Heat and Productivity: Evidence From Flight On-Time Performance

Zhanhan Yu *   Ying Shi †

May, 2024

Abstract

We investigate the impact of high temperatures on productivity using microdata from the U.S. airline industry. By linking high-frequency on-time flight performance measures with meteorological data, we show that higher temperatures significantly reduce airline productivity by increasing cancellation and delay rates and lengthening delay times. Complementary analyses using a sample of transportation workers from the American Time-Use Survey (ATUS) suggest that higher temperatures reduce labor supply (fewer hours worked and greater worker absenteeism) and adversely impact well-being measures such as sleep quality, which may affect on-the-job-productivity.

Keywords: Heat Stress, Productivity, Labor Supply, Air Transportation
JEL Classification: J24, Q54, R41

*Economics Department, Adam Smith Business School, University of Glasgow, Glasgow, UK; Email: yuzhanhan@gmail.com.

†Department of Public Administration and International Affairs, Syracuse University, 426 Eggers Hall, Syracuse, NY 13244-1020, yshi78@syr.edu.

1 Introduction

Extreme temperature events are increasing in frequency, duration, and magnitude across the globe (World Health Organization 2018). Their prevalence amidst a warming planet has spurred research on the economic consequences of rising temperatures. There is mounting evidence based on cross-country and subnational data that higher temperatures reduce economic growth and per capita income (Deryugina and Hsiang 2014; Dell, Jones, and Olken 2012), as well as industrial and agricultural production (Hsiang 2010; Schlenker, Hanemann, and Fisher 2006; Fisher et al. 2012). Understanding the economic consequences of rising temperatures has gained increasing importance for simulating the economic implications of future climate change and for informing policy-making processes in response to global warming (Dell, Jones, and Olken 2014). In order to effectively adapt, it is expedient to build a larger evidence base on the impact of heat on productivity across a range of workplace contexts.

This paper examines the consequences of higher temperatures for productivity in the United States using novel microdata from the airline industry. This industry, alongside the broader transportation and logistics sector, is relatively climate-exposed, making it an opportune setting for exploring how rising temperatures shape productivity.

We build a 15-year longitudinal panel of high-frequency weather data linked to productivity as measured by on-time flight performance. These measures allow us to identify the role of heat stress by leveraging variation in temperatures over time and using a model augmented with a rich set of fixed effects. We find that flights operating during days where temperatures are greater than 35 degrees Celsius (°C) are 30% more likely to be cancelled, 13% more likely to involve a late departure, and experience 21% longer delay time conditional on late departure. The adverse impact of heat extends beyond immediate exposure and persists throughout later periods of the same day when the temperature is cooler. When controlling for contemporaneous temperatures, an additional hour of heat exposure (at temperatures above 35°C) during the day (5am to 6pm) is estimated to increase the departure delay rate and delay time later in the same day by 4% and 3%, respectively. Given that time is a limited yet exceedingly valuable resource, the welfare implications linked to heat-induced time losses (resulting from flight cancellations and delays) are likely to be significant

(Graff Zivin and Neidell 2014). Our study also shows that heat's adverse impacts are decreasing in airport size, with nonhub airports more negatively affected than large and medium hub airports.

We provide suggestive evidence on the mechanisms behind these estimates, with a focus on workers' labor supply and sleep. We use data from the American Time Use Survey (ATUS) linked to daily weather measures to first show that heat reduces hours worked and increases absenteeism. Transportation workers spend 1.2-1.4 fewer hours at work and are significantly more likely to be absent on days with maximum temperatures exceeding 35°C. Additionally, heat exposure decreases workers' sleep time and increases the probability of experiencing sleeplessness. We provide suggestive evidence that the mechanism of sleep quality does not meaningfully influence workers' labor supply. Instead, research on the significant impact of sleep and health on workers' productivity (Bubonya, Cobb-Clark, and Wooden 2017; Gibson and Shrader 2018) are consistent with these mechanisms contributing to decreased on-the-job performance.

This study contributes in several ways to existing literature on the consequences of heat stress for labor output, labor supply, and worker well-being (see, for example, Heal and Park (2016) and Lai et al. (2023) for a review). Earlier studies on these topics tend to focus on the effect of temperatures on task productivity in workplaces with more scope for climate control, such as office environments.[1] More recent causal evidence consistently demonstrate that increasing temperatures negatively impact labor output in middle-income countries such as India and China (Cai, Lu, and Wang 2018; Zhang et al. 2018; Chen and Yang 2019; Adhvaryu, Kala, and Nyshadham 2020; Somanathan et al. 2021; Zhang et al. 2023), or across a larger set of developing economies (LoPalo 2023). This burgeoning literature utilizes a variety of worker- and firm-level output data to demonstrate the adverse consequences of heat in predominantly manufacturing and construction

[1]A meta-review of studies that investigate the relationship between office temperature and work performance, in either the laboratory environment or the field environment, suggests nonlinear decreases in workers' performance when the office temperature is above 25 degrees Celsius (Seppanen, Fisk, and Lei 2006). Effects above the 25°C threshold are documented in studies such as Niemela et al. (2002), which provided evidence from two call centers in Finland that each one-degree Celsius increase in indoor office temperature is associated with a 5-7% decrease in labor productivity, as measured by the average number of telephone calls per active working hour, when the air temperature exceeded 25 degrees Celsius.

industries.[2] While these results generalize to workplaces in developing countries, we provide estimates in the context of an advanced economy for which there is a sparser literature (Cachon, Gallino, and Olivares 2012; Stevens 2017). Cachon, Gallino, and Olivares (2012) show hot days decrease automobile production and Stevens (2017) documents decreased agricultural productivity under heat exposure.[3] Our context is closer to the automobile manufacturing case, given the availability of workplace climate control. In contrast to Cachon, Gallino, and Olivares (2012) who find significant drops in automobile output only for sustained exposure to days-long heat waves, we show that high temperatures in a single day can adversely affect airline productivity.

In examining the U.S. airline industry, we are also shifting away from previous studies focused on industrial and agricultural output to a service-oriented sector. Existing research enrich our understanding of heat stress' effects on individual output in the food and beverage industry (Cai, Lu, and Wang 2018), cloth-weaving (Somanathan et al. 2021), and fruit-picking (Stevens 2017), as well as production line-, plant- or firm-level industrial output, namely in automobile, garment, and steel production (Cachon, Gallino, and Olivares 2012; Adhvaryu, Kala, and Nyshadham 2020; Somanathan et al. 2021). Yet there are few studies focused on service-oriented industries, with the exception of data collection and production (LoPalo 2023) and professional sports (Qiu and Zhao 2021; Burke et al. 2023). Burke et al. (2023) aims to fill a gap on temperature's effects for workers in the service economy in wealthier nations using a global dataset of professional tennis matches. Even so, we advance that more research is needed to understand how heat affects productivity in other service-oriented sectors, particularly those involving collaborative work environments without a clear mapping of individual effort onto output.

[2] Somanathan et al. (2021) use worker- and firm-level output data to show that rising temperatures cause productivity declines in Indian manufacturing plants specializing in cloth weaving, garment sewing, and steel production. Adhvaryu, Kala, and Nyshadham (2020) document similar negative effects using microdata from a large Indian garment firm. For mean daily temperatures above 19°Celsius, there is a large, negative impact on efficiency of approximately 2 points for each one-degree Celsius increase in temperature. In comparison, Somanathan et al. (2021) finds that the effect of a uniform one-degree Celsius increase in daily temperature is a 2% decrease in output for weaving and up to 4-8% decrease for garment production. Among Chinese manufacturing firms, heat exposure adversely impacts both total factor productivity and output (Cai, Lu, and Wang 2018; Zhang et al. 2018; Chen and Yang 2019). A recent study using rich household survey data across 46 developing countries to examine the behavior of interviewers shows that productivity decreases on hot and humid days (LoPalo 2023).

[3] Cachon, Gallino, and Olivares (2012) find an 8 percent decrease in weekly automobile production when exposed to 6-7 days of 90°F+ (32°C) relative to no days at this temperature, while Stevens (2017) shows that agricultural workers (specifically blueberry pickers in California) are 12 percent less productive at 100°F+ (38°C).

Finally, we provide additional U.S.-based evidence on how changes in labor supply and well-being may contribute to the observed productivity effects. The existing literature on labor supply mostly focuses on China and India, with the exception of Graff Zivin and Neidell (2014). Adhvaryu, Kala, and Nyshadham (2020) find the adverse effect on Indian garment production is primarily driven by reductions in productivity per unit labor supplied rather than in the quantity of labor units supplied (worker absenteeism and hours worked). This contrasts somewhat with Somanathan et al. (2021), which find evidence for both channels with magnitudes varying by industry and the presence of climate control in India.[4] Our findings based on extensive U.S. time-use data suggests that both margins of labor supply and on-the-job productivity contribute to the adverse impact of heat exposure. We furthermore document that higher temperatures disturb sleep and rest, which may affect labor productivity via reduced on-the-job performance, although they do not seem to meaningfully affect the labor supply margin. The focus on sleep and well-being contributes to a smaller but growing literature investigating the causal effects of these channels (Obradovich et al. 2017; Mullins and White 2019; Minor et al. 2022).

2 Data

The data used in this paper has two components. In the main analysis, we construct a panel dataset linking flight on-time performance data with hourly climate data to investigate the impact of high temperatures on airline productivity. In addition, we exploit time-use survey data on individual labor supply, absenteeism, and well-being to explore the potential mechanisms behind the estimated effects. In this section, we describe the sources and construction procedures for these two data sets sequentially.

[4]Not all studies find negative or null impacts on labor supply. LoPalo (2023) documents *increases* in the hours worked per day in response to hotter days. Workers likely start their days earlier and log more hours due to strong incentives to maintain similar levels of total daily output.

2.1 Data for the Main Analysis

We measure productivity in the airline industry using flight on-time performance. Information on flight delays and cancellations derives from the Bureau of Transportation Statistics (BTS)'s Airline On-Time Performance (AOTP) Data. It provides detailed information on flights, including origin and destination airports, date of departure, scheduled and actual departure and arrival times, cancellation status, and in particular, the causes of flight cancellations or delays.

We assess the on-time performance of flights across three dimensions: cancellation rate, departure delay rate, and departure delay time. Crucially, we restrict our analyses to cancellations and delays caused by factors over which air carriers exert some control. These include but are not limited to crew issues, baggage and cargo loading, fueling, aircraft cleaning and servicing, maintenance, passenger services, and ramp service.[5] This focus corresponds most closely to our focus on airline productivity, and excludes other categories of delays and cancellations induced by extreme temperatures such as closed runways, general airport conditions, and other consequences that cannot be mitigated through corrective actions or can only be addressed with corrective action from airports or the Federal Aviation Administration.[6] In doing so, we may be underestimating the overall impact of heat on flight operations to focus on causes most related to airline workers. We furthermore exclude delays due to late-arriving inbound aircrafts given that the likely causes are unrelated to recent heat exposure.[7]

[5]BTS classifies the causes of flight delays into five categories: (a) Air Carrier, (b) Extreme weather, (c) National Airspace System (NAS), (d) Security, and (e) Late Arriving Aircraft. The causes of flight cancellations fall into the first four categories only. In our analyses, we focus solely on (a) air carrier-related causes, a category that has been contributing an increasing share of flight delays since 2004 and accounts for 41% of total delay minutes in 2020 (Source: BTS). Note that air carriers only track delays up to the time the plane pushes back from the gate. Delays occurring after pushback are assigned to NAS, which we do not consider in our analyses.

[6]Our outcome measures exclude extreme weather-related causes for delay, including below minimum conditions, deicing aircrafts, earthquakes, extreme temperatures, hail damage, holding at gate for enroute weather, hurricanes, lightning, snow storms, thunderstorms, and tornadoes, as well as NAS-related causes, including but not limited to airport conditions, airport construction, air traffic control, closed runways, volume delays, and air traffic control-related equipment outages, gate holds, ground delays, and ramp traffic.

[7]For flight delays, our focus is on departure rather than arrival delays. Departure delays are less likely to be correlated with other confounding factors, such as weather en route and congestion at the destination airport, compared to arrival delays. Since we lack sufficient data on these confounding factors, using departure delays helps us minimize omitted variable bias in our estimation. Furthermore, many flights that experience departure delays tend to make up for lost time during the flight, resulting in smaller arrival delays compared to departure delays. Consequently, using arrival delays as the performance measure may attenuate the estimated effect of heat on labor productivity.

The AOTP data spans from 1987 to 2021, while information on the causes of delays is only available after 2004. Moreover, we intend to consider only the pre-Covid period to avoid capturing effects driven by the shock of the pandemic. Therefore, we restrict our sample period from January 1, 2004, to December 31, 2019. The AOTP data covers 361 commercial service airports in the contiguous United States and 27 airlines.[8]

Next, we bring in meteorological data on temperature and other weather conditions. Hourly climate data comes from the Automated Surface Observing System (ASOS), made up of airport-based meteorological stations taking minute-by-minute observations to generate weather reports and inputs for the National Weather Service (NWS), the Federal Aviation Administration (FAA), and the Department of Defense (DOD).[9] We retrieve hourly ASOS climate data on air temperature (at 2 meters above the surface), feels-like temperature (also known as apparent temperature), precipitation, snow depth, wind speed and direction, humidity, and visibility from January 1, 2004 to December 31, 2019 for all airports in our sample. Commercial aircraft activities are found to be a major contributor to air quality deterioration at ground-level (Masiol and Harrison 2014; Riley et al. 2021), making it a relevant confounding factor in the context of this paper. To gauge its impact on our results, we collect daily air pollution data from the Air Quality System (AQS) of the Environmental Protection Agency (EPA), focusing on four pollutants: CO, NO_2, PM2.5, and ozone.[10]

We create three outcome variables to measure flight on-time performance: the cancellation rate, departure delay rate, and total departure delay time (in minutes).[11] Both the cancellation and delay rates use the number of scheduled flights as the denominator, such that the rate of on-time

[8]Commercial service airports are publicly owned airports with scheduled air carrier service and at least 2,500 annual enplanements. For more information on airport categories, see Federal Aviation Administration's website https://www.faa.gov/airports/planning_capacity/categories.

[9]ASOS data are downloaded from Iowa Environmental Mesonet (IEM) https://mesonet.agron.iastate.edu/ASOS/.

[10]CO and NO_x are the major air pollutants emitted by airplanes during takeoff, taxiing, and idling (Schlenker and Walker 2016). In addition, like many other mobile sources, aircraft jet engines emit particulates and volatile organic compounds (VOCs) (Federal Aviation Administration 2005). Both VOC, unburned or partially combusted hydrocarbons, and NO_x contribute to ozone formation.

[11]While we focus on flight on-time performance, we acknowledge that there are alternative ways of measuring airline productivity, including passenger enplanement.

departure is one minus the sum of these two rates.[12] It is worth noting that the cancellation rate, departure delay rate, and departure delay time are essentially measured at the flight-level. In other words, they are specific to the particular route (an origin-destination pair), carrier, and day-and-time block pair.[13] Next, we merge flight performance data with hourly climate measures by the date of flight operation, time block, and the 3-digit identifier of the origin and destination airports. In our main analysis, we focus on the sub-sample of months from April to September, because this period typically exhibits higher temperatures, making it most relevant to the aim of this study. The final sample is nationally representative and consists of 350 airports in the contiguous United States.[14] Figure 1 presents the 350 airports in the sample, with the size of each bubble indicating its average annual enplanements (passenger boarding) from 2004 to 2019.

To demonstrate the descriptive correlation, we plot flight on-time performance measures as a function of temperature (in degree Celsius) in Figure 2. Panel (a) shows a positive linear correlation between departure delays (both in terms of rate and time) and temperatures. Panel (b), on the other hand, suggests a likely non-linear correlation between temperatures and the cancellation rate. This finding motivates our decision to flexibly specify a model with temperatures measured at 5-degree Celsius bins in the main analysis, as described below.

We measure temperature exposure in two ways. First, we classify the current temperature into five categories: less than or equal to 20°C (68°F), between 20°C and 25°C (77°F), between 25°C and 30°C (86°F), between 30°C and 35°C (95°F), and above 35°C. Second, we calculate the number of hours when the temperature exceeds 35°C during the day (defined as 5am to 6pm) to measure the cumulative heat exposure for the same day. Our temperature measures are based on the feels-like

[12]An alternative method to calculate these rates is by using the number of flights that actually departed as the denominator. However, this alternative method may overestimate the rate of cancellation or delay, as it computes the rate conditional on departure.

[13]The FAA classifies 18 time-block groups, with 0-5 a.m. forming a single group and each hour from 6 to 23 as individual groups.

[14]Airports with missing climate data, typically small and public-use airports, are excluded from our sample. The 11 excluded airports are Northeast Florida Regional Airport (UST), Phoenix–Mesa Gateway Airport (AZA), Branson Airport (BKG), Tunica Municipal Airport (UTM), McClellan–Palomar Airport (CLD), Glacier Park International Airport (FCA), Hilton Head Airport (HHH), Sawyer International Airport (MQT), University Park Airport (SCE), Pinehurst Regional Airport (SOP), and Concord-Padgett Regional Airport (USA).

temperature, also known as the apparent temperature, instead of the real air temperature.[15] The apparent temperature measures how warm or cool the human body perceives the surrounding air. Because the human body can regulate high and low temperatures through, for example, sweating and insulating, real air temperature does not accurately reflect workers' heat exposure. The apparent temperature takes into account weather factors in the function of body temperature regulation, such as humidity and wind speed.[16]

We present summary statistics for the main analysis in Table 1. The mean cancellation rate is approximately 1%, with an average departure delay rate of 15% and an average departure delay time of approximately 6 minutes. Around 61% of flights operate at temperatures between 20°C and 35°C, while 7% of flights operate at temperatures above 35°C.

2.2 Data for the Exploration of Potential Mechanisms

We rely on data from the American Time Use Survey (ATUS) to explore potential mechanisms. This necessarily limits the scope of what we can determine compared to the ideal case of individual-level worker data obtained from airlines or airport authorities, which are not readily available. We assemble a long panel of time-use data from 2005-2019 that collects detailed information about how individuals spend their time during a diary day.[17] This includes their intertemporal labor supply as measured by hours worked and work absence status, their sleeping activity, such as sleep time and sleeplessness, and various demographic, educational, and employment characteristics. Importantly, it also contains information on the date of the diary day and the respondent's geographic location, allowing us to link it with weather data. In addition to the regular module, we include the Well-being (WB) Module in the ATUS, which is available for the years 2010, 2012, and 2013. All

[15]Throughout this manuscript, unless otherwise noted, all temperature references correspond to feels-like or apparent temperature.

[16]Some recent studies have used an alternative temperature measure—the WetBulb Globe Temperature (WBGT) (e.g., Somanathan et al. 2021; LoPalo 2023). WBGT measures the heat stress in direct sunlight, which takes into account temperature, humidity, wind speed, sun angle, and cloud cover (solar radiation). The apparent temperature differs from WBGT as it is calculated for shady areas (National Weather Services). We use apparent temperature instead of WBGT in this study, not only because it is the best available data for us, but also because it is better suited to our empirical setting. Most airport workers, including ground crew members, work in shady areas and are not exposed to direct sunlight for the majority of their work time.

[17]The sample is randomly selected from a subset of households that have completed their eighth month of interviews for the Current Population Survey (CPS).

respondents interviewed for the regular module in these years were selected for the WB module. The WB module includes questions about respondents' general health, such as how they felt in general compared to a typical day (matching the day of the week of the diary day) and how well-rested they felt.[18] We merge the regular module data with the WB module data, thereby creating a cross-sectional dataset that contains comprehensive information on individuals' intertemporal labor supply, sleep patterns, and subjective health and well-being.

The ATUS dataset also contains information on the industry and occupation of each respondent's main job, enabling us to examine the effect of heat exposure on workers in different industries and occupations. In addition to the full sample, we restrict to workers in the transportation industry, in keeping with the focus on the airline sector. The analyses also examine a sample of workers in transportation and material moving occupations, which comprise 60% of employees in the air transportation industry (Bureau of Labor Statistics 2022).

Next, we collect daily weather data from the Daymet project (Thornton et al. 2020) and link it with the ATUS data set. The Daymet project provides daily weather measures such as minimum and maximum temperatures, precipitation, and day length on a 1 km $\times$ 1 km gridded surface. We map these gridded daily weather parameters to counties using the longitude and latitude coordinates of the centroid of each county.[19]

We restrict our sample to full-time employed individuals. To measure workers' intertemporal labor supply, we create two variables: "Working Time" measures minutes spent on work and work-related activities.[20] "$\mathbb{1}$(Absence Last Week)" denotes a binary indicator which equals one if the respondent was absent from work in the past week. Two additional variables capture respondents' sleep patterns. The first measures the respondents' sleep time during the diary day in minutes,

[18] The WB module also includes information about the respondent's well-being during specific activities, such as work and work-related activities. The well-being measures include how much pain they felt and how tired/sad/stressed/happy they felt. We also utilize these measures to examine whether high temperatures affect these aspects at work and work-related activities. However, due to the small sample size, the estimates are not precisely estimated. These results are available upon request.

[19] We merge the 2005-2019 ATUS data with the Daymet weather data using the county FIPS or the CBSA/MSA code. We are able to identify the geographic location for 80% of the entire sample, where about 56% of them are identified by the county FIPS code, 40% of them are identified by the CBSA/MSA code, and 4% of them are identified by the NECTA code.

[20] For work and work-related activities, we include time spent on activities in categories 0501 and 0502. See the American Time Use Survey Activity Lexicon 2003-2019 for more details.

while the second indicates whether the respondent experienced any sleeplessness.[21] Moreover, we create two dummy variables signaling the respondent's general health. One indicates whether the respondent felt worse than a typical day, while the other indicates whether the respondent felt not well-rested on the diary day.

Similar to the temperature variables defined in the main analysis, we categorize daily maximum temperatures into five bins: less than or equal to 20°C, between 20°C and 25°C, between 25°C and 30°C, between 30°C and 35°C, and above 35°C. We also calculate the number of days within the past week in which the daily maximum temperature falls into each of the five temperature categories.

We present the summary statistics of the variables used to explore potential mechanisms in Table 2. The average working time is 278 minutes, equivalent to about 4.6 hours, while the average absenteeism rate in the past week is 4%. The average sleep time is about 8.5 hours, and approximately 4% of respondents report experiencing sleeplessness. 21% of respondents from the WB module report not feeling well-rested on the diary day.[22] Lastly, 7% of respondents in the sample report feeling that their general health was worse than on a typical day.

3 Heat and Airline Productivity

3.1 Empirical Strategy

We analyze the causal effect of heat exposure on airline productivity using high-frequency flight performance and weather data that leverages temperature variation within the same micro-context such as flight route. First, we consider the contemporaneous impact of heat exposure by estimating

[21]The ATUS data also collect self-reported sleeplessness time. However, we are concerned that this self-reported variable may suffer from non-negligible measurement error, as it relies on respondents' subjective recall rather than reliable equipment monitoring of their sleep periods. Conversely, while individuals may not have an accurate understanding of the exact time of sleeplessness, they should remember whether they experienced sleeplessness. Therefore, we use a dummy indicator to measure the probability of sleeplessness.

[22]Since the sample size of the regular module is about six times larger than that of the WB module, the proportion of respondents reporting sleeplessness is consistent with the proportion of those reporting not feeling well-rested.

the following model:

$$\text{OnTimePerformance}_{ijdh} = \sum_k \beta_k \text{Temp}_{idh}(B_k) + \alpha \mathbf{X}_{idh} + \delta \mathbf{W}_{jdh} + \theta \mathbf{Q}_{id} + \sigma_{ym} + \rho_s + \tau_h + \kappa_{ij} + \phi_c + \zeta_{im} + \varepsilon_{ijdh} \tag{1}$$

where i and j denote the origin airport and destination airport, respectively. d is the day of flight operation, s denotes the day of the week and h indexes the time block of flight departure. The outcome variable is the cancellation rate, departure delay rate, or departure delay time based on air carrier-related causes such as baggage loading, fueling, aircraft cleaning and maintenance, and passenger and ramp services.[23] Temp_{idh} denotes the treatment variable categorized in five bins (B_k): $\leq 20°$C (baseline), (20°C, 25°C], (25°C, 30°C], (30°C, 35°C], and $> 35°$C. We are interested in the coefficient β_k, which gives the effect of temperatures falling in the corresponding bin, relative to the reference temperature of less than or equal to 20°C.

To isolate the effect of heat from the impact of related weather phenomena such as wind gusts and thunderstorms at flight departure, we control for time-varying weather conditions measured at the hour level of origin ($\mathbf{X}_{idh}$) and destination ($\mathbf{W}_{jdh}$) airports, including precipitation, relative humidity, obscuration (visibility), and wind speed. Moreover, we incorporate a rich vector of fixed effects, including month by year (σ_{ym}), day of week (ρ_s), and time block (τ_h) fixed effects to account for seasonal, day-of-week, and hourly patterns governing airlines' on-time performance. Crucially, we control for origin-destination pair (κ_{ij}) fixed effects to consider time-invariant factors that are specific to the route between the origin and destination, and origin-month (ζ_{im}) fixed effects, which absorb time variant unobservables.[24] Given the potential variation in airline productivity across different carriers, our model also includes carrier fixed effects (ϕ_c) to account for carrier-specific confounding factors. Standard errors are clustered at the route, or origin-destination pair, level.

Given the possibility that the adverse impact of heat may be mediated through deteriorating air pollution, we furthermore control for daily local air pollution, measured by CO, NO_2, PM2.5,

[23]As described in Section 2.1, our definition excludes delays due to late-arriving aircraft, security and air traffic control issues, general airport conditions such as closed runways, and extreme weather events. To the extent that heat affects these dimensions of flight operations, we may be underestimating its cumulative impact.

[24]We experiment with adding origin-by-month fixed effects or destination-by-month fixed effects. The results show minimal variation. Therefore, we only report the results with origin-by-month fixed effects in the main text.

and ozone (denoted as $\mathbf{Q}_{id}$).[25] To address the potential endogeneity of air pollution variables, we follow the environmental economics literature and adopt an instrumental variable (IV) strategy.[26] We use atmospheric temperature inversion and its interaction with wind direction as instruments for air pollution. In general, temperature tends to decrease with altitude. However, it increases with altitude during inversion episodes because warmer air at higher altitudes confines cooler air near the surface. As a consequence, this prevents pollutants from rising and dispersing, trapping them close to the ground.[27] This type of instrument produces arguably exogenous variations in air pollution.

In addition to contemporaneous exposure to high temperatures, workers may also suffer from prolonged exposure to heat stress. Its impact may take time to emerge, leading to a lagged effect on labor productivity. For example, workers whose shifts extend to cooler temperature at night could still experience the cumulative impact of heat stress from earlier exposure during the day. To investigate whether this cumulative daytime heat exposure (defined as 5am - 6pm) affects flight on-time performance later in the same day (especially after sunset when the temperature is cooler), we create a discrete variable on same-day cumulative exposure ($\text{CumulativeTemp35}^{\circ}\text{C}_{id}$) which counts the number of hours when temperature exceeds 35°C during 5am - 6pm. We regress each on-time performance outcome for flights operating after 8pm on this same-day-cumulative heat

[25]Following Schlenker and Walker (2016), daily airport-level air pollution is measured by taking the average of monitor readings from all monitors within 100 km of the airport, weighting by the inverse distance between the monitor and the airport. Chen et al. (2023) use flight-level data and granular air pollution measures to show that rising levels of PM2.5 significantly increase flight departure delays.

[26]Wind direction and temperature inversion are two canonical instruments for air pollution in environmental economics studies. For example, Sager (2019), Jans, Johansson, and Nilsson (2018), and Arceo, Hanna, and Oliva (2016), among others, use temperature inversion as an instrument for air pollutants such as PM10, PM2.5, and CO. Deryugina and Hsiang (2014), Schlenker and Walker (2016), and Chen et al. (2023), among others, instrument for air pollutants such as PM2.5, NO, NO_2, and CO with wind directions and wind patterns.

[27]Following Sager (2019), we collect air temperatures at 925hPa pressure level and the surface level at 3am local time for the contiguous U.S. from NASA's MERRA-2 climate reanalysis product (Global Modeling and Assimilation Office 2015). Our temperature inversion variable is defined similarly to Sager (2019) as a continuous variable equal to the temperature difference between 925hPa pressure level and the surface level.

exposure measure using the following model:

$$\begin{aligned}\text{OnTimePerformance}_{ijdh} = & \sum_k \beta_k \text{Temp}_{idh}(B_k) + \chi \text{CumulativeTemp35°C}_{id} \\ & + \alpha \mathbf{X}_{idh} + \delta \mathbf{W}_{jdh} + \theta \mathbf{Q}_{id} + \sigma_{ym} + \rho_s \\ & + \tau_h + \kappa_{ij} + \phi_c + \zeta_{im} + \varepsilon_{ijdh}, \text{ where } h \in \{20, 21, 22, 23\}\end{aligned} \tag{2}$$

Conditioning on the concurrent temperature and weather conditions, χ captures the delayed and cumulative effect of heat stress exposure on the same day.

3.2 The Effect of Contemporaneous Exposure

We first show estimation results corresponding to Equation (1), which regresses flight on-time performance measures on contemporaneous temperature bins. Columns (1) to (3) of Table 3 present the estimated coefficient and the corresponding percentage effect, relative to the sample mean of each outcome, of each temperature group for the cancellation rate, departure delay rate, and delay time, respectively.

Column (1) shows that flights are more likely to be cancelled at higher temperatures relative to those operating below 20 degrees Celsius. The effect magnitudes are 0.18 p.p. at temperatures above 35°C, with the corresponding percentage effects estimated at 30% relative to the sample mean cancellation rate. The estimated effects decrease non-linearly for milder temperatures, with the percentage effects estimated at 18%, 10%, and 6% for temperature bins 30°C-35°C, 25°C-30°C, and 20°C-25°C.

Conditional on flights not being cancelled, results in Columns (2) and (3) suggest that flights operating at high temperatures would experience not only a higher rate of departure delay, but also longer departure delay time. Column (2) shows that compared to flights departing at temperatures below 20°C, the rate of departure delays is between 0.8-2.1 p.p. higher for flights departing during hotter periods. The corresponding percentage effects are estimated at 5%, 7%, 10%, and 13% for the 5-degree temperature bins, respectively. Note that the relative effect magnitudes are not as apparently non-linear for the delay rate measure as compared to the cancellation rate, which accelerates markedly with higher temperatures.

In addition to affecting the probability of departure delays, high temperatures could also affect the length of delay time. The evidence in Column (3) is consistent with this conjecture. Specifically, we find that on average, flights experience a 0.4 minute longer departure delay at temperatures between 25°C and 30°C, compared to flights departing at temperatures below 20°C. It is equivalent to a 8% increase relative to the sample average delay time. The magnitude of the effect increases to 14% (0.8 minutes) when operating at temperatures between 30°C and 35°C, and 20% (1 minute) when operating at temperatures above 35°C.

Compared to existing studies that quantify the impact of high temperatures on worker productivity in manufacturing and service industries, the magnitudes of our estimated heat effects are at least as large, if not greater. For example, Somanathan et al. (2021) find that exposure to an additional hot day in India reduces worker output from 2% to 8%, depending on industry, climate adaptation, and workplace context.[28] Cachon, Gallino, and Olivares (2012) use data on weekly automobile production at 64 facilities in the United States and find that a week with six or more days of heat exceeding 32°C is associated with a reduction in weekly production by 8% on average.

We conjecture that several factors contribute to this difference. First, the effect of high temperatures on flight delays can be partly attributed to the performance of airline and airport crews working in outdoor or semi-outdoor environments. For example, workers involved in baggage loading, fueling, or aircraft maintenance may be affected. Compared to existing studies considering indoor workers, effects could be larger in our context as workers are directly exposed to outdoor environments where climate control is unlikely, making them more vulnerable to heat stress and fatigue. Additionally, since we are not considering individual worker output such as the number of phone calls handled per labor unit per work time, our results may not be directly comparable to existing studies exploring the impact of higher temperatures on worker productivity in contexts where individual effort maps more cleanly onto output.

[28]For example, non-climate controlled garment plants saw a reduction in average daily efficiency by up to 8%, compared to 2% in the weaving industry.

3.3 The Effect of Same Day Cumulative Exposure

Next we explore the effects of same-day cumulative heat exposure. To do so, we estimate Equation (2), where the treatment variable is defined as the number of hours during the period from 5am to 6pm when the temperature exceeds 35°C. We regress the flight on-time performance outcomes measured later in the same day (after 8pm), while controlling for the current temperature. Table 4 shows that the estimated effects of the same-day cumulative exposure are generally much smaller than the effect of same-day contemporaneous exposure. Same-day cumulative exposure to heat has little impact on flight cancellations later in the day. However, there is evidence that the effect of heat exposure persists and impacts departure delays later in the same day. An additional hour of heat exposure (temperature above 35°C) during the day is estimated to increase the departure delay rate starting in the early evening by 0.8 p.p. (equivalent to a 4% increase) and the delay time by 0.2 minutes (equivalent to a 3% increase).[29] This suggests that high temperatures can exert a negative productivity effect that endures several hours after the initial exposure.

3.4 Heterogeneous Impacts

We begin our exploration of the heterogeneous impact of heat by examining whether the effect varies across origin airports of different sizes, as measured by annual passenger boarding. The impact of heat exposure could be amplified at large hub airports given the complexity of flight operations in a high traffic airport, or it may attenuate if larger airports have more resources for climate adaptation and flexibility around staffing or is more efficient in other aspects of airline operations. To investigate this question empirically, we re-run the model of Equation (1) separately for large-hub, medium-hub, small-hub, and nonhub airports. We summarize the estimation results in Figure 3, which plots point estimates and their 95% confidence intervals of the effect of temperatures greater than 35°C, relative to the reference bin of temperatures below or equal to 20°C.

We find that the magnitude of the heat effects decreases with airport size, with nonhub airports being more adversely affected compared to their large hub and medium hub counterparts. For

[29]Note that estimated coefficients for the contemporaneous temperature groups are of the same sign and similar magnitudes compared to the previous table.

example, operating at temperatures above 35°C increases delay rates by 39% for nonhub airports, compared to 7% for large hub airports and 14% for medium hub airports. The effect is significantly more pronounced for nonhub airports on the duration of departure delay time, conditional on experiencing departure delays. The productivity impact of operating at temperatures above 35°C on the delay time, relative to flights departing at temperatures below 20°C, is statistically significantly higher for nonhub airports (66%), compared to large hub airports (16%) and medium hub airports (23%). We observe a similar pattern for the cancellation rate. However, the estimated effects do not statistically differ among the different types of airports.

In further exploring these heterogeneous treatment effects across airport types, we stratify by flight characteristics. Specifically, we examine whether the impact of heat exposure varies between short-haul and medium/long-haul flights. Figure 4 shows that the negative effects on cancellations and delays are primarily driven by short-haul flights. The concentration of shorter flights out of smaller regional hubs likely contributes to the findings by airport size above.

3.5 Robustness

We undertake a number of additional analyses to ensure that our findings are insensitive to the choice of temperature measure and model specifications. First, we replace our use of apparent or feels-like temperature with the real air temperature. The former takes into consideration wind and humidity and is designed to better represent the human body's perception of heat. As such, these two scales can sometimes significantly diverge. Reassuringly, Columns (1)-(3) of Table B.1 find that coefficients are qualitatively unchanged when using actual air temperatures.

Next, we explore whether estimates are robust to alternative IV models and clustering levels. Columns (4) to (6) of Table B.1 show that our results across all three airline on-time performance measures are nearly unchanged when adopting alternative IVs where temperature inversions are interacted with both wind speed and wind direction, rather than only with wind direction. Furthermore, Columns (7) to (9) show that all of our estimates remain statistically significant at the 1% level when standard errors are clustered by origin-month.

A potential concern is that aircraft and other physical equipment may be impacted by extreme heat, with the effects exacerbated for aging or inadequately serviced planes and hardware. We undertook additional analyses by incorporating aircraft age into our models. Specifically, we collected data on the manufacturing year of registered aircraft from the FAA and merged it with our sample using the flight's tail number. Columns (1), (4), and (7) of Table B.2 report results of the preferred model for a sample with non-missing aircraft age data. Columns (2), (5), and (8) include a continuous aircraft age variable. Although we find a small positive correlation between flight on-time performance and the age of the aircraft, the estimated effects of temperature show little change after accounting for aircraft age. Columns (3), (6), and (9) further include controls for carrier fixed effects interacted with aircraft age to account for the possibility that different airlines may have different aircraft maintenance schedules and procedures that vary by hardware age. The results do not change as a result of their inclusion.

Furthermore, we conduct a placebo test by separately shuffling the temperature and outcome variables and reproducing our main results (Table 3). Columns (1) to (3) of Table B.3 present the estimation results using the 5-degree bins derived from the shuffled temperatures as the treatment variable. The estimates show no effect of heat. Next, we separately shuffle the three on-time performance variables and re-estimate Equation (1) using 2SLS. As shown in Columns (4) to (6), the shuffled outcome variables also produce no significant effects. In sum, we find no impact of heat in this exercise, suggesting that the effects we obtained are not coincidental.

Finally, we expand our analysis to a full-year sample spanning from January to December, covering the years 2004 to 2019. Since estimating the preferred model using the entire full-year sample at once requires substantial computational capacity and time, we address this limitation by randomly sampling 50% of the observations from the entire full-year sample and reproducing our main analysis using this random sample.[30] We report the estimation results of the full-year randomized sample in Table B.4 Columns (4) to (6) for the three on-time performance outcomes. For comparison, we estimate the same model using the April-to-September sample and present the results in Columns (1) to (3). We do not find discernible differences between the estimates from

[30]In addition, considering the lower temperatures during the winter season, we categorize temperatures into 8 bins instead of 5: ≤5°C, (5°C, 10°C], (10°C, 15°C] (baseline), (15°C, 20°C], (20°C, 25°C],(25°C, 30°C], (30°C, 35°C], and >35°C.

these two samples.

4 Exploration of Mechanisms

4.1 Conceptual Framework and Extant Literature

There are multiple mechanisms that may underlie our estimated effects of heat on productivity. Most pertinent to our context are reductions in labor supply and on-the-job task performance.[31] Higher temperatures can influence individual decisions to allocate time to work, with labor shortages arising if workers choose not to work or reduce the hours worked on hotter days. Existing research show moderate decreases in labor supply in response to high temperatures, with larger effects concentrated in more climate-exposed industries (Graff Zivin and Neidell 2014). Worker absenteeism and reduced hours can also have spillover effects by imposing extra burdens on colleagues and changing their labor supply via channels such as increased absenteeism (Godøy and Dale-Olsen 2018). We contribute to previous research by investigating whether workers in transportation adjust their labor supply in response to heat.

Another channel through which heat can influence productivity is through on-the-job performance. Higher temperatures can adversely affect the task performance of workers in the airline sector, particularly those with greater climate exposure, such as ground crews.[32] A substantial literature documents the negative impacts of heat on dimensions of health, including reduced physical work capacity, occupational health issues, and increased morbidity and mortality (Deschênes and Greenstone 2011; Heal and Park 2016; Barreca et al. 2016; White 2017; Ebi et al. 2021; Carleton et al. 2022). Occupational exposure to heat stress can have physiological effects such as hyperthermia, and kidney disease or acute kidney injury (Flouris et al. 2018). Heat exposure has also been shown to diminish cognitive performance (Hancock and Vasmatzidis 2003), with a number of papers in

[31] We acknowledge that heat can affect productivity via non-labor channels, particularly in extreme cases in which runway integrity may be compromised or flights are subject to different operating thresholds and must be weight restricted (Coffel and Horton 2015). Closed runways and other events affecting airport operations fall under the category of National Airspace System (NAS)-related delays, which are excluded from our outcome measures of flight delays and cancellations resulting from air carrier-related causes only.

[32] Airport workers can be at particularly high risk of heat stress due to the heat-amplifying effects of asphalt and the need for wearing protective gear (Gelles and Andreoni 2023).

economics documenting the negative effects of extreme temperatures on cognition, test scores, and decision-making (Graff Zivin, Hsiang, and Neidell 2018; Heyes and Saberian 2019; Graff Zivin et al. 2020; Garg, Jagnani, and Taraz 2020; Park et al. 2020; Park 2022). While data constraints and the complexity of airline operations render it difficult to parse out individual contributions to productivity, we provide suggestive evidence by analyzing the impact of high temperatures on workers' rest and well-being. Specifically, we estimate whether heat-exposed individuals experience changes in sleep duration, likelihood of sleeplessness, feeling not well-rested and worse than in a typical day, using an ATUS time-use panel from 2005-2019.[33] Workers' on-the-job performance may be adversely affected if heat exposure undermines sleep quality and rest.

4.2 Empirical Strategy

We adopt models in the manner of Connolly (2008) and Graff Zivin and Neidell (2014) to investigate the effect of high temperatures on worker labor supply, sleep, and well-being:

$$Y_{kct} = \sum_j \delta_j \text{MaxTemp}_{ct}(B_j) + \omega \mathbf{V}_k + \theta \mathbf{Z}_{ct} + f(month, year, dow, c) + \varepsilon_{kct} \tag{3}$$

where Y_{kct} denotes outcome variables such as working and sleep time (both in minutes) and a sleeplessness indicator for individual k on diary day t and geographic unit of residence c. Following the specification in the main analysis, we categorize the daily maximum temperature (MaxTemp_{ct}) into five bins (denoted as B_j): $\leq 20°$C, (20°C, 25°C], (25°C, 30°C], (30°C, 35°C], and $> 35°$C, and set $\leq 20°$C as the reference bin. We control for other time-varying weather attributes ($\mathbf{Z}_{ct}$) that are potentially correlated with the outcome, such as day length and daily precipitation. $\mathbf{V}_k$ is a vector of individual-level covariates as listed in Table 2. $f(month, year, dow, c)$ denotes a set of dummy variables, including day of week dummies to account for differences in schedules throughout the week, and year and month dummy variables to control for seasonal and annual time trends in the outcome. It also includes location dummies that capture all time-invariant observable and unobservable attributes that affect the outcome. The parameter of interest is δ_j, which captures the effect of high temperatures on individuals' hours worked, sleep patterns, and well-being. Moreover,

[33] We conduct our own analyses using ATUS, instead of relying on Graff Zivin and Neidell (2014), to maintain greater temporal overlap with our main sample, examine additional outcomes such as sleep quality, and to focus on transportation workers in particular.

because the absenteeism indicator is measured in the last week, we adopt a slightly different model where the treatment variable of maximum temperatures and weather attributes are also measured at the weekly level. Instead of using a day of week dummy (*dow*), we substitute it with a week dummy.[34] Standard errors in both models are clustered at the state-year level.

4.3 The Effect on Worker Labor Supply, Sleep, and Well-being

Panel A of Table 5 summarizes the estimated effects of heat exposure on working time. The negative effects appear to increase alongside the highest daily temperature among the full sample of employees, although the estimates are statistically insignificant. This is more apparent when restricting to transportation workers as defined by the major industry and occupation codes (Columns 2 and 3). Individuals working in daily maximum temperatures between 20°C and 30°C do not reduce their hours, but those facing maximum temperatures above 35°C decrease work time by approximately 1.2-1.4 hours. The size of these magnitudes relative to the 14-minute decrease for the full sample may reflect the transportation sector's designation as a heat-exposed industry by bodies such as the National Institute for Occupational Safety and Health (NIOSH 1986).[35]

Turning to the effect on work absenteeism, Panel B of Table 5 suggests that the changes in workers' intertemporal labor supply in response to high temperatures are not limited to their hours worked but also manifest in their likelihood of going to work on the same day. Column (1) shows that, on average, having one additional day with a daily maximum temperature above 35°C in the past week yields a statistically significant increase in the probability of work absence of approximately 0.3 p.p. for full-time employed respondents. In the case of transportation workers, the effect becomes even larger at around 0.8-1.1 p.p..

[34]Specifically, we consider the following model:

$$\mathbb{1}(Absence_{kcw}) = \sum_j \alpha_j \sum \mathbb{1}(\text{MaxTemp}_{c,w} \in B_j) + \sigma \mathbf{V}_k + \eta \mathbf{W}_{cw} + f(month, year, w, c) + \varepsilon_{kcw} \tag{4}$$

where w denotes the week before the week of the diary day t. $\sum \mathbb{1}(\text{MaxTemp}_{c,w} \in B_j)$ denotes the count of days with maximum temperature that falls within a certain temperature bin (B_j) in the past week. $\mathbf{W}_{cw}$ denotes weekly weather attributes, including weekly mean daylight and weekly accumulated precipitation. The parameter α_j captures the change in the work absenteeism rate with respect to high temperatures.

[35]Other climate-exposed industries include a) Agriculture, Forestry, Fishing, and Hunting, b) Mining, c) Construction.

Our estimates of the impact of high temperatures on work absenteeism is consistent with the magnitude of findings from Indian manufacturing industries (Adhvaryu, Kala, and Nyshadham 2020; Somanathan et al. 2021). For example, in Somanathan et al. (2021), an additional day above 35°C in the six preceding days causes a 0.5 p.p. increase in the probability of missing work for weavers in India working in a non-climate controlled setting. Our estimates across a range of samples is inclusive of this point estimate.

Table 6 presents the estimated heat effects on workers' sleep duration and quality. We find suggestive evidence that high temperatures negatively affect workers' sleep by reducing their average sleep time and increasing their likelihood of sleeplessness. For example, Column (1) of Panel A shows that on average, individuals sleep 9 minutes less on hotter days when daily maximum temperature exceeds 35°C, compared to days when the temperature does not exceed 20°C. Sleep among workers in the transportation industry decreases by 20 minutes when the temperature is between 25-30°C, while the effects of even hotter days are in the expected negative direction but not precisely estimated. In Panel B, Column (1) shows that on average, workers are more likely to experience sleeplessness on hotter days when daily maximum temperature is above 35°C, by about 2 p.p.. Among transportation workers, the negative consequences of heat exposure for this aspect of sleep quality is already apparent for daily maximum temperatures between 25-30°C, and continues to increase the incidences of sleeplessness for days above 30°C.[36]

In Table 7, we report the estimation results of two well-being outcomes, one indicating whether the respondent felt not well-rested (Column 1), with the other indicating whether the respondent felt worse than a typical day (Column 2). Column (1) shows that individuals are 6 p.p. more likely to report that they felt not well-rested on days with daily maximum temperature above 35°C, although the effect is only marginally significant. This finding echoes the result in Panel B of Table 6 that individuals are more likely to experience sleeplessness on hotter days. There is no statistically significant evidence that individuals become more likely to feel worse in terms of general health than typical.

[36]One caveat is that our choice of maximum temperatures may not accurately capture individuals' precise indoor temperature exposure during the night. We furthermore do not observe climate control at home. These omissions likely lead to underestimates of the negative effect of heat on worker sleep, to the extent that climate control and other adaptive strategies can moderate the adverse consequences of heat on sleep.

The analyses using ATUS data focus on the transportation sector. Small samples prevent us from disaggregating further to examine the effects of heat exposure on air transportation workers only. To do so, we turn to a supplementary dataset - Severe Injury Reports from the Occupational Safety and Health Administration - to examine how high temperatures affect workplace injuries across detailed industry classifications.[37] Heat can adversely impact workers' occupational health by impairing cognition and concentration, thereby increasing the likelihood of workplace injuries (Park, Pankratz, and Behrer 2021). Table B.5 shows that higher temperatures exert nonlinear effects on the likelihood of at least one severe injury, with magnitudes rising quickly for the top temperature bin of 35°C in the full sample (Column 1) and across transportation industry subsectors (Columns 2-4). The evidence indicates that heat's effects on air transportation is at least comparable to the aggregated transportation industry in this dimension of labor response to heat.

Taken together, the evidence shows that heat can lead to a decrease in labor supply, resulting in fewer hours worked and an increase in absenteeism among transportation workers. Furthermore, our results suggest that heat negatively affects workers' sleep (both duration and quality) and well-being. Previous studies have established a strong correlation between sleep, well-being, and workers' labor productivity (Bubonya, Cobb-Clark, and Wooden 2017; Gibson and Shrader 2018). Thus, we provide indirect evidence suggesting that heat is likely to be negatively correlated with workers' on-the-job performance through channels that result in poorer sleep and declines in well-being.[38]

[37]Injury reports are mandated for all severe work-related injuries involving in-patient hospitalization, amputation, or loss of an eye. These reports contain detailed industry codes along with information on the time and place of each incident. We merge all incidents with corresponding weather data at the county-day level from Daymet from 2015-2019, then estimate models that examine the likelihood of any incident occurring as a function of heat and a rich set of covariates that incorporate time-varying weather patterns and county and time fixed effects.

[38]Another possibility is that poorer sleep can also affect workers' intertemporal labor supply, reducing hours worked and increasing the likelihood of work absenteeism. To investigate this conjecture, we conduct a complementary analysis by regressing hours worked on temperature variables, adopting the model in Equation (3), while including additional controls for the duration of sleep time and a dummy variable indicating the occurrence of sleeplessness. Due to data limitations, we are unable to conduct a similar analysis for the work absenteeism rate. We report the estimation results in Appendix Table B.6. Comparing Panel A of Table 5 and Table B.6, we find that the corresponding estimates are similar and not significantly different from each other. We interpret these results as supporting evidence that poorer sleep likely has a limited impact on hours worked and labor supply.

5 Exploration of Adaptive Strategies

In this section we consider the possibility that workers and airlines have differing abilities to acclimatize to high temperatures depending on the region's usual climate conditions. To gauge whether heat effects vary across average temperatures, we borrow the classification of climate zones from the International Energy Conservation Code (IECC 2015), which defines various climate regions based on average temperatures, precipitation, and related temperature-based metrics (International Code Council 2015).[39] Zones 1 and 2 comprise the hottest zones with an average temperature of approximately 30°C. This group includes airports in some of the warmest areas of the Southeast and Southwest including Phoenix, Houston, Miami, and Orlando. This is followed by Zone 3 with an average temperature of 24°C covering significant portions of southern and southwestern states. The coldest zones 6 and 7 range from eastern Washington to the Rockies, Minnesota, and Wisconsin all the way to the northeastern states of New York, Vermont, New Hampshire, and Maine.[40]

Figure A.2 shows the estimated results across climate zones for temperatures greater than 35°C, relative to the reference bin of temperatures less than or equal to 20°C.[41] The estimated heat effect on cancellation rates is similar across climate zones, while differences for flight delays are discernible across zones. Both the delay probability and duration in Zones 3 and 4 show greater vulnerability to heat exposure, whereas the hottest climate regions in our sample (Zones 1 and 2) are relatively less affected. We conjecture that factors such as the infrequent occurrence of higher temperatures in milder zones and the adoption of adaptive strategies by workers and airlines in the hottest regions may contribute to the smaller effect. Flight cancellations in cooler areas (Zones 6 and 7) are not differentially affected by heat, but these regions experience lower delay rates than Zones 3 and 4.

To explore potential variation in adaptation across regions, we examine effects on the labor

[39] Panel (a) of Figure A.1 illustrates the distribution of climate zone for the contiguous U.S. according to the definition of IECC 2015, while Panel (b) plots the number of airports for each climate zone along with their average daily temperatures from April to September.

[40] We consolidate airports in Zone 1 and Zone 2 into one group and airports in Zone 6 and Zone 7 into one group, because i) a relatively small share of airports are located in Zone 1 and Zone 7 (1% in Zone 1 and 5% in Zone 7), as shown in Figure 1b and ii) the average temperatures during our sample period are similar between Zone 1 and Zone 2, as well as between Zone 6 and Zone 7.

[41] We estimate a version of Equation (1) that includes the temperature variables interacted with climate zone dummies, as well as temperature bins interacted with airport type, flight length, and weather covariates.

supply and sleep quality of workers from different climate zones.[42] The upper panels of Figure A.3 illustrate that worker hours and absenteeism rates in hotter regions (Zone 1 and Zone 2) are insensitive to high temperatures exceeding 35°C, while workers in Zone 4 had fewer hours worked and those in Zone 3 showed elevated absenteeism rates. We find similar patterns when shifting to sleep patterns, with workers in Zone 3 exhibiting a significant reduction in sleep time, and those in Zones 3 and 4 experiencing an uptick in incidences of sleeplessness. The lack of any notable shifts along these margins for those residing in the hottest climate zones suggests that adaptive measures may extend outside of the workplace context into other aspects of the built environment, such as residential homes.

6 Conclusion

Rising global temperatures underscore the urgency of establishing the impact of heat on productivity across workplace contexts. In this paper, we investigate the effects of heat on productivity in a U.S. sector that is climate-exposed: the airline industry. By utilizing granular data on flight on-time performance linked with hourly meteorological variables, and employing a model augmented with a rich set of fixed effects, we find statistically significant evidence that high temperatures increase the cancellation rate, departure delay rate, and departure delay time of flights. The negative effect on flight on-time performance is not limited to immediate exposure but also persists through later periods during the same day. Our estimates remain robust across different model specifications, and alternative air pollution instruments and temperature measures.

In addition, we find that nonhub airports are particularly vulnerable to rising temperatures relative to their medium and large hub counterparts, and this may be driven by the concentration of shorter flights out of smaller airports. Our finding of heterogeneous effects across different enplanement and flight characteristics underscores how climate change disproportionately affects certain locations.

Supplemental analyses employing time use data illuminate potential mechanisms behind the

[42]Due to the small sample size of the well-being data, we are unable to explore the heterogeneous impacts on this aspect.

effects of heat stress. We find that higher temperatures reduce workers' intertemporal labor supply, with pronounced effects among transportation workers, suggesting that declines in airlines' on-time performance can be partially attributed to reduced hours and higher absenteeism. We also find negative impacts on sleep duration and quality as well as measures of well-being. These relatively under-studied channels of heat stress can contribute to erosion in labor productivity, namely through deteriorating on-the-job-performance.

This paper's focus on a service-based industry in the United States expands existing evidence on the consequences of heat exposure to non-manufacturing sectors that are vulnerable to the changing climate. Adaptation via climate control is expensive or infeasible in many similar contexts, and alternative long-term adaptive strategies may be necessary. These topics are fertile grounds for future research.

References

Adhvaryu, Achyuta, Namrata Kala, and Anant Nyshadham. 2020. The light and the heat: productivity co-benefits of energy-saving technology. *The Review of Economics and Statistics* 102:779–792.

Arceo, Eva, Rema Hanna, and Paulina Oliva. 2016. Does the effect of pollution on infant mortality differ between developing and developed countries? evidence from mexico city. *The Economic Journal* 126:257–280.

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

Bubonya, Melisa, Deborah A. Cobb-Clark, and Mark Wooden. 2017. Mental health and productivity at work: does what you do matter? *Labour economics* 46:150–165.

Bureau of Labor Statistics. 2022. *Air Transportation - May 2022 OEWS Industry-Specific Occupational Employment and Wage Estimates.*

Burke, Marshall, Vincent Tanutama, Sam Heft-Neal, Miyuki Hino, and David Lobell. 2023. Game, Sweat, Match: Temperature and Elite Worker Productivity.

Cachon, Gerard, Santiago Gallino, and Marcelo Olivares. 2012. Severe weather and automobile assembly productivity. *SSRN Electronic Journal.*

Cai, Xiqian, Yi Lu, and Jin Wang. 2018. The impact of temperature on manufacturing worker productivity: Evidence from personnel data. *Journal of Comparative Economics* 46:889–905.

Carleton, Tamma, Amir Jina, Michael Delgado, Michael Greenstone, Trevor Houser, Solomon Hsiang, Andrew Hultgren, et al. 2022. Valuing the Global Mortality Consequences of Climate Change Accounting for Adaptation Costs and Benefits*. *The Quarterly Journal of Economics* 137:2037–2105.

Chen, Xiaoguang, Luoye Chen, Wei Xie, Nathaniel D. Mueller, and Steven J. Davis. 2023. Flight delays due to air pollution in China. *Journal of Environmental Economics and Management* 119:102810.

Chen, Xiaoguang, and Lu Yang. 2019. Temperature and industrial output: Firm-level evidence from China. *Journal of Environmental Economics and Management* 95:257–274.

Coffel, E., and R. Horton. 2015. Climate Change and the Impact of Extreme Temperatures on Aviation. *Weather, Climate, and Society* 7:94–102.

Connolly, Marie. 2008. Here comes the rain again: weather and the intertemporal substitution of leisure. *Journal of Labor Economics* 26 (January): 73–100.

Crocker, D. 2005. *Dictionary of aviation.* A Bloomsbury reference book. Bloomsbury.

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* 4:66–95.

———. 2014. What do we learn from the weather? the new climate-economy literature. *Journal of Economic Literature* 52:740–98.

Deryugina, Tatyana, and Solomon M. Hsiang. 2014. *Does the Environment Still Matter? Daily Temperature and Income in the United States.* Working Paper.

Deschênes, Olivier, and Michael Greenstone. 2011. Climate Change, Mortality, and Adaptation: Evidence from Annual Fluctuations in Weather in the US. *American Economic Journal: Applied Economics* 3:152–185.

Ebi, Kristie L, Anthony Capon, Peter Berry, Carolyn Broderick, Richard de Dear, George Havenith, Yasushi Honda, et al. 2021. Hot weather and heat extremes: health risks. *The Lancet* 398 (August): 698–708.

Federal Aviation Administration. 2005. *Aviation & emissions: a primer.* Office Environment / Energy, January.

Fisher, Anthony C., W. Michael Hanemann, Michael J. Roberts, and Wolfram Schlenker. 2012. The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather: comment. *American Economic Review* 102:3749–3760.

Flouris, Andreas D, Petros C Dinas, Leonidas G Ioannou, Lars Nybo, George Havenith, Glen P Kenny, and Tord Kjellstrom. 2018. Workers' health and productivity under occupational heat strain: a systematic review and meta-analysis. *The Lancet Planetary Health* 2 (December): e521–e531.

Garg, Teevrat, Maulik Jagnani, and Vis Taraz. 2020. Temperature and Human Capital in India. *Journal of the Association of Environmental and Resource Economists* 7:1113–1150.

Gelles, David, and Manuela Andreoni. 2023. How Extreme Heat Affects Workers and the Economy. *The New York Times.*

Gibson, Matthew, and Jeffrey Shrader. 2018. Time use and labor productivity: the returns to sleep. *Review of Economics and Statistics* 100:783–798.

Global Modeling and Assimilation Office, (GMAO). 2015. *MERRA-2 tavg3_3d_asm_Nv: 3d, 3-Hourly, Time-Averaged, Model-Level, Assimilation, Assimilated Meteorological Fields.* V. 5.12.4.

Godøy, Anna, and Harald Dale-Olsen. 2018. Spillovers from gatekeeping – Peer effects in absenteeism. *Journal of Public Economics* 167:190–204.

Graff Zivin, Joshua, Solomon M. Hsiang, and Matthew Neidell. 2018. Temperature and Human Capital in the Short and Long Run. *Journal of the Association of Environmental and Resource Economists* 5:77–105.

Graff Zivin, Joshua, and Matthew Neidell. 2014. Temperature and the allocation of time: implications for climate change. *Journal of Labor Economics* 32:1–26.

Graff Zivin, Joshua, Yingquan Song, Qu Tang, and Peng Zhang. 2020. Temperature and high-stakes cognitive performance: Evidence from the national college entrance examination in China. *Journal of Environmental Economics and Management* 104:102365.

Hancock, P. A., and I. Vasmatzidis. 2003. Effects of heat stress on cognitive performance: the current state of knowledge. *International Journal of Hyperthermia* 19:355–372.

Heal, Geoffrey, and Jisung Park. 2016. Reflections—temperature stress and the direct impact of climate change: a review of an emerging literature. *Review of Environmental Economics and Policy* 10 (July 1, 2016): 347–362.

Heyes, Anthony, and Soodeh Saberian. 2019. Temperature and Decisions: Evidence from 207,000 Court Cases. *American Economic Journal: Applied Economics* 11:238–265.

Hsiang, Solomon M. 2010. Temperatures and cyclones strongly associated with economic production in the caribbean and central america. *Proceedings of the National Academy of sciences* 107:15367–15372.

International Code Council. 2015. *International energy conservation code.* 2015th ed. International Energy Conservation Code. International Code Council.

Jans, Jenny, Per Johansson, and J. Peter Nilsson. 2018. Economic status, air quality, and child health: evidence from inversion episodes. *Journal of health economics* 61:220–232.

Lai, Wangyang, Yun Qiu, Qu Tang, Chen Xi, and Peng Zhang. 2023. The Effects of Temperature on Labor Productivity. *Annual Review of Resource Economics* 15.

LoPalo, Melissa. 2023. Temperature, Worker Productivity, and Adaptation: Evidence from Survey Data Production. *American Economic Journal: Applied Economics* 15:192–229.

Masiol, Mauro, and Roy M. Harrison. 2014. Aircraft engine exhaust emissions and other airport-related contributions to ambient air pollution: a review. *Atmospheric Environment* 95:409–455.

Minor, Kelton, Andreas Bjerre-Nielsen, Sigga Svala Jonasdottir, Sune Lehmann, and Nick Obradovich. 2022. Rising temperatures erode human sleep globally. *One Earth* 5:534–549.

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.

Niemela, Raimo, Mika Hannula, Sari Rautio, Kari Reijula, and Jorma Railio. 2002. The effect of air temperature on labour productivity in call centres—a case study. *Energy and Buildings,* 6.

NIOSH. 1986. *Working in hot environments.* US Department of Health / Human Services CincinnatiOH.

Obradovich, Nick, Robyn Migliorini, Sara C. Mednick, and James H. Fowler. 2017. Nighttime temperature and human sleep loss in a changing climate. *Science Advances* 3.

Park, R. Jisung. 2022. Hot temperature and high-stakes performance. *Journal of Human Resources* 57:400–434.

Park, R. Jisung, Joshua Goodman, Michael Hurwitz, and Jonathan Smith. 2020. Heat and learning. *American Economic Journal: Economic Policy* 12:306–39.

Park, R. Jisung, Nora Pankratz, and A. Patrick Behrer. 2021. *Temperature, workplace safety, and labor market inequality.* IZA Discussion Paper.

Qiu, Yun, and Jinhua Zhao. 2021. Adaptation and the distributional effects of heat: evidence from professional archery competitions. *Southern Economic Journal* 88:1149–1177.

Riley, Karie, Rich Cook, Edward Carr, and Bryan Manning. 2021. A systematic review of the impact of commercial aircraft activity on air quality near airports. *City and environment interactions* 11:100066.

Sager, Lutz. 2019. Estimating the effect of air pollution on road safety using atmospheric temperature inversions. *Journal of Environmental Economics and Management* 98:102250.

Schlenker, Wolfram, W. Michael Hanemann, and Anthony C. Fisher. 2006. The impact of global warming on US agriculture: an econometric analysis of optimal growing conditions. *Review of Economics and statistics* 88:113–125.

Schlenker, Wolfram, and W. Reed Walker. 2016. Airports, air pollution, and contemporaneous health. *The Review of Economic Studies* 83:768–809.

Seppanen, Olli, William J Fisk, and Q H Lei. 2006. *Effect of temperature on task performance in office environment.* Lawrence Berkeley National Laboratory, Berkeley, CA.

Somanathan, E., Rohini Somanathan, Anant Sudarshan, and Meenu Tewari. 2021. The impact of temperature on productivity and labor supply: evidence from indian manufacturing. *Journal of Political Economy* 129:1797–1827.

Stevens, Andrew. 2017. Temperature, Wages, and Agricultural Labor Productivity.

Thornton, M.M., R. Shrestha, Y. Wei, P.E. Thornton, S. Kao, and B.E. Wilson. 2020. *Daymet: daily surface weather data on a 1-km grid for north america, version 4.*

White, Corey. 2017. The Dynamic Relationship between Temperature and Morbidity. *Journal of the Association of Environmental and Resource Economists* 4:1155–1198.

World Health Organization. 2018. *Heat and Health [Fact sheet].* Technical report.

Wragg, David W. 1973. *A dictionary of aviation.* UK: Osprey Publishing Ltd.

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

Zhang, Wei, Ning Ding, Yilong Han, Jie He, and Na Zhang. 2023. The impact of temperature on labor productivity——evidence from temperature-sensitive enterprises. *Frontiers in Environmental Science* 10.

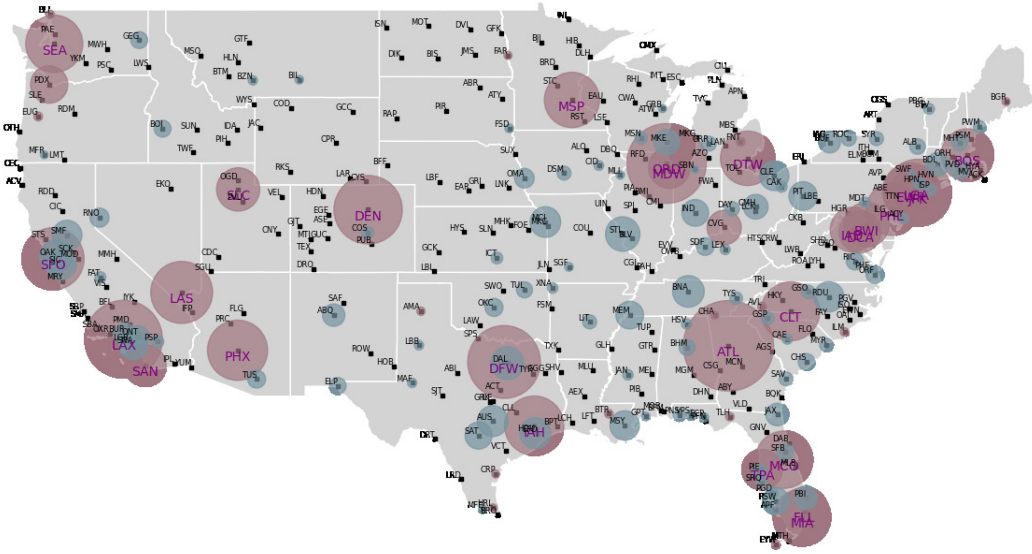
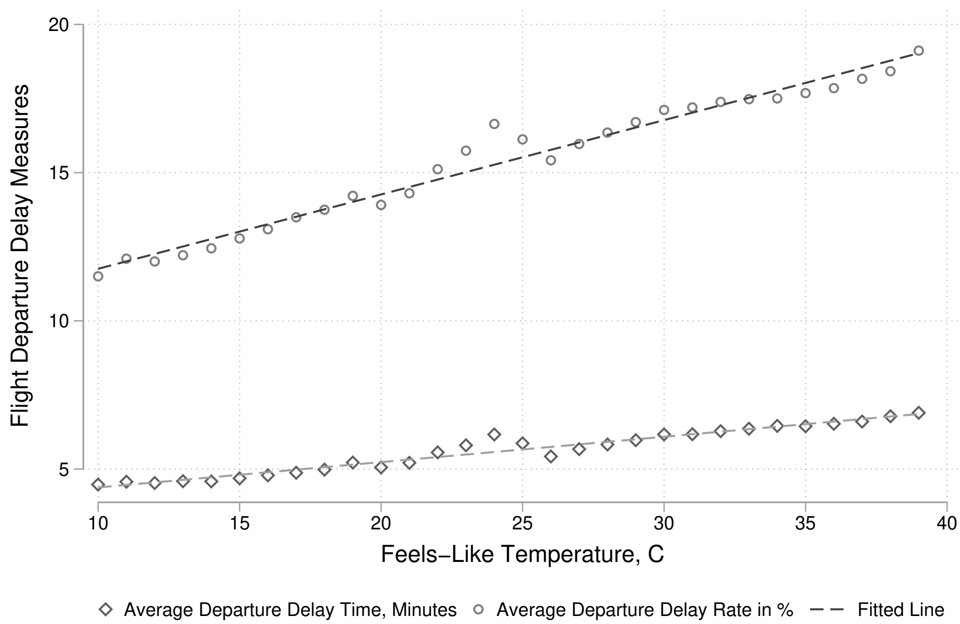
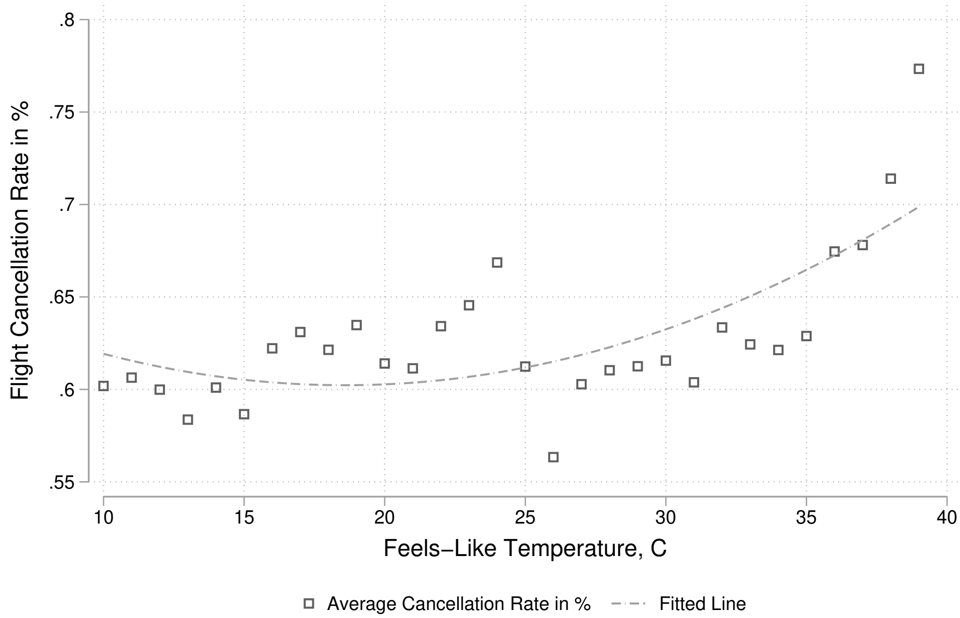


Figure 1: Airports in Sample and Their Average Annual Enplanements, 2004-2019
Notes: This bubble map summarizes airports in the sample and their average annual passenger boarding (enplanements) over 2004-2019. The size of the bubble indicates the share of airport's annual enplanements. We use red to flag large hub airports, defined by FAA as airports receiving 1 percent or more of the annual U.S. commercial enplanements.



(a) Departure Delay Measures



(b) Cancellation

Figure 2: Association Between Temperatures and Flight On-Time Performance
Notes: This figure plots the correlation between feels-like temperatures (in Celsius) and the average flight departure delay time (in minutes), departure delay rate, and cancellation rate, over flights departed at given temperature levels.

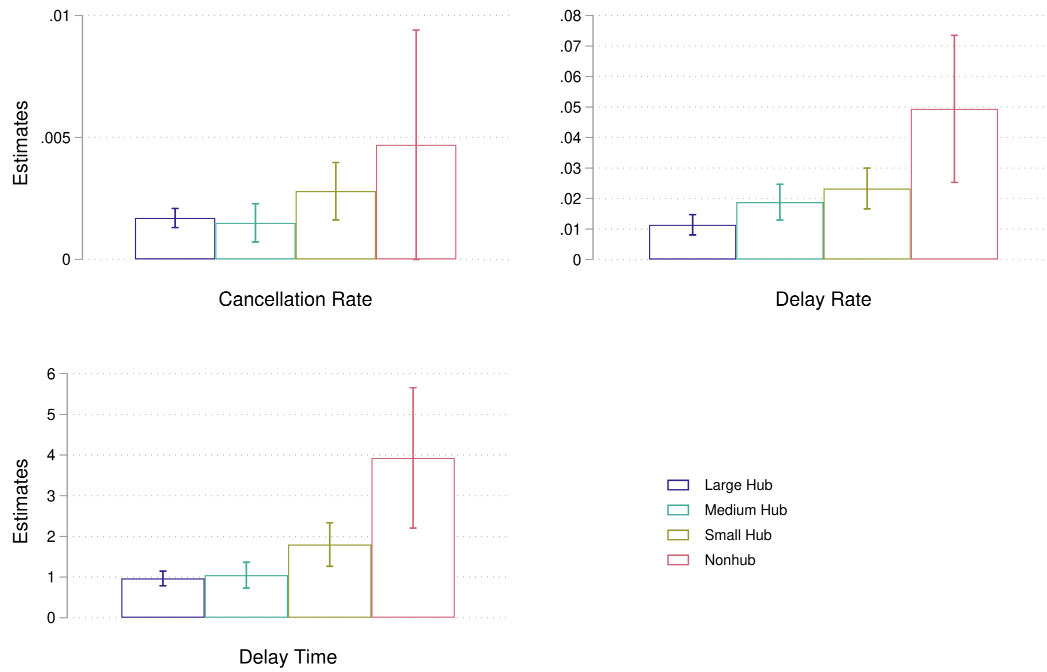


Figure 3: Heterogeneous Effects by Origin Airport Type

Notes: This figure summarizes the point estimates and their 95% confidence intervals of the effect of temperatures falling in the bin $>35°C$, relative to the reference bin of temperatures below 20°C by the origin airport type and for outcomes of the cancellation rate, departure delay rate, and delay time, respectively.

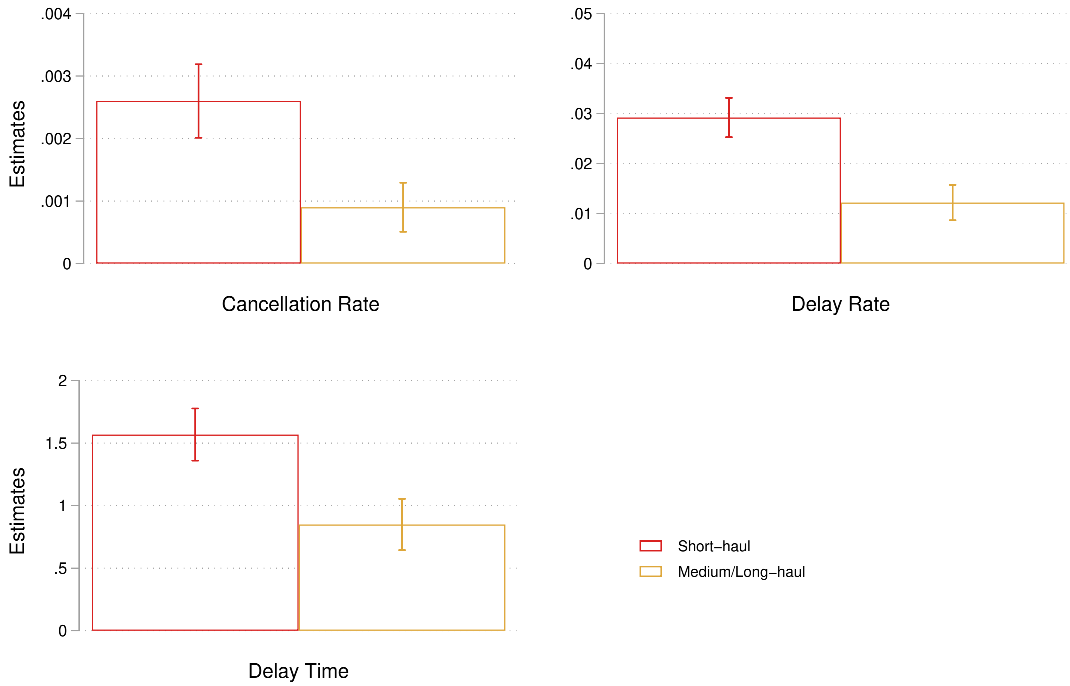


Figure 4: Heterogeneous Effects by Flight Length

Notes: This figure summarizes the point estimates and their 95% confidence intervals of temperatures falling in the treatment bin of greater than 35°C, relative to the reference bin of temperatures below 20°C, for outcomes cancellation rate, departure delay rate, and delay time, separately for short-haul and medium/long-haul flights. Following Wragg (1973) and Crocker (2005), we define medium/long-haul flights as those with a distance greater than 1000 km and short-haul flights as those with a distance less than 1000 km.

Table 1: Summary Statistics: Main Analysis

	N	Mean	Std.	Min	Max
Flight On-Time Performance					
Cancellation Rate	49330008	0.01	0.08	0	1
Departure Delay Rate	49330008	0.15	0.36	0	1
Departure Delay Time (Minutes)	49330008	5.56	26.28	0	2450
Temperature					
Temperature Groups (°C)					
≤ 20	49330008	0.32	0.47	0	1
$\in (20,25]$	49330008	0.25	0.43	0	1
$\in (25,30]$	49330008	0.21	0.41	0	1
$\in (30,35]$	49330008	0.15	0.35	0	1
> 35	49330008	0.07	0.26	0	1
Hours of Temp >35°C, 5am-6pm	49329563	0.45	1.53	0	14
Covariates					
Origin					
Precipitation Indicator	49330008	0.08	0.26	0	1
One hour precipitation (inches)	49330008	0.00	0.03	0	5
Wind Speed (mph)	49329940	8.84	5.07	0	266
Pressure altimeter (inches)	49326109	29.98	0.16	1	55
Relative Humidity (%)	49330008	57.98	21.80	1	167
Temperature Inversion (925hPa − Surface)	49330008	-0.19	2.88	-6	20
Air Pollution					
CO	39730941	0.34	0.18	0	4
NO_2	40918950	10.84	6.83	0	77
ozone	48041152	0.03	0.01	0	0.10
PM2.5	38352614	10.25	6.00	0	287
Destination					
Precipitation Indicator	49330008	0.07	0.26	0	1
One hour precipitation (inches)	49330008	0.00	0.03	0	8
Wind Speed (mph)	49329933	8.87	5.08	0	266
Pressure altimeter (inches)	49326484	29.98	0.15	1	40
Relative Humidity (%)	49330008	58.01	21.78	1	116

Notes: Table 1 presents summary statistics of variables in the sample for the main analysis. It spans months April to September and contains data on cancellation rate, departure delay rate, and departure delay time aggregated by origin-destination pair, carrier, date, and time block.

Table 2: Summary Statistics: Labor Supply, Well-being, and Temperature

	N	Mean	Std.	Min	Max
Labor Supply:					
Working Time (Minutes)	85623	277.69	271.53	0	1380
1(Absence Last Week)	85623	0.04	0.19	0	1
Sleeping:					
Sleep Time on a Diary Day (Minutes)	85623	507.19	129.86	0	1428
Any Sleeplessness on a Diary Day	85623	0.04	0.20	0	1
Well-being:					
Not Well-Rested	16320	0.21	0.41	0	1
Worse-Than a Typical Day	16320	0.07	0.25	0	1
Weather:					
Daily Max Temperature (°C) $\in (20,25]$	69281	0.16	0.37	0	1
Daily Max Temperature (°C) $\in (25,30]$	69281	0.20	0.40	0	1
Daily Max Temperature (°C) $\in (30,35]$	69281	0.15	0.36	0	1
Daily Max Temperature (°C) > 35	69281	0.04	0.21	0	1
Minimum Temperature (°C)	69281	7.45	10.07	-38	33
Maximum Temperature (°C)	69281	20.24	10.68	-20	47
Accumulated Precipitation (mm/day)	69281	3.07	7.89	0	185
Day Length (s/day)	69281	43064.51	6702.25	28921	57432
# Days Max Temp $\in (20,25]$ Last Week	69281	1.13	1.56	0	7
# Days Max Temp $\in (25,30]$ Last Week	69281	1.38	1.91	0	7
# Days Max Temp $\in (30,35]$ Last Week	69281	1.09	1.97	0	7
# Days Max Temp >35 Last Week	69281	0.32	1.22	0	7
Weekly Avg Precipitation Last Week	69281	21.12	27.48	0	473
Weekly Avg Day Length Last Week	69281	43039.83	6706.37	29053	57468
Covariates:					
Diary day a holiday	85623	0.02	0.13	0	1
Male	85623	0.54	0.50	0	1
Married	85623	0.57	0.50	0	1
Has Child <18	85623	0.51	0.50	0	1
Age	85623	43.26	12.08	15	85
% Age >65	85623	0.03	0.17	0	1
Paid Hourly	85623	0.46	0.50	0	1
% Reside in Urban Area	85623	0.84	0.37	0	1
% Hispanic	85623	0.14	0.35	0	1
% Black	85623	0.13	0.33	0	1
% Asian	85623	0.04	0.20	0	1
% < High School	84954	0.06	0.23	0	1
% High School Graduate	84954	0.23	0.42	0	1
% Some College	84954	0.55	0.50	0	1

Notes: Table 2 shows the summary statistics of variables of the full-time employed sample. Labor supply data comes from ATUS regular module 2005-2019. Well-being data comes from ATUS Well-being module 2010, 2012, and 2013.

Table 3: The Same-Day Contemporaneous Effect of Temperature on Flight On-Time Performance

	(1)	(2)	(3)
	Cancellation Rate	*Delay Rate*	*Delay Time*
Temp $\in$ (20°C, 25°C]	0.0004***	0.0084***	0.294***
	(0.00007)	(0.0005)	(0.0266)
Temp $\in$ (25°C, 30°C]	0.0006***	0.0117***	0.426***
	(0.00009)	(0.0008)	(0.0404)
Temp $\in$ (30°C, 35°C]	0.0011***	0.0159***	0.783***
	(0.00012)	(0.0011)	(0.0587)
Temp >35°C	0.0018***	0.0205***	1.184***
	(0.00016)	(0.0014)	(0.0758)
Percentage Effects (in %)			
Temp $\in$ (20°C,25°C]	6.25	5.36	5.20
Temp $\in$ (25°C,30°C]	9.55	7.45	7.53
Temp $\in$ (30°C,35°C]	17.67	10.13	13.85
Temp >35°C	29.67	13.06	20.94
N	29005757	29005757	29005757
R^2	0.0002	0.012	0.005

Notes: Standard errors in parentheses. * $p < 0.10$, **$p < 0.05$, *** $p < 0.01$. All models controls for origin-destination pair, carrier, month by year, day of week, and time block fixed effects, and time-varying weather covariates, including a binary indicator and quadratic polynomials of hourly precipitation, a factor variable measuring visual obscuration, and continuous variables for humidity, and wind speed. Our model also controls for destination weather conditions and origin by month fixed effects. Percentage Effect (in %) indicates the percentage effect relative to the corresponding baseline sample mean. Standard errors are clustered at the origin-destination pair level.

Table 4: The Same-Day Cumulative Effect of Temperature on Flight On-Time Performance

	(1)	(2)	(3)
	Cancellation Rate	*Delay Rate*	*Delay Time*
Hours of Temp >35°C, 5am-6pm	0.0001	0.0080***	0.2416***
	(0.0001)	(0.0005)	(0.0290)
Temp ∈ (20°C, 25°C]	0.0005***	0.0216***	0.6836***
	(0.0002)	(0.0012)	(0.0650)
Temp ∈ (25°C, 30°C]	0.0009***	0.0162***	0.3199***
	(0.0003)	(0.0019)	(0.1050)
Temp ∈ (30°C, 35°C]	0.0015***	0.0125***	0.2204
	(0.0004)	(0.0024)	(0.1364)
Temp >35°C	0.0017***	0.0179***	0.4786**
	(0.0005)	(0.0036)	(0.2035)
Percentage Effects (in %)			
Hours of Temp >35°C, 5am-6pm	1.72	3.53	3.33
N	3345949	3345949	3345949
R^2	0.0004	0.006	0.0009

Notes: Standard errors in parentheses. * $p < 0.10$, **$p < 0.05$, *** $p < 0.01$. All models controls for origin-destination pair, carrier, month by year, day of week, and time block fixed effects, and time-varying weather covariates, including a binary indicator and quadratic polynomials of hourly precipitation, a factor variable measuring visual obscuration, and continuous variables for humidity, and wind speed. Our model also controls for destination weather conditions and origin by month fixed effects. Percentage Effect (in %) indicates the percentage effect relative to the corresponding baseline sample mean. Standard errors are clustered at the origin-destination pair level.

Table 5: The Effect of Temperature on Worker Labor Supply

	(1)	(2)	(3)
		Transportation	
	All	Industry	Occupation
Panel A.		Working Time	
Max Temp ∈ (20°C, 25°C]	1.275	-23.44	-7.954
	(3.858)	(18.32)	(20.48)
Max Temp ∈ (25°C, 30°C]	-2.981	1.136	-20.40
	(4.173)	(23.72)	(25.37)
Max Temp ∈ (30°C, 35°C]	-6.658	-34.50	-60.89**
	(5.635)	(27.60)	(30.95)
Max Temp >35°C	-13.60	-69.45*	-82.35*
	(8.512)	(40.91)	(44.49)
Sample Mean	277.7	294.9	297.7
R^2	0.370	0.411	0.420
Panel B.		1(Absence Last Week)	
Days Max Temp ∈ (20°C, 25°C] Last Week	0.0002	0.0037	0.0016
	(0.0008)	(0.0031)	(0.0032)
Days Max Temp ∈ (25°C, 30°C] Last Week	0.0007	0.0069**	0.0047
	(0.0007)	(0.0027)	(0.0032)
Days Max Temp ∈ (30°C, 35°C] Last Week	0.00061	0.0081**	0.0037
	(0.0010)	(0.0041)	(0.0041)
Days Max Temp > 35°C Last Week	0.0028*	0.0113**	0.0078
	(0.0015)	(0.0056)	(0.0060)
Sample Mean	0.038	0.042	0.039
R^2	0.034	0.214	0.198
N	68750	3508	3175

Notes: Standard errors in parentheses. * p<0.10 ** p<0.05 *** p<0.01. Column (1) includes the entire sample of full-time employed individuals. Column (2) includes full-time employed individuals who work in the transportation industry, based on the major industry code for the respondent's main job. Column (3) includes full-time employed individuals who work in transportation and material moving occupations, as classified by the major occupation code for the main job. Standard errors are clustered at the state-year level.

Table 6: The Effect of Temperature on Sleep Time and Sleeplessness

	(1)	(2)	(3)
		Transportation	
	All	Industry	Occupation
Panel A.	Sleep Time on a Diary Day (in Minutes)		
Max Temp ∈ (20°C, 25°C]	-1.535	2.935	-5.775
	(2.190)	(10.69)	(11.66)
Max Temp ∈ (25°C, 30°C]	-6.411**	-19.88*	-16.780
	(2.544)	(11.85)	(13.78)
Max Temp ∈ (30°C, 35°C]	-6.865**	-6.476	-14.230
	(3.251)	(15.01)	(17.17)
Max Temp >35°C	-9.197**	-14.39	7.226
	(4.367)	(22.11)	(26.54)
Sample Mean	507.2	505.4	515.9
R^2	0.124	0.234	0.282
Panel B.	Any Sleeplessness on a Diary Day		
Max Temp ∈ (20°C, 25°C]	-0.0006	0.0035	0.0096
	(0.0037)	(0.0159)	(0.0152)
Max Temp ∈ (25°C, 30°C]	0.0069	0.0476**	0.0403**
	(0.0048)	(0.0205)	(0.0167)
Max Temp ∈ (30°C, 35°C]	0.0051	0.0328	0.0534**
	(0.0058)	(0.0262)	(0.0226)
Max Temp >35°C	0.0180**	0.0409	0.0572*
	(0.0083)	(0.0392)	(0.0295)
Sample Mean	0.04	0.04	0.04
R^2	0.027	0.204	0.252
N	68750	3508	3175

Notes: Standard errors in parentheses. * p<0.10 ** p<0.05 *** p<0.01. Column (1) includes the entire sample of full-time employed individuals. Column (2) includes full-time employed individuals who work in the transportation industry, based on the major industry code for the respondent's main job. Column (3) includes full-time employed individuals who work in transportation and material moving occupations, as classified by the major occupation code for the main job. Standard errors are clustered at the state-year level.

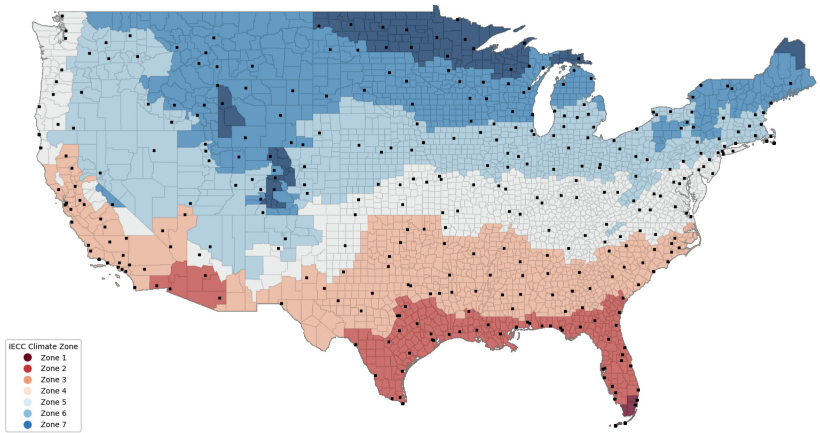
Table 7: The Effect of Temperature on General Well-Being

	(1)	(2)
	Not Well-Rested	Worse-Than Typical
Max Temp $\in$ (20°C,25°C]	-0.0056	-0.0072
	(0.0151)	(0.0112)
Max Temp $\in$ (25°C,30°C]	0.0067	0.0036
	(0.0163)	(0.0141)
Max Temp $\in$ (30°C,35°C]	0.0413	0.0011
	(0.0263)	(0.0169)
Max Temp >35°C	0.0632*	0.0097
	(0.0360)	(0.0171)
Sample Mean	0.207	0.068
R^2	0.066	0.101
N	13167	13167

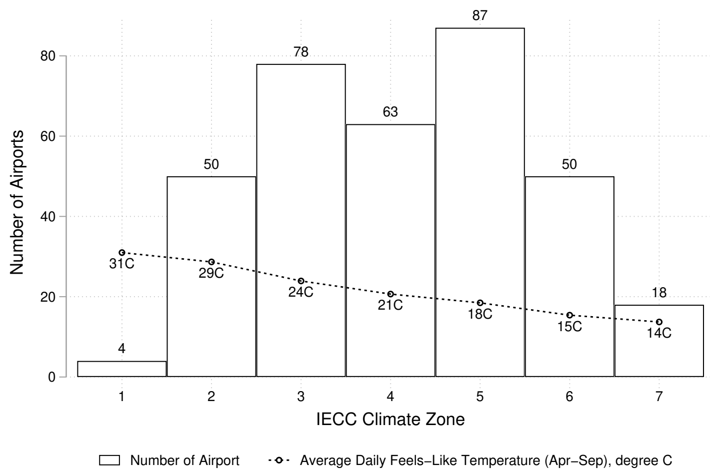
Note: Standard errors in parentheses. * p<0.10 ** p<0.05 *** p<0.01. The shown results are for all full-time employed respondents. Following the instruction in the data codebook of the ATUS Well-being Module, we use the WB respondent-level final weights in the estimation of Columns (1)-(2). Standard errors are clustered at the state-year level.

Online Appendix

A Figures



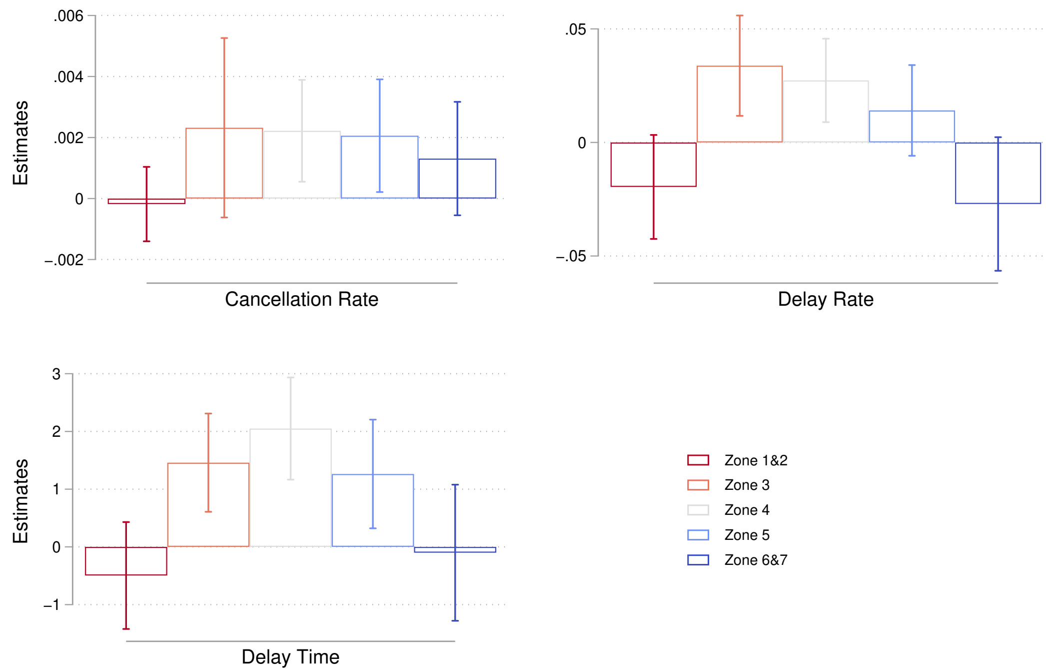
(a) U.S. Climate Zone



(b) Airport Distribution and Average Temperature

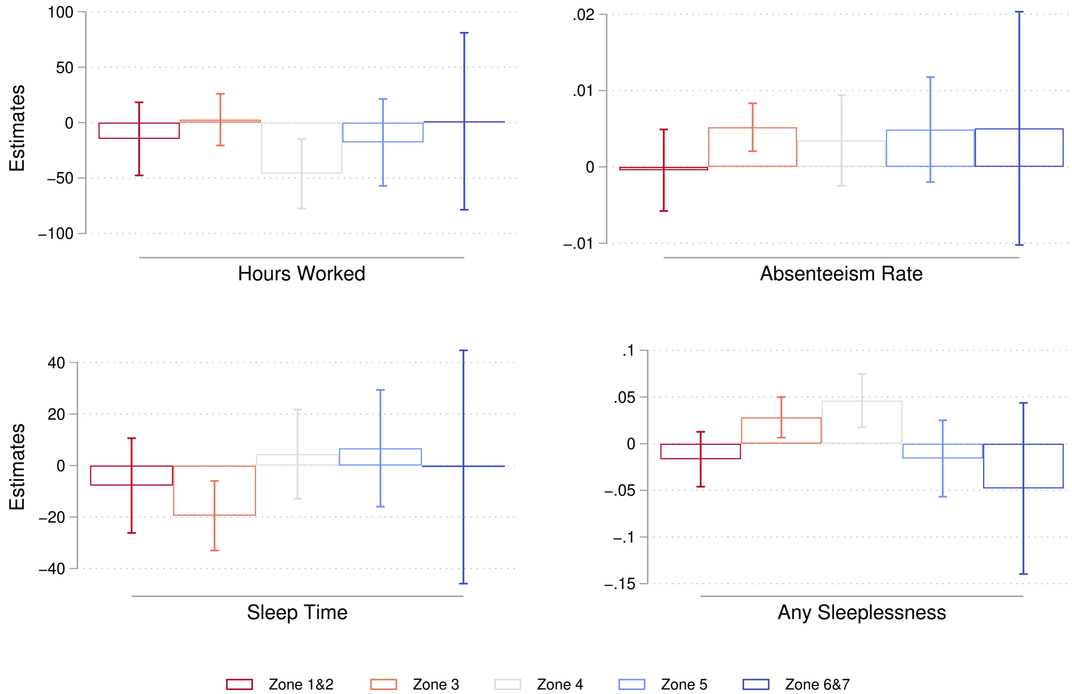
A.1: U.S. Climate Zone and Airports

Notes: This figure presents airports in our sample by climate zone and their distribution across climate zones. Panel (a) illustrates the climate zone of the contiguous U.S. according to the IECC 2015 definition, while Panel (b) shows a histogram of airports by climate zone and the average temperature for each climate zone over months April to September in Celsius.



A.2: Heterogeneous Effects on Flight On-Time Performance by Climate Zone

Notes: This figure summarizes the effect of temperatures falling in the bin >35°C, relative to the reference bin of temperatures less than or equal to 20°C, by climate zone and for outcomes of the cancellation rate, departure delay rate, and delay time, respectively. We estimate a version of Equation (1) that includes the temperature variables interacted with climate zone dummies, as well as temperature bins interacted with airport type, flight length, and weather covariates.



A.3: Heterogeneous Effects on Labor Supply, Sleep, and Well-being by Climate Zone
Notes: This figure plots the estimated effect of temperatures above 35°C, relative to the reference bin of temperatures below or equal to 20°C, by climate zone for outcomes hours worked, absenteeism rate, sleep time, and sleeplessness indicator, respectively.

B Tables

B.1: Robustness Check (1/4) - Alternative Model Specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Real Air Temperatures			*Alternative IV Model*			*Alternative Clustering Level*		
	Cancellation Rate	Delay Rate	Delay Time	Cancellation Rate	Delay Rate	Delay Time	Cancellation Rate	Delay Rate	Delay Time
Temp ∈ (20°C, 25°C]	0.0004***	0.0089***	0.295***	0.0004***	0.0087***	0.307***	0.0004**	0.0084***	0.294***
	(0.0001)	(0.0005)	(0.0273)	(0.0001)	(0.0005)	(0.0270)	(0.0002)	(0.0021)	(0.0793)
Temp ∈ (25°C, 30°C]	0.0006***	0.0106***	0.387***	0.0006***	0.0120***	0.439***	0.0006***	0.0117***	0.426***
	(0.0001)	(0.0008)	(0.0424)	(0.0001)	(0.0008)	(0.0409)	(0.0002)	(0.0026)	(0.108)
Temp ∈ (30°C, 35°C]	0.0010***	0.0151***	0.762***	0.0011***	0.0161***	0.790***	0.0011***	0.0159***	0.783***
	(0.0001)	(0.0012)	(0.0631)	(0.0001)	(0.0011)	(0.0593)	(0.0003)	(0.0033)	(0.147)
Temp >35°C	0.0014***	0.0248***	1.022***	0.0018***	0.0207***	1.195***	0.0018***	0.0205***	1.184***
	(0.0002)	(0.0020)	(0.105)	(0.0002)	(0.0014)	(0.0765)	(0.0005)	(0.0048)	(0.200)
Percentage Effects (in %)									
Temp ∈ (20°C, 25°C]	5.78	5.66	5.22	6.35	5.52	5.43	6.25	5.36	5.20
Temp ∈ (25°C, 30°C]	9.65	6.75	6.84	9.72	7.64	7.76	9.55	7.45	7.53
Temp ∈ (30°C, 35°C]	16.48	9.62	13.48	17.83	10.25	13.97	17.67	10.13	13.85
Temp >35°C	24.00	15.80	18.08	29.83	13.18	21.14	29.67	13.06	20.94
N	29005757	29005757	29005757	29005757	29005757	29005757	29005757	29005757	29005757
R^2	0.0002	0.012	0.005	0.0002	0.011	0.005	0.0002	0.012	0.005

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table reports the robustness estimation results of different models. Columns (1)-(3) use the real air temperature as the treatment variable, Columns (4) to (6) adopts an alternative IV model employing temperature inversions interacted with wind speed and temperature inversions interacted with 45-degree binned wind directions as IVs for air pollution, and Columns (7) to (9) cluster the standard errors at the origin-by-month level. Standard errors are clustered at the origin-destination pair level for Columns (1) to (6).

B.2: Robustness Check (2/4) - Controlling for Aircraft Age

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Cancellation Rate			Departure Delay Rate			Departure Delay Time		
Temp ∈ (20°C, 25°C]	0.0003***	0.0003***	0.0003***	0.0084***	0.0084***	0.0084***	0.296***	0.296***	0.296***
	(0.0001)	(0.0001)	(0.0001)	(0.0005)	(0.0005)	(0.0005)	(0.0282)	(0.0282)	(0.0282)
Temp ∈ (25°C, 30°C]	0.0004***	0.0004***	0.0004***	0.0114***	0.0114***	0.0114***	0.420***	0.419***	0.419***
	(0.0001)	(0.0001)	(0.0001)	(0.0008)	(0.0008)	(0.0008)	(0.0427)	(0.0427)	(0.0427)
Temp ∈ (30°C, 35°C]	0.0006***	0.0006***	0.0006***	0.0155***	0.0155***	0.0155***	0.786***	0.784***	0.784***
	(0.0001)	(0.0001)	(0.0001)	(0.0011)	(0.0011)	(0.0011)	(0.0620)	(0.0620)	(0.0621)
Temp >35°C	0.0012***	0.0012***	0.0012***	0.0196***	0.0195***	0.0193***	1.183***	1.181***	1.174***
	(0.0001)	(0.0001)	(0.0001)	(0.0014)	(0.0014)	(0.0014)	(0.0800)	(0.0799)	(0.0801)
Aircraft Age		0.0001***			0.0011***			0.0467***	
		(0.0000)			(0.0000)			(0.0019)	
Percentage Effects (in %)									
Temp ∈ (20°C, 25°C]	7.34	7.34	7.26	5.34	5.34	5.32	5.17	5.17	5.17
Temp ∈ (25°C, 30°C]	8.82	8.82	8.77	7.22	7.22	7.22	7.34	7.32	7.32
Temp ∈ (30°C, 35°C]	16.08	16.03	16.11	9.81	9.81	9.81	13.74	13.70	13.70
Temp >35°C	29.15	28.89	28.89	12.41	12.34	12.22	20.67	20.64	20.52
N	26012256	26012256	26012256	26012256	26012256	26012256	26012256	26012256	26012256
R^2	0.0001	0.0002	0.0003	0.013	0.013	0.013	0.005	0.006	0.006
Aircraft Age × Carrier FEs			✓			✓			✓

Note: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Percentage Effect (in %) indicates the percentage effect relative to the corresponding baseline sample mean. Standard errors are clustered at the origin-destination pair level. For reference, Columns (1), (4), and (7) report the result of Equation (1) without controlling for aircraft age. Columns (2), (5), and (8) include a continuous variable for aircraft age. Columns (3), (6), and (9) add carrier fixed effects interacted with the continuous aircraft age variable.

B.3: Robustness Check (3/4) - Placebo Test

	(1)	(2)	(3)	(4)	(5)	(6)
	Randomized Temperature			*Randomized Outcome*		
	Cancellation Rate	Delay Rate	Delay Time	Cancellation Rate	Delay Rate	Delay Time
Temp ∈ (20°C, 25°C]	-0.00002	0.00037	0.00862	-0.00001	-0.0003	0.00355
	(0.00004)	(0.00018)	(0.0130)	(0.00005)	(0.00024)	(0.0174)
Temp ∈ (25°C, 30°C]	-0.00001	-0.00002	-0.00350	0.00009	-0.00043	0.0182
	(0.00004)	(0.00018)	(0.0135)	(0.00007)	(0.00035)	(0.0247)
Temp ∈ (30°C, 35°C]	-0.00002	-0.00003	0.00235	0.000176	-0.000440	0.0264
	(0.00004)	(0.00021)	(0.0150)	(0.00010)	(0.00046)	(0.0326)
Temp >35°C	-0.00005	0.00023	-0.00175	0.00014	-0.00007	0.0430
	(0.00006)	(0.00027)	(0.0197)	(0.00012)	(0.00057)	(0.0416)
N	29005757	29005757	29005757	29005757	29005757	29005757
R^2	0.0002	0.012	0.005	0.00000	0.00000	0.00000

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table reports the estimation results for shuffled treatment (Columns 1 to 3 and outcome variables (Columns 4 to 6), adopting the preferred model Equation (1) and 2SLS. Standard errors are clustered at the origin-destination pair level.

B.4: Robustness Check (4/4) - Results of Randomized Full-Year Sample

	(1)	(2)	(3)	(4)	(5)	(6)
	April - September Sample			*Full-Year Sample*		
	Cancellation Rate	Delay Rate	Delay Time	Cancellation Rate	Delay Rate	Delay Time
Temp < 5°C	0.0009***	0.0208***	0.7014***	0.0018***	0.0248***	0.8661***
	(0.0002)	(0.0009)	(0.0543)	(0.0001)	(0.0006)	(0.0325)
Temp ∈ (5°C, 10°C]	0.0006***	0.0068***	0.1581***	0.0004***	0.0057***	0.2178***
	(0.0001)	(0.0006)	(0.0397)	(0.0001)	(0.0005)	(0.0302)
Temp ∈ (15°C, 20°C]	0.0002**	0.0037***	0.1113***	0.0001**	-0.0004	0.0303
	(0.0001)	(0.0006)	(0.0304)	(0.0001)	(0.0005)	(0.0267)
Temp ∈ (20°C, 25°C]	0.0005***	0.0112***	0.3746***	0.0003***	0.0025***	0.1712***
	(0.0001)	(0.0008)	(0.0415)	(0.0001)	(0.0007)	(0.0404)
Temp ∈ (25°C, 30°C]	0.0007***	0.0139***	0.4853***	0.0005***	0.0046***	0.2883***
	(0.0001)	(0.0011)	(0.0551)	(0.0001)	(0.0011)	(0.0591)
Temp ∈ (30°C, 35°C]	0.0012***	0.0176***	0.8275***	0.0008***	0.0079***	0.5952***
	(0.0002)	(0.0014)	(0.0718)	(0.0002)	(0.0014)	(0.0803)
Temp >35°C	0.0019***	0.0218***	1.2174***	0.0013***	0.0120***	0.9533***
	(0.0002)	(0.0016)	(0.0876)	(0.0002)	(0.0017)	(0.0948)
N	29005757	29005757	29005757	25707944	25707944	25707944
R^2	0.0002	0.012	0.005	0.0002	0.011	0.005

Notes: Standard errors in parentheses. * $p < 0.10$, **$p < 0.05$, *** $p < 0.01$. Results are estimated using a sample that randomly draws 50% of the observations from a full-year sample spanning January to December for the years 2004 to 2019. All models controls for origin-destination pair, carrier, month by year, day of week, and time block fixed effects, and time-varying weather covariates, including a binary indicator and quadratic polynomials of hourly precipitation, a factor variable measuring visual obscuration, and continuous variables for humidity, and wind speed. Our model also controls for destination weather conditions and origin by month fixed effects.

B.5: The Contemporaneous Effect of Temperature on Workplace Injuries

	(1)	(2)	(3)	(4)
		Any Injury		
	All	Transportation	Air Transport	Truck Transport
Temp $\in$ (20°C, 25°C]	0.000385***	0.0000211	0.00000945	0.00000885
	(0.000129)	(0.0000272)	(0.0000106)	(0.0000198)
Temp $\in$ (25°C, 30°C]	0.000674***	0.0000273	0.00000819	0.0000243
	(0.000148)	(0.0000315)	(0.0000120)	(0.0000225)
Temp $\in$ (30°C, 35°C]	0.00198***	0.000157***	0.0000230*	0.000105***
	(0.000177)	(0.0000369)	(0.0000136)	(0.0000273)
Temp > 35°C	0.00295***	0.000128*	0.0000377	0.000150***
	(0.000324)	(0.0000697)	(0.0000297)	(0.0000504)
Percentage Effects (in %)				
Temp $\in$ (20°C, 25°C]	4.41	5.14	13.65	4.17
Temp $\in$ (25°C, 30°C]	7.72	6.66	11.82	11.44
Temp $\in$ (30°C, 35°C]	22.72	38.28	33.25	49.60
Temp > 35°C	33.84	31.22	54.42	70.87
N	5673925	5673925	5673925	5673925

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Sample combines county-day Daymet weather information with data on severe injuries from the Occupational Safety and Health Administration from 2015-2019. Outcome variables are indicators for any injury across all industries (Column 1), and separately by industry type (Columns 2-4). All models control for other weather conditions such as day length, daily precipitation, and daily minimum temperatures, and separate county, year, and day of week fixed effects. Percentage Effects (in %) indicate the percentage effect relative to the corresponding baseline sample mean. Standard errors are clustered at the county-year level.

B.6: More on The Effect of Temperature on Worker Labor Supply

	(1)	(2)	(3)
		Transportation	
	All	Industry	Occupation
Panel A.		Working Time	
Max Temp $\in$ (20°C, 25°C]	0.242	-21.25	-11.03
	(3.538)	(17.27)	(18.65)
Max Temp $\in$ (25°C, 30°C]	-6.954*	-11.69	-28.28
	(3.665)	(21.82)	(22.44)
Max Temp $\in$ (30°C, 35°C]	-10.99**	-38.18	-65.91**
	(5.090)	(26.62)	(28.47)
Max Temp $>$35°C	-19.01**	-78.55**	-72.57*
	(7.785)	(37.88)	(38.95)
Sample Mean	277.7	294.9	297.7
N	68750	3508	3175
Sleeplessness Dummy	✓	✓	✓
Sleep Time Control	✓	✓	✓

Notes: Standard errors in parentheses. * p<0.10 ** p<0.05 *** p<0.01. Column (1) includes the entire sample of full-time employed individuals. Column (2) includes full-time employed individuals who work in the transportation industry, based on the major industry code for the respondent's main job. Column (3) includes full-time employed individuals who work in transportation and material moving occupations, as classified by the major occupation code for the main job. Standard errors are clustered at the state-year level.