WORKING PAPER · NO. 2018-69

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

E. Somanathan, Rohini Somanathan, Anant Sudarshan, and Meenu Tewari

AUGUST 2018

Becker Friedman Institute
FOR ECONOMICS AT UCHICAGO

1126 E. 59th St, Chicago, IL 60637
Main: 773.702.5599
bfi.uchicago.edu

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 on hot days declines by 2 to 4 percent per degree celsius. Sustained heat also increases worker absenteeism. Using a national panel of manufacturing plants, we find similar temperature effects on output and show that these can be fully accounted for by reductions in the productivity of labor. Estimated effect sizes are consistent with studies that rely on country GDP panels.

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. We are grateful to Mehul Patel for field assistance and to Sheekha Verma for research assistance. This paper has benefited from several helpful comments from anonymous referees and conference attendees at the NBER Summer Institute, the North-East Universities Development Conference, the Indian Statistical Institute Economics Department, and the University of Gothenburg.

†Corresponding Author.

1 Introduction

Recent research has uncovered a negative relation between temperature and aggregate national output, especially in developing countries. High temperatures have been shown to reduce crop yields and also appear to 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 use data from India at different levels of aggregation to quantify the extent to which high temperatures affect output through reductions in the productivity of labor. There are two channels through which temperature might affect workers. They may produce less while at work and also be absent more often. We separately identify both of these effects using data on individual workers from selected firms in three industries: cloth weaving, garment sewing, and steel infrastructural products. We find that productivity drops by 2 to 4 percent per degree celsius on a hot day. Hot spells increase absenteeism among salaried workers but not those with piece rate contracts. Climate control in the workplace mitigates contemporaneous productivity declines but not absenteeism.

After estimating these effects, we examine how temperature influences factory output using a 15-year nationally representative panel of manufacturing plants. We find that the value of output declines in years with more hot days and that changes in the output elasticity of labor, in response to high temperatures, can entirely account for this effect. 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 annual output

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

[2] Hsiang (2010) discusses heat stress, Hsiang, Burke, and Miguel (2013) identify a temperature-conflict relationship and Burgess et al. (2011) study effects on mortality.

declines by about 3 percent per degree celsius. This effect size closely matches predictions using our plant data and is comparable to existing country-level estimates of the response of contemporaneous manufacturing GDP to high temperatures (Dell, Jones, and Olken, 2012; Burke, Hsiang, and Miguel, 2015).

Notwithstanding the magnitude 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. We also use a survey of 150 diamond cutting plants to study adaptation investments *within* plants and we find that air-conditioning is selectively used in processes which are both labor intensive and critical in determining diamond quality. Turning to our national plant panel, although we do not have direct information on the use of climate control, we find that temperature effects on output fall slightly over time, possibly the result of investments in adaptation.

After presenting our main results, we consider two alternatives to heat stress as an explanation for the effect of temperature on labor productivity: natural disasters and conflict. For the years covered by our plant panel, we collect data on workdays lost in all recorded industrial disputes as well as all instances of flooding. 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 magnitudes and 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 Mechanisms

The physics of how temperature affects human beings is straightforward. Heat generated while working must be dissipated to maintain body temperatures and avoid heat stress. The efficiency with which this happens depends primarily on ambient temperature but also on humidity and wind speed. If body temperatures cannot be maintained at a given activity level, it may be necessary to reduce the intensity of work (Kjellstrom, Holmer, and Lemke, 2009; Iso, 1989). Several indices of ambient weather parameters have been used to measure the risk of heat stress. Most widely accepted is the Wet Bulb Temperature or WBT (Parsons, 1993; Iso, 1989). 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):

$$WBT = 0.567T_A + 0.216\rho + 3.38 \tag{1}$$

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

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

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. These temperatures are not considered unsafe from the point of view of occupational safety and commonly occur in many developing countries.[3]

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).[4] More recently, Adhvaryu, Kala, and Nyshadham (2014) 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 changes in total worker absences, within firms, in response to high temperatures.

[3] 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).

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

3 Data Sources

Our labor and output data are at three levels of aggregation: the worker or 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 referencing such a financial year, we use the initial calendar year.

3.1 Worker Data

We collected data on the output and attendance of workers from selected firms in three industries: cloth weaving, garment sewing, and the production of large infrastructural steel products. Our three cloth-weaving factories are located in the industrial city of Surat in the state of Gujarat, in Western India. Our garment factories are managed by a single firm, and located in several cities in North and South 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. Textiles and Garments respectively employ 12 and 7 percent of factory workers in India and the Bhilai steel mill is the largest producer of steel rails in the world.[5]

Cloth Weaving: For the three cloth-weaving factories, we gathered daily data on the meters of cloth woven and the attendance of 147 workers employed during the financial year starting April 2012. A worker in each of these factories operates about 6 mechanized looms

[5]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.

producing woven cloth.[6] 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 and the median daily production per worker was 125 meters.[7]

Garment Sewing: For garment sewing, we have data from eight factories owned by a single firm producing garments for foreign apparel brands. Six of the factories are in the National Capital Region (NCR) around Delhi, the other two are in Hyderabad and Chhindwara in South and Central India respectively. In each plant, production is organized in sewing lines of 10-20 workers with each line creating part or all of a clothing item.[8] 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.

Our productivity measure relates to the entire sewing line. The 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 operations 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 daily attendance records for

[6]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 dying and printing firms. In the Appendix, Figure A.1 provides a photograph of the production floor in one of these units.

[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).

[8]Lines are usually stable in their composition of workers, although the garment manufactured by a given line changes based on production orders. In the Appendix, Figure A.1 provides a photograph of a typical sewing line.

2744 workers over the same time period. To restrict attention to regular, full-time employees, we study absenteeism within a stable cohort of 2700 workers present for at least 600 days over our full two-year period.

These garment factories also provide us an opportunity to study the impact of climate control investments on productivity. During the period for which we have data, the firm was in the process of installing cooling and dehumidifying equipment on its shopfloors. This equipment had been installed in five of the manufacturing units in the NCR before 2012 but the sixth unit in the NCR did not have 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 for workers in co-located factories within the same firm, assigned to shopfloors with and without climate control. Since we observe both absenteeism and attendance we are able to separately estimate the effects of climate control on each of these.

Steel Production: 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.[9] 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. When a bloom is successfully shaped, it is said to have been *rolled* and the number of blooms rolled in an eight-hour shift is our output measure.

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

[9] Figure A.1 provides a photograph of the shop-floor.

over a shorter period of 857 working days between February 2000 and March 2003.[10]

Unlike the weaving and garment manufacture factories, 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 and the combination of automation and climate control mediates the effect of outside temperatures on output.

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.[11] The ASI provides annual data on output, 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.[12]

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 (2014), we create an unbalanced panel of 69643 plants over 1998 to 2012.[13] We match plants to temperature and rainfall at the level of the district.[14]

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

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

[12]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.

[13]Section A.4 in the Appendix provides details.

[14]There are 529 districts with at least one plant surveyed. Figure A.1 in the Appendix shows the geographic distribution of ASI plants and locations of our micro-data sites.

3.3 District Panel Manufacturing GDP

The Planning Commission of India has made available data on district-level manufacturing sector GDP over a 12 year period from 1998 to 2009.[15] 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.

[15] These figures include output from plants in the ASI panel, with appropriate weights. They also include estimates of unregistered manufacturing and smaller factories not covered by the ASI.

Table 1: Summary of Worker and Firm Data

Data Source	Location	Unit (# of obs)	Dependent Variables	Time Period	Climate Control
Cloth Weaving	Surat	Worker (147)	Meters of cloth woven, Worker Attendance	365 days	No
Garment Sewing	NCR, Hyderabad, Chhindwara	Sewing Line (103)	Operations completed, Worker Attendance	730 days	Partial (74 lines)
Steel Products	Bhilai	Shift-Team (9)	Blooms rolled, Team Absences	3339 days (Production), 857 days (Attendance)	Yes
Annual Survey of Industry	National	Plant (69,643)	Value of output	15 years	Not Observed

3.4 Weather Data

To identify the impacts of heat stress on output, we would like to use temperature data from close to the workplace while also accounting for humidity. Together these allow us to estimate the ambient wet bulb temperature, as discussed in Section 2. Ideally, if hourly temperatures and humidity measures were also available, we could average these to estimate the WBT experienced during working hours only. This type of information is rarely available over long time periods and multiple locations so we approximate this ideal as best we can, given data limitations.

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 geographically more precise measure for locations near a weather station. The second is best suited to averaging over space.

In the case of our worker data, we know the precise factory locations and can use public weather stations where available. These provide daily humidity levels and allow us to estimate wet bulb temperatures. Temperature highs occur during the day and are arguably the best proxy for hot weather while at work. However, over our study period, the most proximate weather stations contain a large number of missing observations for daily maximum temperatures, especially in the NCR. Therefore we use daily mean temperatures, rather than daily highs, to calculate a daily wet bulb temperature measure using (1).

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 rely on the IMD gridded dataset of daily temperature and precipitation and use an inverse distance weighted average of grid points within 50 km of the plant to assign daily weather 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. We describe this process in greater detail in the Appendix.

Because we observe only annual output in our panel of manufacturing plants, we aggregate daily temperatures thus obtained up to the annual level in two ways: A simple average of the daily maximum temperature over the year, and a non-linear measure consisting of the number of days in the year falling within different temperature bins. We create five bins using the daily maximum temperature expressed in celsius, $\{(0, 20], (20, 25], (25, 30], (30, 35], (35, 50]\}$. The vector $\mathbf{N} = (N^1, N^2, N^3, N^4, N^5)$ specifies the number of days in each of these bins and summarizes the temperature distribution over the year. This vector can be separately calculated for every district and each year.

3.5 Climate control within diamond-cutting 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, (iv) cutting and (v) polishing. Although these factories are small and labor intensive, similar to 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 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 Estimation and Results

4.1 Worker Productivity

Our simplest specification relates output to temperature using a piece-wise linear function of WBT:

$$log(Y_{id}) = \alpha_i + \gamma_M + \gamma_Y + \gamma_W + \beta_k WBT_{id} \times \mathrm{D_k} + \theta R_{id} + \epsilon_{id}. \tag{2}$$

Our output measure varies across the three types of firms. For cloth weaving, Y_{id} is the meters of cloth produced by worker i on day d. For the steel mill, it is the number of rectangular blooms rolled in each shift i on day d. For garment plants, it is the average over the day of the hourly output of each sewing line. In this last case we also control for the line-specific target output set by the firm, as described in Section 3.

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. Recall that 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 responds in part to demand, so we also include fixed-effects for each month and year (γ_M, γ_Y). Day of week fixed effects (γ_W) are included because some workers have a regular day off. R_{id} is the relevant measure of rainfall.

We allow for non-linear effects of heat-stress by interacting the daily wet bulb temperature, WBT_{id}, with indicator variables D_k for four temperature ranges: $\{[0°C - 20°C)$, $[20°C$-$25°C)$, $[25°C$-$27°C)$, $[27°C - 35°C]\}$. We therefore estimate the marginal effect of a degree change in WBT on output, within each of these intervals.[16]

Table 2 presents our estimates. Column 1 is for the steel mill. Columns 2 and 3 offer a within-firm comparison of garment factories in the NCR with different levels of climate control. Estimates from climate-controlled plants are shaded. Column 4 presents data from garment factories located elsewhere in India. Column 5 is for cloth output from weaving factories. Column 6 uses output in meters rather than logged meters for the cloth-weaving firms. Since weaving workers are paid piece-rates, this coefficient can be used to compute the expected loss in wages resulting from higher temperatures.

The most systematic declines in productivity are observed for the highest temperature bin. Above 27 degrees, a one degree change in WBT is associated with productivity declines ranging from 3.7 percent for sewing lines in the milder climate of South and Central India, to about 8 percent for sewing lines and weaving workers in the hotter regions of Delhi and Gujarat. For plants with climate control, we do not find systematic temperature effects on output.

In addition to this piece-wise linear specification, we estimate a more flexible model of output as a function of restricted cubic splines in WBT, with four knots at the 20th, 40th, 60th and 80th percentiles of the temperature distribution at each location. Figure 1 shows the predicted impact of temperature on output using this specification. Output at 25 degrees is normalized to 100 percent. The pattern of these results is very similar to those in Table 2.

[16] We have chosen these breakpoints to facilitate a comparison of our estimates with others in the literature (Hsiang, 2010).

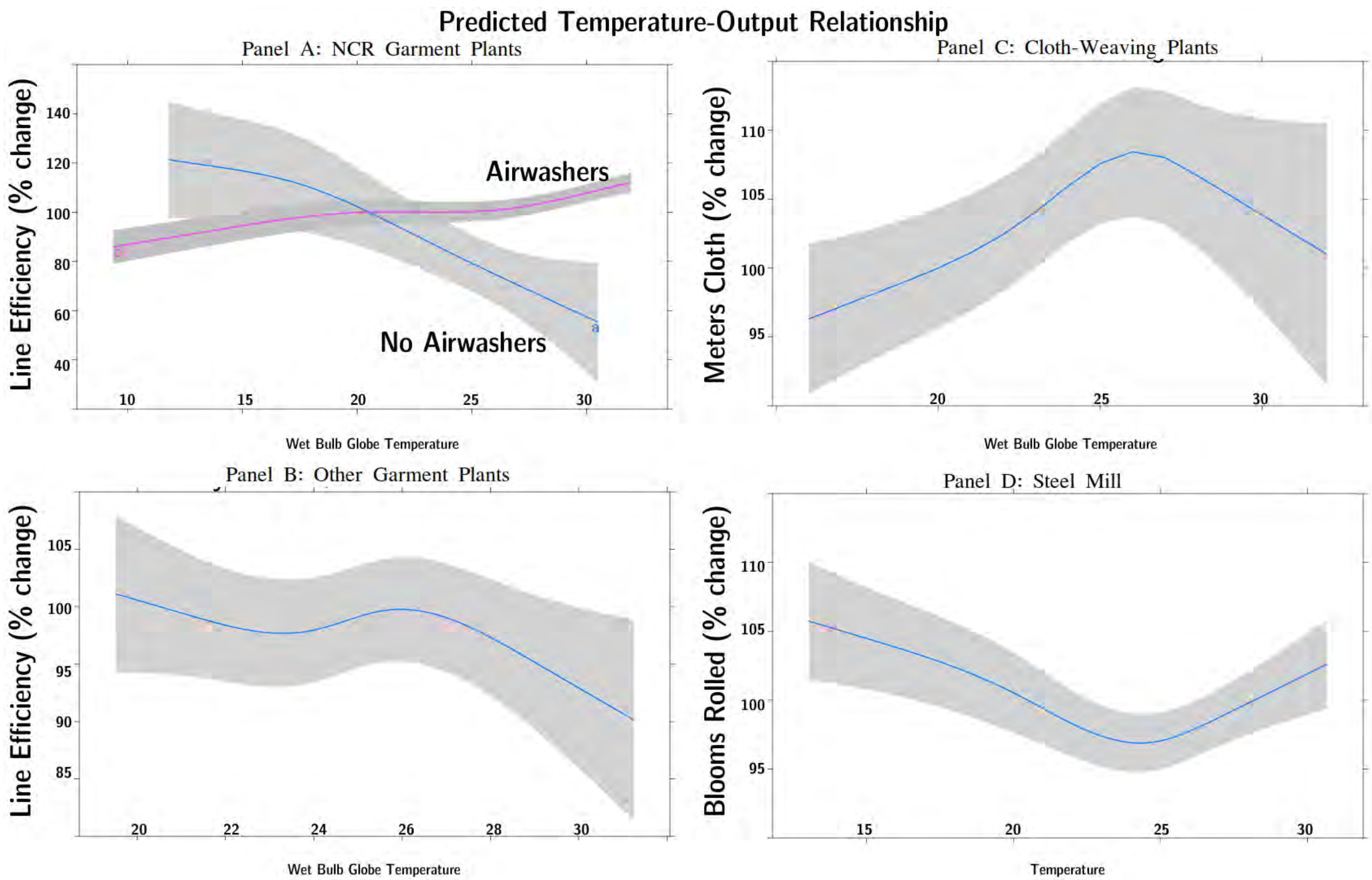

Figure 1: Restricted cubic spline models of the impact of temperature on logged output shown with 90 percent bootstrapped confidence intervals. The output at 25 degrees is normalized to 100 percent. In panels A and B, we control for the target outputs set by the firm. We have a small number of additional observations from plants with air-washers, relative to those without this equipment. This extends the temperature range over which the corresponding spline in Panel A is estimated.

We see from Figure 1, Panels A and B, that the clearest evidence in support of the heat-stress hypothesis comes from a within-firm comparison of garment sewing lines with and without climate control. For sewing lines on shop-floors without access to climate control, high temperatures are associated with a drop in efficiency at all locations while the NCR factories which have climate control, show no such declines.[17] In the cloth-weaving factories (Panel C), our estimates are less precise and statistically significant only in the highest WBT bin. For the mechanized steel mill (Panel D), the response function looks more complicated. Even though workers sit in climate-controlled cabins, the production process involves the heating and casting of steel which may be directly influenced by ambient temperatures. We return to the question of interactions of capital equipment with temperature in Section 4.2.

Note that for garment lines and weaving workers, we estimate effects on output per degree rise in *wet bulb* temperatures. We see from (1) that holding humidity constant, a one degree rise in temperature corresponds to a 0.567 degree rise in wet bulb temperatures. The effects of temperature alone, with no humidity correction, are therefore about half the estimates in Table 2. These are comparable to estimates from other studies discussed in Section 2.

In addition to contemporaneous effects of heat exposure on output, there may also be lagged effects. We examine these using ten-day histories of WBT. Results are in Table A.1 in the Appendix. We find no clear evidence that temperature lags are important.

[17] High temperatures may affect productivity directly if they are associated with more frequent power outages. All the factories in our dataset have a power backup, so this is unlikely to be a concern. Also, if outages were driving our results, we should expect to see this effect in plants with and without climate control.

Table 2: Effect of Wet Bulb Temperature on Output

	Dependent variable					
	Steel Products	Garment Manufacture			Cloth Weaving	
	log(blooms)	log(output)	log(output)	log(output)	log(meters)	meters
	(1)	(2)	(3)	(4)	(5)	(6)
(1) rainfall	0.001***	0.083**	0.044	−0.067	0.006	1.512
	(0.0002)	(0.030)	(0.192)	(0.035)	(0.008)	(0.958)
(2) log(target output)		0.796***	0.421***	0.525***		
		(0.034)	(0.126)	(0.044)		
(3) WBT:[0-20)	−0.008*	0.014***	−0.026***	−0.15	0.001	0.530
	(0.003)	(0.004)	(0.007)	(0.097)	(0.008)	(0.596)
(4) WBT:[20-25)	−0.0002	−0.014**	−0.064***	−0.004	0.008	1.700**
	(0.005)	(0.007)	(0.020)	(0.009)	(0.008)	(0.813)
(5) WBT:[25-27)	0.011*	0.029**	−0.149***	0.004	−0.012	−0.417
	(0.006)	(0.014)	(0.026)	(0.020)	(0.013)	(1.091)
(6) WBT:[27-35]	0.016	0.001	−0.087***	−0.037**	−0.077**	−6.722**
	(0.011)	(0.007)	(0.024)	(0.016)	(0.033)	(2.803)
Number of Plants	1	5	1	2	3	3
Number of Observations	9,172	23,827	621	6,073	53,655	53,655
Climate Control	Y	Y	N	N	N	N

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Observations are the total number of worker-days (weaving), line-days (garments) or shift-days (steel). Parentheses provide robust standard errors clustered at worker, line, or shift-team level. All models include fixed effects for workers (weaving), or lines (garments), or shift-teams (steel) as well as month and day-of-week fixed effects. Shaded columns represent sites with climate control. The coefficients on WBT:[25-27) should be interpreted as the change in the dependent variable for a degree increase in WBT in the range 25°C to 27°C

Worker Absenteeism

In the semi-tropical, low-income environments we study, sustained heat could induce fatigue among workers and their families, resulting in time off work. We obtained detailed histories of worker attendance for the six garment factories in the NCR and all three weaving factories. For the steel plant, we have information on the number of daily absences in each team.

For workers in the garment and cloth-weaving sites, we estimate a linear probability model using a specification similar to (2), with Y_{id} now an indicator for the worker i being present on day d. For the steel mill, we use the log of team absences as the dependent variable. The coefficient on temperature in this case represents the percentage increase in absences within a team rather than changes in the probability that an individual worker is absent.

As with worker productivity, we estimate models using contemporaneous WBT, and WBT averages over the preceding ten days. Table 3 summarizes our results. In contrast with worker productivity, absenteeism is clearly influenced by lagged temperature exposures. For the highest WBT bin, a one degree increase in the ten-day average raises the probability of a garment worker being absent by 10 percent and the number of team absences in the steel mill by 2 percent. In the case of cloth-weaving workers, who are paid a piece rate, we see no change in absenteeism.

Table 3: Absenteeism Response to Contemporaneous and Lagged Temperatures

	Garment Workers Probability of Absence				Weaving Workers Probability of Absence				Steel Workers Number of Team Absences			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
WBT	0.02***				0.00				0.32***			
	(0.00)				(0.01)				(0.02)			
WBT_{10}		0.03***				0.00					0.72***	
		(0.00)				(0.01)					(0.07)	
Q1:WBT			0.04***				−0.03			0.44***		
			(0.00)				(0.02)			(0.06)		
Q2:WBT			−0.04***				−0.02			0.60***		
			(0.00)				(0.04)			(0.06)		
Q3:WBT			0.03***				0.14**			0.24**		
			(0.01)				(0.07)			(0.10)		
Q4:WBT			0.07***				0.01			−0.06		
			(0.01)				(0.09)			(0.47)		
Q1:WBT_{10}				0.03***				−0.02				0.57**
				(0.01)				(0.02)				(0.27)
Q2:WBT_{10}				−0.01**				−0.03				0.64***
				(0.01)				(0.04)				(0.14)
Q3:WBT_{10}				0.01				0.08				0.24
				(0.02)				(0.07)				(0.30)
Q4:WBT_{10}				0.10***				0.01				1.98***
				(0.01)				(0.09)				(0.46)

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Parentheses provide robust standard errors clustered at worker or team level. All models have worker, month and day-of-week fixed effects and control for rainfall. WBT refers to the contemporaneous wet-bulb temperature, WBT_{10} refers to the ten-day average and Q_i is the ith quartile of WBT, WBT_{10}

These initial results lead us to further explore the structure of the temperature-absenteeism relationship. To do this, we generalize our definition of exposure and use total absences in each of the three types of plants: garments, cloth-weaving, and steel.

We can define heat exposure on day d as:

$$E_d = \theta_0 T_d + \theta_1 T_{d-1} + ... + \theta_K T_{d-K}$$

With equal weights θ_i, this simplifies to the temperature average that we have already used. More generally, weights could be a function of both temperature levels and lags:

$$E_d = \sum_{k=1}^{K} \theta(T,k) T_{d-k} \tag{3}$$

This allows, for example, ten days at 30^{o}C to result in a different response than eight days at 28^{o}C and two days at 38^{o}C, even though both produce the same ten-day average.

Given an accumulated exposure level E_d, the total number of absences on day d in a population with characteristics X_d can be modeled as:

$$log(A_d) = \alpha + \beta E_d + \gamma X_d + \epsilon_d. \tag{4}$$

Gasparrini (2014) shows how this class of models can be estimated using restricted cubic splines in temperatures and lags. One spline describes exposure as a function of temperature, while the other describes exposure as a function of the lag. We estimate this model for each of the three industries – garments, weaving, and steel, and then use the fitted models to predict the percentage changes in total absences under different WBT trajectories.

Figure 2 illustrates two scenarios. The left column shows the predicted change in the log of daily absences for a 1°C increase in WBT over a 27°C reference, for a duration ranging from 1 to 10 days. In the right column, we plot predicted absenteeism for different temperature levels sustained over a ten-day period.

We find that sustained high temperatures increase absenteeism for both steel and garment workers.[18] As with the simpler linear models in Table 3, effects on daily-wage weaving workers are not systematically different from zero. In contrast to our results on worker productivity, we find that temperature affects absenteeism even if the work-place has climate control. This is consistent with evidence from the United States which finds that temperature influences time-allocation and labor-leisure trade-offs (Zivin and Neidell, 2014).

We include month fixed effects in all our empirical models and our estimates therefore represent short-run temperature impacts. We do this to isolate the role of heat stress from other mechanisms and also because the result of long-run increases in temperature cannot be identified separately from other seasonal factors. In the Appendix, we show that there are seasonal changes in the availability of casual workers during high temperature months (Figure A.2).

[18] Absenteeism following sustained heat exposure may occur due to illness. 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 heat through different mechanisms, for example, increased breeding of pathogens and disease vectors.

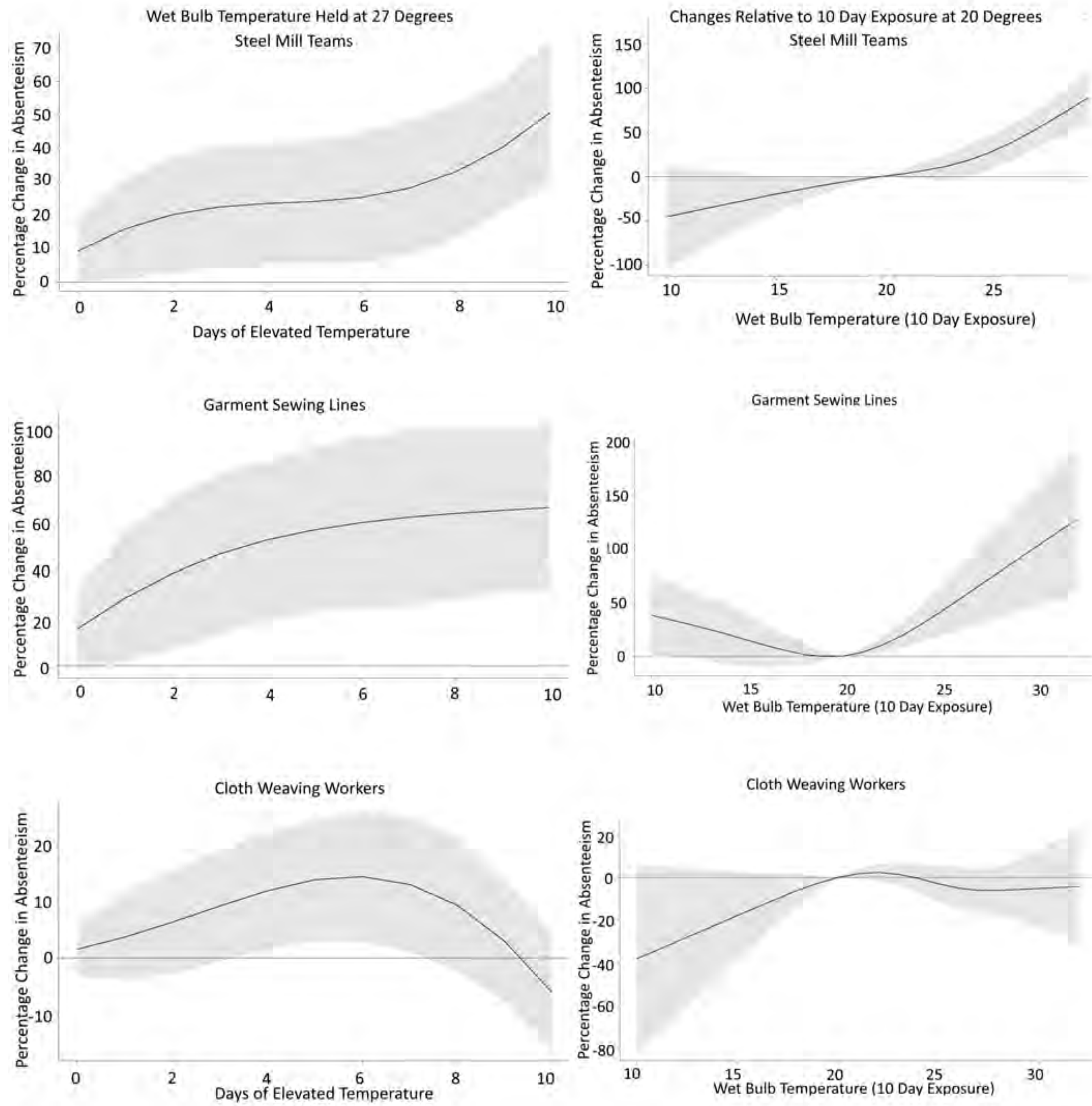

Figure 2: The impact of wet-bulb temperature on absenteeism. Natural cubic splines with three degrees of freedom shown with 90 percent bootstrapped confidence intervals. All models include controls for rainfall as well as month and day of week fixed effects. The left column shows the predicted change in the logarithm of daily absences for a 1°C increase in WBT over a 27°C reference, for a duration ranging from 1 to 10 days. In the right column, we plot predicted absenteeism for different temperature levels sustained over a ten-day period.

4.2 Aggregate Plant Output

Our worker data establish that high temperatures lower productivity. We now use our panel of manufacturing plants to examine whether there are discernible temperature effects on the value of plant output, and whether these can be attributed to a decline in the productivity of labor.

We begin with the following reduced-form specification:

$$log(Y_{it}) = \alpha_i + \gamma_t + \beta T_{it} + \theta R_{it} + \epsilon_{it}. \tag{5}$$

Y_{it} is the value of output for plant i in year t and α_i and γ_t are plant and year fixed effects. T_{it} is the average daily maximum temperature for the district in which a plant is located and R_{it} is the average daily rainfall. As discussed in Section 3.4, we use temperature without humidity corrections due to data limitations.

In our second reduced-form specification, the value of annual output depends on the distribution of temperatures over the year. Specifically, we count the number of days with maximum temperatures falling in each of five bins: $\{(0,20],(20,25],(25,30],(30,35],(35,50]\}$ and denote the number in bin j by N^j. These bin counts sum to 365. We then estimate:

$$log(Y_{it}) = \alpha_i + \gamma_t + \sum_{j=2}^{5} \beta_j N_{it}^j + \theta R_{it} + \epsilon_{it}. \tag{6}$$

N^1 is the omitted bin in this specification and results from both (5) and (6) are reported in Table 4.[19] A 1°C increase in the average maximum temperature results in a 3.2% decrease in output. Coefficients from (6) are best interpreted using specific counter-factual scenarios,

[19]We use robust standard errors clustered at the district level to account for spatial and serial correlation.

because changes in the temperature distribution can affect all bins. To illustrate, these estimates imply that a shift of 10 days from the fourth to the fifth bin leads to a 0.3 percent decline in the value of output.

To study sector-specific temperature effects, we modify (5) by interacting average temperature with an indicator for each 2-digit manufacturing sector.[20] Figure 3 shows sector-wise mean estimates and confidence intervals. Output is lower in a hotter year, for most sectors, to varying degrees.

[20] 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.

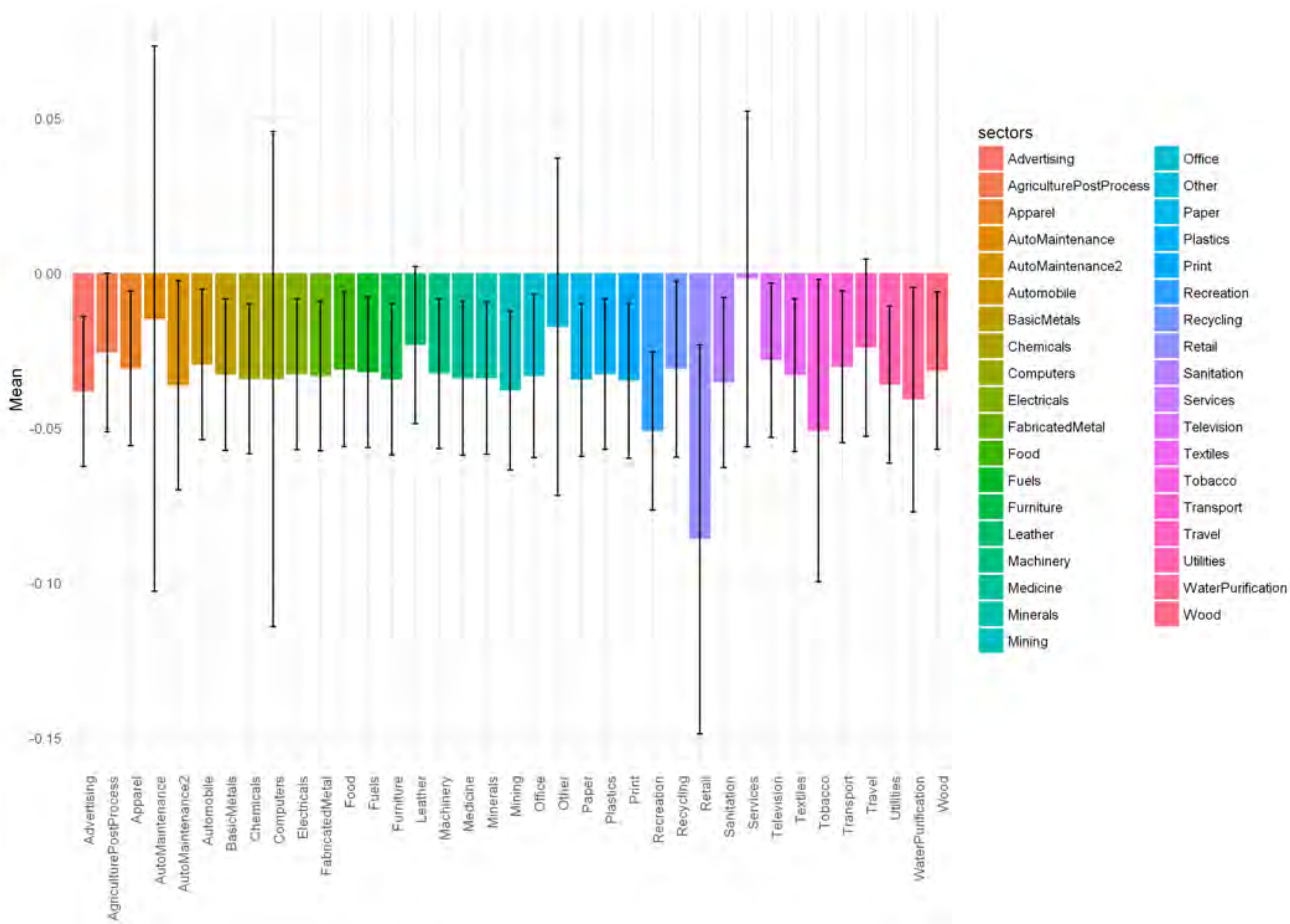

Figure 3: Sector-wise percentage change in output for a one degree increase in annual averages of daily highs.

Table 4: Temperature Effects on Log Plant Output Value

	Log Output	Log Output
	(1)	(2)
T_{max}	-0.0318**	
	(0.0144)	
N^2		-0.0025***
		(0.00061)
N^3		-0.00196***
		(0.00073)
N^4		-0.00243***
		(0.00078)
N^5		-0.00272***
		(0.00085)
rainfall	-0.0011	-0.0019
	(0.00361)	(0.00364)

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Robust standard errors clustered at district level provided in parentheses. All models have plant and year fixed effects. $(N^1, N^2, N^3, N^4, N^5)$ are days in $\{(0, 20], (20, 25], (25, 30], (30, 35], (35, 50]\}$ and reported coefficients are relative to N^1.

These estimates suggest an inverse, non-linear relationship between temperature and plant output, akin to temperature effects on worker productivity. To understand the role of labor in the plant data, we write down a simple production function in log form relating the value of output to capital and labor:

$$y = \alpha(\mathbf{N}) + \omega(\mathbf{N})k + \beta(\mathbf{N})l \tag{7}$$

Here y, k, l are logged values of output, capital and labor respectively and $\mathbf{N}$ is a vector of the number of days in five temperature bins. The ASI reports the net value of plant equipment and machinery at the start of each year and we use this as our measure of capital. Our labor measure is the number of workers. The terms $\omega(\mathbf{N})$ and $\beta(\mathbf{N})$ are the output elasticities of capital and labor that depend on the temperature distribution during the year. Temperature effects acting through all other inputs are captured by the residual $\alpha(\mathbf{N})$.

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

Thus we have,

$$\alpha(\mathbf{N}) = \alpha_o + \sum_{j=1}^{5} \alpha_j N^j$$

$$\omega(\mathbf{N}) = \omega_o + \sum_{j=1}^{5} \omega_j N^j$$

$$\beta(\mathbf{N}) = \beta_o + \sum_{j=1}^{5} \beta_j N^j$$

Making these substitutions in (7) we obtain

$$y = \alpha_o + \sum_{j=1}^{5} \alpha_j N^j + \omega_o \cdot k + \sum_{j=1}^{5} \omega_j N^j k + \beta_o \cdot l + \sum_{j=1}^{5} \beta_j N^j l \tag{8}$$

Temperature effects on labor and capital may also depend on the levels of these inputs. Firms that employ a large number of workers may see higher output losses from heat stress, or conversely invest more in climate control. To allow for such heterogeneity, we create dummies D^q and E^q, representing the qth deciles of capital and labor respectively and then estimate

$$\begin{aligned} y_{it} = \sum_{q=2}^{10} \left(\omega_{oq} D_{it}^q + \sum_{j=2}^{5} \omega_{jq} N_{it}^j D_{it}^q \right) + \sum_{q=2}^{10} \left(\beta_{oq} E_{it}^q + \sum_{j=2}^{5} \beta_{jq} N_{it}^j E_{it}^q \right) \\ + \left(\sum_{j=2}^{5} \alpha_j N_{it}^j \right) + \theta R_{it} + \alpha_i + \gamma_t + \epsilon_{it} \end{aligned} \tag{9}$$

Here y_{it} is the log of the value of output for plant i and year t. ω_{jq} and β_{jq} correspond to ω_j, β_j in (8), estimated now for each decile q. As in all our previous models, we include fixed effects for each plant, α_i, each year, γ_t, and a control for rainfall R_{it}.

We can use estimates from (9) to identify the effect of temperature acting through capital, labor, and the residual. Given changes ΔN^j in the temperature bins, for any fixed D^q_{it}, E^q_{it}

$$\sum_{q=2}^{10}\left(\sum_{j=2}^{5}\omega_{jq}\Delta N^j_{it}D^q_{it}\right) \tag{10}$$

and

$$\sum_{q=2}^{10}\left(\sum_{j=2}^{5}\beta_{jq}\Delta N^j_{it}E^q_{it}\right) \tag{11}$$

are the predicted effects of temperature on output, operating through the capital and labor channels respectively, and

$$\sum_{j=2}^{5}\alpha_j\Delta N^j_{it} \tag{12}$$

is the residual effect.

This decomposition can be carried out for any vector of changes ΔN^j. We would like to choose a vector ΔN^j which corresponds to a one-degree change in average maximum temperatures, thus facilitating a comparison with the estimates from (5) as provided in Table 4.

There is no unique way to map annual temperatures to the number of days in different bins. Our annual panel of district level temperatures provides bin values, N^j_{it}, and average maximum temperatures, T^{max}_{it}, for each district i and year t between 1998 and 2012. We regress each N^j_{it} on T^{max}_{it}, controlling for district fixed effects. The coefficient on T_{max} in these regressions can be interpreted as the change in bin values for a one-degree increase in temperature, estimated using variation observed over the years spanned by our plant level panel.

These estimated coefficients imply that for every one-degree rise in T_{max} there are 5.89 fewer days in Bin 1 ($T_{max} \leq 20$), 5.73 fewer days in Bin 2 ($T_{max} \in (20, 25]$), 11.76 fewer days in Bin 3 ($T_{max} \in (25, 30]$), 1.93 more days in Bin 4 ($T_{max} \in (30, 35]$) and 21.28 more days in Bin 5 ($T_{max} > 35$). As we might expect, the temperature distribution shifts to the right and within a rounding error, the bin counts provided by point estimates also add up to 365.

Figure 4 shows the results of this decomposition exercise.[21] We find that the negative impact of temperature is driven almost entirely by declines in labor productivity. Output losses attributable to capital, the residual, and their combination are statistically indistinguishable from zero. Comparing the reduced-form results from Column 1 of Table 4 with the total effect size in Figure 4 suggest that our choices over ΔN^j map very well to a one-degree change in average maximum temperatures.[22]

We also find evidence that the importance of the labor channel varies with the number of workers. Figure 5 plots the effect of temperature on output, through changes in the productivity of labor, for different labor deciles. Point estimates become more negative as the number of workers increases.[23]

[21]This model contains a large number of coefficients, so its implications are best understood in graphical form. For reference, Table A.2, Column 2 lists a subset of estimated coefficients.

[22]Our results are robust to more flexible specifications. Figure A.2 in the Appendix reports results with the residual modeled using quadratics in bins.

[23]Note that these estimates are net of plant fixed effects and therefore do not simply represent comparisons of plants with larger or smaller average output.

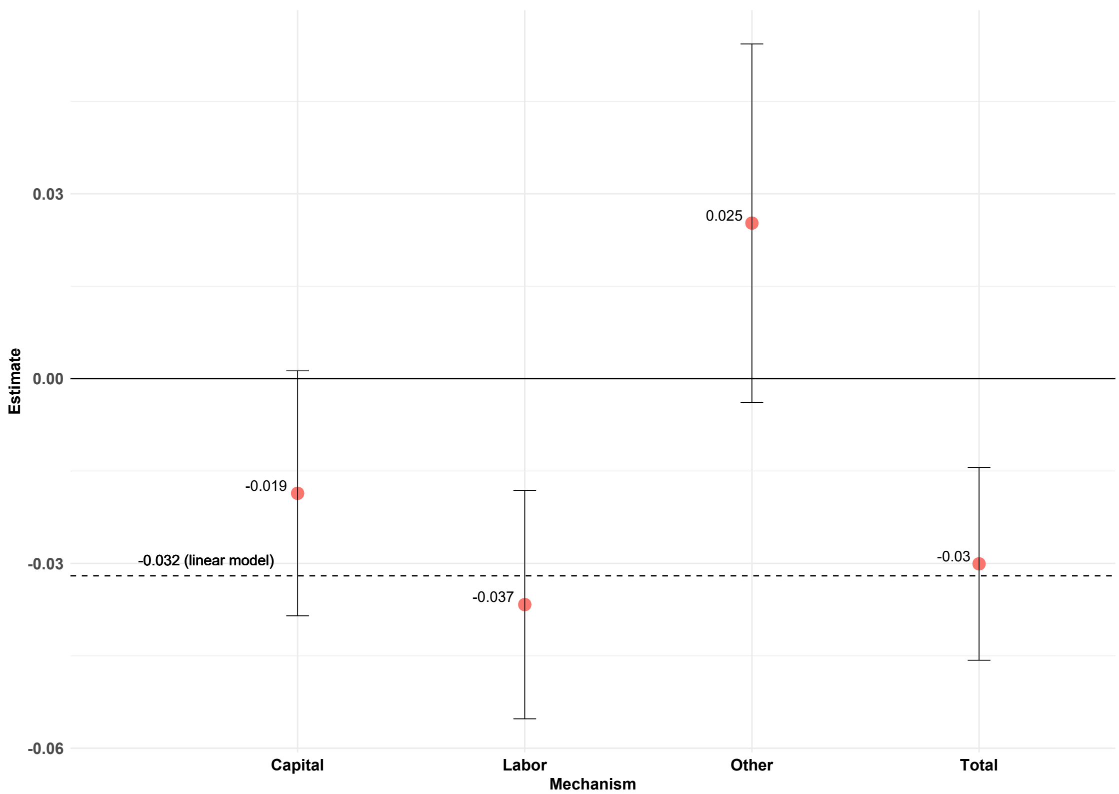

Figure 4: Marginal effect of temperature on log output decomposed into labor, capital and a residual as in (9). 90 percent confidence intervals with robust standard errors clustered at district level. Dotted line plotted at the total reduced-form temperature effect on output from (5) and solid line is at zero. Predictions use $\Delta N = (-5.89, -5.73, -11.76, 1.93, 21.28)$ where $(N^1, N^2, N^3, N^4, N^5)$ are days in $\{(0, 20], (20, 25], (25, 30], (30, 35], (35, 50]\}$.

Before concluding this section, it is worth noting that a concern with using production function regressions of this form is potential bias induced by the endogeneity of labor (Ackerberg, Caves, and Frazer, 2006; Levinsohn and Petrin, 2003). An upward bias on the labor coefficient may become significant if plants can quickly vary the number of workers.

In the present instance, we believe this is unlikely to be a significant concern. India has a notoriously inflexible labor market in the manufacturing sector. Indeed, 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 any workers.

Beyond the specifics of the Indian context, in the production function estimated in (9), we discretize labor and capital into bins. This makes both inputs relatively inflexible and our identification is primarily based off year-to-year variations in temperature.[24] A more direct piece of evidence comes from a comparison between the total temperature effect from (9) and the reduced-form estimate from (5), which contains no potentially endogenous regressors. The two are almost identical.

[24] Some plants at the boundaries of decile cut-offs do switch bins. We estimate the conditional correlation between temperature and the labor measure in (9) and find this to be approximately zero.

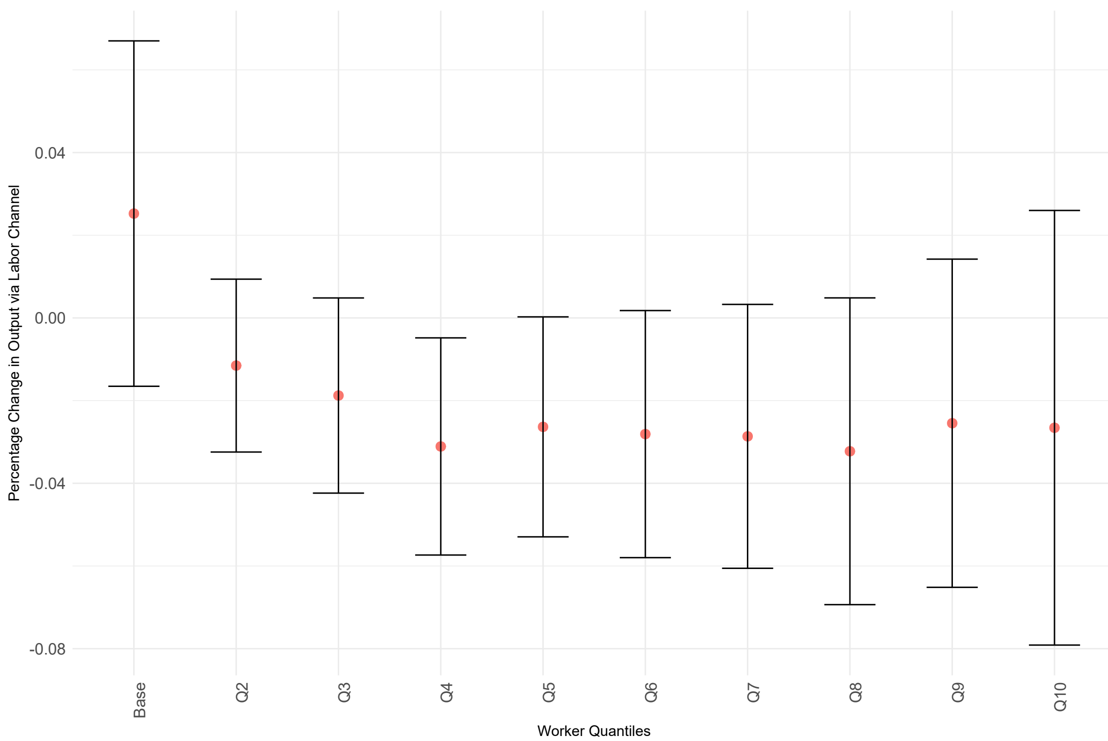

Figure 5: Percentage change in plant output attributable to temperature induced changes in the productivity of labor, split by deciles. Predictions use $\Delta N = (-5.89, -5.73, -11.76, 1.93, 21.28)$ where $(N^1, N^2, N^3, N^4, N^5)$ are days in $\{(0, 20], (20, 25], (25, 30], (30, 35], (35, 50]\}$. 90 percent confidence intervals with robust standard errors clustered at district level

5 Aggregation to Macro-level Estimates

Our worker and plant output data both point to the importance of labor productivity in explaining the negative effects of high temperatures. In this section, we show that our predicted temperature effects on workers and plants are consistent with each other, and with estimates based on district-level manufacturing output. We also compare our results at different levels of aggregation with estimates of GDP declines in hot years, obtained from country-level studies. These comparisons suggest that the labor productivity channel is large enough to explain the entire temperature-output relationship observed at the country level.

First, we place an upper bound on how much of the temperature-output effect seen at the plant level can be accounted for by the temperature-output effect estimated at the worker level. Since the worker-level effect must be operating through the labor channel, this is a way of bounding the labor channel. We begin by taking an average of temperature effects from our worker data for factory sites *without* climate control. These are the cloth-weaving units and a subset of the garment factories. The log of output is non-linear in WBT, so we average over all estimates from the two *highest* WBT bins.[25] Using (1), we note that the effect of a $1°C$ increase in temperature on the log of output equals 0.567 times the corresponding effect size per degree WBT. This gives us an estimated effect size of $\hat{\alpha}$ = -0.034. Although this comes from a handful of sites, it is comparable to effects observed across different lab and field settings, as summarized in Section 2.

At the plant level, we assume that the marginal worker is less productive by $\hat{\alpha}$ when the temperature increases by $1°C$. We estimate

$$log(Y_{it}) = \alpha_i + \gamma_t + \beta l_{it} + \theta R_{it} + \epsilon_{it}. \tag{13}$$

[25] These are the 6 estimates provided in rows 5-6 in columns 3-5 of Table 2. We appropriately combine the variance of these individual estimates to produce the variance of the mean. Over 85 percent of days in our plant sample have a maximum temperature above 25 degrees celsius, so employees would normally find themselves coming to work in warm weather.

Denoting the estimated labor elasticity by $\hat{\beta}$, we note that the percentage change in output when the temperature rises by one degree is $(1 + \hat{\alpha})\hat{\beta}l - \hat{\beta}l = \hat{\alpha}\hat{\beta}l$. Evaluated at the mean value of l, we find that a one-degree increase in temperature corresponds to a change in plant output of -6%.

The direct estimate of the temperature effect on output, obtained from (5) and reported in Table 4, is -0.0318. This is smaller than our prediction based on $\hat{\alpha}$, but comparable in that the 90 percent confidence intervals for the two estimates overlap. The difference in point estimates is not surprising since $\hat{\alpha}$ assumes no climate control in plants. These comparisons are illustrated in the first three bars of Figure 6, and suggests that the impact of temperature on worker productivity is large enough to explain the entire response of plant output to higher annual temperatures.

Above the level of the the plant, we have data on the manufacturing GDP of each district. Using the production function for plants estimated in (9), we predict district-specific temperature effects for a one degree rise in temperature by repeating the procedure used to generate Figure 4, using the population of plants in each district. We calculate the mean of these individual district estimates to obtain the average percentage change in district manufacturing GDP for a one degree increase in temperature. Because district manufacturing GDP includes output from smaller plants not sampled in the ASI, the accuracy of this prediction depends on the extent to which the response of these smaller manufacturing units mimics the population of factories included in the ASI. We also directly estimate the effect of temperature on district manufacturing GDP using a panel specification analogous to (5). The fifth and sixth bars in Figure 6 compare these predicted and estimated changes and we see that the two are very similar.[26]

The last two bars provide existing estimates from two recent country-level studies (Dell,

[26] We are unable to study temperature effects using GDP figures for Indian states because these data are interpolated in several years and therefore unreliable.

Jones, and Olken, 2012; Burke, Hsiang, and Miguel, 2015). These numbers are derived using annual average temperatures for several countries across the world, observed over long periods of time. The earliest temperature-country observations used in both papers are from 1950. Since they come from different empirical models, the country-level numbers are not directly comparable with our estimates. Nevertheless they provide a useful benchmark and are statistically indistinguishable from the effect sizes we obtain at lower levels of aggregation.

To summarize, our various results, predicted and directly estimated, are remarkably similar. From the worker up to the country level, output appears to decrease between 2 to 4 percent per degree celsius. Plant and district level estimates are also similar to those obtained from country-level studies. In combination with the evidence presented in Figure 4, this suggests that reduced labor productivity could explain the temperature-output relationship at more aggregate levels.

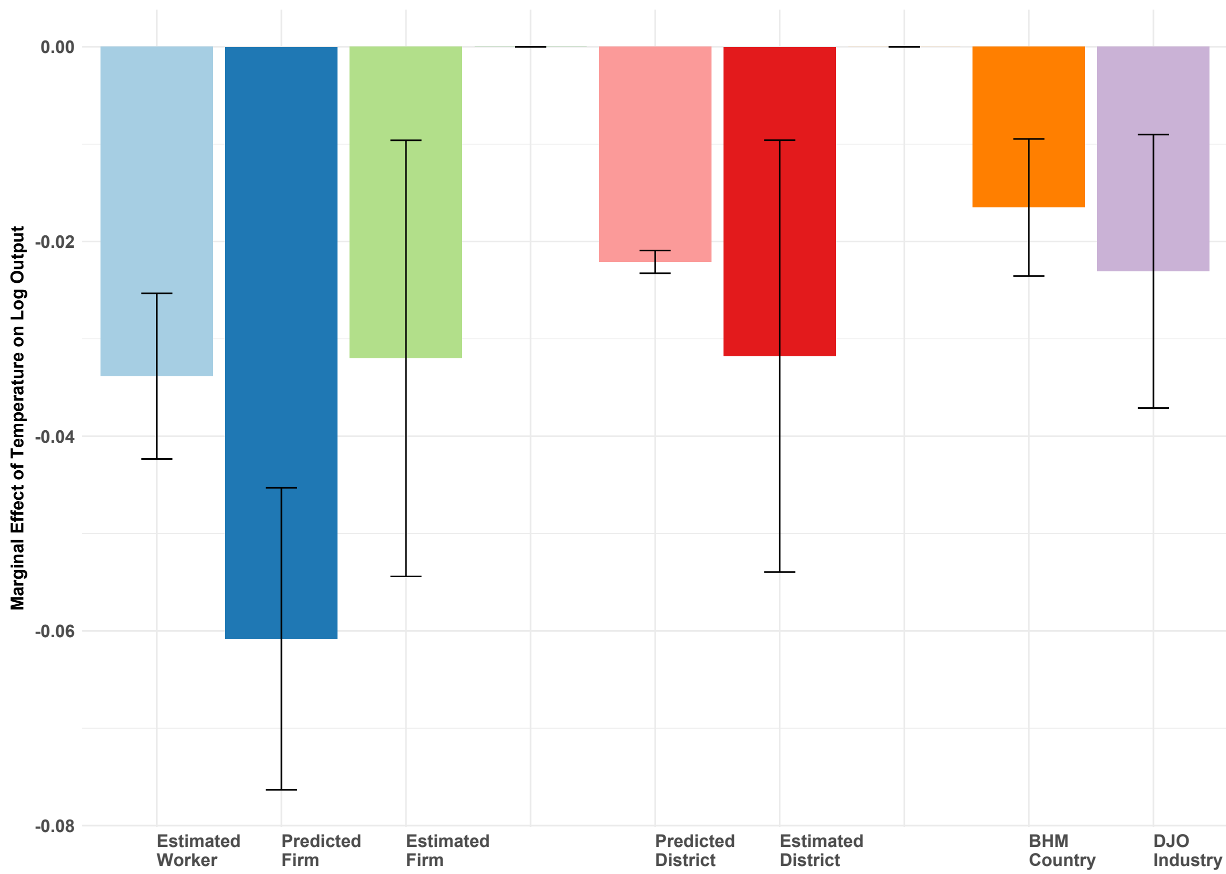

Figure 6: Marginal effect of temperature on log output at different levels of production with 90 percent confidence intervals. Legend entry **BHM Country** refers to the contemporaneous effect of temperature on country output reported in Burke, Hsiang, and Miguel (2015). Legend entry **DJO Industry** refers to the contemporaneous effect of temperature on industrial sector output reported in Dell, Jones, and Olken (2012). All other estimates come from data in this paper.

6 Adaptation and Energy Use

The full implications of the negative relationship between temperature and labor productivity will depend on how the manufacturing sector adapts to higher temperatures. In the medium-term, firms can mitigate the effects of temperature through climate control. Longer-run adaptation may include automation, the relocation of plants, and changes in the composition of manufacturing. Microdata from our garment manufacturing firm show that climate control effectively mitigates productivity losses on hot days. However, firm decisions to invest in climate control depend on the costs of cooling, relative to the expected output losses resulting from heat stress. None of our three cloth-weaving plants, for example, had invested in climate control.

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. We also present results from 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. This survey allowed us to study the selective adoption of air-conditioning *within* plants. We find that this is more likely to be used in processes that are labor-intensive and contribute most to diamond quality. Finally, we return to our national plant panel and show that the effects of a degree-rise in temperature seem to be falling over a 15-year period. While this does not establish adaptation, it is consistent with it.

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 2 to 4 per cent per degree rise in temperature, these investments are unlikely to be profitable for firms with small price mark-ups.

The diamond cutting plants we study are located in the same city but have much higher value-added. Each plant uses five main processes: (i) sorting and grading, (ii) planning and marking, (iii) bruting, (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 use a logit model of the likelihood of air conditioning as a function of process characteristics: labor intensity, mechanization, and 'importance' in determining output quality. The first of these variables is measured by the share of the plant's workers engaged in the process, the second by the share of the plant's machines, and the third is a self-reported assessment by management on a scale of 1 to 5. We find that diamond polishing units in Surat choose to preferentially cool high-value and labor-intensive processes. Table A.8 in the Appendix contains these results.

Adaptation to high temperatures at the national level depends on decisions by a heterogeneous population of plants of different sizes and in different sectors. While we do not directly observe climate-control investments in our plant data, we can ask whether temperature effects decline in our plant panel over the 15-year period during which we observe plant output. To do this, we modify (6) to include a full set of interactions of temperature bin counts with a continuous time variable. We find that output become less responsive to temperature over time but change is relatively slow. The negative effect on output from an additional day in the fourth and fifth temperature bins reduces by about 6 to 8 percent per year. Column 1 of Table A.7 in the Appendix has coefficient estimates.

7 Alternative Explanations

We have argued that heat stress causes declines in labor productivity and this can explain reductions in manufacturing output due to high temperatures. In this section we consider alternative explanations that have been suggested in the literature. These include an increased probability of 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 temperature effects on output at higher levels of aggregation and so we assemble two additional data sets to test for their importance.

Industrial disputes are a relevant measure of conflict for manufacturing plants. Data for these are available from an annual publication by India's Ministry of Labour and Employment titled 'Statistics on Industrial Disputes, Closures, Retrenchments and Lay-Offs'. Because all episodes are not equally severe, this handbook also reports the total number of workday minutes lost due to industrial disputes in each state. We use the log of this number as a proxy for the annual exposure to conflict. We use data for all recorded incidents between 2000 and 2013, except for the two years of 2001 and 2002.[27]

Among natural disasters in India, floods are the most widespread and are directly linked to hurricanes and heavy rain. We tabulate every recorded instance of flooding in India from 1997 to 2013. Our data comes from the Dartmouth Flood Observatory Archive which records flooding incidents along with affected areas, severity, duration, and damages. The dataset is built using a combination of information from remote sensing, news stories, government releases, and ground instruments. The Dartmouth Flood Observatory defines the magnitude of a flood as log(Duration x Severity x Affected Area). We calculate the total magnitude of all floods for every year within each state.

[27] Data for these two years, and for years before 2000, were not available.

Both these measures are at the state level and for each year, we assign plants the values for the state in which they are located. We then estimate modified versions of (6) and (8), including a full set of interactions between these variables and temperature bins. We find that floods and conflict have quantitatively small and statistically insignificant impacts on the value of plant output. Additionally, coefficients on the labor-temperature interaction terms remain unchanged.[28] Coefficient estimates are reported in Table A.3 in the Appendix.

We conclude that these mechanisms cannot explain our results. Although severe conflicts or natural disasters could create large negative shocks to aggregate output, such events may be rare or absent during our study period.

We consider three other factors that may influence plant output, without necessarily changing the productivity of labor: power outages, input price changes, and agricultural spillovers.

The large plants surveyed in the ASI are typically served by dedicated distribution lines with scheduled load-shedding. Thus temperature-induced power outages may not be a significant concern. Nevertheless, we formally test this hypothesis by constructing a state-level measure of monthly outages using data from India's Central Electricity Authority, and then regressing logged output on temperature and outages. Comparing Table 4 with Table A.5 in the Appendix suggests that outages have no appreciable effect on either output or the effect of temperature.

For price changes, we gather data on plant input prices from the ASI and regress these on temperature. We find no evidence that temperature has significant effects on prices after controlling for year fixed effects. These results are in the Appendix in Table A.4. 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 minimized by storage and regional markets. Finally, it seems unlikely that agricultural spillovers are very large because

[28] Compare Table A.3, Columns (3) and (4) with Table A.2, Column (1).

we observe negative temperature effects across sectors, even for activities with no obvious connection to agriculture (Figure 3).

8 Conclusions

This paper has described the impact of temperature on the productivity and attendance of workers, and the output of manufacturing plants. 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. Plant-level temperature effects also match closely with estimates from studies examining country-level manufacturing GDP. Our objective has been to show that non-agricultural GDP may decline at high temperatures, largely due to the physiological effect that heat has on human beings both through heat stress, and perhaps also through increased morbidity.

This result has fundamental implications for how we should think about the costs of climate change going forward. The evidence we uncover on the effectiveness of climate control and also its limited adoption, suggests that 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. Our main finding, identifying the importance of the labor channel, has significant distributional implications. We might expect adaptation to be directed towards more productive workers, adding high value, who may also be richer.

Although our focus throughout this paper has been on the manufacturing sector, the potential ramifications of our findings are wider. The conclusion that a physiological mechanism is economically important implies that these effects may be significant in all labor-intensive activities where climate control is expensive or infeasible. Temperature impacts on worker productivity may be pronounced and widespread in sectors such as agriculture and construction across the world, where exposure is higher and adaptation possibilities more limited. 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." .

Adhvaryu, Achyuta, Namrata Kala, and Anant Nyshadham. 2014. "The Light and the Heat: Productivity Co-benefits of Energy-saving Technology." *Unpublished* .

Allcott, Hunt, Allan Collard-Wexler, and Stephen O'Connell. 2014. "How Do Electricity Shortages Affect Productivity? Evidence from India." *Unpublished* .

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. 2011. "Weather 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.

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.

Gasparrini, Antonio. 2014. "Modeling exposure-lag-response associations with distributed lag non-linear models." *Stat. Med.* 33 (5):881–899.

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.

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.

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. The steel mill shown in Panel A uses smelting, casting and forging processes, all of which are capital intensive. In the garment manufacturing factories shown in Panel B, workers are arranged in lines, with each person repeatedly carrying out a specific task. The cloth weaving workers walk up and down in the aisles between looms (Panel C), adjusting alignment, restarting feeds when interrupted, and making other necessary corrections.

A.2 Lagged Effects of Temperature on Worker Output

Table A.1 presents results from a modification of (2) where we replace the contemporaneous WBT and rainfall values, with their respective averages over the last ten days. We see no clear relationship between how much a worker produces on a given day, and exposure in previous days. This contrasts with our results on worker absenteeism, summarized in Table 3

Figure A.1: Production floor images from the steel mill (Panel A), garment sewing plants (Panel B), and cloth-weaving plants (Panel C)

Table A.1: Worker Output Response to 10 day Temperature History

	Cloth Weaving	Garment Sewing
	Meters	Log(Efficiency)
	(1)	(2)
R_{10}	0.03571	0.47483**
	(0.02700)	(0.23990)
$Q1 : WBT_{10}$	−0.02293**	0.00861
	(0.00918)	(0.01107)
$Q2 : WBT_{10}$	−0.00486	−0.04774
	(0.00996)	(0.03969)
$Q3 : WBT_{10}$	−0.01458	0.00683
	(0.02742)	(0.03869)
$Q4 : WBT_{10}$	−0.04396	0.02508*
	(0.05727)	(0.01361)

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Robust standard errors clustered at worker (1) or line (2) level are reported in parentheses. Estimates are based on factories without climate control and all models have worker, month, and day-of-week fixed effects. WBT_{10} denotes the ten-day wet bulb temperature average and Q_i is the ith quartile of WBT_{10}. R_{10} denotes the ten-day rainfall average.

A.3 Seasonal Patterns in Worker Absenteeism

We carried out interviews with managers of various cloth-weaving firms in Surat including, but not restricted to, the three plants from where we obtain worker-level data. Managers claimed that during the hottest months, daily wage workers preferred to go home to their villages and lived off income from the National Rural Employment Guarantee Scheme, rather than working under more strenuous conditions at the factory. Some owners reported that they were considering the possibility of temporarily raising wages through a summer attendance bonus, while others felt this would render profit margins too small to make operation worthwhile.

Figure A.2 plots worker attendance data from the small cloth-weaving firms and the garment sewing factories. We see seasonal reductions in the attendance of daily wage cloth weaving workers (Panel A), concentrated in high temperature months. These patterns are absent for the garment sewing workers who have long term employment contracts (Panel B). Although many factors differentiate the two types of work settings, it is plausible that formal employment contracts reduce the costs to taking an occasional day of leave, but significantly increase the opportunity cost of 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. This is an area that would benefit from further research.

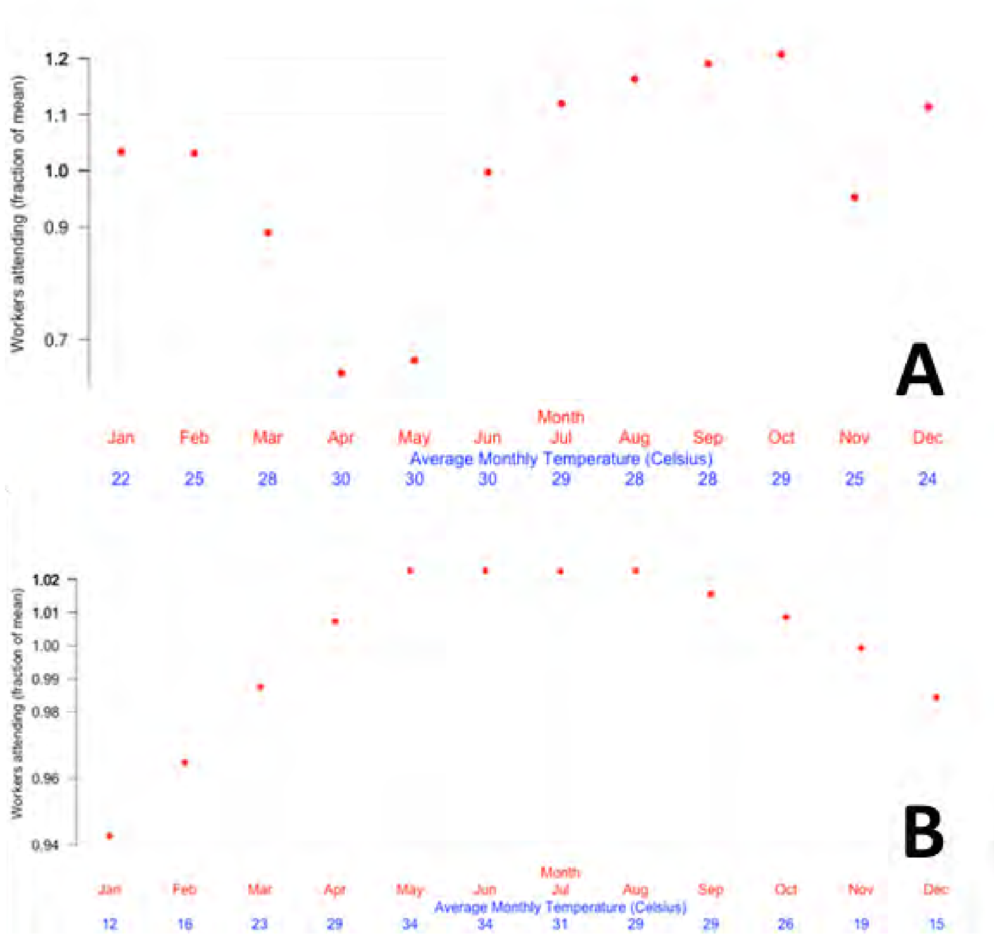

Figure A.2: Worker attendance by month for daily wage workers in cloth-weaving factories (Panel A) and salaried workers in garment plants (Panel B)

A.4 Annual Survey of Industry Data Cleaning

This section describes how our 15-year panel is constructed from the Annual Survey of Industry datasets, made available for purchase by the Indian government.

Between 1998 and 2007, the Annual Survey of Industry is 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 purchase both forms of the data and merge the two to obtain a panel containing district identifiers. This merging can be done on any suitable subset of the other observed characteristics, which remain identical.

We use the state code, NAICS code, year of starting operations, and value of output to complete this matching.

From 2008 until 2012 only a cross-section *without* district identifiers is available. At the time of writing, the latest survey with micro-data available for sale was for financial year 2012-2013.

Since there is only one version of the dataset available, the procedure above cannot be used. However, observations in these years can still be used to expand the panel by matching plants on *time-invariant* characteristics. These include the state location, the sector 4 digit NAICS codes and the year of starting operations. For each survey between 2008 to 2012, we first list plants that are uniquely identified based on these variables.

We then search in each year of our panel (running from 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 thus enters the panel. Note that 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, but a small amount of additional matches may be obtained by 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. A few additional matches may be obtained using this fact by comparing observations in the last year of the panel, to observations in the following year.

After the panel is constructed we carry out a few data cleaning operations listed below.

1. We remove observations where values of output, workers, total wages, or cash on hand are less than or equal to zero.

2. The ASI dataset also contains observations with implausibly high or low reported values of these variables. For instance there are plants with reported annual output less than a few dollars. We drop the top 2.5 percent and bottom 2.5 percent of values of output, number of workers, wages, or cash on hand. This is done to transparently eliminate these outliers. Incidentally, this process also removes a small number of manufacturing units that report having less than 10 workers employed. Such observations represent a discrepancy between the criterion used to select the survey sample and reported data by plants.

3. We drop plants where the reported state or district changes over the panel duration.

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

Our final sample has 69,643 manufacturing plants distributed all over India (Figure A.1) and spanning all major sectors (Figure 3). These plants are then matched to district temperature and precipitation measures as described in the text.

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.

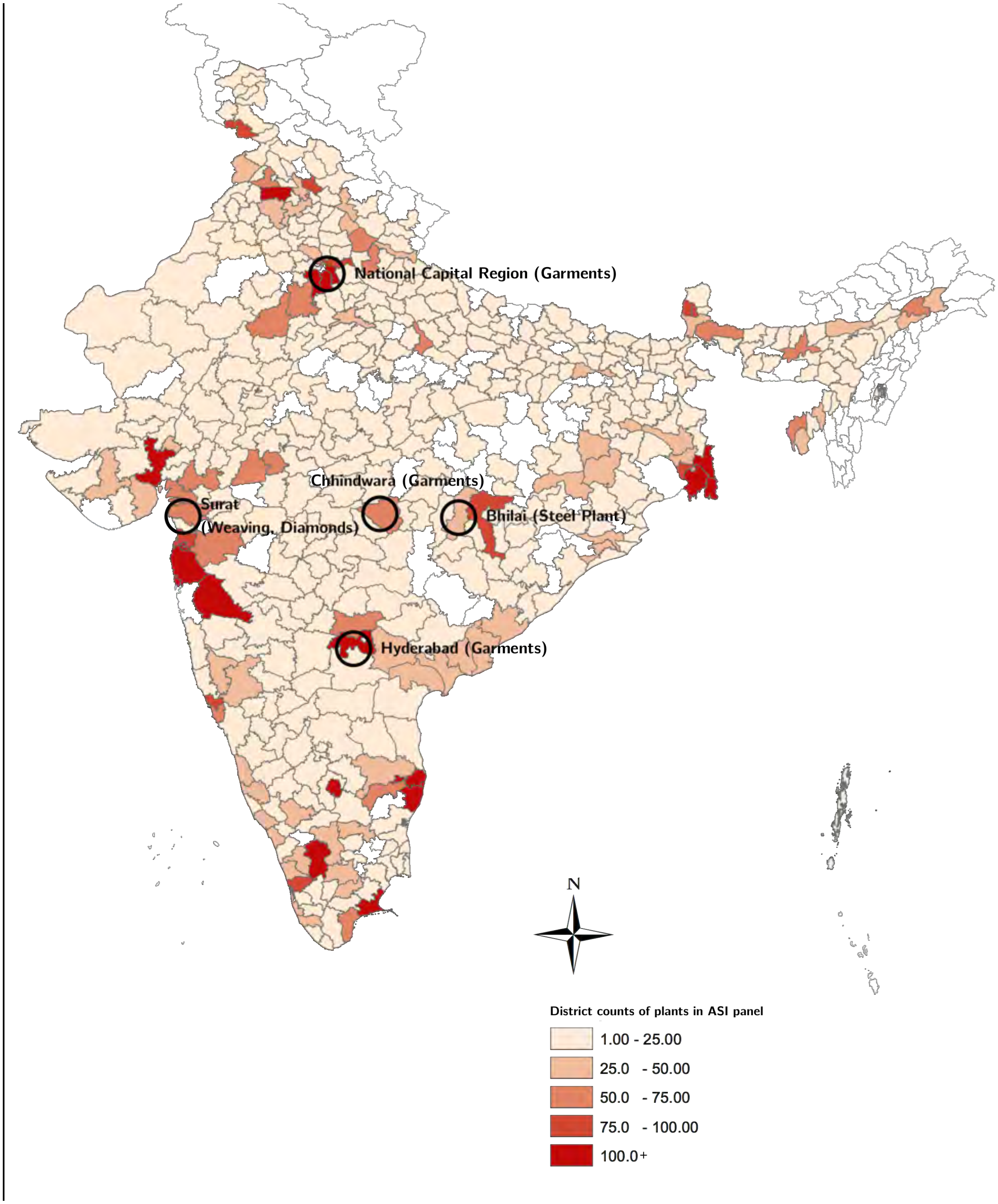

Figure A.1: Distribution of ASI plants over Indian districts, and location of micro-data sites

A.5 Temperature Labor Interactions

Column 1 of Table A.2 provides results from estimating (8). This is a simpler version of (9), where we do not allow for heterogeneity in labor and capital deciles. Column 2 provides a subset of coefficients from estimating (9).

A.6 Alternative Explanations and Robustness Checks

Floods: To create a variable corresponding to the flood exposure of a plant, we begin by assembling a data-set of all flooding incidents in India, between 1998 and 2013, along with the states affected by these floods. This data is obtained from the Dartmouth Flood Observatory. Because all flood incidents are not equally severe, the Dartmouth Flood Observatory uses a standard measure of the 'magnitude' of the flood defined as **Flood Magnitude =LOG(Duration x Severity x Affected Area)**. For each year, and each state, we calculate the total magnitude of all flooding. This state-level measure is then used as a proxy for the flood exposure of all plants in a state.

Conflict: India's Ministry of Labor and Statistics reports a variety of statistics corresponding to labor disputes every year (as part of a publication entitled 'Statistics on Industrial Disputes, Closures, Retrenchments and Lay-Offs'). Amongst these measures is the total number of workday minutes lost every year due to industrial disputes in each state. We use the log of this annual, state level measure to proxy for the exposure of plants to conflict.

To test whether conflicts and/or flooding incidents may explain our results, we include these variables in a set of regression equations as below. These correspond to our specifications in (5), (6), (8).

$$y_{it} = \alpha_i + \gamma_t + \beta T_{it} + \theta R_{it} + \omega_1 M_{it} + \omega_2 C_{it} + \epsilon_{it}. \tag{14}$$

Table A.2: Interaction of Labor and Capital with Temperature Bins

	Log Output Value	
	(1)	(2)
N^2	0.01330***	0.00524**
	(0.00395)	(0.00250)
N^3	0.01443***	0.00458**
	(0.00329)	(0.00199)
N^4	0.01611***	0.00470**
	(0.00344)	(0.00200)
N^5	0.01408***	0.00406**
	(0.00343)	(0.00183)
l	1.02147***	
	(0.15020)	
k	0.42824***	
	(0.08162)	
$l : N^2$	−0.00197***	
	(0.00051)	
$l : N^3$	−0.00160***	
	(0.00042)	
$l : N^4$	−0.00145***	
	(0.00042)	
$l : N^5$	−0.00162***	
	(0.00044)	
$k : N^2$	−0.00065**	
	(0.00028)	
$k : N^3$	−0.00080***	
	(0.00022)	
$k : N^4$	−0.00097***	
	(0.00023)	
$k : N^5$	−0.00081***	
	(0.00024)	
E^5		2.72543***
		(0.54153)
D^5		1.39263**
		(0.65202)
$E^5 : N^2$		−0.00605***
		(0.00221)
$E^5 : N^3$		−0.00525***
		(0.00153)
$E^5 : N^4$		−0.00562***
		(0.00156)
$E^5 : N^5$		−0.00525***
		(0.00146)
$D^5 : N^2$		−0.00382
		(0.00247)
$D^5 : N^3$		−0.00266
		(0.00185)
$D^5 : N^4$		−0.00294*
		(0.00178)
$D^5 : N^5$		−0.00274
		(0.00179)

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Column (1) reports results from Equation (8) and Column (2) from Equation (9). Robust standard errors clustered at district level are reported in parentheses. All models include plant and year fixed effects. l and k denote the log of the number of plant workers and the log of the capital measure. D^5 and E^5 are dummies for the 5th quantile of capital and labor. The full model corresponding to Column (2) contains coefficients corresponding to all quantiles of labor and capital. $(N^1, N^2, N^3, N^4, N^5)$ are days in $\{(0, 20], (20, 25], (25, 30], (30, 35], (35, 50]\}$ and all coefficients are relative to N^1.

$$y_{it} = \alpha_i + \gamma_t + \sum_{j=2}^{5} \beta_j N_{it}^{j} + \theta R_{it} + \omega_1 M_{it} + \omega_2 C_{it} + \epsilon_{it}. \tag{15}$$

$$y_{it} = \alpha_i + \gamma_t + \sum_{j=2}^{5} \alpha_j N_{it}^{j} + \omega_o \cdot k + \sum_{j=2}^{5} \omega_j N_{it}^{j} k + \beta_o \cdot l + \sum_{j=2}^{5} \beta_j N_{it}^{j} l + \sum_{j=2}^{5} \gamma_j N_{it}^{j} \cdot M_{it} \tag{16}$$

$$y_{it} = \alpha_i + \gamma_t + \sum_{j=2}^{5} \alpha_j N_{it}^{j} + \omega_o \cdot k + \sum_{j=2}^{5} \omega_j N_{it}^{j} k + \beta_o \cdot l + \sum_{j=2}^{5} \beta_j N_{it}^{j} l + \sum_{j=2}^{5} \gamma_j N_{it}^{j} \cdot C_{it} \tag{17}$$

where M_{it} is the total magnitude of flooding in year t for the state in which plant i is located. C_{it} is the log of total minutes lost to industrial disputes in year t for the state in which plant i is located. All other variables are as defined in (5), (6) and (8).

The estimates from these models are provided in Table A.3 below.

Table A.3: Temperature Effects on Log Plant Output Controlling for Disputes and Floods

	Log Output Value			
	(1)	(2)	(3)	(4)
rainfall	−0.00086	−0.00194	−0.00124	−0.00074
	(0.00374)	(0.00364)	(0.00355)	(0.00310)
T_{max}	−0.03184**			
	(0.01447)			
N^2		−0.00251***	0.01337***	0.01384***
		(0.00061)	(0.00394)	(0.00402)
N^3		−0.00196***	0.01443***	0.01460***
		(0.00073)	(0.00330)	(0.00335)
N^4		−0.00243***	0.01609***	0.01651***
		(0.00078)	(0.00345)	(0.00342)
N^5		−0.00272***	0.01411***	0.01412***
		(0.00085)	(0.00345)	(0.00348)
l			1.02172***	1.02304***
			(0.15029)	(0.15018)
$N^2 : l$			−0.00196***	−0.00197***
			(0.00051)	(0.00051)
$N^3 : l$			−0.00160***	−0.00160***
			(0.00042)	(0.00042)
$N^4 : l$			−0.00145***	−0.00145***
			(0.00042)	(0.00042)
$N^5 : l$			−0.00162***	−0.00163***
			(0.00044)	(0.00044)
Floods	−0.00022	−0.00027	0.00521	
	(0.00070)	(0.00072)	(0.01796)	
Disputes	0.00187	0.00192		0.01066
	(0.00304)	(0.00299)		(0.05614)
N^2:*Floods*			−0.00001	
			(0.00007)	
N^3:*Floods*			−0.00002	
			(0.00005)	
N^4:*Floods*			−0.00001	
			(0.00005)	
N^5:*Floods*			−0.00002	
			(0.00005)	
N^2:*Disputes*				−0.00005
				(0.00021)
N^3:*Disputes*				−0.00001
				(0.00015)
N^4:*Disputes*				−0.00004
				(0.00015)
N^5:*Disputes*				0.00000
				(0.00015)

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Robust standard errors clustered at district level reported in parentheses. Models have plant and year fixed effects. l is the log of the number of plant workers, *Floods* denotes the total magnitude of flood exposure and *Disputes* is time lost due to industrial conflict expressed in units of ten (8hr) workdays. $(N^1, N^2, N^3, N^4, N^5)$ are days in $\{(0, 20], (20, 25], (25, 30], (30, 35], (35, 50]\}$ and all coefficients are relative to N^1. Capital-temperature interaction coefficients not shown for brevity

A.7 Price Shocks

If higher temperatures induce changes in the prices of input materials, plant output value may also change. To check whether this may influence our reduced-form results, we first create a variable containing the price of the primary input for each plant. The primary input in the ASI is the input with the largest total value consumed.

Next, for the subset of plants where the listed primary input does not change, we regress price on linear and binned specifications of temperature as below:

$$P_{it} = \alpha_i + \gamma_t + \beta T_{it} + \theta R_{it} + \epsilon_{it}. \tag{18}$$

$$P_{it} = \alpha_i + \gamma_t + \sum_{j=2}^{5} \beta_j N_{it}^{j} + \theta R_{it} + \epsilon_{it}. \tag{19}$$

Here the dependent variable is the input price for plant i in year t. As before α_i, γ_t, are plant and year fixed effects. T_{it} and R_{it} are the average of daily highs and rainfall over year t for plant i. N_{it}^{j} are the number of days in the temperature bin j, as defined in the main text.

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. Changes in relative prices in the economy over the long term will result from general equilibrium effects that lie outside the scope of this paper.

Table A.4: Temperature Effects on Price

	Price of Primary Input	
	(1)	(2)
rainfall	-72.47	-23.15
	(903.73)	(874.73)
T_{max}	-813.44	
	(1759.91)	
N^2		92.04
		(106.38)
N^3		106.20
		(88.83)
N^4		97.73
		(102.36)
N^5		71.36
		(115.72)

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$ $^{*}p < 0.1$. Robust standard errors clustered at district level are reported in parentheses. All models have plant and year fixed effects. Prices are in Indian Rupees, $(N^1, N^2, N^3, N^4, N^5)$ are days in $\{(0, 20], (20, 25], (25, 30], (30, 35], (35, 50]\}$, and all coefficients are relative to N^1.

A.8 Power Outages

We do not have a direct measure of electricity supply at the plant level. However large plants are generally 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. This does not imply that large plants receive 24x7 supply, only that unscheduled, temperature-dependent outages may be relatively rare. It is this type of outage that is of relevance in our context.

Notwithstanding these facts, we also investigate the possible impacts of outages by constructing an annual, state-level panel of the difference between average monthly electricity supply and an imputed measure of average monthly electricity demand, as calculated by the Central Electricity Authority of India. This is commonly used as a measure of supply shortfalls and the required data is made publicly available by India's Central Electricity Authority in an annual publication called the Load Generation Balance Report.

Introducing this outage proxy into (5) and (6) does not change our results. Compare coefficients in Table A.5 with those in Table 4.

Table A.5: Temperature Effects on Log Plant Output

	Log Output Value	
	(1)	(2)
rainfall	-0.00066	-0.00202
	(0.00381)	(0.00357)
T_{max}	-0.02986	
	(0.01898)	
N^2		-0.00279^{***}
		(0.00077)
N^3		-0.00199^{**}
		(0.00089)
N^4		-0.00251^{***}
		(0.00093)
N^5		-0.00268^{**}
		(0.00108)
outages	-0.00003	-0.00003
	(0.00004)	(0.00004)

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$ $^{**}p < 0.1$. Robust standard errors clustered at district level are reported in parentheses. All models have plant and year fixed effects. $(N^1, N^2, N^3, N^4, N^5)$ are days in $\{(0, 20], (20, 25], (25, 30], (30, 35], (35, 50]\}$ and coefficients are relative to N^1.

A.9 Leads in Temperature Variables

As a robustness check, we also estimate (5) and (6) by replacing averaged or binned annual temperatures, with one year leads of average daily highs or the vector of days in temperature bins. Most coefficients become statistically insignificant with smaller point estimates, and standard errors are significantly higher. This is consistent with leads having little explanatory power, except through serial correlation with the contemporaneous temperature variables. Table A.6 reports results.

Table A.6: Log Plant Output Value Regressed on Leads in Annual Temperatures

	Log Output	Log Output
	(1)	(2)
T_{max} (lead)	-0.0243	
	(0.0168)	
N^2 (lead)		-0.0015
		(0.0012)
N^3 (lead)		-0.0021
		(0.0014)
N^4 (lead)		-0.0025*
		(0.0014)
N^5 (lead)		-0.0028*
		(0.0015)
rainfall	0.0072	0.0055
	(0.0044)	(0.0041)

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Robust standard errors clustered at district level provided in parentheses. All models have plant and year fixed effects. $(N^1, N^2, N^3, N^4, N^5)$ are days in $\{(0,20], (20,25], (25,30], (30,35], (35,50]\}$ and reported coefficients are relative to N^1.

A.10 Adaptation Evidence from ASI Plant Panel

Table A.7 examines how the effect of temperature on plant output changes over time by running the following regression:

$$y_{it} = \alpha_i + \gamma_t + \sum_{j=2}^{5} \beta_j N_{it}^j \times t + \theta R_{it} + \epsilon_{it}. \tag{20}$$

Where y_{it} is the logged value of output for plant i in year t and α_i and γ_t are plant and year fixed effects. N_{it}^j are the number of days in our different temperature bins, omitting the lowest bin. R_{it} is the average rainfall for the district in which a plant is located. We find evidence of decreasing temperature sensitivity over time, as shown in Column 1 of Table A.7.

We also examine how the temperature-output relationship varies across firms with electricity consumption that is above or below the median. To do so, we modify (6) by interacting temperature bins with a dummy variable that takes the value 1 when plant electricity consumption is above the median. We examine heterogeneity by electricity use because climate control is an electricity intensive technology. Our results are provided in Column 2 of Table A.7, showing that the output of plants with high electricity use 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.

Table A.7: Temperature Sensitivity of Output with Electricity Consumption and Time

	Log Output Value	
	(1)	(2)
rainfall	-0.00198	-0.00190
	(0.00352)	(0.00343)
N^2	-0.00449^{***}	-0.00734^{***}
	(0.00150)	(0.00125)
N^3	-0.00490^{***}	-0.00523^{***}
	(0.00120)	(0.00107)
N^4	-0.00442^{***}	-0.00575^{***}
	(0.00122)	(0.00110)
N^5	-0.00520^{***}	-0.00572^{***}
	(0.00128)	(0.00116)
N^2:time	0.00030^{*}	
	(0.00017)	
N^3:time	0.00044^{***}	
	(0.00012)	
N^4:time	0.00031^{**}	
	(0.00012)	
N^5:time	0.00040^{***}	
	(0.00014)	
D_e		-2.74975^{***}
		(0.52839)
$N^2 : D_e$		0.00996^{***}
		(0.00188)
$N^3 : D_e$		0.00677^{***}
		(0.00142)
$N^4 : D_e$		0.00699^{***}
		(0.00144)
$N^5 : D_e$		0.00661^{***}
		(0.00149)

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Robust standard errors clustered at district level are reported in parentheses. All models have plant and year fixed effects. D_e is an indicator that takes the value 1 when electricity consumption is above the median. $(N^1, N^2, N^3, N^4, N^5)$ are days in $\{(0, 20], (20, 25], (25, 30], (30, 35], (35, 50]\}$and all coefficients are relative to N^1.

A.11 Adaptation in Diamond Firms

This section presents results using data from a survey of 150 diamond cutting factories located in the city of Surat. These factories were drawn randomly from a list of all plants registered with the local diamond industry association.

We estimate a logit model using data from 750 processes in these 150 firms to describe the probability of using air-conditioning within a process as a function of (i) labor-intensity (ii) mechanization and (iii) importance in determining final diamond quality. The first of these variables is measured by the share of the firm's workers engaged in the process, the second by the share of the plant's machines used, and the third is a self-reported assessment by management on a scale of 1 to 5.

We estimate models with and without plant fixed effects. In the latter case we additionally include the number of workers (a proxy for size) as a control variable. Specifically the probability model with fixed effects is as follows

$$P_{ip}^{AC} = \frac{1}{1 + \exp(\beta_o + \beta_1 D_{ip} + \beta_2 m_{ip} + \beta_3 w_{ip} + FE_i)} \tag{21}$$

Here P_{ip}^{AC} is the probability that plant i has air conditioning installed in process p. D_{ip} is a dummy variable that takes the value 1 when a process is rated as highly important (ranking of 5 on a scale from 1 to 5) and 0 otherwise. m_{ip} is the share of the machines in plant i deployed in process p. w_{ip} is the share of the workers in plant i deployed in process p.

Alternatively the model can be estimated without fixed effects but including a measure of size (the total number of workers). This significantly simplifies the interpretation of the logit probability model.

$$P_{ip}^{AC} = \frac{1}{1 + \exp(\beta_o + \beta_1 D_{ip} + \beta_2 m_{ip} + \beta_3 w_{ip} + \beta_f s_i)} \tag{22}$$

Here s_{ip} is the size of plant i measured by the total number of workers employed.

Table A.8 summarizes our results. We find that diamond polishing units in Surat choose to preferentially cool high value and labor-intensive processes, consistent with decisions to optimally allocate air conditioning. Larger firms are more likely to use air conditioning than small plants.

Table A.8: Air Conditioning Investments in Diamond Firms

	AC Presence Indicator	
	(1)	(2)
Workers		0.00135**
		(0.00064)
Importance	3.07407***	1.27207***
	(0.41638)	(0.20769)
Worker Share	15.88670***	7.66071***
	(2.43892)	(1.47864)
Machine Share	−17.53345***	−9.48015***
	(1.91772)	(1.08042)

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{**}p < 0.1$. Model 1 is estimated with firm fixed effects and Model 2 has no fixed effects. *Workers* denotes the number of workers, *Worker Share* is the fraction of total workers employed in a process, *Machine Share* is the fraction of total machines in a process and *Importance* is the manager rating on a scale of 1 (not important) to 5 (very important), describing how much a process contributes to final diamond quality.

A.12 Decomposition of Temperature Effects on Plant Output

Figure A.2 is a modification of Figure 4 in the main text. Here we model the residual term in (9) using quadratics in temperature bins.

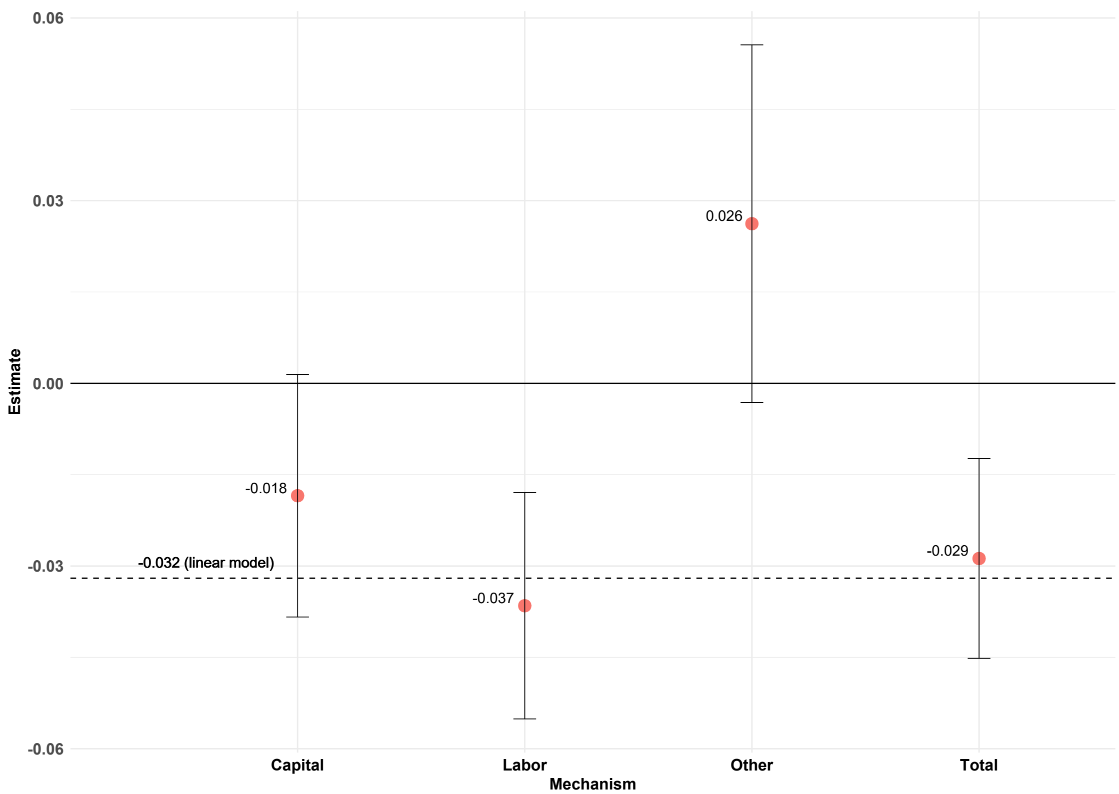

Figure A.2: Marginal effect of temperature on output decomposed into labor, capital and a residual modeled by quadratics in temperature bins. 90 percent confidence intervals with robust standard errors clustered at district level. Dotted line plotted at the total temperature effect on output from (5) and solid line is at zero. Predictions use $\Delta N = (-5.89, -5.73, -11.76, 1.93, 21.28)$ where $(N^1, N^2, N^3, N^4, N^5)$ are days in $\{(0, 20], (20, 25], (25, 30], (30, 35], (35, 50]\}$.