Temperature, Particulate Matter, and Mental Health in India

Kavita Sardana* Jamie T Mullins† Onupurba Das ‡

April 2024

Abstract

Growing evidence from the emerging and developed world indicates a link between environmental conditions and mental wellbeing, however, little is known about the relationship between this relationship in the developing world where environmental conditions are often much harsher and the demands on mental health can be substantial. In this paper we investigate the link between exposure to high temperatures and particulate matter and self-reported mental health in India data from a massive, nationwide survey linked to conditions at each respondent's residence in the time leading up to their enumeration. While, we find that exposure to higher temperatures and concentrations of particulate matter in the lead up to enumeration is predictive of poorer self-reported mental well-being, we also identify an adaptive spillover effect whereby harms from PM exposure are lower-magnitude on hot days.

*Ashoka University; kavita.sardana@ashoka.edu.in.
†University of Massachusetts Amherst; jmullins@umass.edu
‡University of Massachusetts Amherst; odas@umass.edu

1 Introduction

Mental health is increasingly recognized as a major dimension of general human health and well-being, with mental and addictive disorders estimated to drive from a fifth to a third of all years lived with disability globally (Vigo, Thornicroft, and Atun 2016; Rehm and Shield 2019). Focusing more narrowly, it is estimated that as of 2019, approximately 5% of adults globally suffer from depression (WHO 2023). This rate is higher among those over 55 globally (5.7%), and among those over 55 living in India (7.1%) (WHO 2023; IHME 2019). In this examination, we will consider how ambient environmental conditions impact self-reported measures of depression symptoms among older Indians.

As the importance and prevalence of mental health concerns and depression specifically have become ever more apparent, it has also become increasingly clear that environmental factors play a significant role in mental health. For instance, a series of investigations have linked realizations of high-temperatures to increased incidences of negative mental health outcomes ranging from self-assessments of mental well-being (Obradovich et al. 2018; Mullins and White 2019) to suicides in both developed (Mullins and White 2019; Burke et al. 2018) and emerging (Burke et al. 2018; Carleton 2017) economies. Air pollution too has been increasingly linked to a range of mental health problems, including anxiety, depression, and suicide (Chen, Oliva, and Zhang 2018; Braithwaite et al. 2019; Gu et al. 2020; Persico and Marcotte 2022; Molitor, Mullins, and White 2023; Zhang et al. 2024).

Even as evidence has mounted that temperatures and air pollution independently compromise mental well-being, it is important to recognize that environmental exposures do not happen in isolation. In addition to assessing the individual relationships between self-reported depressive symptomatic and both temperatures and PM, we therefore also evaluate cross-exposure impacts of simultaneous exposures to high temperatures and high ambient PM levels.

The investigations in this paper, contribute to the understanding of the short-term associations between mental health outcomes and two environmental conditions, namely, air pollution and high temperatures. Our analysis relies on three primary data sources that allows us to measure self-reported mental health status, ambient $PM_{2.5}$ concentrations, average temperatures for each states in India from 2017-2019. Specifically, we use LASI data on mental health based on the ten-question, Center for Epidemiological Studies Depression (CES-D) scale, which focuses on the 7-day period leading up to the survey. Mean temperature averaged over 7 days and monthly average $PM_{2.5}$ ambient concentrations.

Our empirical strategy relies on matching measures of mental health state and week including the date of the survey to temperature at fine spatial and temporal scales (state and week including the date of the survey) and air quality (state and month). We then apply panel fixed effect regression models, with rich fixed effects to include location by year fixed effects. Though we estimate our results to allow for the relationship between environmental conditions and mental health outcome more flexibly to allow for non-linearities, our main results are derived from the evidence of linearity between the outcome variable of interest and the environmental conditions.

We find that the incidence of depressive symptoms among older Indians increases in both temperatures and ambient $PM_{2.5}$ levels. This is the first evidence that these environmental stressors are harmful to mental well-being among older adults in India, and the first evidence that both high temperatures and PM levels are harmful to mental health identified simultaneously in any population. We are then able to go further and consider non-additive effects of simultaneous exposures to these two environmental stressors. We we find somewhat larger increases in depressive symptoms in response to higher ambient $PM_{2.5}$ concentrations when temperatures are low than when temperatures are high. This may be because some adaptive measures to easily observable high temperature increases also serve to moderate the harm from less-observable rises in $PM_{2.5}$, a sort of "spillover adaptation" which has not been previously identified.

The remaineder of the paper is organized as follows. The next section provides details on the setting and data of our investigation. Section 3 lays out our empirical approach and Section 4 describes our results. We discuss our findings and conclude in Section 5

2 Data

This investigation is based on linking environmental conditions to self-reported measures of mental health. In particular, we link weather data and measures of fine particulate matter on the day of and period preceding survey enumeration from Wave 1 of the Longitudinal Aging Study of India (LASI). LASI is a survey of people aged 45 and older and their partners, living in private households in India. The sample is representative of Indias 29 states and 6 union territories and four selected metropolitan cities (Delhi, Kolkata, Mumbai, and Chennai). This investigation relies solely on data from Wave 1, collected between 2017 and 2019, which is all that is currently available.

LASI recruited 42,949 households (96% participation rate) and 72,250 respondents (87% participation rate). The sample was recruited using a stratified, multi-stage cluster sampling of all households in India. The specifics of sampling approach for the study are available in (Arokiasamy et al. 2012). A household and an individual survey are complete by or for each respondent. The household survey characterizes participants physical environment and household finances. This survey is completed by any knowledgeable household member aged 18 years or older. The individual survey is completed only by age-eligible household members and their spouses unless a proxy respondent was needed. This survey characterizes participant demographics, health includ-ing mental health and health behaviors. All survey questions are translated into the local dialect (e.g., Telegu, Malayalam) and interviews are done in the language of respondents choice.

2.1 Weather Data

Onto each LASI survey response we merge weather data based on the date of enumeration and latitude and longitude of the respondent's residence. Weather data is from the ERA5-Land reanalysis data product maintained by the European Centre for Medium-Range Weather Forecasts (Muñoz

Sabater et al. 2019). We begin with the hourly observations from this dataset on temperature, dew-point temperature, precipitation, barometric pressure, and cloud cover for points on a $0.1° \times 0.1°$ over all of India. We aggregate the hourly values to the daily level, capturing minimum, mean, and maximum values for each variable except precipitation for which hourly values are summed to a daily total. We next create a series of bin indicators for each weather variable, which capture whether daily values of each variable fell in specified ranges. For instance, daily mean temperatures – which are the focus of our investigation of weather conditions in this study – are grouped in to 2°C wide bins, ranging from < 20°C in 2°C increments 20-22°C to >38°C.

Similar bins are created for each measure of daily weather, and these indicators are then summed across the day of LASI enumeration and the preceding six days. This procedure results in a series of counts of the number of days in the week leading up to and including the day of LASI enumeration in which weather conditions fell in each bin-range. This approach allows us to flexibly estimate (or control for) the effects of weather conditions at the place of residence over the week leading up to the observed survey responses. See Appendix Section for additional details, bin ranges, and summary statistics.

2.2 Air Quality Data

Air pollution data from monitors is sparse in India during the 2017-2019 period of our study, especially outside major cities. We therefore utilize modelled data on ambient $PM_{2.5}$ concentrations at the monthly level from the Satellite-Based Application For Air Quality Monitoring and Management at National Scale (SAANS). This data provides monthly average $PM_{2.5}$ concentrations for all of India on a 1km grid (Dey et al. 2020).

We estimate the average $PM_{2.5}$ concentration at the place of residence of each LASI respondent for 30-day period leading up to the day of the survey. This is done by taking the weighted average of the $PM_{2.5}$ concentrations from the nearest gridcell in the SAANS data for the calendar month of the survey completion and the prior month with weights assigned based on how much of the 30-day period falls in each calendar month.

We also create six indicator variables which take on the value 1 if the monthly $PM_{2.5}$ measure falls in 25 $\mu g/m^3$ bins ranging from 0-25$\mu g/m^3$ up to >125$\mu g/m^3$.

2.3 Outcomes

Our main outcomes of interesting are self-reported mental health measures collected by LASI. In particular, we examine scores from the 10-question Center for Epidemiological Studies Depression (CES-D) scale. This instrument asks respondents to consider how frequently each of 10 depression-related symptoms were experienced during the 7-day period leading up to the survey. Options are: rarely or none of the time (less than 1 day), some or a little of the time (1-2 days), occasionally or moderate amount of the time (3-4 days), most or all of the time (5-7 days). Responses are coded with integer values from 1-4, with higher values indicating more frequent experiences of negative symptoms. Summary statistics of some relevant demographic and mental health variables and relevant demographic are provide in Table 1.

2.4 Privacy Considerations

In order to protect the identities and privacy of LASI respondents, state is the most specific information provided for respondent place of residence. By special request, LASI data administrators provided us with weather and $PM_{2.5}$ conditions linked to each LASI response based on exact date of enumeration and latitude and longitude of residence, however values of these variables were only provided after they were aggregated over extended time periods (seven days for weather conditions and thirty days for $PM_{2.5}$) in order to prevent *ex post* location identification.

Table 1: Summary Statistics

	(1) Mean	(2) StdDev.	(3) Min	(4) Max	(5) Obs
Female	0.578	0.494	0	1	66,287
Age	57.801	11.592	18	116	66,287
Caste					
Scheduled Caste(SC)	0.169	0.375	0	1	66,287
Scheduled Tribe(ST)	0.163	0.370	0	1	66,287
Others(O)	0.386	0.487	0	1	66,287
No caste(NC)	0.282	0.450	0	1	66,287
Rural	0.657	0.474	0	1	66,287
Access	0.664	0.472	0	1	66,287
ADL	0.202	0.757	0	5	66,274
IADL	0.404	0.815	0	3	66,274
CES-D (higher is more often)	9.466	4.028	0	30	66,287
Trouble concentrating	1.521	0.783	1	4	66,284
Feel depressed	1.545	0.785	1	4	66,285
Feel Tired	1.842	0.88	1	4	66,286
Feel afraid	1.376	0.722	1	4	66,285
Feel satisfied	2.02	0.985	1	4	66,284
Feel alone	1.486	0.797	1	4	66,285
Feel unusually bothered	1.626	0.835	1	4	66,283
Feel everything was effort	1.756	0.923	1	4	66,281
Feel hopeful about future	2.135	1.007	1	4	66,281
Feel happy	2.5	0.985	1	4	66,283

3 Empirical Approach

To establish the short-term causal impact of weather and air pollution on our measures of mental health, we adopt a panel fixed effect methodology based on location-by-month and location-by-year fixed effects in all specifications. While we control for a range of environmental conditions, our focus will be on the effects of temperature and $PM_{2.5}$ concentration levels. Our main specification is:

$$Y_{irsdmy} = f(Temp_{rdmy}) + g(PM_{rdmy}) + X_{rdmy} + \delta_{sm} + \delta_{sy} + \epsilon_{irsdmy} \tag{1}$$

Y_{irsdmy} is the outcome variable, and for this investigation is most commonly the aggregate CES-D score per individual respondent, i, living at residential location r, in state s, collected on day, d, of month m, in year, y. Our specification includes a function of the temperatures in the seven days leading up to the day of enumeration and a function of the average $PM_{2.5}$ concentration in the 30 days leading up to the day of enumeration. Our main estimates are based on a function of temperatures which includes counts of the number of days (of the last seven) with maximum temperatures in each of six 4°C bins with the bin for <20°C omitted as the baseline category. In other specifications, the mean of the daily maximum temperature over the last seven days enters linearly. For $PM_{2.5}$, our main specification includes indicator variables for which of the six 25 $\mu g/m^3$-wide bins the 30-day mean concentration was in. The lowest pollution category, 0-25$\mu g/m^3$, is omitted.

All specifications control flexibly for weather conditions including precipitation, dew point temperature, barometric pressure, and cloud cover using counts of days on which conditions fell into various bins. Additionally, state-by-month and state-by-year fixed effects control for time-constant local conditions, local seasonality, and state-and-year-specific idiosyncratic shocks. Because we are limited by data use restrictions to the coarse geographic scale of state for our fixed effects, we also control for a series of individual- specific respondent characteristics including gender, 5-year age cohort, caste, and education level (<5, 5-9, or >9 years), as well as household-specific characteristics including: annual household consumption quintile and rural residence.

Identifying variation is therefore based on deviations in conditions at the residence of LASI respondents on the day of (and days leading up to) enumeration relative to norms in the state of residence for the month, once individual demographic characteristics and statewide shocks are accounted for. Because the date of enumeration for each LASI respondent was determined based on logistical considerations long in advance, the deviation of conditions at the respondent's residence on the day of enumeration from local, seasonal norms, is likely to be as good as random for each respondent. Based on this identifying assumption, we interpret our estimates as the causal effects of recent, daily temperatures and monthly $PM_{2.5}$ on the self-assessed mental health outcomes we consider.

4 Findings

4.1 Baseline Impacts

The coefficient estimates for temperature and $PM_{2.5}$ bins are shown graphically in Figure in 1 (and reported in Table 1). CESD-10 scores generally increase as both temperatures and concentrations of $PM_{2.5}$ increase. Because CESD scores measure the frequency of depressive symptoms, our estimates provide causal evidence that higher temperatures and poorer air quality negatively impact mental well-being. Because all these estimates derive from a single regression, our estimates are evidence that the temperature effects are clear even when controlling for $PM_{2.5}$ levels and visa versa.

The magnitude of our estimates suggest that if one of the last seven days had a maximum temperature between 32-36°C rather than being <20°C, CESD-10 scores would be 0.173 points (or 4.3% of a standard deviation) higher. Our estimates for $PM_{2.5}$ imply that if average concentrations over the preceding 30-days were over $125\mu g/m^3$ rather than being below $25\mu g/m^3$, CESD-10 scores would be 0.74 points (or 18.4% of a standard deviation) higher. Coefficients for most bins but are significant at the 1% level, and all are significant at the 5% level.

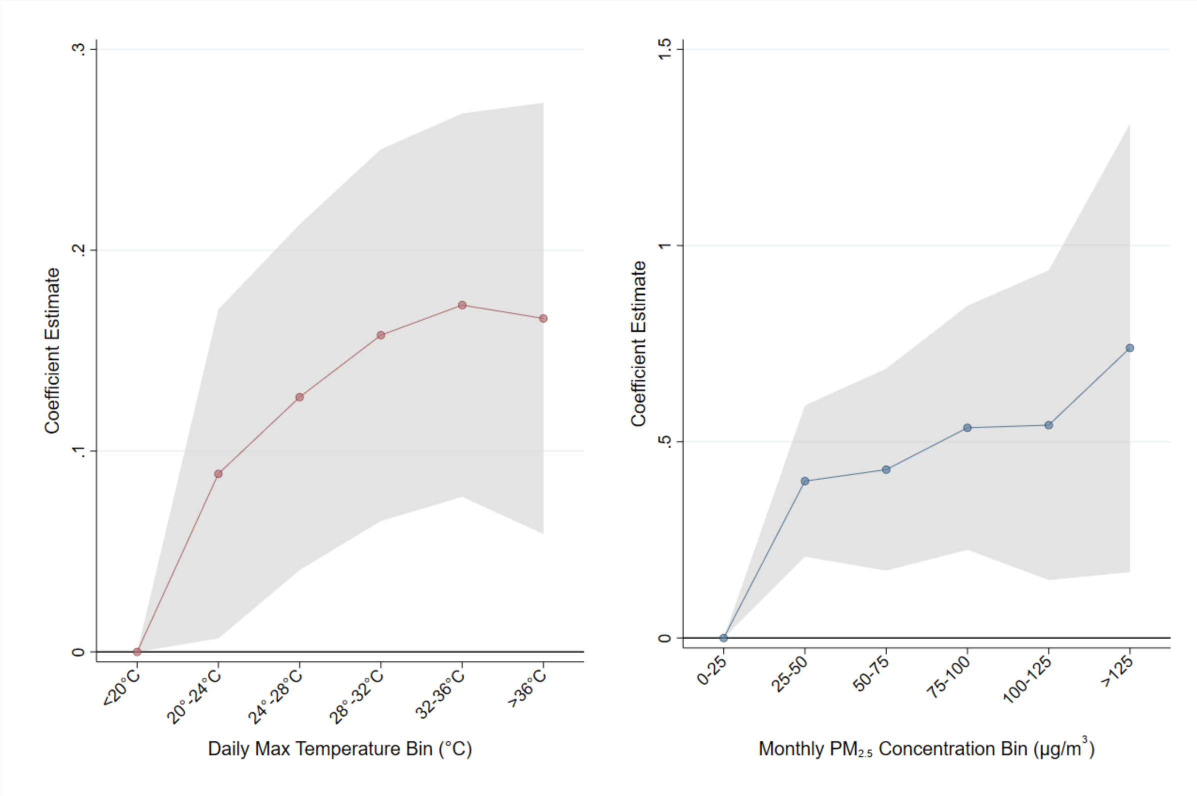

Figure 1: Baseline Estimates with Maximum temperature and PM Bins

Notes: Shaded areas represent 95% confidence intervals based on standard errors clustered by household. All estimates are from a single regression with the CESD-10 score as the outcome. The temperature coefficients can be interpreted as the effect of one additional day in the relevant bin, relative to a day below 20°C . The $PM_{2.5}$ coefficients can be interpretted as the effect of the 30-day average concentration falling in the indicated range rather than below $25ug/m^3$.

Table 2: Estimates for Figure 1

	(1) CESD-10 Score
Panel A - Max Temperature Bins	
$20°$-$24°C$	0.08864**
	(0.04178)
$24°$-$28°C$	0.12684***
	(0.04389)
$28°$-$32°C$	0.15770***
	(0.04722)
$32°$-$36°C$	0.17267***
	(0.04869)
$> 36°C$	0.16604***
	(0.05478)
Panel B - $PM_{2.5}$ Bins	
$PM_{2.5}$ 25-50$\mu g/m^3$	0.39984***
	(0.09839)
$PM_{2.5}$ 50-75$\mu g/m^3$	0.42907***
	(0.13122)
$PM_{2.5}$ 75-100$\mu g/m^3$	0.53573***
	(0.15861)
$PPM_{2.5}$ 100-125$\mu g/m^3$	0.54240***
	(0.20135)
$PM_{2.5} > 125\mu g/m^3$	0.73917**
	(0.29139)
N	66,287

*- p<0.10 **- p<0.05 ***-p<0.01

Notes: Reports the estimates used to construct Figure 1. All estimates are from a single regression with the CESD-10 score as the outcome. Standard errors in parenthesis are clustered by household. Regression includes state-by-month and state-by-year fixed effects as well as the full set of weather and demographic controls described in the text.

For both temperature and $PM_{2.5}$, the dose-response relationship with CESD-10 scores appears more-or-less linear across the range of conditions in our sample. One notable exception to the overall linearity of the estimated relationships is the coefficient for the highest temperature bin (>36°C), which takes on a smaller value than the coefficient of the next hottest bin. One potential explanation for this, is the fact that LASI survey enumerators had leeway within the assigned day regarding the exact time of enumeration. During the hottest stretches, it is more likely that enumerators would seek to schedule surveys during cooler times of the day, such as the morning. Such strategic intra-day heat avoidance, if it occurred more frequently during the hottest periods, could explain attenuated estimates for the highest temperature bin that we observe. We also note however, that (Carleton 2017) also identifies a distinctly concave dose response relationship between temperatures and suicide in India.

Given general the linear character of our estimated dose response relationships, we next estimate models with simple linear terms which simplify interpretation and further investigation. In 3, column 1 reports the estimates for the coefficients on the linear measures of the 7-day average of daily maximum temperatures and the 30-day average ambient $PM_{2.5}$ concentration. These coefficients suggest that a 1°C increase in the 7-day average of the daily max temperature is associated with a 0.023 point increase in the CESD-10 score, while a $10\mu g/m^3$ increase in the 30-day average $PM_{2.5}$ level is associated with a magnitude increase in CESD-10 scores of 0.044 points. Column 2 adds an interaction between the linear temperature and $PM_{2.5}$ measures providing more precise estimates and yielding a negative coefficient on the interaction term. This suggests that the effects of temperature and air pollution are not simply additively separable as is implicitly assumed by most estimation approaches, but instead that higher levels of temperature may somehow moderate damages from $PM_{2.5}$ and/or visa versa.

We next decompose the linear estimates of maximum daily temperature and $PM_{2.5}$ by the level of the other exposure measure by interacting the linear measure of each with bin-indicators for the other. Estimates are depicted in Figure 2 and show that while an increase in temperature appears to

Table 3: Temperature Measure Comparison

	(1) CESD-10 Score	(2) CESD-10 Score	(3) CESD-10 Score	(4) CESD-10 Score
Max Daily Temp ($^{\circ}C$)	0.02309*	0.05841***		
	(0.01273)	(0.01828)		
Min Daily Temp ($^{\circ}C$)			0.05844***	0.07994***
			(0.01370)	(0.01828)
$PM_{2.5}(\mu g/m^3)$ /10	0.04433*	0.23900***	0.05343**	0.12856***
	(0.02510)	(0.07533)	(0.02519)	(0.04962)
Interaction		-0.00706***		-0.00449*
		(0.00256)		(0.00252)
N	66,287	66,287	66,287	66,287

*- p<0.10 **- p<0.05 ***- p<0.01

Notes: Standard Errors clustered by household

have a fairly consistent impact on CES-D scores across $PM_{2.5}$ levels (righthand panel), the harmful impact of increased-$PM_{2.5}$ levels are larger at low temperatures. Put another way, increases in air pollution appear somewhat less harmful when it is hot out. It could be that avoidance behaviors taken in the face of observed high temperatures – staying indoors, using air conditioning, and/or avoiding strenuous activities, for example – may provide some adaptive benefits against the harms of air pollution exposure. We don't see much evidence of such multi-exposure adaptation arising at high levels of $PM_{2.5}$, perhaps because changes in air pollution are less salient, or typical avoidance behaviors – wearing masks for instance – may not ameliorate harm from increased temperatures.

In any case, this evidence of intra-exposure adaptive spillover suggests a more complicated picture for projecting future damages from air pollution. Our results suggest that not only air quality, but also temperatures and perhaps available means of adaptation to temperature need also be considered in a careful implementation of such an exercise.

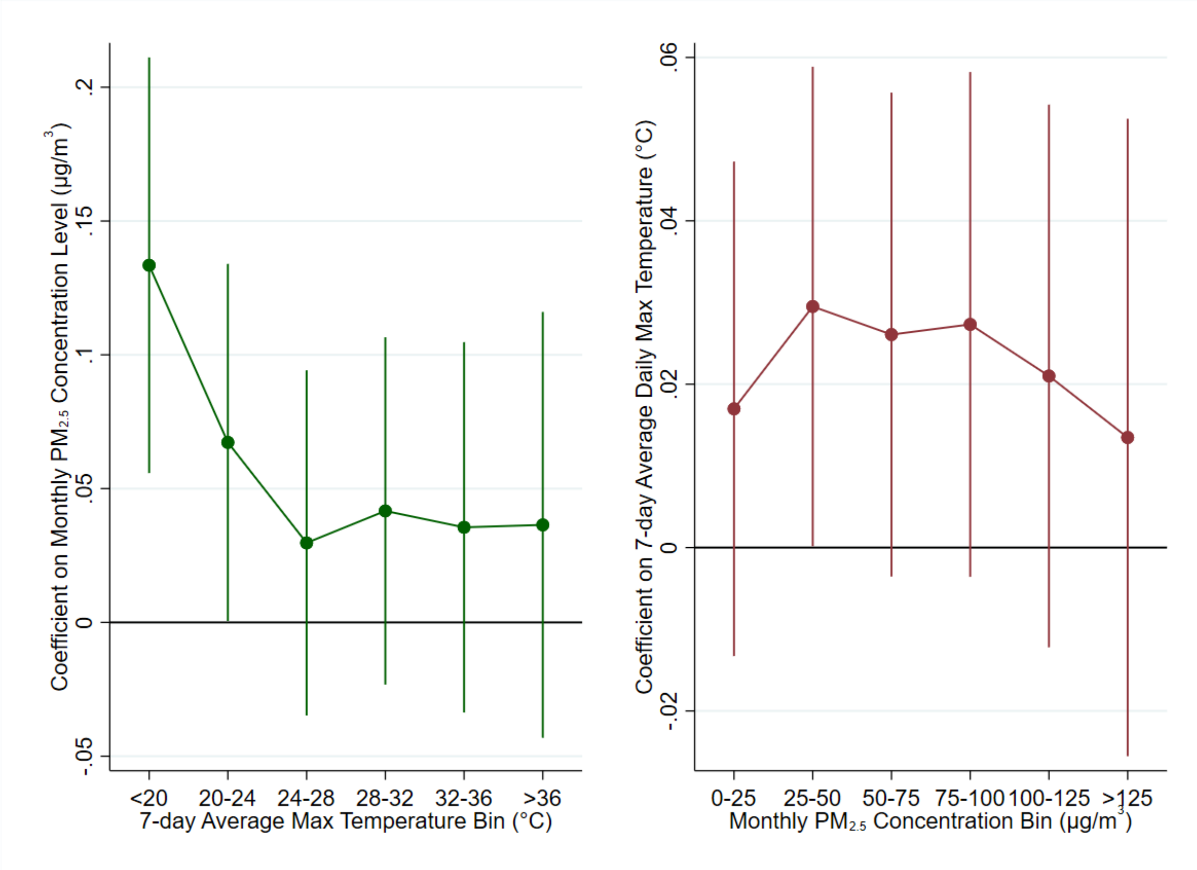

Figure 2: Heterogeneity Results by Level of Other Exposure

Notes: Left panel shows coefficients on the linear measure of $PM_{2.5}$ concentration separately for each temperature bin. Right panel shows coefficients of 7-day average daily maximum temperature by $PM_{2.5}$ level. Estimates come from a single regression with the CESD-10 score as the outcome. 95% confidence intervals are based on standard errors clustered by household. Regressions include state-by-month and state-by-year fixed effects as well as the full set of weather and demographic controls described in the text.

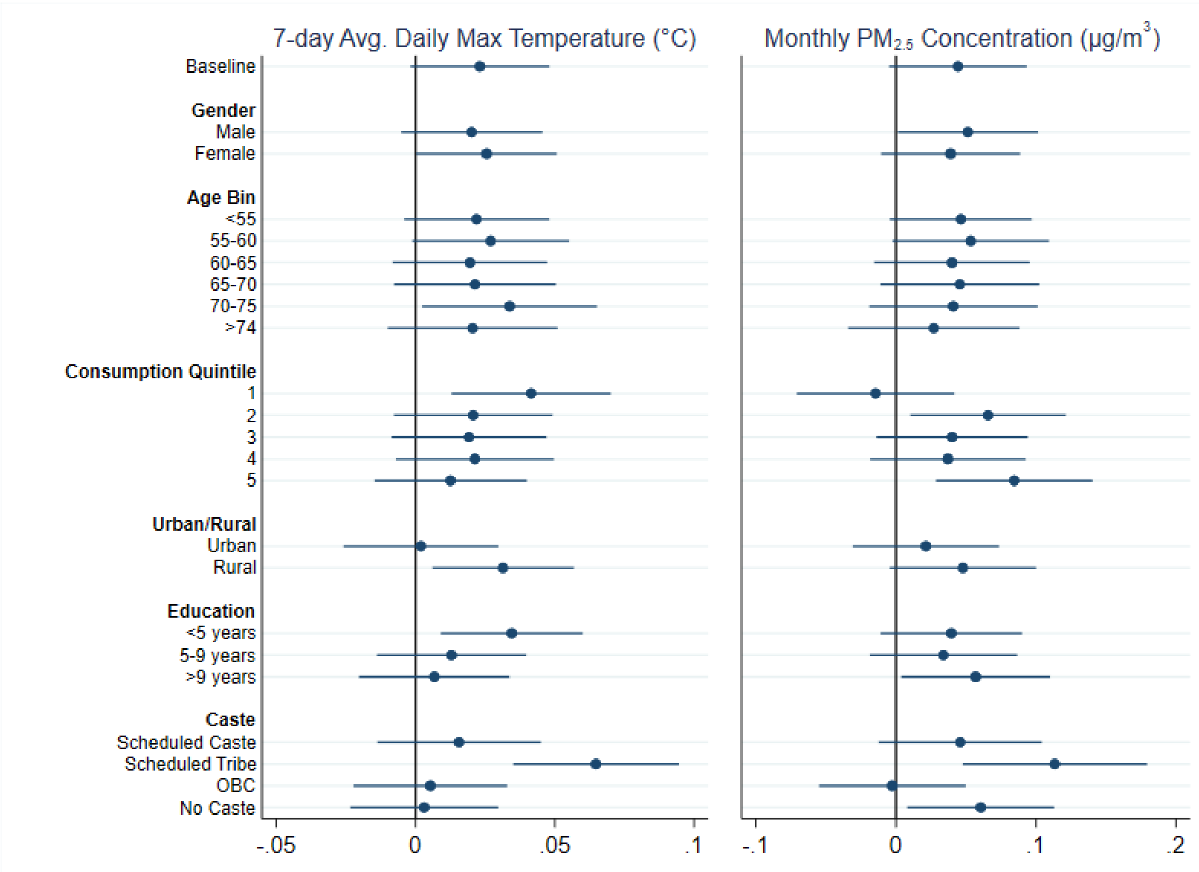

Figure 3: Heterogeneity Results with respect to Linear Temperature and PM Levels

Notes: Coefficients plotted for interaction of indicators for each demographic category with the linear measure of temperature in the left-hand graph and the linear measure of $PM_{2.5}$ in the right-hand graph. Estimates for each demographic grouping are from a single regression with the CESD-10 score as the outcome. 95% confidence intervals are based on standard errors clustered by household. Regressions include state-by-month and state-by-year fixed effects as well as the full set of weather and demographic controls described in the text.

4.2 Heterogeneity Results

Figure 3 reports estimates by socio-economic subgroup for the main linear measures of temperature and $PM_{2.5}$, when no interaction term is included. Overall, we don't see much heterogeneity in the responses of CESD-10 scores to the linear measures of temperature and $PM_{2.5}$, by group. Coefficients are generally similar between men and women, and across age groups. There is some suggestive evidence that the least affluent (those in the first consumption quintile) and the least educated are more senesitive to temperature increases, perhaps because these groups have the least access to means of avoidance of high temperatures. Additionally, we see some evidence rural respondents are more sensitive to increases in both temperature and air pollution. This is consistent with findings of (Molitor, Mullins, and White 2023) regarding the higher responsiveness of suicides in the United States to both wildfire smoke and higher temperatures in rural areas, and may be because of higher likelihood of exposures to outside conditions among rural populations in both cases. Finally we see that respondents identifying as among Scheduled Tribes appear more susceptible to harm from both higher temperatures and $PM_{2.5}$. Given that this group tends to be less affluent and concentrated in more rural areas, this is generally consistent with our other

5 Discussion and Conclusion

While there is growing evidence that environmental conditions impact mental health and well-being, environmental exposures never happen in isolation. This investigation provides the first evidence of non-additievly separable effects of environmental exposures on mental health. Specifically, we find somewhat larger increases in depressive symptoms in response to higher ambient $PM_{2.5}$ concentrations when temperatures are low than when temperatures are high. This may be because some adaptive measures to easily observable high temperature increases also serve to moderate the harm from less-observable rises in $PM_{2.5}$, a sort of "spillover adaptation" which has not been previously identified.

Whether or not this adaptation story is correct, the significant interactive relationship between high temperatures and $PM_{2.5}$ concentrations identified here underscores the necessity of accounting for temperatures when considering changes in PM levels and *vice versa*. Practically, this means accounting for climate-change-induced temperature changes when considering the benefits of pollution rregulation and factoring in anticipated PM levels when calculating expected damages from climate change (or avoided harm from mitigation efforts). Accounting for cross-factor interactions in harm is especially consequential when considering scenarios in which both temperatures and air quality may be impacted, such as the adoption of electric vehicles or the transition from fossil fuels to renewable sources of electricity generation. Our results show that any such accounting which simply sums the benefits of improved air quality and temperature decreases will inappropriately inflate achievable benefits or avoidable damages.

We all operate in complex exposure environments. While we are unable to characterize the full interaction space with this investigation, we hope our identification of one important interactive relationship in the environment-mental-health dose-response function spurs further investigation of exposure interactions and careful consideration of such in projection exercises.

References

Arokiasamy, Perianayagam, David Bloom, Jinkook Lee, Kevin Feeney, and Marija Ozolins. 2012. "Longitudinal aging study in India: Vision, design, implementation, and preliminary findings." In *Aging in Asia: Findings from new and emerging data initiatives.* National Academies Press (US).

Braithwaite, Isobel, Shuo Zhang, James B Kirkbride, David PJ Osborn, and Joseph F Hayes. 2019. "Air pollution (particulate matter) exposure and associations with depression, anxiety, bipolar, psychosis and suicide risk: a systematic review and meta-analysis." *Environmental health perspectives* 127 (12): 126002.

Burke, Marshall, Felipe González, Patrick Baylis, Sam Heft-Neal, Ceren Baysan, Sanjay Basu, and Solomon Hsiang. 2018. "Higher temperatures increase suicide rates in the United States and Mexico." *Nature Climate Change* 8 (8): 723–729.

Carleton, Tamma A. 2017. "Crop-damaging temperatures increase suicide rates in India." *Proceedings of the National Academy of Sciences* 114 (33): 8746–8751.

Chen, Shuai, Paulina Oliva, and Peng Zhang. 2018. *Air pollution and mental health: evidence from China.* Technical report. National Bureau of Economic Research.

Dey, Sagnik, Bhavesh Purohit, Palak Balyan, Kuldeep Dixit, Kunal Bali, Alok Kumar, Fahad Imam, et al. 2020. "A satellite-based high-resolution (1-km) ambient PM2. 5 database for India over two decades (2000–2019): applications for air quality management." *Remote Sensing* 12 (23): 3872.

Gu, Hejun, Weiran Yan, Ehsan Elahi, and Yuxia Cao. 2020. "Air pollution risks human mental health: an implication of two-stages least squares estimation of interaction effects." *Environmental Science and Pollution Research* 27:2036–2043.

Institute for Health Metrics and Evaluation. 2019. "Global Burden of Disease Study 2019 Reference Life Table."

Molitor, David, Jamie T Mullins, and Corey White. 2023. "Air pollution and suicide in rural and urban America: Evidence from wildfire smoke." *Proceedings of the National Academy of Sciences* 120 (38): e2221621120.

Mullins, Jamie T, and Corey White. 2019. "Temperature and mental health: Evidence from the spectrum of mental health outcomes." *Journal of Health Economics* 68:102240.

Muñoz Sabater, Joaqun, et al. 2019. "ERA5-Land hourly data from 1981 to present." *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)* 10 (10.24381).

Obradovich, Nick, Robyn Migliorini, Martin P Paulus, and Iyad Rahwan. 2018. "Empirical evidence of mental health risks posed by climate change." *Proceedings of the National Academy of Sciences* 115 (43): 10953–10958.

Persico, Claudia, and Dave E Marcotte. 2022. *Air quality and suicide.* Technical report. National Bureau of Economic Research.

Rehm, Jürgen, and Kevin D Shield. 2019. "Global burden of disease and the impact of mental and addictive disorders." *Current Psychiatry Reports* 21:1–7.

Vigo, Daniel, Graham Thornicroft, and Rifat Atun. 2016. "Estimating the true global burden of mental illness." *The Lancet Psychiatry* 3 (2): 171–178.

WHO. 2023. *Depressive disorder (depression),* March.

Zhang, Peng, Tamma Carleton, Liguo Lin, and Maigeng Zhou. 2024. "Estimating the role of air quality improvements in the decline of suicide rates in China." *Nature Sustainability,* 1–10.

Temperature, Particulate Matter, and Mental Health in India

Kavita Sardana, Jamie T. Mullins, and Onupurba Das

Supplemental Appendix for Online Publication

SA1 Data

SA1.1 Pollution and weather data

Hourly data is aggregated to daily mean, minimums, and maximums temperatures. Temperature data are grouped in to 2°C wide bins, ranging from 20-22°C to >38°C. Number of days in each temperature bin are summed over day of interview and preceding 6 days in the sample. The independent variable of interest is therefore count of days for which a state had maximum temperature in each bin in 7-days. In alternative specification for all outcomes we also use a simple specification that uses continuous mean/maximum/minimum temperature averaged over 7 days in place of temperature bins.

SA1.1.1 Relative Humidity

The data source on humidity is derived from the ERA5 reanalysis data product from the European Centre for Medium-Range Weather Forecasts. This contains hourly data on dewpoint for points on a $0.1^\circ \times 0.1^\circ$ grid for India over the period. Hourly data is aggregated to daily average dewpoint. Dewpoint data are grouped in to 2°C wide bins, ranging from 8-10°C to >28°C. Number of days in each average dewpoint bin are summed over day of interview and preceding 6 days in the sample. The independent variable of interest is therefore count of days for which a state had an average dewpoint in each bin in 7-days.

SA1.1.2 Precipitation

The data source on precipitation is derived from the ERA5 reanalysis data. This contains hourly data on precipitation. Hourly data is aggregated to daily mean precipitation. The units of precipitation are depth in metres. Precipitation data are grouped in to 10 m wide bins, ranging from 0, >0-10m to >40m. Number of days in each precipitation bin are summed over day of interview and preceding 6 days in the sample. The independent variable of interest is therefore count of days for which a state had an average precipitation in each bin in 7-days.

SA1.1.3 Barometric Pressure

The pressure within the Earths atmosphere is derived from the ERA5 reanalysis data. Barometric data are grouped in to 5000 Pascal width bins, ranging from under 80000 Pascal, up to 100,000 Pascal. The independent variable of interest is therefore count of days for which a state had an average barometric pressure in each bin in 7-days including the date of the interview.

SA1.1.4 Total Cloud Cover

Bin indicators for 20 percentage point bins in share of daytime with Cloud Cover.

SA2 Robustness

If we are to take the estimates for temperature and $PM_{2.5}$ seriously, it is important to know that the coefficients on the included controls are also reasonable. Figure SA2 presents the coefficients on the demographic controls from our main regression. We see that in general:

1. females report more depressive symptoms than males,
2. depressive symptoms increase across age
3. those in more affluent households report fewer depressive symptoms,
4. depressive symptoms decrease in education level, and
5. those with the highest status in the caste system (i.e.: "No Caste") report the least depressive symptoms.

All of these associations are in line with typical findings, lending credibility to the estimates for temperature and $PM_{2.5}$ from the same regression.

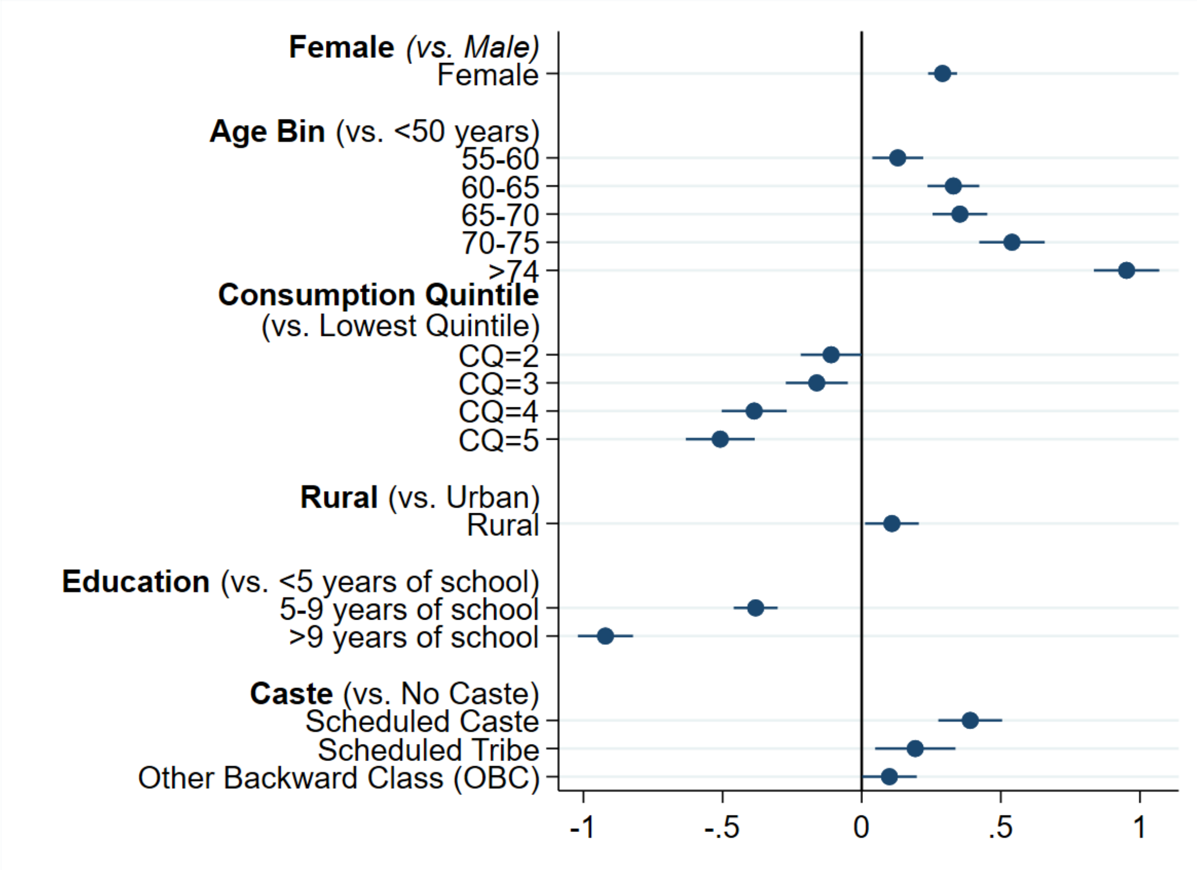

Figure SA1: Coefficient Estimates on Demographic Controls from Main Regression

Notes: Whiskers represent 95% confidence intervals based on standard errors clustered by household. All estimates are from the same, single regression with the CESD-10 score as the outcome from which the estimates in Figure 1 are taken. All demographic characteristics enter as indicator variables with the omitted value for each characteristic indicated in paranthesis next the relevant group heading. Consumption is the preferred measure of affluence by the LASI staff.

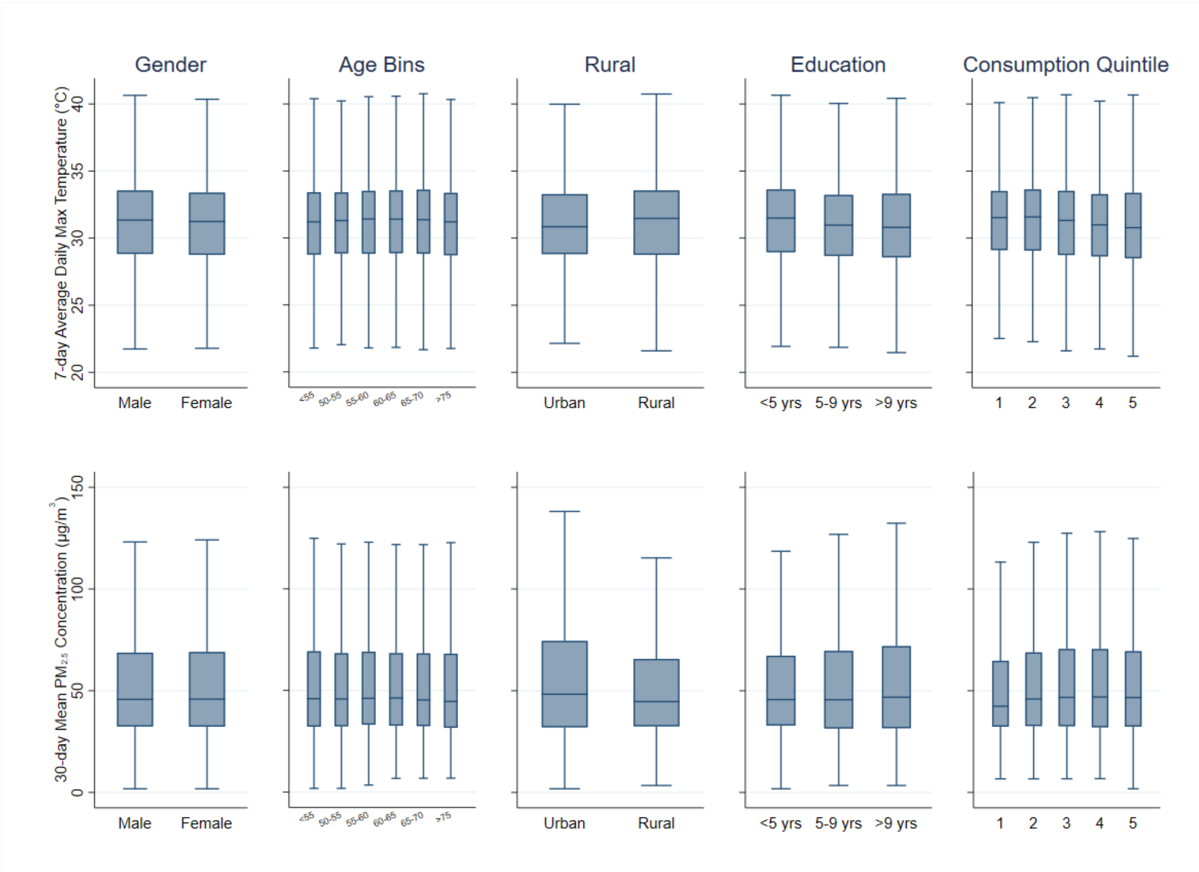

Figure SA2: Mean Exposures by Demographic Characteristics

Notes: Midline represents level of median, box bottom and top indicate 25th and 75th percentiles respectively, and whiskers show 1.5 times the inter-quartile range subtracted and added to the bottom and top of the box respectively.

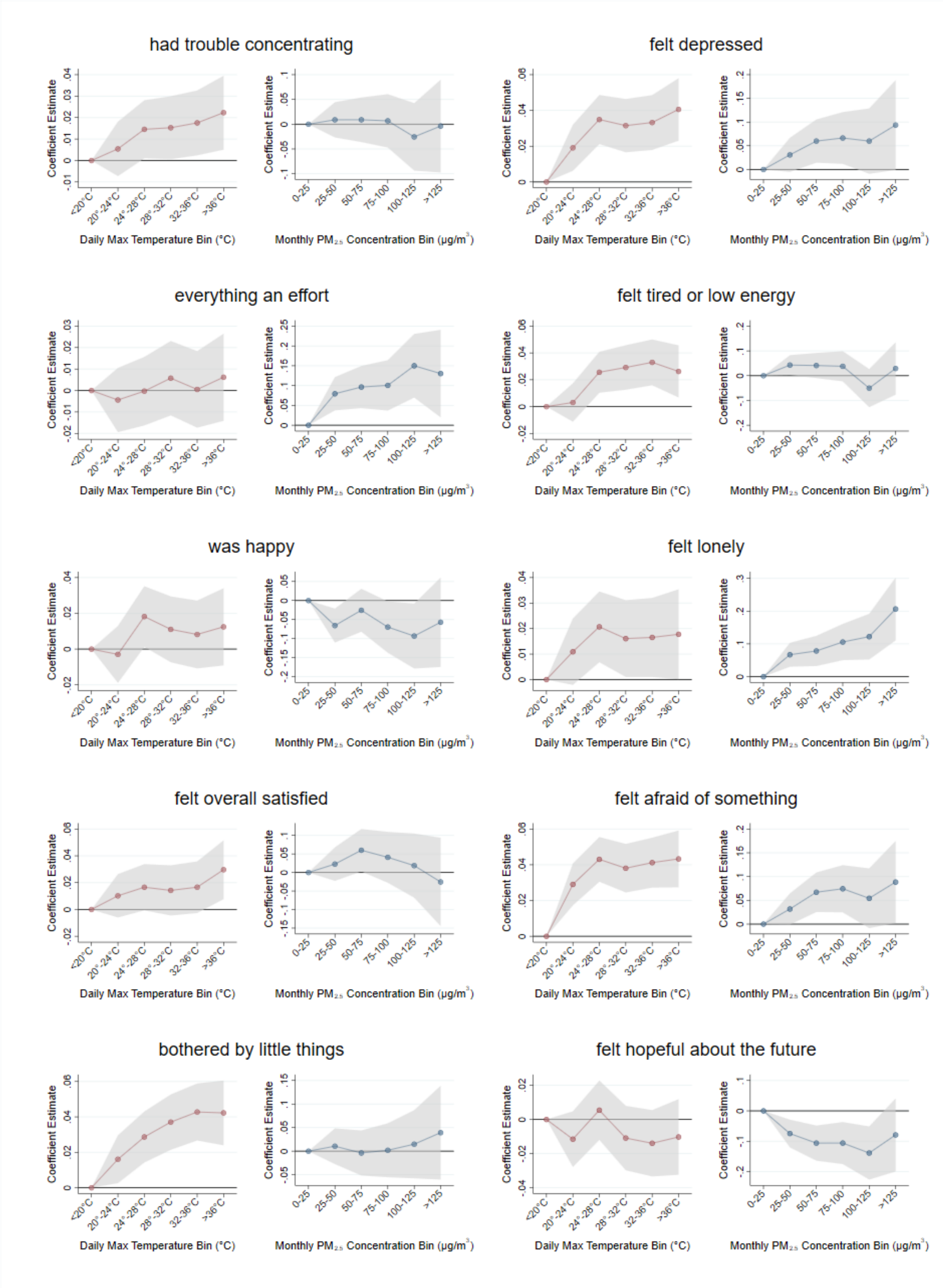

Figure SA3: Baseline Estimates by Individual CESD Question