivity. There are alternative explanations for temperature induced declines in output in the literature. Climatic changes may result in increased temperature-induced conflict (Hsiang, Burke, and Miguel, 2013), and output losses owing to natural disasters (Kahn, 2005). Neither of these are likely to influence our worker-level results because they occur on time-scales that are much longer than a day. They could, however, mediate the temperature effects on output that we observe in our annual manufacturing panel. Other factors that may influence plant output, without necessarily changing the productivity of labor include power outages, input price changes, and agricultural spillovers. We consider some of these alternative explanations.

We gather data on industrial strikes (as a measure of conflict) and on the timing of floods (for natural disasters). We then estimate modified versions of (3) controlling for these potential confounders. Our estimates of temperature effects on output remain almost unchanged (Table 4, Column 3). Details on data sources, measurement, and coefficient estimates are reported in Appendix Section A.6.

To proxy for power outages, we use data on supply shortfalls from India's Central Electricity Authority. As with conflict and flood variables, we find that accounting for these shortfalls has no significant effect on our results (Table 4, Column 4). We also examine whether

[27] Since climate control uses a lot of electricity, we also look at heterogeneity in temperature response by the electricity intensity of output. We find that plants with above median levels of electricity intensity respond more weakly to high temperatures (Table A.5, Column 2).

input prices change with temperature, using data on the price of the input with the largest expenditure share, as reported in the ASI (Table A.4). We find no evidence of temperature effects on input prices in our data. It may be that most changes in prices are captured by the year fixed-effects in our models, and any local price shocks from local temperature fluctuations are insulated through storage.

Finally, to examine the role of agricultural spillovers, we provide sector-wise estimates of temperature effects by estimating a model in which we include interactions of average annual maximum temperature with indicators for 2-digit manufacturing sectors. We observe negative temperature effects across sectors, even for activities with no obvious connection to agriculture (Appendix Figure A.6).

8 Conclusions

This paper estimates the impact of temperature on the output of manufacturing plants. We also use selected factory settings to separately study effects on the daily productivity and attendance of workers. We show that, in the absence of climate control, worker productivity declines on hot days. For absenteeism, we find effects of contemporaneous and lagged temperatures even for workers in factories with climate control, suggesting that workplace adaptation alone is insufficient to mitigate all the effects of heat. In the annual plant-level national panel, we find that the effect of temperature on the value of plant output appears to be driven in large part by its effect on the output elasticity of labor.

Our estimates from both worker and annual plant data are comparable to those found in studies of country-level manufacturing GDP. This suggests that heat stress, through its effects on productivity, time-allocation and morbidity, is an important underlying cause for the declines in non-agricultural GDP at high temperatures.

The evidence we uncover on the effectiveness of climate control and on its limited adoption, has implications for how we should think about the costs of climate change going forward. Research into low-cost technologies to protect workers from ambient temperatures may have significant social value. In the long term, there are other ways in which the industrial sector might respond to high temperatures. These include increasing automation and shifting away from labor-intensive sectors in hot parts of the world. These adaptive responses may have significant distributional implications. If directed towards more productive workers, they will tend to increase wage inequality.

Although our focus throughout this paper has been on the manufacturing sector, the potential ramifications of our findings are wider. Our conclusion that a physiological mechanism is economically important implies that these effects may be significant in labor-intensive activities across the world, such as construction and agriculture, where heat exposure is high and adaption through climate control is expensive or infeasible. Observed productivity losses in agriculture that have been attributed by default to plant growth responses to high temperatures may in fact be partly driven by lower labor productivity. These possibilities are yet to be researched.

References

Ackerberg, Daniel, Kevin Caves, and Garth Frazer. 2006. "Structural identification of production functions." *MPRA Paper* 38349.

Adhvaryu, Achyuta, Namrata Kala, and Anant Nyshadham. 2019. "The Light and the Heat: Productivity Co-benefits of Energy-saving Technology." *Review of Economics and Statistics (forthcoming)* .

Allcott, Hunt, Allan Collard-Wexler, and Stephen O'Connell. 2016. "How Do Electricity Shortages Affect Productivity? Evidence from India." *American Economic Review* 106 (3):587–624.

Auffhammer, M, V Ramanathan, and J R Vincent. 2006. "From the Cover: Integrated model shows that atmospheric brown clouds and greenhouse gases have reduced rice harvests in India." *Proceedings of the National Academy of Sciences* 103 (52):19668–19672.

Berman, Eli, Rohini Somanathan, and Hong W Tan. 2005. "Is Skill-biased Technological Change Here Yet?: Evidence from Indian Manufacturing in the 1990's." *Ann. Econ. Stat.* 79-80:299–321.

Burgess, Robin, Olivier Deschenes, Dave Donaldson, and Michael Greenstone. 2017. "Weather, climate change and Death in India." *Unpublished* .

Burke, Marshall, Solomon M Hsiang, and Edward Miguel. 2015. "Global non-linear effect of temperature on economic production." *Nature* 527 (7577):235–239.

Chen, Xiaoguang and Lu Yang. 2017. "Temperature and industrial output: Firm-level evidence from China." *Journal of Environmental Economics and Management* .

Das, Sanghamitra, Kala Krishna, Sergey Lychagin, and Rohini Somanathan. 2013. "Back on the Rails: Competition and Productivity in State-Owned Industry." *Am. Econ. J. Appl. Econ.* 5 (1):136–162.

Dell, Melissa, Benjamin F Jones, and Benjamin A Olken. 2012. "Temperature Shocks and Economic Growth: Evidence from the Last Half Century." *American Economic Journal: Macroeconomics* 4 (3):66–95.

Fiorese, Elena, Franco Bonollo, Giulio Timelli, and Lars Arnberg. 2015. "New Classification of Defects and Imperfections for Aluminum Alloy Castings." *International Journal of Metalcasting* 9 (1):55–66.

Gupta, Ridhima, E Somanathan, and Sagnik Dey. 2017. "Global warming and local air pollution have reduced wheat yields in India." *Climatic change* 140 (3-4):593–604.

Hsiang, S M, M Burke, and E Miguel. 2013. "Quantifying the Influence of Climate on Human Conflict." *Science* 341 (6151):1235367–1235367.

Hsiang, Solomon M. 2010. "Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America." *Proc. Natl. Acad. Sci. U. S. A.* 107 (35):15367–15372.

Huntington, Ellsworth. 1915. *Climate and civilization.* Harper & Bros.

Iso. 1989. "Hot environments – Estimation of the heat stress on working man, based on the WBGT-index (wet bulb globe temperature)." *Technical Report (International Standards Organization)* .

Kahn, Matthew E. 2005. "The Death Toll from Natural Disasters: The Role of Income, Geography, and Institutions." *Rev. Econ. Stat.* 87 (2):271–284.

Kjellstrom, Tord, Ingvar Holmer, and Bruno Lemke. 2009. "Workplace heat stress, health and productivity: An increasing challenge for low and middle-income countries during climate change." *Glob. Health Action* 2: Special:46–51.

Kovats, R Sari and Shakoor Hajat. 2008. "Heat stress and public health: a critical review." *Annu. Rev. Public Health* 29:41–55.

Kumar, Hemanshu and Rohini Somanathan. 2009. “Mapping Indian Districts Across Census Years, 1971-2001.” .

Lemke, Bruno, Bruno and Tord Kjellstrom, Tord. 2012. “Calculating Workplace WBGT from Meteorological Data: A Tool for Climate Change Assessment.” *Ind. Health* 50 (4):267–278.

Levinsohn, James and Amil Petrin. 2003. “Estimating Production Functions Using Inputs to Control for Unobservables.” *Rev. Econ. Stud.* 70 (2):317–341.

Lobell, David B, Wolfram Schlenker, and Justin Costa-Roberts. 2011. “Climate trends and global crop production since 1980.” *Science* 333 (6042):616–620.

Mackworth, N H. 1946. “Effects of heat on wireless telegraphy operators hearing and recording Morse messages.” *Br. J. Ind. Med.* 3:143–158.

Mendelsohn, R and A Dinar. 1999. “Climate Change, Agriculture, and Developing Countries: Does Adaptation Matter?” *World Bank Res. Obs.* 14 (2):277–293.

Parry, Jonathan. 1999. “Lords of Labour: Working and Shirking in Bhilai.” *Contributions to Indian Sociology* 33 (1):107–140.

Parsons, K C. 1993. *Human Thermal Environments*. Informa UK (Routledge).

Schlenker, Wolfram and Michael J Roberts. 2009. “Nonlinear temperature effects indicate severe damages to US crop yields under climate change.” *Proceedings of the National Academy of sciences* 106 (37):15594–15598.

Seppanen, Olli, William J Fisk, and David Faulkner. 2003. “Cost Benefit Analysis of the Night-Time Ventilative Cooling in Office Building.” Tech. rep., Lawrence Berkeley National Laboratory.

Wyndham, C H. 1969. “Adaptation to heat and cold.” *Environ. Res.* 2:442–469.

Zhang, Peng, Olivier Deschenes, Kyle Meng, and Junjie Zhang. 2018. "Temperature effects on productivity and factor reallocation: Evidence from a half million chinese manufacturing plants." *Journal of Environmental Economics and Management* 88:1–17.

Zivin, Joshua Graff and Matthew Neidell. 2012. "The Impact of Pollution on Worker Productivity." *Am. Econ. Rev.* 102 (7):3652–3673.

———. 2014. "Temperature and the Allocation of Time: Implications for Climate Change." *J. Labor Econ.* 32 (1):1–26.

Appendix: For Online Publication

A.1 Microdata Sites

Figure A.1 shows the shop-floors of each of our worker microdata sites. In the garment manufacturing factories shown in Panel A, workers are arranged in lines, with each person repeatedly carrying out a specific task. For example, one worker may repeatedly sew on buttons, further down the line another person may finish the collar and so on.

In the cloth weaving plants of Panel B, workers walk up and down in the aisles between looms, adjusting alignment, restarting feeds when interrupted, and making other necessary corrections. One worker typically covers about 6 machines.

The steel mill shown in Panel C uses smelting, casting and forging processes, all of which are capital intensive. Worker tasks and teams have been already discussed in some detail in Section 3.

A.2 Seasonal Patterns in Worker Absenteeism

We carried out open-ended interviews with the owner-managers of several cloth-weaving firms in Surat in addition to the three factories in our data sample. We found that these units were similar in that there was no climate control and workers were paid daily wages based on output with no full-time contracts. Several owners spoke of workers being less willing to work in their factories during the hot summer months. Many return to their home villages and rely on income from India's National Rural Employment Guarantee Scheme. This safety net is lower paid than factory wages so this narrative suggests that the disutility from heat exposure exceeds this difference in wages. Thus one response to sustained high temperatures (as opposed to the occasional hot day) may be to shift to other occupations. Some owners reported that they were considering a summer attendance bonus to keep workers while others felt this bonus would depress profit margins too much to make it worthwhile.

The identification strategies used in this paper do not allow us to utilize long-run temperature variation since there are other seasonal factors that may be associated with attendance. However worker attendance in the cloth weaving and garment plants we study can be plotted over the year. Figure A.2 shows seasonal reductions in the attendance of daily wage cloth weaving workers, concentrated in high temperature months. These patterns are absent for the garment sewing workers who are both paid more, and have long term employment contracts.

Many factors differentiate the two types of work settings, but it is plausible that formal employment contracts reduce the costs of taking an occasional day of paid leave, but significantly increase the risks associated with switching occupations for extended periods of time. When accounting for longer term responses to temperature, formal employment contracts might therefore do better at retaining labour. These responses to sustained and longer-term temperature increases are an area that would benefit from further research.

A.3 Effect of wet bulb temperatures on worker output

In this section we replicate Figure 1 using wet bulb temperatures in place of maximum temperatures. The specifications used mirror those in the main paper with the exception that the cut-off points for bins differ from those used in drawing Figure 1.

The measurement of WBT requires specialized instruments but it can be approximated by combining temperature and relative humidity. We use a formula provided by Lemke and Kjellstrom (2012):

$$WBT = 0.567T_A + 0.216\rho + 3.38 \tag{7}$$

where T_A is air temperature in degrees celsius and ρ is water vapour pressure which is calculated from relative humidity, RH as follows:

$$\rho = (RH/100) \times 6.105 \exp\left(\frac{17.27T_A}{237.7 + T_A}\right)$$

As the equation makes clear, wet bulb temperatures are a non-linear function of temperature. Numerically, this measure is smaller than ambient temperatures and is also a compressed scale. That is to say, a one degree increase in wet bulb temperatures corresponds to a greater than one degree increase in maximum temperatures. For these reasons we cannot use the same cut-offs here as with maximum temperatures. Instead we use the cutoffs specified below:

Garment workers in the NCR: 0,15,17,19,21,23,25,27,100 Garment workers in Hyderabad and Chhindwara: 0,23,25,27,29,31,100 Weaving workers in Surat: 0,20,21.5,23,24.5,26,27.5,100 Steel workers in Bhilai: 0,21,23,25,27,29,100

In order to calculate wet bulb temperatures for all sites we need to obtain humidity measures. Except for workers in the NCR, we use reanalysis measures purchased from Weather Online in all cases. For NCR plants we use information from a weather station in the Indira Gandhi International Airport (Station 33934 in the National Climatic Data Center dataset of weather stations across the world).

On the one hand we expect there is significant measurement error in the WBT estimated this way partly because humidity varies quite significantly over small scales. On the other hand human physiology would suggest that WBT is a better indicator of heat stress than dry temperatures. It is therefore useful to see how these results compare to those obtained using maximum temperatures.

As Figure A.3 shows, we find the same patterns of output response to heat as we do when using maximum temperatures. If anything, standard errors are smaller and effect sizes slightly larger which is consistent with what we would expect when using a non-linear scale that better approximates heat stress.

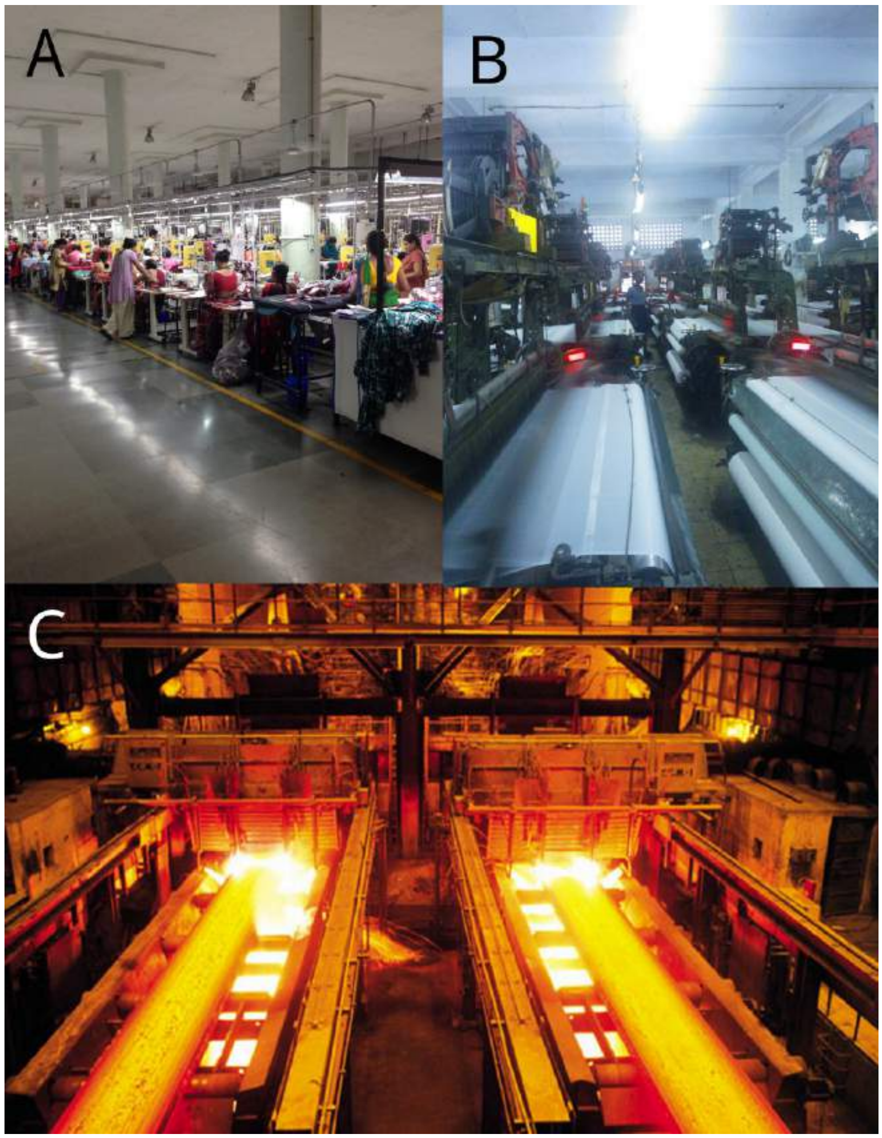

Figure A.1: Garment sewing plants (Panel A), Cloth-weaving plants (Panel B), Production floor of steel mill (Panel C)

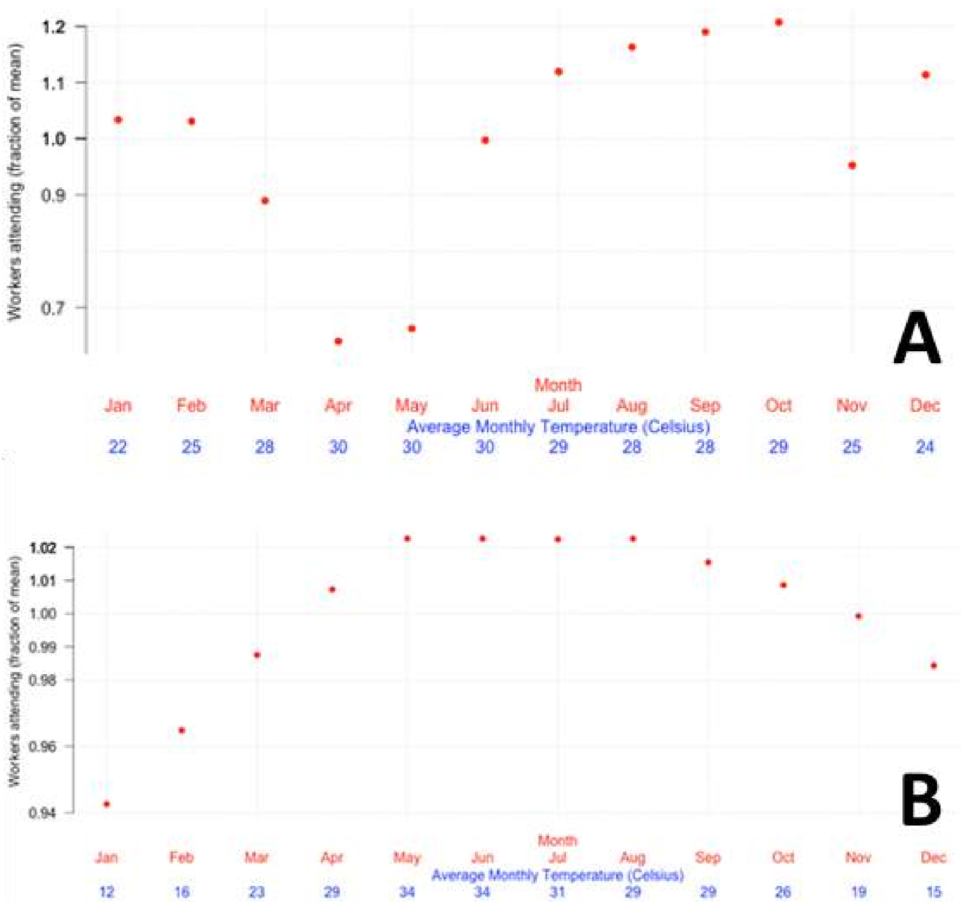

Figure A.2: Worker attendance by month for daily wage workers in cloth-weaving factories (Panel A) and salaried workers in garment plants (Panel B)

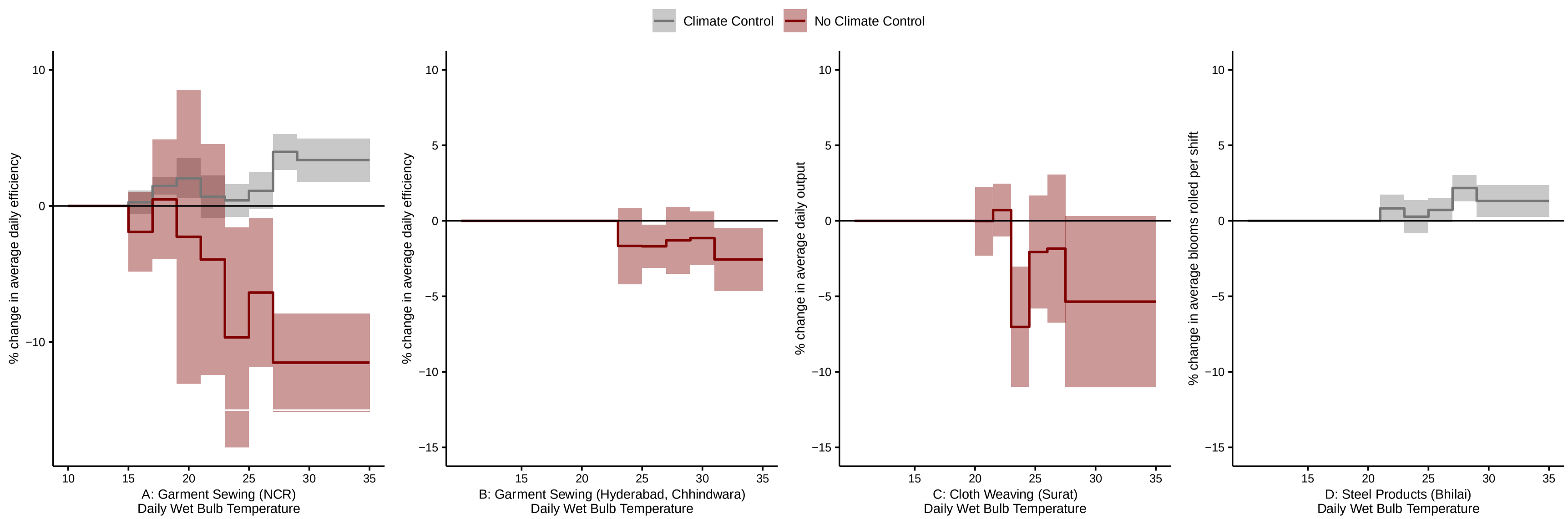

Figure A.3: *The effect of wet bulb temperature on worker output.* Estimates are percentage changes in daily output - averaged over a week - for a day in the week moving to a hotter temperature bin from the coolest (omitted) bin. Shaded areas represent 90 percent confidence intervals using robust standard errors clustered at the worker level. The number of bins varies across worker types, reflecting differences in observed wet bulb temperatures. Panel A: Garment sewing lines in NCR, Panel B: Garment sewing lines in Hyderabad, Panel C: Cloth weaving in Surat, Panel D: Steel mill in Bhilai. The output variable for garment plants (A,B) is defined as the logarithm of the 'efficiency' measure of each sewing line. In Panel C, the output variable is defined as the inverse hyperbolic sin transformation of the meters of cloth woven by a worker. In Panel D, the output variable is the logarithm of blooms rolled by a team of workers.

A.4 Annual Survey of Industry Data Cleaning

This section describes how our 15-year panel is constructed from the Annual Survey of Industry (ASI) datasets purchased from the Indian Central Statistical Office.

Between 1998 and 2007, ASI data are available in two forms. The first is a panel with plant identifiers and no district identifiers, while the second is a cross section with district identifiers but no plant identifiers. For these years we purchased and merged these two datasets to obtain a panel containing district identifiers. We use the state code, NAICS code, year of starting operations, and value of output to complete this matching.

At the time of writing, the latest survey with micro-data available for sale was for financial year 2012-2013. From 2008 until 2012 only a cross-section *without* district identifiers is available so the above procedure cannot be used. For these years, we first list plants that are uniquely identified based on *time-invariant* characteristics. These include the state location, the four-digit industry codes (NAICS) and the year operations started. We then search in each year of our panel (1998-2007) for matches based on these three characteristics. All such matches are associated with a firm identifier. When there is only a unique match in the panel, the corresponding observation from the 2008-2012 surveys is accordingly assigned this firm identifier and enters the panel. This matching process requires searching over all years in the panel because plants are not necessarily surveyed every year. In cases where these time-invariant characteristics do not identify a unique plant in the non-panel years (2008-2012), or do not match to a unique plant in the panel years (1998-2007), the corresponding observation is given a new firm identifier.

Most matches are completed this way. A few additional matches were obtained using two additional variables: the start-of-year cash on hand, and the end-of-year cash on hand. For any plant surveyed in successive years t and t+1, the end of year balance in year t must be the same as the start of year balance in year t+1.

After constructing the panel, we performed the following data cleaning operations:

1. Removed observations where values of output, workers, total wages, value of capital, or total value of inputs is less than or equal to zero. We also drop observations with missing values for these variables.

2. The ASI dataset also contains observations with implausibly high or low reported values of output. For instance there are plants with reported annual output less than a few dollars. We drop the top 2 percent and bottom 2 percent of values of output. This is done to transparently eliminate these outliers.

3. As with output, we also have a smaller number of implausible values in the number of reported workers (over 10000 workers in a factory), total value of the capital measure (less than 2 USD), and value of total wages aggregated across all employees (less than 300 USD per year). We drop such values, together making up only 0.5 percent of our data. Our results are robust to ignoring this step.

4. We drop plants where the reported state or district changes over the panel duration. Misreported locations will induce significant errors when assigning temperatures to these plants.

5. We drop plants observed only once in the panel.

Our final sample has 53,015 manufacturing plants distributed all over India (Figure A.4) spanning different industry sectors. These plants are matched to district temperature and precipitation measures as described in the text. The figure also shows the sites for our daily worker data.

To calculate district average temperatures, we use a gridded dataset sold by the Indian Meteorological Department. The resolution of the original temperature grid is at the 1° level. We first create a finer grid by linear interpolation down to 0.083° (5 arc-minutes), and then average over all points falling within district polygon boundaries.

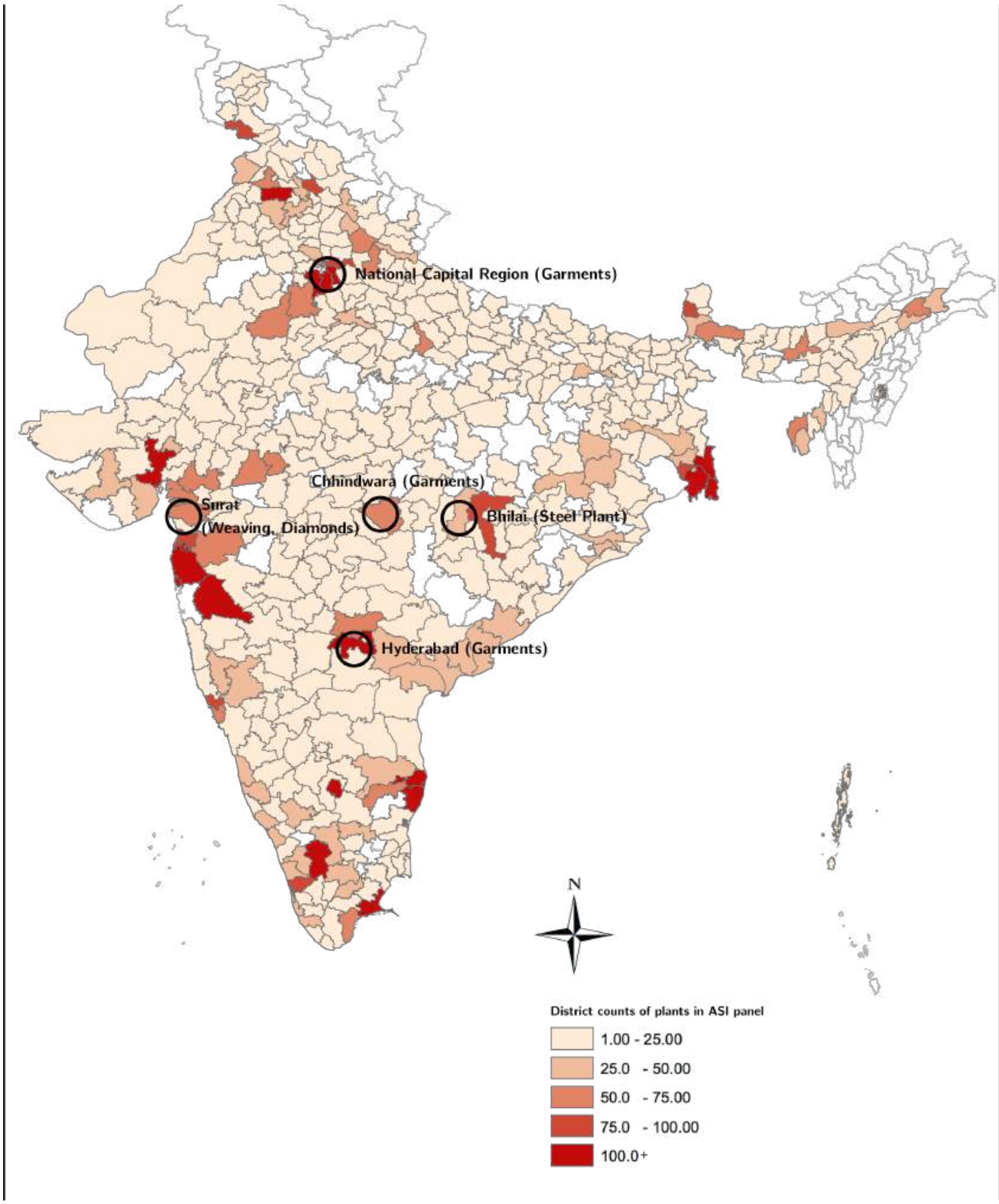

Figure A.4: Distribution of ASI plants over Indian districts, and location of micro-data sites

A.5 Additional specifications using plant data

In addition to the results from our main empirical specification presented in Section 4.2, we estimate models that include interactions of each factor input in turn with temperature bins, as well as alternative ways of controlling for unobserved time shocks. Results from these are in Table A.1.

Column 1 provides OLS estimates from a production function specification *without* any temperature interactions. Columns 2 and 3 present the specification from (3) additionally control for state-specific quadratic time-trends and state-year dummies respectively. Column 4 controls only for labor, and Column 5 additionally for labor-temperature interactions. Columns 6 and 7 do the same for capital and capital-temperature interactions. We continue to see the negative effects of temperature, with our estimates less precise in some of these models.

Although we are interested primarily in the interaction of factor inputs with temperature, these OLS estimates, and the ones in the main body of the paper, are unbiased only if the specific measures of labor and capital are relatively unresponsive to short-term shocks. As we discuss in Section 4.2, we believe this is a reasonable assumption. In the main text (Table X) we also present Levinsohn-Petrin first-stage estimates that allow for flexible labor inputs.

The Levinsohn-Petrin estimator uses a control function approach to removing bias. The control function is a flexible function of the capital measure and variable intermediate inputs m that are correlated unobserved productivity shocks. As is common in the literature, we let m be the total value of the ten largest inputs used by the plant.

Then we estimate the model below, where ϕ is the sum of a fully interacted two degree polynomial in capital k and the material input measure m, and a similar polynomial in capital k and temperature bins T^j.

$$y_{it} = \alpha_i + \gamma_t + \beta_o \cdot l + \sum_{j=2}^{5} \beta_j T^j l + \phi(k_{it}, m_{it}, T^j_{it}) + \theta R_{it} + \epsilon_{id} \tag{8}$$

The coefficients on labor and the labor-temperature interactions, β_j, can be estimated using OLS. These are of primary interest to us and are correspondingly reported in the main paper.

In addition to the models in Table A.1 we also reproduce Table 2 in the main text using a different set of bin-cutoffs. Our results are provided in Table A.2 and reflect similar patterns to our original specification.

A.6 Alternative Explanations

Power Outages

High temperatures are often accompanied by power outages in India so it is legitimate to ask whether these outages could be partly responsible for the temperature effects we observe.

We investigate the possible impact of outages by using annual, state-level measures of supply shortfalls published by India's Central Electricity Authority in its annual *Load Generation Balance Report*. This measure is the difference between an imputed measure of average monthly electricity demand and average monthly electricity supply. This difference takes a negative value in a little over two percent of our observations and is zero in a few others. These cases denote years with no shortfalls. We set the negative observations equal to zero and take the logarithm of this truncated difference as a proxy for outages. To handle our zero observations we add a small constant (1) before taking logarithms. Note that the mean value of the difference is 387 MWh and zeros are very uncommon.

We find that introducing this logged outage proxy into (3), our main specification, leaves our temperature effects intact. These are shown in Column 2 of Table A.3.

Table A.1: Effect of temperature on plant output: Additional Specifications

	logoutput						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
l	0.60526***			0.66521***	0.86117***		
	(0.00655)			(0.00808)	(0.09566)		
k	0.13645***					0.21999***	0.20433**
	(0.00318)					(0.00585)	(0.05674)
T^2		−0.00093*	0.00063	−0.00111**	0.00256**	−0.00118**	0.00008
		(0.00048)	(0.00107)	(0.00035)	(0.00097)	(0.00042)	(0.00211)
T^3		−0.00130**	−0.00028	−0.00099**	0.00147	−0.00129**	−0.00205
		(0.00056)	(0.00111)	(0.00039)	(0.00103)	(0.00047)	(0.00229)
T^4		−0.00160***	−0.00071	−0.00099**	0.00081	−0.00140**	−0.00094
		(0.00059)	(0.00119)	(0.00041)	(0.00108)	(0.00049)	(0.00237)
T^5		−0.00183***	−0.00127	−0.00110**	0.00003	−0.00168***	−0.00499*
		(0.00067)	(0.00120)	(0.00044)	(0.00118)	(0.00055)	(0.00259)
$l \times T^2$					−0.00098***		
					(0.00027)		
$l \times T^3$					−0.00067**		
					(0.00027)		
$l \times T^4$					−0.00052*		
					(0.00027)		
$l \times T^5$					−0.00036		
					(0.00029)		
$k \times T^2$							−0.00009
							(0.00015)
$k \times T^3$							0.00005
							(0.00016)
$k \times T^4$							−0.00003
							(0.00016)
$k \times T^5$							0.00022
							(0.00018)
State-Trends	N	Y	N	N	N	N	N
State-Year FE	N	N	Y	N	N	N	N
Observations	179,107	179,107	179,107	179,107	179,107	179,107	179,107

Notes: Data on the value of output and inputs are at the plant level from the Annual Survey of Industry. *, **, *** denote estimates significant at 10, 5, 1 percent level. Standard errors are corrected for serial and spatial correlation following Conley (2008). Coefficient on rainfall omitted for brevity. All models include plant and year fixed effects. Temperature bins are defined as $\{(0,20],(20,25],(25,30],(30,35],(35,50]\}$ and T^j is the number of days in the jth bin. T^1 is the omitted bin in all models.

Table A.2: Effect of hot days on plant output (Alternative Bins

Panel A							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Day in Bin 2	-0.00135***	-0.00088*	-0.00140***	-0.00159***	-0.00158**	-0.00012	-0.00023
	(0.00044)	(0.00053)	(0.00045)	(0.00056)	(0.00066)	(0.00037)	(0.00014)
Day in Bin 3	-0.00208***	-0.00157***	-0.00205***	-0.00239***	-0.00093*	-0.00029	-0.00037*
	(0.00054)	(0.00059)	(0.00054)	(0.00057)	(0.00052)	(0.00035)	(0.00022)
Day in Bin 4	-0.00173***	-0.00139**	-0.00179***	-0.00206***	-0.00150***	0.00039	-0.00055*
	(0.00051)	(0.00057)	(0.00051)	(0.00055)	(0.00056)	(0.00266)	(0.00033)
Day in Bin 5	-0.00216***	-0.00178***	-0.00221***	-0.00252***	-0.00150*	-0.00132*	-0.00114*
	(0.00056)	(0.00062)	(0.00056)	(0.00060)	(0.00079)	(0.00069)	(0.00069)
Mean shift by 1 degree	-0.01556**	-0.01580**	-0.01646**	-0.01774**	-0.01762*	-0.01887*	-0.01670*
	(0.00696)	(0.00683)	(0.00697)	(0.00772)	(0.00977)	(0.00979)	(0.01003)
Projected warming 2075-80	-0.05197**	-0.05046**	-0.05338**	-0.05795**	-0.04466	-0.07727*	-0.06680*
	(0.02106)	(0.02133)	(0.02118)	(0.02338)	(0.04290)	(0.04063)	(0.04012)
Plant FE	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y
State-Year Polynomial	N	Y	N	N	N	N	N
Conflict and Flood Controls	N	N	Y	N	N	N	N
Outage Controls	N	N	N	Y	N	N	N

Notes: Rows 1-4 present the effect on log output of an additional day in temperature bin T^j relative to T^1 for different specifications relating output to temperature. Bins are defined as $\{(0,20],(20,24],(24,28],(28,32],(32,50]\}$ and T^j is the number of days in the jth bin. Row 5 presents the effect of a one degree increase in the temperature of every day in the year. Row 6 presents the effect of shifting outwards the temperature distribution for each year in our data to simulate a warmer future temperature distribution averaged over 2075-2080, projected using the Hadley GEMS2 climate model under the RCP 8.5 scenario. Column 1 provides estimates from our preferred specification in (X). Columns 2-4 add additional controls for state-year quadratic time trends, floods, conflict, and outages. Column 5 presents estimates from a model where temperature is described by the number of degree days falling in the temperature ranges described above. Column 6 presents estimates from a model where daily output is assumed to depend on a quadratic polynomial of daily temperature. Column 7 presents estimates from a model where daily output is assumed to depend linearly on temperature. Data on the value of output and inputs are at the plant level from the Annual Survey of Industry.*, **, *** denote estimates significant at 10, 5, 1 percent level. Standard errors are corrected for serial and spatial correlation following Conley (2008).

These results are not surprising because large plants are typically served by dedicated high voltage (33kV) grid feeders with fixed supply schedules. When load shedding is unavoidable, these feeders are generally shed last, so that only large grid disruptions will percolate down to plants served by high voltage lines. As a result, unscheduled, temperature-dependent outages are relatively rare.

Floods and Conflict

Among natural disasters in India, floods are the most widespread and directly reduce industrial output. For example, in the industrial city of Surat (the site of our weaving workers and diamond firms), there were floods in 1998, 2006, and 2013. Likewise, industrial disputes are a relevant measure of conflict for manufacturing plants. Both of these can severely disrupt manufacturing activity.

For floods, we use data from the Dartmouth Flood Observatory Archive. These data combine remote sensing information, news stories, government releases, and ground instruments to measure the severity, duration, and damage caused by each flooding incident. The magnitude of each flood is defined as log(Duration x Severity x Affected Area). For each year, and each state, we use the total magnitude of all flooding as a proxy for the flood exposure of all plants in a state.

For conflict, we use the total number of workday minutes lost every year due to industrial disputes in each state. These data are published by India's Ministry of Labor and Statistics as part of its annual publication *Statistics on Industrial Disputes, Closures, Retrenchments and Lay-Offs.* This statistic takes only non-zero values and has a skewed distribution. We use the logarithm of lost minutes as a proxy for the annual exposure to conflict for each plant. [28]

[28] Our data are complete from 2003 onwards and missing for a few years before that.

We modify (3) to include these measures as additional regressors. Denoting by M_{it} and C_{it} our measures of flooding and conflict for year t in the state in which plant i is located, we estimate:

$$y_{it} = \alpha_i + \gamma_t + \sum_{j=2}^{5} \beta_j T_{it}^j + \theta R_{it} + \omega_1 M_{it} + \omega_2 C_{it} + \epsilon_{it}. \quad (9)$$

Our estimates are in Column 3 of Table A.3. The estimated coefficients on temperature bins are very similar to those from (3), reported in the main text in Table 4 Column 2, and reproduced here in Column 1.

Table A.3: Temperature effects on output controlling for outages, floods and disputes

	Log Plant Output		
	(1)	(2)	(3)
T^2	-0.00139^{**}	-0.00173^{**}	-0.00145^{**}
	(.00044)	(.000494)	(0.00044)
T^3	-0.00164^{***}	-0.00183^{***}	-0.00163^{***}
	(.00049)	(.00053)	(0.00049)
T^4	-0.00189^{***}	-0.00225^{***}	-0.00193^{***}
	(.00051)	(.00056)	(0.00051)
T^5	-0.00217^{***}	-0.00238^{***}	-0.00223^{***}
	(.00056)	(.00063)	(0.00057)
log outages		-0.01283^{**}	
		(0.00450)	
M			-0.00055
			(0.00049)
C			-0.00092
			(0.00176)
N	181442	177,916	177,916

Notes: Data on the value of output are from the Annual Survey of Industry. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors are corrected for serial and spatial correlation following Conley (2008). All models include plant and year fixed effects. Rainfall coefficients omitted for brevity. M denotes the total magnitude of flood exposure and C is the log of total minutes lost to industrial disputes. (1) reproduces estimates from Table 4, Column 2 in the main text. T^1, T^2, T^3, T^4, T^5 are days in (0,20], (20,25], (25,30], (30,35],(35,50] bins and all bin coefficients are relative to T^1

A.7 Price Shocks

High temperature years could raise input prices and thereby induce firms to reduce output. These higher prices could result from temperature-induced productivity shocks in other sectors. The Annual Survey of Industry reports prices of the ten most important inputs for each plant, based on total expenditure shares. We regress the logarithm of price of the *primary* input for each plant (the one with the largest expenditure among all inputs) on linear and binned specifications of temperature as below:

$$log(P_{it}) = \alpha_i + \gamma_t + \beta T_{it}^a + \theta R_{it} + \epsilon_{it}. \tag{10}$$

$$log(P_{it}) = \alpha_i + \gamma_t + \sum_{j=2}^{5} \beta_j T_{it}^j + \theta R_{it} + \epsilon_{it}. \tag{11}$$

Table A.4 reports results from both specifications. We find no evidence that temperature influences the price of plant input materials. This does not imply that long-term changes in the number of hot days in a year will leave prices unaffected, only that this factor cannot explain the results in this paper.

Table A.4: Temperature Effects on Price

	Price of Primary Input	
	(1)	(2)
T^a	-0.0237	
	(0.029)	
T^2		0.0005
		(.0017)
T^3		0.0004
		(.0020)
T^4		0.0002
		(.0020)
T^5		0.0012
		(.0021)
Observations	$144,531$	$144,531$

Notes: Data on input prices are available from the Annual Survey of Industry. $^{***}p < 0.01$; $^{**}p < 0.05$ $^{*}p < 0.1$. Standard errors are corrected for serial and spatial correlation following Conley (2008). All models have plant and year fixed effects. Rainfall coefficients omitted for brevity. Prices are in Indian Rupees, $(T^1, T^2, T^3, T^4, T^5)$ are days in $\{(0, 20], (20, 25], (25, 30], (30, 35], (35, 50]\}$, and all coefficients are relative to T^1.

A.8 Leads and Lags in Temperature Variables

As a robustness check, we also estimate a variant of (3) including both lags and leads of the highest temperature bin. We find that only the contemporaneous temperature bin explains any variation in output, with both lags and leads statistically indistinguishable from zero. Figure A.5 shows estimates for the coefficient on the highest temperature bin for three lagged years, the contemporaneous year, and three years in the future.

A.9 Sector-wise Temperature Effects

To study heterogeneous impacts across manufacturing sectors, we regress plant output on annual average maximum temperatures interacted with an indicator for each 2-digit manufacturing sector:[29]

$$y_{it} = \alpha_i + \gamma_t + \beta_o T^a_{it} + \beta_k (T^a_{it} \times S_k) + \omega_k S_k + \theta R_{it} + \epsilon_{it}. \tag{12}$$

S_k is a sector dummy and other variables are defined as before. Figure A.6 shows sector-wise temperature effects on output plotted with 90 percent confidence intervals. Plotted estimates in A.6 are the sector effects $\beta_o + \beta_k$. Temperature effects on output are consistently negative, to varying degrees.

We also use an alternative method to evaluate the robustness of our results to the composition of plants in our sample. There are 89 sectors represented in our plant panel at the 3-digit level. From this list, we draw with replacement a sample of 40 sectors. Using the subset of plants belonging to these sectors, we evaluate our main specification (3). We repeat this 100 times and in Figure A.7 we plot a histogram of the effect of an additional day in the highest temperature bin. Each of these estimates can be viewed as being drawn from an economy with a different, and less diverse, manufacturing sector composition than India as a whole. Across all but one draw we see negative temperature effects on output, varying from close to zero to a reduction of slightly under 0.5 percent of output, with a mean close to the overall effect size of -0.22 percent reported in the paper.

[29] We use the ISIC system to define sectors. This is the same as the industrial classification used in India up to the 4-digit level.

A.10 Adaptation

A.10.1 Evidence from ASI Plant Panel

We introduce time trends into (3) to examine the changing relationship between temperature and plant output changes over the study panel:

$$y_{it} = \alpha_i + \gamma_t + \sum_{j=2}^{5} \beta_j T_{it}^j + \sum_{j=2}^{5} \delta_j T_{it}^j \times t + \theta R_{it} + \epsilon_{it}. \tag{13}$$

Estimates are in Table A.5. As before, T_{it}^j are the number of days in our different temperature bins, omitting the lowest bin. We now also interact these with t, a continuous time trend. We find evidence of decreasing temperature sensitivity over time, as shown in Column 1 of Table A.5.

We also examine how the temperature-output relationship varies by the electricity intensity of the plant. We do this because climate control is electricity intensive and, as we have seen in our worker data, has the potential to eliminate heat stress at work. We define electricity intensity as the ratio of electricity purchased to the number of employees and create a dummy variable that takes the value 1 when plant electricity intensity is above the median. We include this as an additional regressor in (3) and also allow for interactions with the temperature bins. These results, in Column 2 of Table A.5, show that the output of electricity-intensive plants is less sensitive to temperature. Note that these results are net of plant fixed effects and therefore do not simply represent a comparison of larger and smaller plants but rather increases in electricity use within plants.

Table A.5: Variation in Temperature Sensitivity over Time and by Electricity Intensity

	Log Plant Output	
	(1)	(2)
T^2	-0.43543	-0.00218^{***}
	(0.33903)	(0.00079)
T^3	-0.64529^{***}	-0.00251^{***}
	(0.21473)	(0.00070)
T^4	-0.40682^{*}	-0.00281^{***}
	(0.21175)	(0.00074)
T^5	-0.75642^{***}	-0.00312^{***}
	(0.23529)	(0.00079)
$T^2 \times t$	0.00022	
	(0.00017)	
$T^3 \times t$	0.00032^{***}	
	(0.00011)	
$T^4 \times t$	0.00020^{*}	
	(0.00011)	
$T^5 \times t$	0.00038^{***}	
	(0.00012)	
elec intensity		-0.51733^{**}
		(0.23857)
$T^2 \times$elec intensity		0.00166^{*}
		(0.00091)
$T^3 \times$elec intensity		0.00172^{***}
		(0.00066)
$T^4 \times$elec intensity		0.00185^{***}
		(0.00066)
$T^5 \times$elec intensity		0.00195^{***}
		(0.00071)
Observations	179,107	175,604

Notes: Data on the value of output are from the Annual Survey of Industry. *, **, *** denote estimates significant at 10, 5, 1 percent level. Standard errors are robust and clustered at the district level. All models include plant and year fixed effects. Rainfall coefficients omitted for brevity. *elec intensity* is a dummy indicating whether a plant has an electricity intensity that is higher than the median.

A.10.2 Costs and benefits in cloth weaving firms

Although we find cloth weaving workers are less productive on hot days, we do not see these firms invest in climate control. This is in contrast to our garment sewing firm, which gradually introduced climate control in all its plants. One reason for this might be that weaving workers have limited value added since they operate early in the garment supply chain. In this section, we use rough estimates of energy and wage costs from these plants to do a back-of-the-envelope cost-benefit analysis of climate control.

Our three cloth-weaving firms collectively produce a median daily output of about 7200 meters of cloth and workers are paid INR 2.0 per meter, implying a median daily wage bill of about INR 14,400. Cooling the shop-floors of all three factories would require an air conditioning load of roughly 24 tonnes or 84 KW. At the time of our data collection, electricity tariffs for industry in the state of Gujarat were about 5 INR per KWh. Assuming 8 hours of operation, daily air conditioning costs would be INR 3360. The costs of climate control would therefore be about about 23 percent of the total wage bill. Given our estimates of a reduction in productivity by about 2 percent per degree rise in temperature, these investments are unlikely to be profitable for firms with small price mark-ups. For such firms, the negative effects of increasing temperatures may not be mitigated by technological solutions.

A.10.3 Selective climate control in diamond firms

This section presents results using data from a survey of 150 diamond cutting factories located in the city of Surat. These firms have much higher value-added than our cloth-weaving firms and we find they exhibit different behavior even though they are in the same city. The surveyed factories were drawn randomly from a list of all plants registered with the local diamond industry association. Each plant uses five main processes: (i) sorting and grading, (ii) planning and marking cuts in the stone (iii) bruting (rounding a diamond), (iv) cutting, and (v) polishing. These vary in the amount of labor they use, and in their contribution to overall diamond value.

We observe the presence or absence of air-conditioning in the different rooms in which these activities take place. With 5 processes in each of the 150 firms, we have 750 observations. We estimate the probability of using air conditioning as a function of firm and process characteristics. For firm characteristics we use firm size as measured by the number of workers, as well as the age of the firm in years. For process characteristics we use labor intensity (defined as the share of total workers engaged in the process), mechanization (defined as the share of total machines used in the process), and an 'importance' rating. The importance rating is a self-reported assessment by management on a scale of 1 to 5. Based on this rating we construct an importance dummy which takes the value 1 when the manager rating is 5.

The middle three processes (marking, rounding and cutting) are largely manual. On average, they accounted for less than 10 per cent of the machines used. Cutting, for instance, was rated one of the most important process in determining final quality but accounted for only 4.5 percent of the machines in the firm. Over 98 percent of firms in our sample used climate control in the rooms where diamond cutting occurred. This contrasts with polishing, also an important process, but accounting for 63 percent of machines used. Only 33 percent of rooms where polishing occurred were climate controlled.

Figure A.8 shows marginal effects for a standard logit model. These estimates are consistent with firms choosing to optimally allocate air conditioning investments to protect workers. These estimates suggest that if the share of workers in a process were to fall by 10 percent, the probability of observing climate control for that process reduces by 0.17. Correspondingly other processes would see an increase in the probability of air-conditioning, as these workers are moved elsewhere. A process characterized as very important by the firm management has a 0.27 higher probability of climate control. Replacing labor by machines significantly reduces the probability of observing climate control investments.

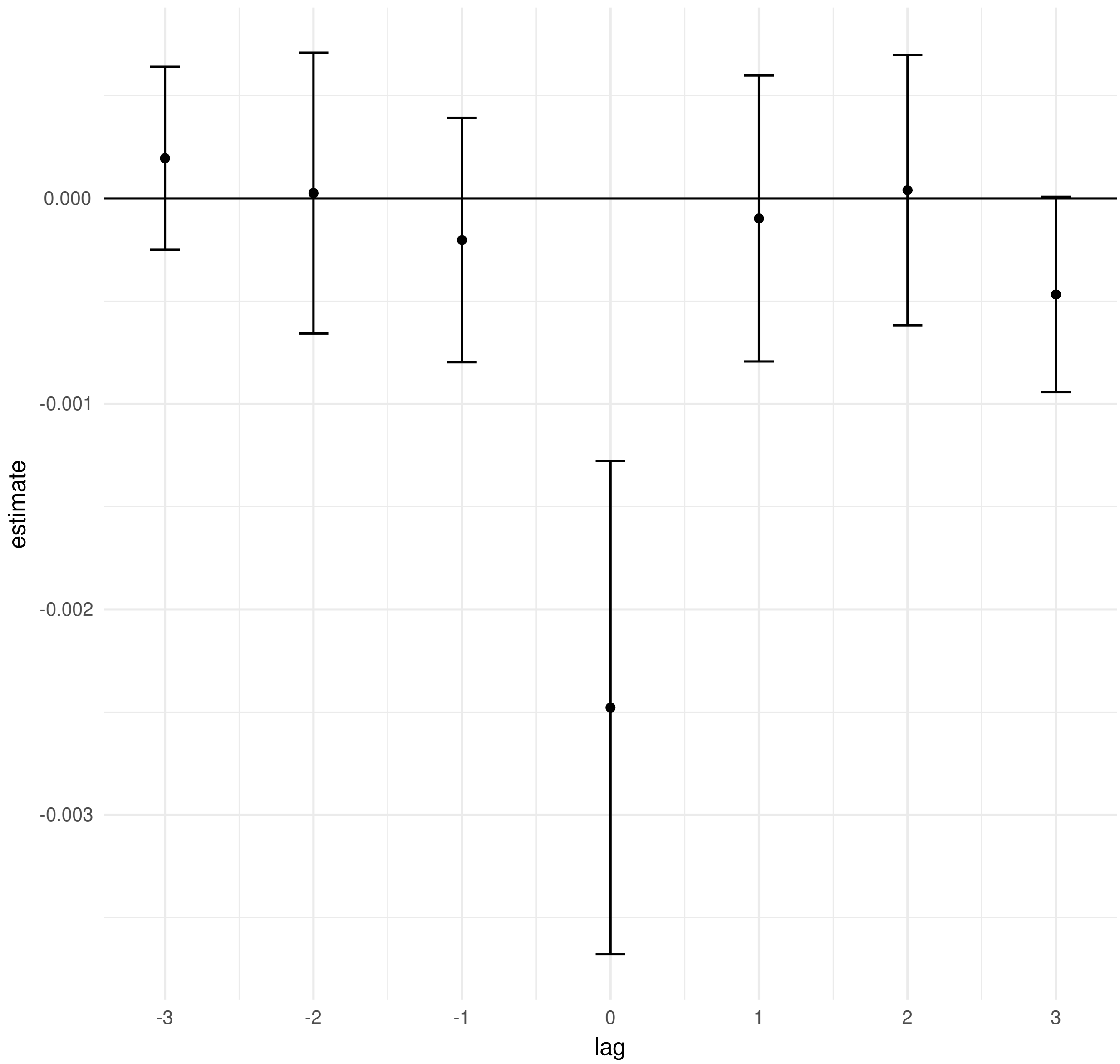

Figure A.5: Effect of lagged temperature bins on plant output.

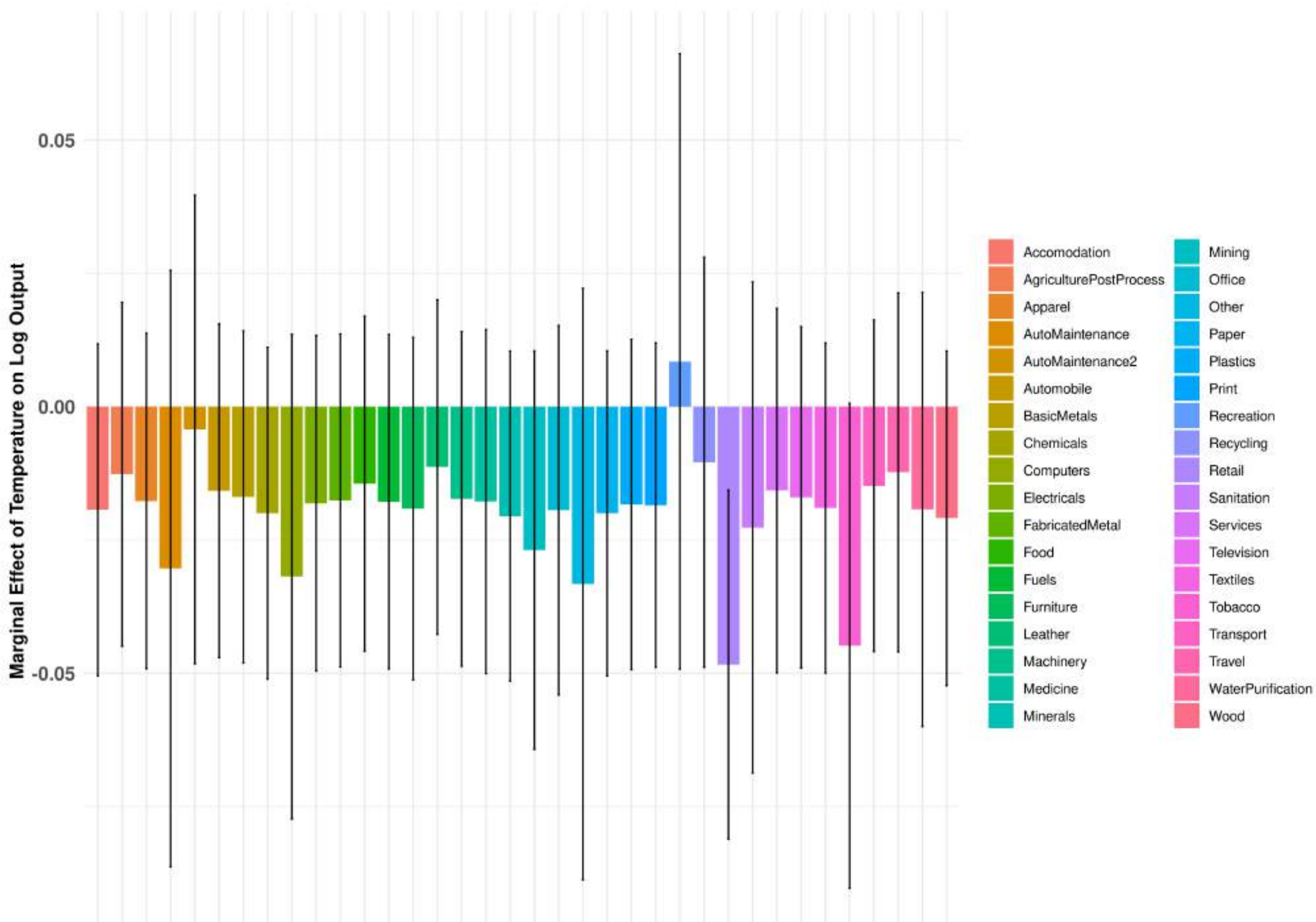

Figure A.6: Sector-wise percentage change in output for a one degree increase in annual averages of daily highs.

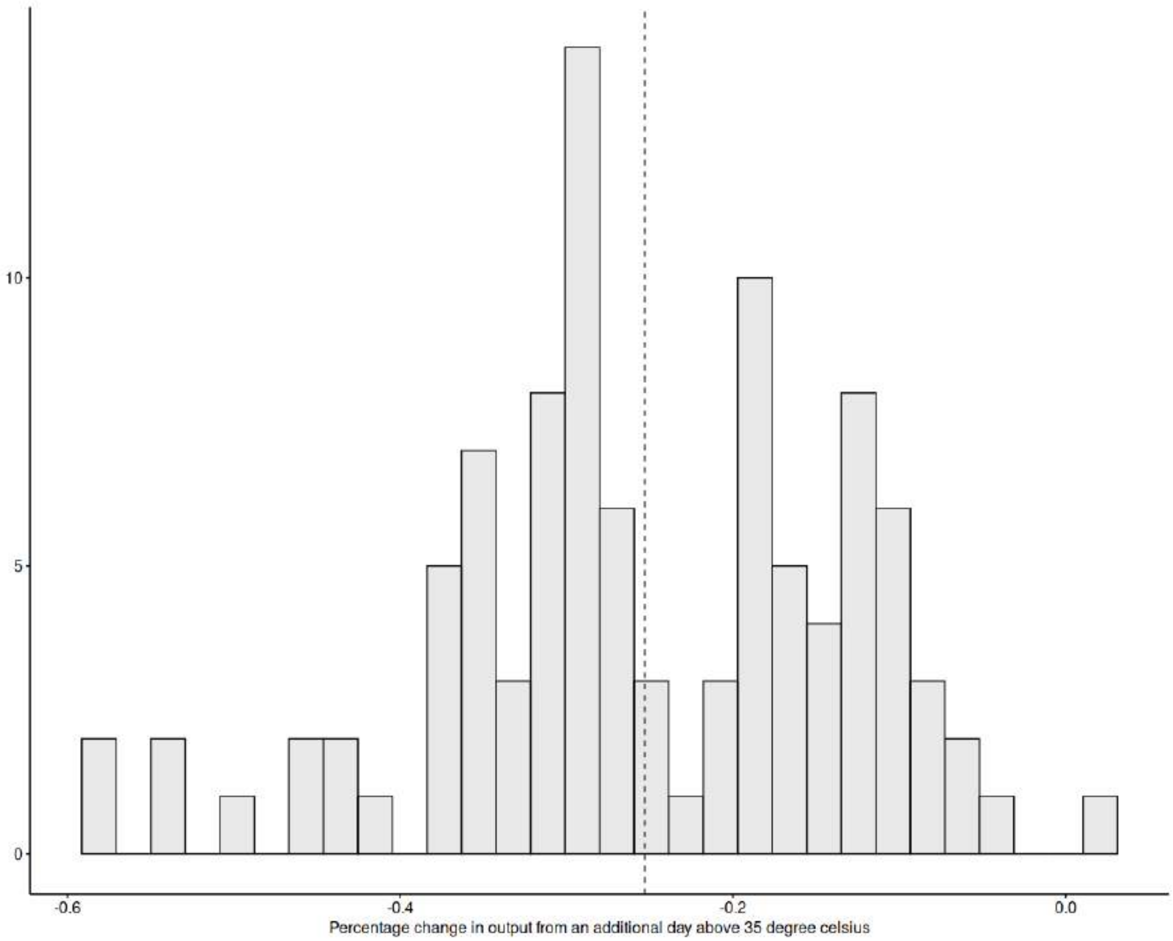

Figure A.7: Temperature effect on output for different manufacturing sector compositions.

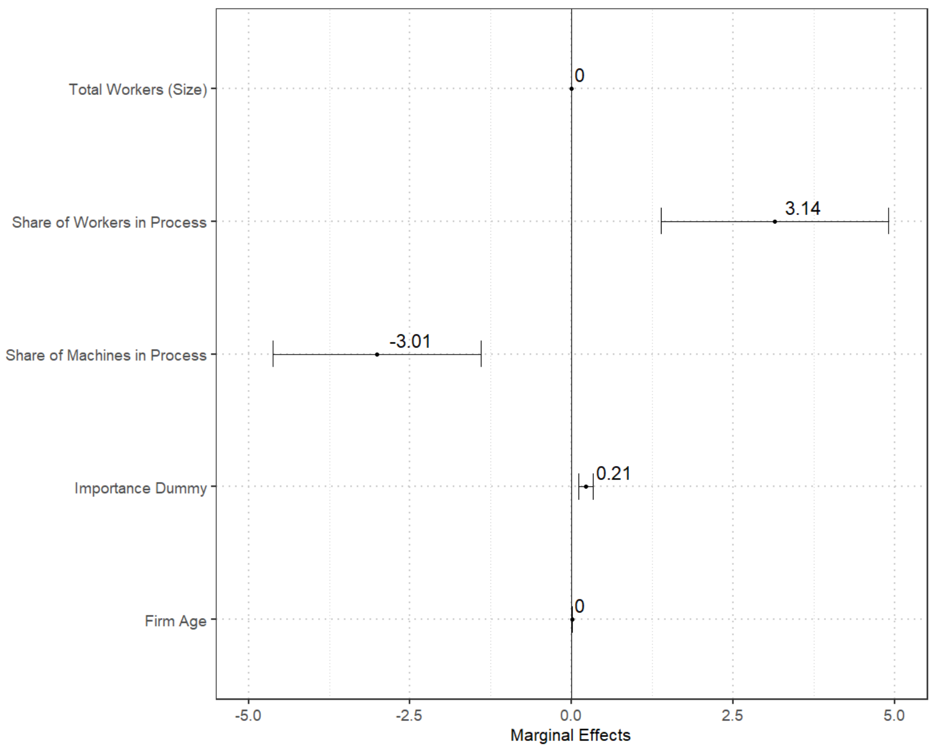

Figure A.8: Average marginal effects for logit model describing diamond firm decisions to selectively invest in climate control for different processes.