









Impact Factor: **1.848**5-Year Impact Factor: **1.965**

 Contents
  PDF
  More
  Citations
  Share
  Information, rights and permissions
  Metrics and citations
  Figures and tables

Abstract

Using a sample of more than 0.3 million marathon runners of 56 race events in China in 2014 and 2015, we estimate the air pollution elasticity of finish time to be 0.041. Our causal identification comes from the exogeneity of air pollution on the race day because runners are required to register for a race a few months in advance and we control for confounding factors. Including individual fixed effects also provides consistent evidence. Our study contributes to the emerging literature on the effect of air pollution on short-run productivity, particularly on the performance of athletes engaging in outdoor sports.

An emerging literature finds a sizable, negative effect of air pollution on short-run labor productivity ([Adhvaryu, Kala, & Nyshadham, 2014](#); [Chang, Graff Zivin, Gross, & Neidell, 2016a, 2016b](#); [Fu, Viard, & Zhang, 2017](#); [Graff Zivin and Neidell, 2012](#); [Lichter, Pestel, & Sommer, 2017](#)). This study contributes to this literature by estimating the causal effect of air pollution on marathon runners' performance (finish time) using a sample of more than million runners in 37 cities and 56 race events in China in 2014 and 2015. Our causal identification relies mainly on the exogeneity of air quality on the race day because runners are required to register for a race a few months in advance and air quality on the race day can be considered random. This identification strategy has been implemented in a few environmental studies (e.g., [Lavy, Ebenstein, & Roth, 2014](#); [Lichter et al., 2017](#); [Park, 2016](#)). We estimate the air pollution elasticity of finish time to be 0.0408. This effect is economically significant because of large variations in air quality across Chinese cities. For example, an average full-marathon runner will need 12 more minutes to cross the finish line if he or she were to run the Beijing Marathon in 2014 when the air was severely polluted, compared with running on a day with average air quality.

PDF

Help

Privacy

The related literature can be grouped into two strands.¹ The first is a large literature documenting a harmful effect of air pollution on human health. Common air pollutants include particulate matter 2.5 ug or less in diameter (PM_{2.5}), particulate matter 10 ug or less in diameter (PM₁₀), sulfur dioxide (SO₂), nitrogen dioxide (NO₂), carbon monoxide (CO), and ozone (O₃). Long-run exposure to these pollutants can lead to cardiopulmonary diseases, respiratory infections, lung cancer, infant morbidity, asthma, and reduced life expectancy ([Chay & Greenstone, 2003](#); [Y. Chen, Ebenstein, Greenstone, & Lie, 2013](#); Environmental Protection Agency [EPA], 2004; [Neidell, 2004](#); [Zhang, Chen, & Zhang, 2018](#)). More relevant in our setting are the effects of short-run exposure to ambient air pollution. These include decreased lung function, irregular heartbeat, increased respiratory problems, nonfatal heart attacks, and angina.² Air pollution can also lower cognitive ability ([Lavy et al., 2014](#); [Marcotte, 2017](#)), increase anxiety, and have other negative psychological effects ([Bullinger, 1989](#); [S. Chen, Oliva, & Zhang, 2018](#); [Pun, Manjouride, & Suh, 2017](#)). In addition, the sports health literature provides evidence for a negative effect of air pollution on athletes' health and performance ([Chimenti et al., 2009](#); [Rundell, 2012](#)).

The second strand of literature focuses on the effect of air pollution on short-run labor productivity. [Graff Zivin and Neidell \(2012\)](#) find that ozone reduces the productivity of outdoor fruit pickers in California. [Chang, Graff Zivin, Gross, and Neidell \(2016a\)](#) find that PM_{2.5} reduces the productivity of indoor pear packers in California. [Adhvaryu, Kala, and Nyshadham \(2014\)](#) identify that PM_{2.5} reduces hourly productivity of workers in a garment factory in India. [He, Liu, and Salvo \(2018\)](#) find that PM_{2.5} and SO₂ reduce the output of textile workers at two firms in Henan and Jiangsu Provinces, China. [Chang, Graff Zivin, Gross, and Neidell \(2016b\)](#) identify the negative effects of air pollution on the productivity of workers at two call centers in Shanghai and Nantong, China. [Archsmith, Heyes, and Saberian \(2016\)](#) find that CO and PM_{2.5} negatively affect the productivity of professional baseball umpires in the United States. [Fu, Viard, and Zhang \(2017\)](#) provide more comprehensive evidence that air pollution decreases the labor productivity of manufacturing firms, using a nationwide longitudinal firm survey sample capturing 90% of manufacturing output in China.

The closest related paper is by [Lichter, Pestel, and Sommer \(2017\)](#). They find that a 1% increase in the concentration of PM₁₀ leads to a 0.02% decrease in professional soccer players' performance (measured by the number of passes in a match) in Germany, an elasticity one half the size of our estimate but still comparable. Their causal identification

PDF

Help

Privacy

takes advantage of the exogeneity of match scheduling which is controlled by the German Football League and beyond the control of individual teams and players. Different from our setting, in that an individual runner's performance is mainly determined by individual effort, football is a team sport and a free rider problem may arise in that a tiring player may have an incentive to reduce his or her own effort and rely on other team players' effort.³ If air pollution strengthens this free riding problem, this will amplify and, therefore, overestimate the negative effect of air pollution on players' performance. Regardless of the different settings, our study complements theirs by identifying a similar, robust, negative effect of air pollution on marathon runners' performance.

Our findings have a few important implications for professional athletes who engage in outdoor sports, for city governments organizing outdoor mega events, and for the growing running industry. Our estimates show that the negative effect of air pollution on top runners is also sizable: a top 10 full-marathon runner will need 4.8 more minutes to finish the race if she or he were to run the 2014 Beijing Marathon compared with running on a day with average air quality in China. This suggests that professional athletes who compete in outdoor sports for awards (such as participating in the Olympic Games) should consider the negative impact of air pollution on their performance ([Florida-James, Donaldson, & Stone, 2011](#); [Lippi, Guidi, & Maffulli, 2008](#)).

Many city governments organize various mega events, such as the Olympic Games, world or nationwide exhibitions, sporting events, or music concerts, to promote media exposure and urban development ([Andranovich, Burbank, & Heying, 2001](#)). Since air pollution has significant, negative effects on the short-run health and productivity of people, city governments need to consider the costs and benefits of hosting outdoor mega events on polluted days. A lesson can be learned from the 34th Beijing International Marathon held on October 19, 2014. The average air pollution index was 289 on the race day and 320 during the race hours, which is considered heavily polluted and healthy people should avoid outdoor activities.⁴ However, the organizer did not reschedule the race. Many runners quit and many of the remaining 30,000 runners ran the race donning all kinds of facemasks.⁵ Our empirical evidence reminds hosting cities as well as participants of outdoor mega events that air quality needs to be taken into account.

Our findings are also informative for the growing running industry. The number of runners in China is estimated to be about 10 million, including runners running outdoors and in gyms (<http://sports.sina.com.cn/run/2016-06-08/doc-afxsvex3635108.shtml> [in Chinese])

PDF

Help

More
Privacy

and more cities rush to organize running races including marathon races. The running industry, including producers and retailers of running gears, running clubs, and race organizations, is growing rapidly

(<http://www.nielsen.com/cn/en/insights/news/2016/business-opportunity-looms-as-marathon-mania-sweeps-across-china.html>). Our study suggests that the industry

stakeholders need to consider the negative effect of air pollution on runners and the ripple effects on event management and sales of running products.

The rest of the article is organized as follows: Section 2 describes the data, Section 3 specifies the econometric models and discusses the causal identification issues, Section 4 reports the results, and Section 5 concludes.

Data

Our data for marathon runners and races are downloaded from www.runchina.org.cn, which is maintained by the Chinese Athletic Association (CAA). This website publishes finish time data for each runner, for all the full-marathon and half-marathon races hosted in China since 2014. The CAA certifies the running routes of these races. We have collected the 2014 and 2015 data. The individual-level data include runner name, gender, age-group, the name of the race, and the net time (the difference between the time of crossing the finish line and the time of leaving the start line). There are 37 cities and 56 race events. Some race events organize both half- and full-marathon races, while others organize only half- or full-marathon races. There are 90 races in total, including 47 half- and 43 full-marathon races. Each city hosted only one race event each year during the sample period except for Hangzhou, which hosted two race events in 2015 (on November 1 and November 29, respectively). In our sample, 19 cities hosted one race event, 17 cities hosted two, and 1 hosted three. [Figure 1](#) maps all the cities in our sample.

PDF

Help

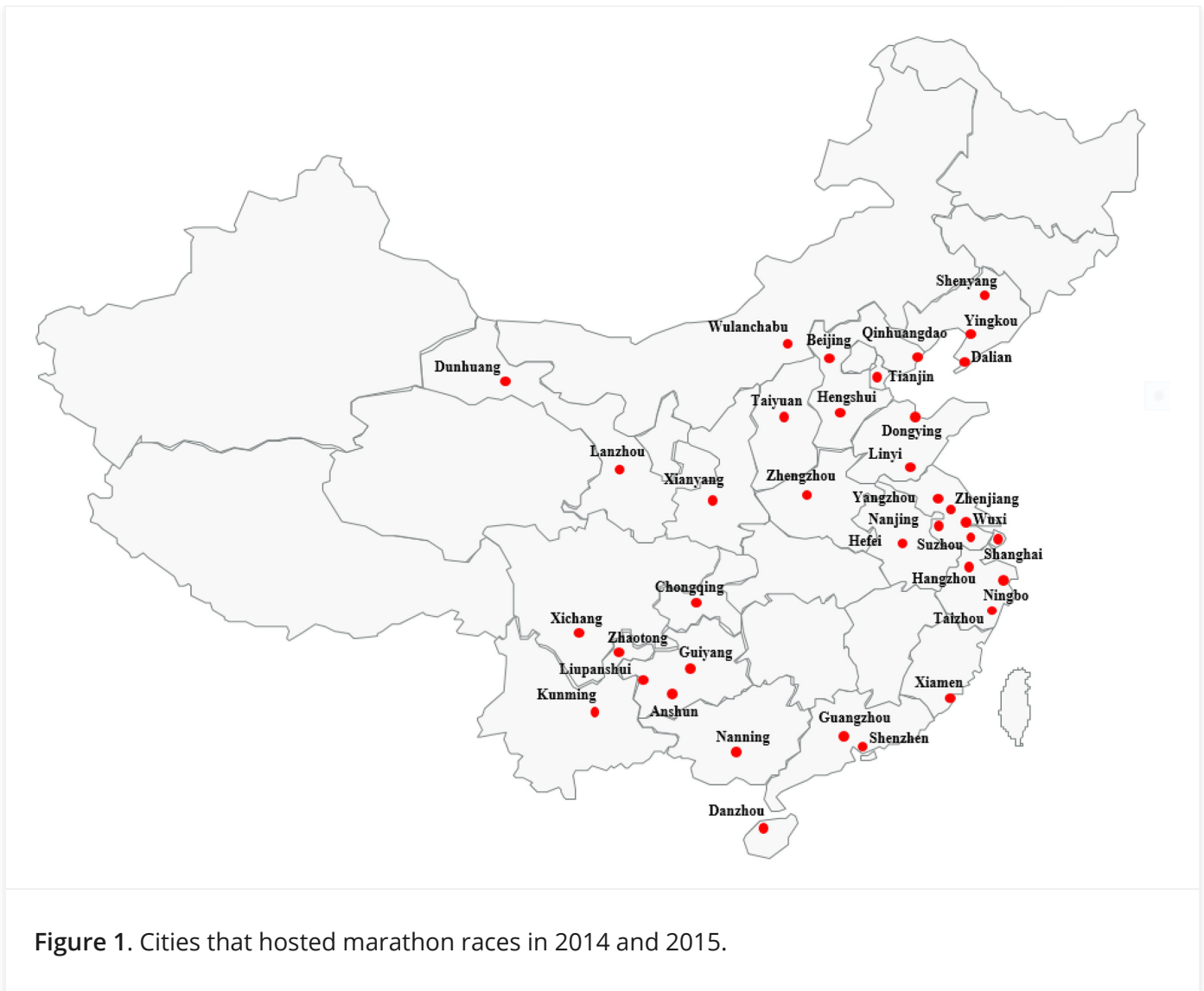


Figure 1. Cities that hosted marathon races in 2014 and 2015.

The daily air quality index (AQI) data at the city level are downloaded from the website of the Ministry of Environmental Protection of China (<http://datacenter.mep.gov.cn/>). The daily AQI for a city is the maximum of the six pollutant indexes based on hourly data from multiple monitoring stations in that city. These six pollutants are PM_{2.5}, PM₁₀, SO₂, NO₂, CO, and O₃. From the same website, we have also obtained the hourly data for the concentrations of each of these six pollutants for 46 race events.

The AQI ranges between 0 and 500 and a larger value means worse air quality. A day with the AQI below 100 is considered a “blue sky day” and has no health implications on healthy people (but sensitive people will be affected when the AQI is between 50 and 100). An AQI above 100 has progressively negative effects on health (see Table A1 in the Online Appendix).

PDF

Help

The daily weather condition data are drawn from the Global Weather Database provided by Bloomberg. We select four variables that most likely affect a runner's performance on the race day: precipitation (in centimeters), average temperature (in Celsius), average wind speed (kilometers per hour), and relative humidity (in percentage). For example, conditioning on air quality, wind speed can affect the runners' movement positively at their backs and negatively into their faces. The daily weather variables also likely correlate with the daily air quality (e.g., strong wind may blow pollutants away from a city); omitting the weather variables will bias the estimates of air pollution effects.

Our final sample includes 314,341 domestic runners. [Table 1](#) reports the summary statistics for the key variables. For full-marathon runners, the variations in finish time are large, ranging between 8,301 s (2 hr 18 min and 21 s, 2:18:21 for short) and 24,337 s (6:45:37) with a mean of 16,581 s (4:36:21) and a standard deviation of 2,701 s (0:45:01). This is consistent with the distribution of world marathon races documented in [Allen, Dechow, Pope, and Wu \(2017\)](#) with a mean of 4:26:33 and a standard deviation of 0:59:11 based on a sample of about 10 million runners. A similar pattern holds for half-marathon runners' finish time. About 19% of runners are females and 50% of runners are young people (aged between 18 and 34).

Table 1. Summary Statistics.

Variable	Mean	Standard Deviation	Minimum	Maximum	Sample Size
Panel 1: Runner characteristics					
Finish time (seconds)	12,775	4,732	3,941	24,337	314,341
Finish time for full-marathon runners (seconds)	16,581	2,701	8,301	24,337	172,523
Finish time for half-marathon runners (seconds)	8,147	1,311	3,941	20,712	141,818
Full-marathon runner (dummy)	0.55	0.50	0	1	314,341
Female (dummy)	0.19	0.39	0	1	314,341

PDF
Help

Privacy

Variable	Mean	Standard Deviation	Minimum	Maximum	Sample Size
Age 18–34 (dummy)	0.50	0.50	0	1	314,341
Age 35–39 (dummy)	0.15	0.36	0	1	314,341
Age 40–44 (dummy)	0.15	0.36	0	1	314,341
Age 45–49 (dummy)	0.10	0.30	0	1	314,341
Age 50–54 (dummy)	0.05	0.22	0	1	314,341
Age 55 or above (dummy)	0.05	0.23	0	1	314,341
Panel 2: Air quality and weather condition on the race day					
Air quality index	102.19	58.86	28.00	289.00	314,341
Precipitation (cm)	0.148	0.33	0	2.2	314,341
Temperature (Celsius)	16.53	4.05	6	25	314,341
Wind speed (kilometers per hour)	11.36	8.73	3.52	68.04	314,341
Relative humidity (%)	70.53	18.29	5.56	97.47	314,341
Panel 3: Average air quality and pollutant concentration (ug/m ³) during race hours					
Air quality index	101.05	66.59	15.22	319.57	314,341
PM _{2.5}	73.89	60.89	5.89	268.57	271,296
PM ₁₀	104.79	77.98	12.78	347.29	271,296
SO ₂	21.83	15.57	6.10	94.89	271,296
NO ₂	45.34	24.70	10.33	116.44	271,296
O ₃	54.42	31.75	10.40	200.00	271,296
CO	1.20	0.61	0.36	2.67	271,296

The AQI also shows a large variation across cities and days, ranging between 28 and 289 with a standard deviation of 59. The average AQI during race hours has an even larger v

Privacy

ranging between 15 and 320 with a standard deviation of 67. These large variations in air quality across races help estimate the pollution effect precisely.

Some medical studies fail to find a correlation between pollutants and marathon runners' performance in the United States and some European countries because the concentrations of pollutants on race days rarely exceed the health limits set by the U.S. EPA or the World Health Organization ([Helou et al., 2012](#); [Marr & Ely, 2010](#)). However, Panel 3 of [Table 1](#) shows that the pollutant concentrations in Chinese cities in general far exceed the health limits. For example, the standard set by the WHO for PM_{2.5}, PM₁₀, and SO₂ concentrations is 25, 50, and 20 ug/m³ for the 24 hr mean; however, their means during the race hours in our sample are about 74, 105, and 22 ug/m³, respectively, suggesting harmful effects on runners.⁶

Model Specification and Causal Identification

To estimate the effect of air quality on marathon runners' performance, we specify the following baseline cross-sectional model:

$$\ln(\text{Finishtime}_{ijt}) = \alpha_j + \beta_1 \ln(\text{AQI}_{jt}) + \beta_2 W_{jt} + \beta_3 X_i + \epsilon_{ijt}, \quad (1)$$

where the dependent variable is the logarithm of Finishtime_{ijt} referring to the net finish time (in seconds) of runner i who ran a race in city j on day t . α_j denotes city fixed effect. AQI_{jt} is the AQI on the race day in a city hosting the race. W_{jt} is a vector of weather condition variables including temperature, wind speed, relative humidity, and precipitation.⁷ X_i is a vector of runner's demographic variables including a dummy variable indicating female and five dummy variables for five age-group categories: aged 35–39, 40–44, 45–59, 50–54, and 60 or above; the default age-group is 18–34. β_1 , β_2 , and β_3 are the coefficient vectors to be estimated and ϵ_{ijt} is the error term.

Since we have 2 years' data, we also include a dummy variable indicating year 2015. Ideally, we would also like to control for seasonal effects by including 11 monthly dummies, but most races concentrate in a few months with moderate temperature and there is no race in February and only one race in July; therefore, we include 5 bimonthly dummies.⁸ We also include a dummy indicating whether a runner finished a full marathon or a half marathon. To match with the available daily weather data, we use daily AQI in our baseline models and also use average AQI during race hours as robustness checks.

To identify the causal effect of air quality on a marathon runner's finish time, we rely mainly on the exogeneity of air quality on the race day. In general, runners are required to register for a race a few months in advance. For example, the Beijing Marathon requires registration 2 months in advance; the Wuhan Marathon, 3 months. While a runner can anticipate the average air quality of a city in a particular season or month, it is unlikely to predict precisely the air quality on the race day. This implies that air quality on the race day can be treated as random and exogenous to runners. Note that predictable average air quality of a city in a particular season is controlled for by city fixed effects and bimonthly dummies. Therefore, the coefficient β_1 can be interpreted as the causal effect of air pollution on runners' finish time.

This causal identification strategy has been implemented in the environmental economics literature. For example, [Lichter et al. \(2017\)](#) estimate the effect of PM₁₀ concentration on professional soccer players' performance in Germany using the exogeneity of match scheduling as the identification—the scheduling is controlled by the German Football League and beyond the control of teams and players, implying that air quality on the match day is exogenous to players. [Lavy, Ebenstein, and Roth \(2014\)](#) estimate the negative effect of air pollution during exam periods on Israeli students' test scores. [Park \(2016\)](#) estimates the negative effect of high temperature during exam periods on New York students' test scores. Our research design complements these studies. There are a few other identification issues worth discussion.

First, each certified marathon route is different in terms of geographic features such as altitude, surface, flatness, curvature, and landscape along the course. Since these characteristics hardly change over time, they are subsumed into city fixed effects.

Second, some runners may choose a particular city or a particular season to run a race based on their preferences or other unobserved characteristics. This concern is also taken care of by the inclusion of city fixed effects and bimonthly dummies.

Third, it is possible that there are other unobserved personal characteristics that correlate with air quality on the race day, biasing our estimate of the key coefficient β_1 . For example, some runners may have spent more time training themselves, which helps them better adapt to air pollution; some runners may simply have different genes that affect their performance on a polluted day; some runners may have different reference points in finish time which may provide different psychological incentives ([Allen, Dechow, Pope, & Wu, 2017](#)). We can address this issue by constructing individual panel data and including

PDF

Help

Privacy

fixed effects in the model. Specifically, we drop runners with the same name, gender, and age-group showing up in the same race because these must be different persons. Then, we treat runners with the same name, gender, and age-group as the same person. This generates a runner panel data set and we reestimate Model (1) by including runner fixed effects and cluster the standard errors at both the runner level and the city level.

Fourth, runners in the same race may be affected by event-specific factors. For example, some races are better organized or invite top runners, generating stronger peer effects ([Aral & Nicolaidis, 2017](#)). This implies that finish time of runners in the same race may be correlated. We cluster the standard errors at the race level.

There is, however, one issue we cannot address. Some runners may quit the race (or quit during the race) when they know the air quality on the race day is bad. This “avoidance behavior” creates a sample selection problem.⁹ Unfortunately, we cannot access the registration data; therefore, we cannot gauge the sample selection bias using methods such as Heckman’s two-step consistent estimator. The quitters are likely to be a mixture of both fast and slow runners, so the sample selection bias is very likely to be small. In Robustness Checks section, we provide an indirect test to support this. Furthermore, our individual runner panel data model does not suffer sample selection bias since we compare the effects on the same runner across races.

Runners may exert more or less effort deliberately during a race when they know that the air quality is bad. This endogenous behavioral adjustment may bias our estimates either downward or upward ([Graff Zivin & Neidell, 2013](#)). We argue that either case is unlikely for marathon running. If runners try to slow down hoping to breathe in less pollutants, they will take a longer time to finish and will be exposed to pollution longer; in addition, a longer time will lower their rank, damaging their financial awards or pride. If runners try to speed up to finish the race earlier, they will inhale more pollutants due to intense lung functioning and most probably will not be able to sustain the accelerated pace—after all, a full marathon has 42.195 km! More importantly, if they could have run faster, why didn’t they do so?

Results

Cross-Sectional Results

[Table 2](#) presents the results of estimating Model (1) using the full sample. All columns include city fixed effects and bimonthly dummies and the standard errors are clustered at

the race level. Column 1 excludes the weather condition and demographic variables and the estimated pollution effect on finish time is significantly positive with an elasticity of .0273. Column 2 adds weather variables and the coefficient of $\ln(\text{AQI})$ becomes larger, .0408, and more significant. In terms of the effects of weather conditions on runners' performance, high temperature slows down running speed, as commonly found in many marathon studies because high temperature increases heat stress and may lead to hyperthermia, dehydration, and loss of electrolytes ([Helou et al., 2012](#); [Maughan, 2010](#); [Spellman, 1996](#); [Vihma, 2010](#)). More precipitation (raining) is also negatively associated with running speed. Relative humidity helps increase speed, consistent with the finding in [Vihma \(2010\)](#) and [Helou et al. \(2012\)](#).¹⁰ Although column 2 of [Table 2](#) shows that high wind speed tends to increase running speed, this result should be interpreted with caution since wind direction may change over the course of the race and time and we do not have the wind direction data.¹¹ Weather conditions may have nonlinear effects on runners' performance. As a robustness check, we also add the quadratic terms of the weather variables and report the results in column 2 of Table A3 in the Online Appendix. The coefficient of $\ln(\text{AQI})$ is very close (.0471) and remains statistically significant. Although the coefficient of temperature is not significant and the coefficient of temperature squared is significant only at the 10% level, they are jointly significant at the 1% level. The implied optimal temperature for runners is 10 °C, which is consistent with the findings in the literature that the optimal temperature for marathon runners is generally between 10 and 12 °C ([Helou et al., 2012](#); [Maughan, 2010](#)). Both the coefficient of the linear term (positive) and the coefficient of the quadratic term (negative) of wind speed are statistically significant and the implied "worst" wind speed is 17 km/hr. This means that wind speed either below or above 17 km/hr helps increase speed. We have no good explanation for this result because the effect of wind speed can be very complicated: Head wind (wind blowing toward the runner) increases resistance and therefore reduces speed but promotes cooling; tail wind (wind following the runner) can propel a runner forward but exacerbate the cooling problem which may lead to hypothermia ([Davies, 1980](#); [Pugh, 1971](#); [Spellman, 1996](#)). In addition, wind speed and direction may change over the race course during race hours. Therefore, the effect of wind speed is rather indeterminate. The linear and quadratic terms of precipitation and humidity variables are insignificant. Since our key estimate of pollution effect is robust to linear and quadratic weather controls, in the rest of analysis, we employ the models with only linear weather controls.

PDF

Help

Table 2. Full Sample Results.

Variable	1	2	3	4
ln(Air quality index)	.0273** (.0114)	.0408*** (.0030)	.0408*** (.0026)	
Air quality index				2.7311*** (0.4476)
Full-marathon dummy	.6731*** (.0057)	.6731*** (.0056)	.6944*** (.0052)	8,216.7570*** (90.2228)
Year 2015 dummy	.0098* (.0051)	.0202*** (.0026)	.0171*** (.0023)	247.6191*** (49.6329)
Precipitation		.0282*** (.0074)	.0280*** (.0066)	134.2860 (81.8291)
Temperature		.0049*** (.0006)	.0047*** (.0006)	85.9956*** (10.0497)
Wind speed		-.0007*** (.0002)	-.0007*** (.0002)	-3.5602 (4.3090)
Relative humidity		-.0008*** (.0002)	-.0008*** (.0019)	-2.6216 (3.3811)
Female			.0919*** (.0030)	988.7489*** (36.0599)
Age 35–39			-.0257*** (.0034)	-383.6603*** (57.8257)
Age 40–44			-.0453*** (.0041)	-663.6623*** (76.2672)
Age 45–49			-.0516*** (.0049)	-758.7814*** (90.5278)
Age 50–54			-.0489*** (.0054)	-727.1522*** (96.0032)
Age 55 or above			-.0309*** (.0050)	-473.3934*** (82.2286)
Adjusted R^2	.8299	.8306	.8417	.8107

Note. The dependent variable for columns 1–3 is ln(Finish time). The dependent variable for column 4 is finish time. All models also include city fixed effects and bimonthly dummies. Standard error

PDF

Help

Privacy

clustered at the race level and reported in the parentheses. Sample size: 314,341.

"*" indicates significance at the 10%. "***" indicates significance at the 5% level. "****" indicates significance at the 1% level.

Column 3 further adds gender and age-group dummies and the estimated pollution effect is identical to column 2. This suggests that air pollution on the race day is orthogonal to observed individual characteristics, implying that the correlation between air pollution and unobserved individual characteristics is likely to be very small too (Oster, 2016). This is our preferred specification since we have included all the possible control variables in our data.

Female runners on average take 9.63% more time to finish a marathon race.¹² Compared with young runners aged between 18 and 34, older runners run faster.¹³ This is somewhat surprising and we do not have a good explanation for this. One possible interpretation could be due to sample selection that older runners on average are more highly motivated, spend more time on training, and have gained more running experience. However, this pattern does not hold for top runners and we will provide the evidence in the next subsection.

Column 3 also shows that a 1% increase in AQI causes a .0408% increase in the finish time of a marathon runner. Put a different way, doubling the AQI increases finish time by 2.8%.¹⁴ Evaluating at the mean AQI of 102 and mean finish time of 16,581 s (4:36:21) for a full marathon, a 10% increase in AQI will increase finish time by about 1.1 min. This effect seems small but actually not. Suppose a runner takes 16,581 s to finish a full marathon on a day with average air quality (AQI is 102), this runner will need 12 more minutes to finish the 2014 Beijing Marathon during which the AQI is 289. For the best full-marathon runner (a young male runner) in our sample, the finish time is 8,301 s (2:18:21) and the AQI on the race (March 15, 2015) in that city (Wuxi) is 105. If he were to run the Beijing Marathon in 2014, a back-of-the-envelope calculation suggests that he would need 5.8 more minutes to cross the finish line. Including the quadratic terms of the weather variables generates very similar results (column 3 of Table A3).

Column 4 replicates the model in column 3 but assumes a linear relationship between the AQI level and finish time. A one-unit increase in AQI increases finish time by about 2.7 s. Evaluating at the sample mean (102 for AQI, 12,775 s for finish time) for column 4, the implied elasticity is .0218. A one standard deviation increase in air pollution (59) slows down a runner's finish time by 2.7 min. A runner who can finish a full marathon on a day

PDF

Help

Privacy

average AQI of 102 will need about 8.5 more minutes to finish the Beijing Marathon in 2014, which is in the ballpark compared with the log-log models.

Results for Full-Marathon and Half-Marathon Samples

We also estimate Model (1) for the full-marathon and half-marathon subsamples. [Table 3](#) reports the results based on the full-marathon runner sample. Column 1 uses the full sample of full-marathon runners and the estimated pollution elasticity of finish time is .0274, one third smaller than the baseline estimate of .0408 but still sizable. This elasticity is slightly larger for male runners (.0290) and young runners (.0352) and moderately smaller for female runners (.0163) and old runners (.0214). We do not have a clear answer yet why the pollution elasticity varies across demographic traits.¹⁵ However, the positive, significant estimates of the pollution elasticity of finish time confirm that overall, air pollution negatively affects runners' performance.

Table 3. Results for Full-Marathon Runners.

Variable	1	2	3	4	5	6
	Full Sample	Male	Female	Young	Old	Top 10
ln(Air quality index)	.0274*** (.0034)	.0290*** (.0032)	.0163*** (.0051)	.0352*** (.0053)	.0214*** (.0037)	.0228* (.0122)
Age 35-39	-.0376*** (.0044)	-.0377*** (.0047)	-.0363*** (.0035)			.0231*** (.0031)
Age 40-44	-.0603*** (.0046)	-.0609*** (.0050)	-.0553*** (.0034)			.0331*** (.0058)
Age 45-49	-.0690*** (.0052)	-.0703*** (.0058)	-.0601*** (.0040)			.0322*** (.0077)
Age 50-54	-.0641*** (.0057)	-.0648*** (.0063)	-.0588*** (.0037)			.0331*** (.010)
Age 55 or above	-.0386*** (.0058)	-.0389*** (.0061)	-.0357*** (.0049)			.0747*** (.0167)
Female dummy	Yes	No	No	Yes	Yes	Yes

Variable	1	2	3	4	5	6
	Full Sample	Male	Female	Young	Old	Top 10
Adjusted R^2	.1188	.1000	.0985	.1042	.0748	.7961
Sample size	172,523	149,630	22,893	75,538	96,985	837

Note. The dependent variable is $\ln(\text{Finish time})$. All models also include weather condition variables on the race day, year 2015 dummy, bimonthly dummies, and city fixed effects. Standard errors are clustered at the race level and reported in the parentheses. There are 43 full-marathon races; one race has missing data for top runners.

“*” indicates significance at the 10% level. “**” indicates significance at the 5% level. “***” indicates significance at the 1% level.

Top runners are generally professional athletes competing for cash awards. Column 6 shows that for the top 10 runners in each race, the pollution elasticity is slightly smaller, .0228, but still statistically significant at the 10% level even with a much smaller sample. Column 7 includes the top 20 runners in each race and the estimated pollution elasticity is .0286 and significant at the 1% level, confirming that even for well-trained professional athletes, air pollution also exerts a negative effect on their performance. Column 8 includes the top 30 runners and the effect is almost identical.

For the subsample of top 10 full-marathon runners, the mean finish time is 11,220 s (3:10) and the mean AQI is 95. If a top 10 runner were to run the Beijing Marathon in 2014, the runner would need 4.8 more minutes to finish the race. For top runners, 4.8 min can significantly decrease the probability of qualifying for other international races or breaking the world record. The economic cost of even 1 min is also huge for top runners. The Beijing Marathon offers the top eight runners cash awards, based on both their ranks and a finish time threshold. Even if air pollution affects all runners equally and the relative ranks stay the same, top runners may have to take a time longer than the threshold to cross the finish line. The best male runner whose finish time is also less than 2 hr 9 min would be awarded US\$40,000; the second best, US\$20,000 if his finish time is less than 2 hr 10 min.¹⁶ For top runners, the opportunity cost of 1 min is US\$20,000!

PDF

Help

Privacy

[Table 3](#) also presents the age effects across subsamples. Although for the full sample of full-marathon runners (column 1), as well as for the male and female full-marathon runners (columns 2–3), old runners generally run faster than do young runners, this pattern is reversed for top runners. Columns 6–8 show that among top 30 runners, finish time generally increases with age, consistent with the findings in the sports medicine and physiology literature ([Lara, Salinero, & Del Coso, 2014](#); [Trappe, 2007](#); [Zavorsky, Tomko, & Smoliga, 2017](#)). In general, elite marathon runners achieve their best performance between age 25 and 35 and their performance declines with aging as the cardiovascular capacity declines with aging ([Trappe, 2007](#); [Zavorsky et al., 2017](#)).

[Table 4](#) reports the estimate results for half-marathon runners, parallel to the columns in [Table 3](#). The overall pattern is very similar except that the pollution elasticity is uniformly larger (even for top runners), ranging between .0263 and .0519, possibly because half-marathon runners run more aggressively and are affected more by air pollution. We also identify the same age-related decline in performance for top half-marathon runners (the results are not reported but are available upon request).

Table 4. Results for Half-Marathon Runners.

Variable	1	2	3	4	5	6
	Full Sample	Male	Female	Young	Old	Top 10
ln(Air quality index)	.0494*** (.0105)	.0518*** (.0099)	.0422*** (.0114)	.0519*** (.0111)	.0455*** (.0116)	.0263** (.0110)
Female dummy	Yes	No	No	Yes	Yes	Yes
Age categories	Yes	Yes	Yes	No	No	Yes
Adjusted R^2	.1453	.0693	.0738	.1396	.1402	.7379
Sample size	141,818	105,362	36,456	79,975	61,843	940

PDF

Help

Privacy

Note. The dependent variable is $\ln(\text{Finish time})$. All models also include weather condition variables on the race day, year 2015 dummy, bimonthly dummies, and city fixed effects. Standard errors are clustered at the race level and reported in the parentheses. There are 47 half-marathon races.

"*" indicates significance at the 10% level. "**" indicates significance at the 5% level. "***" indicates significance at the 1% level.

Nonlinear Effect

Different degrees of air pollution severity may have different impacts on runners' performance. This nonlinear effect is identified in [Table 5](#). Column 1 replicates the baseline results (same as column 3 of [Table 2](#)). Column 2 replaces $\ln(\text{AQI})$ by a dummy variable indicating whether the AQI is above 100 or not. Compared with a blue sky day ($\text{AQI} \leq 100$), a non-blue sky day increases a runner's finish time by 3.73%. Column 3 uses two dummy variables: AQI between 100 and 200 (slightly and moderately polluted) and AQI above 200 (severely or heavily polluted). Compared with a blue sky day, running on a slightly or moderately polluted day increases finish time by 2.94%; running on a severely or heavily polluted day increases finish time further by about 3 times—a 7.95% increase. This shows that worse air quality imposes a progressively negative impact on runners, consistent with the finding in [Lichter et al. \(2017\)](#) that a higher concentration of PM_{10} exceeding the European Union limit has a stronger effect on soccer players' performance.¹⁷

Table 5. Nonlinear Effect of Air Pollution.

Variable	1	2	3	4
$\ln(\text{Air quality index})$.0408*** (.0026)			
Air quality index > 100 (dummy)		.0367*** (.0088)		
$100 < \text{Air quality index} \leq 200$ (dummy)			.0290*** (.0065)	
$50 < \text{Air quality index} \leq 100$ (dummy)				.0437* (.0261)

PDF

Help

Privacy

Variable	1	2	3	4
100 < Air quality index ≤ 150 (dummy)				.0617*** (.0176)
150 < Air quality index ≤ 200 (dummy)				.0569*** (.0172)
Air quality index > 200 (dummy)			.0765*** (.0099)	.0972*** (.0151)
Adjusted R^2	.8417	.8415	.8416	.8416

Note. The dependent variable is $\ln(\text{Finish time})$. All models also include weather condition variables on the race day, female dummy, age category dummies, year 2015 dummy, bimonthly dummies, and city fixed effects. Standard errors are clustered at the race level and reported in the parentheses. Sample size: 314,341.

"*" indicates significance at the 10% level. "**" indicates significance at the 5% level. "***" indicates significance at the 1% level.

The nonlinear effect is also confirmed in column 4 where four dummy variables are used. Compared with excellent air quality (AQI below 50, the default category), even good air quality (AQI between 50 and 100) has a significantly negative effect on runners (4.47%); slight pollution (AQI between 100 and 150) increases finish time by 6.36%; moderate pollution (AQI between 150 and 200) increases finish time by 5.86%; and severe or worse pollution (AQI above 200) increases finish time by 10.21%.

The nonlinear pollution effect, in terms of demographic heterogeneity, can also be found across runners. We use quantile regressions to estimate the top and bottom 10 percent finish time conditional on the same set of independent variables as in column 3 of [Table 4](#). Table A4 in the Online Appendix presents the quantile regression results for full- and half-marathon runners. Note that "q10" in column 1 denotes the runners with finish times below the bottom 10 percentile, meaning that they are the top 10% (fast) runners. For full-marathon runners, air pollution has a 3 times larger effect on top 10% (fast) runners than on bottom 10% (slow) runners, regardless of gender. It is interesting to see that in Panel 2 for half-marathon runners, the pattern is reversed: Air pollution has a larger effect on bottom 10% (slow) runners; we have no good explanation for why the half-marathon results are the opposite of the full-marathon results.

PDF

Help

Runner Fixed-Effect Model Results

Since unobserved individual runners' characteristics may bias our estimated pollution elasticity, we estimate panel data models with runner fixed effects. To create a sample of repeated runners, we first drop the runners with the same name, gender, and age-group in the same full- or half-marathon race. This amounts to 3.58% of our total observations (11,264 of the 314,341). We then treat the remaining runners with the same name, gender, age-group, and always running the full (or half) marathon in different races as the same runner.¹⁸ The results of panel data models with runner fixed effects are reported in [Table 6](#). Since error terms may be correlated within a repeated runner and within a city, we cluster standard errors at both the individual runner and the city levels ([Cameron & Miller, 2015](#)).¹⁹

Table 6. Runner Fixed-Effect Model Results.

Variable	1	2	3	4	5	6
	2 Races	2-3 Races	2-4 Races	≥2 Races	≥2 Races, Male	≥2 Races, Female
ln(Air quality index)	.0280*** (.0030)	.0283*** (.0028)	.0287*** (.0022)	.0269*** (.0018)	.0274*** (.0021)	.0230*** (.0030)
Adjusted R^2	.9395	.9346	.9298	.9167	.9121	.9384
Sample size	65,726	100,073	121,969	169,750	140,913	28,837

Note. The dependent variable is ln(Finish time). All models also include weather condition variables on the race day, year 2015 dummy, bimonthly dummies, city fixed effects, and individual runner fixed effects. Standard errors are clustered two-way at the individual runner and the city levels and reported in the parentheses.

"*" indicates significance at the 10% level. "**" indicates significance at the 5% level. "***" indicates significance at the 1% level.

PDF

Help

Column 1 of [Table 6](#) uses a sample of runners who have run just two full-marathon races or two half-marathon races; in this case, the dummy variable indicating full marathon is subsumed into the runner fixed effect. For these runners, the pollution elasticity of finish time is .028. Column 2 uses a sample of runners who have run two or three full-marathon or half-marathon races. The estimated pollution elasticity is almost identical and significant at the 1% level. Column 3 expands the sample to include the runners who have also finished four races, and the estimates are very close. Column 4 includes all repeated runners and the coefficient of $\ln(\text{AQI})$ is still significant at the 1% level although the magnitude is slightly smaller (0.027).²⁰ Columns 5–8 show that the negative impact is mainly on male runners and young runners, consistent with [Tables 3](#) and [4](#).

These estimates are smaller than the baseline results (.0408; column 3 of [Table 2](#)), possibly because these runners have better training or better experience. Since the runner fixed-effects control for all the unobserved, time-unvarying, individual-specific characteristics that may affect finish time, we consider these within-runner estimates the “causal” effect and also the lower bound of the pollution effect on runners' performance. These effects are still economically sizable. Taking column 4 as the example, evaluating at the mean finish time of 16,089 s (4:28:09) for full-marathon runners and mean AQI of 111 in this sample, if a runner were to run the Beijing Marathon in 2014, he or she will need 7 more minutes to cross the finish line.

Robustness Checks

The above analysis uses the daily AQI to match the available daily weather condition variables. Since runners are exposed to the outdoor pollution mainly during the race hours, it is important to check whether average air pollution during the race hours causes similar negative effects on runners. In our sample, all races are scheduled on a weekend day and start as early as 7 a.m. and close as late as 2:30 p.m. Therefore, we use the average of hourly AQI from 6 a.m. to 3 p.m. to replace the daily AQI and reestimate all models.²¹ It turns out that the negative effects of air pollution during the race hours are very similar to the estimates using the daily AQI, albeit slightly smaller. The main reason is that the correlation between the average AQI on the race day and the average AQI during the race hours is very high: The correlation coefficient is .974 and statistically significant at the 1% level. As a demonstration, we present the full sample results in the Online Appendix Table A5, which is parallel to [Table 2](#).

PDF

Help

Taking our preferred specification as the example, column 3 of Table A5 shows an elasticity of .0262 using the average AQI during the race hours. This is smaller than .0408 using the average AQI of the race day but estimated more precisely with a smaller standard error. When the finish time is used as the dependent variable, the coefficient of average AQI during the race hours is 2.0234, close to 2.731 estimated using the average AQI of a race day. It is worth noting that the coefficients of demographic variables and full-marathon dummy are almost identical to those in [Table 2](#) but are estimated more precisely. This further suggests that air pollution on the race day is orthogonal to observed individual characteristics and choice and is also unlikely to be correlated with unobserved individual characteristics.

Many medical studies find a correlation between different types of pollutants and athletes' health and sport performance, using laboratory or field survey data ([Carlisle & Sharp, 2001](#); [Chimenti et al., 2009](#)). To identify the effects of different types of pollutants, we reestimate the column 3 model in Table A5 by replacing average AQI during the race hours with the average concentration of one of the six pollutants during the race hours: PM_{2.5}, PM₁₀, SO₂, NO₂, CO, and O₃.²² The results are presented in [Table 7](#). Except for O₃, the negative effect of each pollutant concentration is similar in magnitude: The elasticity ranges between .0265 and .0501, comparable to the elasticity of .0262 using the AQI. The coefficient of the level of ozone concentration is $-.0218$ and is statistically significant, which is counterintuitive. This is most likely because, different from other pollutants, ozone's adverse health effects occur from short-term exposure to only high levels and its concentrations are affected by NO_x, volatile organic compound emissions, and temperature, in a complex photochemistry process ([Deschênes, Greenstone, & Shapiro, 2017](#); [Lippman, 2009](#)). We therefore use a different model specification for ozone concentration; specifically, we create a dummy variable set to one if the ozone concentration is above 70 ug/m³ (EPA, 2015 standard) and estimate the baseline model in column 7 of [Table 7](#).²³ The result suggests that a high-level concentration of ozone increases runners' finish time by about 6.17%.

Table 7. Effects of Pollutant Concentration During a Race.

Variable	1	2	3	4	5	6
ln(Air quality index)	.0262*** (.0015)					

PDF

Help

Privacy

Variable	1	2	3	4	5	6
ln(PM _{2.5})		.0270*** (.0018)				
ln(PM ₁₀)			.0265*** (.0025)			
ln(SO ₂)				.0459*** (.0044)		
ln(NO ₂)					.0501*** (.0072)	
ln(CO)						.0410*** (.0059)
O ₃ above 70 ug/m ³						
Sample size	314,341	271,296	271,296	271,296	271,296	271,296
Adjusted R ²	.8417	.8410	.8409	.8410	.8408	.8409

Note. The dependent variable is ln(Finish time). All models also include weather condition variables on the race day, female dummy, age category dummies, year 2015 dummy, bimonthly dummies, and city fixed effects. All pollutant variables and air quality index are defined as the average of hourly data between 6 a.m. and 3 p.m. on the race day. Standard errors are clustered at the race level and reported in parentheses.

"*" indicates significance at the 10% level. "**" indicates significance at the 5% level. "***" indicates significance at the 1% level.

PDF

Help

Because of high correlations between these pollutant concentrations, it is difficult to isolate the contribution of each pollutant concentration. Regardless, a horse race model including all types of pollutant concentrations shows that either PM_{2.5} or PM₁₀ concentration is still statistically significant with a larger magnitude, suggesting that particulate matter is more harmful to runners.

Our final robustness check provides an indirect test to show that sample selection is likely to be a very minor issue.²⁴ Many runners travel to attend races. Runners are likely to form their expectations on air quality on the race day based on the air quality before the race day since daily air quality is serially correlated.²⁵ If air quality is excellent on the day before a race, runners are more likely to prepare for participation; and if the air quality on the race day turns out to be bad and many runners quit, then the sample selection bias would be large. The opposite can happen too. Based on this logic, we create a dummy variable “Similar” set to one if the absolute value of the difference between the AQI on the race day and the AQI on the day before a race is less than 15.²⁶ We then interact this dummy variable “Similar” with $\ln(\text{AQI})$ or air quality level on the race day and reestimate all the models in [Table 2](#). The results are presented in Table A6 (Online Appendix): The main effects of air quality stay almost identical, as in [Table 2](#). The coefficients of the interaction terms in all columns are positive, suggesting that good runners may have withdrawn from a race to avoid running on polluted days; however, the coefficients are not statistically significant in three of the four models and the one that is significant has a very small magnitude. Taking together, these suggest that the sample selection bias is likely very small.

Conclusion

Using a sample of more than 0.3 million marathon runners of 56 race events hosted by Chinese cities in 2014 and 2015, we estimate the air pollution elasticity of finish time to be .0408. This shows that air pollution has a nonnegligible, negative effect on runners' performance. Our causal identification uses the exogeneity of air quality on the race day because runners are required to register for a race a few months in advance and the air quality on the race day can be considered random. Based on a panel data set of runners who ran more than one race, our estimates with runner fixed effects confirm the negative impact of air pollution on runner performance albeit in a smaller magnitude (with an elasticity of .028).

Our study contributes to the emerging literature on the effects of air pollution on short-run productivity. Our findings remind city governments that the negative effect of air pollution on the health and performance of participants should be taken into account when organizing outdoor mega events. For example, on heavily polluted days, if weather conditions permit, a city government may use cloud seeding to increase rainfall to wash out part of the pollutants. Our findings are also informative for professional athletes who

PDF

Help

Privacy

compete for awards in outdoor sports games such as football, running, and biking and for workers whose jobs require intensive physical activities and long exposure to ambient air pollution.

Acknowledgments

We thank the editor, two anonymous referees, Xiaojia Bao, Elena Irwin, Xiaoyan Lei, Carl Lin, Christopher Magee, Shunfeng Song, Todd Sorensen, Junjie Zhang, Xi Zhu, Dilek Uz, seminar, and conference participants at Bucknell University, Dongbei University of Finance and Economics, Shanghai University of Finance and Economics, Kunshan Duke University, Ohio State University, Peking University, Renmin University, Tsinghua University, University of Nevada-Reno, the 2016 China Economists Society Conference, the 2016 North American Regional Science Congress for very helpful comments and Zexin Ye for excellent research assistance. We acknowledge the financial support from the National Natural Science Foundation of China (Grant #71773096 and Grant #71873109).

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: National Natural Science Foundation of China (Grant #71773096 and Grant #71873109).

[PDF](#)[Help](#)

ORCID iD

Shihe Fu 

Footnotes

1. The third strand of literature is on the negative shocks to athlete human capital accumulation, see [Gong, Sun, and Wei \(2018\)](#). At this stage, we are unable to quantify the long-run effect of air pollution on runners due to data availability constraint.

[Privacy](#)

[GO TO FOOTNOTE](#)

2. For more details, refer to the Environmental Protection Agency (EPA) websites. For example, <https://www.epa.gov/pm-pollution>, <https://www.epa.gov/so2-pollution>, and <https://www.epa.gov/co-pollution>

[GO TO FOOTNOTE](#)

3. Based on the team-game day data from the German Bundesliga, [Weimar and Wicker \(2017\)](#) find that individual players' effort, measured by total distance run per player and per match, has a significant, positive effect on team performance, and [Wicker, Prinz, Weimar, Deutscher, and Upmann \(2013\)](#) find that players' effort has insignificant effect on players' market value. This undercompensation suggests a possible free-riding incentive. In addition, the summary statistics of the data in [Weimar and Wicker \(2017\)](#) shows substantial variations in effort across players, suggesting there are some "lazy" players in some games.

[GO TO FOOTNOTE](#)

4. Table A1 in the Online Appendix summarizes the health implications of the air quality index (AQI).

[GO TO FOOTNOTE](#)

5. Many news media reported the event (see, e.g., <http://www.foxnews.com/world/2014/10/19/beijing-marathon-runners-wear-masks-to-combat-smog-from-pollution.html>).

[GO TO FOOTNOTE](#)

6. The mean concentration of ozone is below the World Health Organization (WHO) limit of 100 $\mu\text{g}/\text{m}^3$ for the 8 hr mean. The WHO air quality guidelines for different types of pollutants are online at http://www.who.int/phe/health_topics/outdoorair/outdoorair_aqg/en

[GO TO FOOTNOTE](#)

7. Our main models include only linear terms of weather control variables, including the quadratic terms of weather variables generate very similar results and increase the

PDF

Help

Privacy

R^2 by only .0001.

[GO TO FOOTNOTE](#)

8. When monthly dummy variables are included, four would be dropped due to collinearity with race or year fixed effects. Therefore, we prefer bimonthly dummy variables. Using quarterly dummy variables generates very similar results.

[GO TO FOOTNOTE](#)

9. The registration fees are relatively small and nonrefundable. In general, the fees for domestic runners range between RMB100 and 200 (US\$16–32 based on the current exchange rate). Registered runners usually receive a complimentary T-shirt, which has a market value close to the registration fee. This means the financial penalty in quitting a race is almost zero.

[GO TO FOOTNOTE](#)

10. However, very high humidity, especially combined with high temperature, can increase heat stress ([Spellman, 1996](#)) and therefore reduce running speed.

[GO TO FOOTNOTE](#)

11. Although many marathon race courses in China make a single loop, some are not. For example, Xiamen and Beijing Marathon courses are close to a half circle.

[GO TO FOOTNOTE](#)

PDF

Help

12. The coefficient of female dummy variable is .0919; using the formula $\exp(.0919) = 1.09626$ shows that female runners on average take 9.63% more time. We apply the same calculation for the coefficients of dummy variables in the rest of the article.

[GO TO FOOTNOTE](#)

13. This pattern holds when we interact the age categorical variables with the full-marathon dummy variable. The coefficients of age categorical variables and interaction terms are all negative and significant except one. Overall, the absolute values of the coefficients are slightly increasing with age. This shows that old people run faster in both half- and

Privacy

marathon races. The coefficient of $\ln(\text{AQI})$ is almost identical.

[GO TO FOOTNOTE](#)

14. For $x\%$ change in AQI, the change in finish time is approximated by $.0408x\%$ if x is a small number. For a large x , we apply this formula $(1 + x\%)^{.0408} - 1$ to approximate the percent change in finish time. For example, doubling AQI (100% increase in AQI) leads to $2^{.0408} - 1 = .028$ or a 2.8% change in finish time. We thank one anonymous referee very much for

pointing out this calculation and have applied this formula in the rest of the article wherever applicable.

[GO TO FOOTNOTE](#)

15. For example, male and young runners may run more aggressively. [Lichter, Pestel, and Sommer \(2017\)](#) find that PM_{10} has a larger effect on soccer players who are midfielders and defenders because those positions require more active physical activities than strikers.

[GO TO FOOTNOTE](#)

16. See Table A2 in the Online Appendix for the award scheme for the 2014 Beijing Marathon. Other races in China follow this scheme very closely.

[GO TO FOOTNOTE](#)

17. [Lavy, Ebenstein, and Roth \(2014\)](#) find a nonlinear effect of air pollution on students' scores and [Aragón, Miranda, and Oliva \(2017\)](#) find a nonlinear effect on workers' labor supply (work hours).

[GO TO FOOTNOTE](#)

18. It is possible that in the panel data sample, some repeated runners are in fact different runners. For example, a very common name may result in a "fake" runner with many races. [Table 6](#) shows that the estimate results are very stable for subsamples of runners who ran only 2, 2-3, and 2-4 races as well as for all repeated runners. This demonstrates that the issue of duplicates is minimal.

PDF

Help

Privacy

[GO TO FOOTNOTE](#)

19. Clustering standard errors at the city level or at the individual runner level generates similar significance levels: All the coefficients of $\ln(\text{AQI})$ in [Table 6](#) are statistically significant at the 1% level.

[GO TO FOOTNOTE](#)

20. Recreational or amateur runners are generally advised to run no more than two full-marathon races each year to prevent injuries (see discussions at www.runnersworld.com).

[GO TO FOOTNOTE](#)

21. Hourly AQI data for the morning period are missing for six races so we use their daily AQI as the proxy.

[GO TO FOOTNOTE](#)

22. Pollutant concentrations during morning hours are missing for 10 races so the sample size becomes smaller.

[GO TO FOOTNOTE](#)

23. The U.S. EPA sets the ozone standard to be 70 ug/m^3 8 hr mean in 2015 (<https://www.epa.gov/ozone-pollution/2015-national-ambient-air-quality-standards-national-ozone>).

[GO TO FOOTNOTE](#)

24. We thank the referee very much for suggesting this additional test for sample selection bias.

[GO TO FOOTNOTE](#)

25. In our sample, the correlation between the AQI on the race day and the AQI on the day before a race is about .8 and significant at the 1% level, suggesting a high serial cor

PDF

Help

Privacy

[GO TO FOOTNOTE](#)

26. This ensures a relatively balanced sample split: about 40% of observations with *Similar* equal to 1.

[GO TO FOOTNOTE](#)

Notes

References

Adhvaryu A., Kala N., Nyshadham A. (2014). *Management and shocks to worker productivity: Evidence from air pollution exposure in an Indian garment factory* (Working Paper). Ann Arbor, MI: University of Michigan.

[+ SHOW REFERENCES](#)[Google Scholar](#)

Allen E., Dechow P., Pope D., Wu W. (2017). Reference-dependent preferences: Evidence from marathon runners. *Management Science*, 63, 1657–1672.

[+ SHOW REFERENCES](#)[Crossref](#)

PDF

Help

[Google Scholar](#)

Andranovich G., Burbank M., Heying C. (2001). Olympic cities: Lessons learned from mega-event politics. *Journal of Urban Affairs*, 23, 113–131.

[GO TO REFERENCE](#)

Privacy

[Crossref](#)[ISI](#)[Google Scholar](#)

Aragón F., Miranda J. J., Oliva P. (2017). Particulate matter and labor supply: The role of caregiving and non-linearities. *Journal of Environmental Economics and Management*, 86, 295–309.

[GO TO REFERENCE](#)[Show all references](#)

Biographies

Mengmeng Guo is an associate professor of economics at Southwestern University of Finance and Economics. Her research fields include environmental economics and applied econometrics.

Shihe Fu is a professor of Economics at Xiamen University. His research fields are urban economics, labor economics, and environmental economics.

Supplementary Material

Supplemental Material

Please find the following supplemental material visualised and available to downlo

[Privacy](#)

PDF

Help

Figshare in the display box below. Where there are more than one item, you can scroll through each tab to see each separate item.

Please note all supplemental material carries the same license as the article it is here associated with

- [Supplemental Material](#)

Online Appendix:

Table A1: Air quality index and its health implication

Air quality index	Air quality	Health implications for healthy people
0-50	Excellent	No health implication.
51-100	Good	Sensitive people may be slightly affected.
101-150	Slightly polluted	Slight irritation may occur.
151-200	Moderately polluted	Irritation may occur; should reduce outdoor exercises.
201-300	Severely polluted	Noticeably affected; should reduce outdoor activities.
>300	Heavily polluted	Reduced endurance in activities; should avoid outdoor activities.

Source: Ministry of Environmental Protection of China, 2012, Technical Regulation on Ambient Air Quality Index (on trial), HJ633-2012, China Environmental Science Press.

PDF

Help



figshare



Share



File (jse_supple.pdf)

DOWNLOAD

Privacy

[VIEW FULL TEXT](#)

.

[DOWNLOAD PDF](#)

Related content

Similar articles:



Available access

[When Falling Just Short is a Good Thing: The Effect of Past Performance on Improvement](#)[Show details](#) 

Available access

[Can Professional Football Players Adapt to Air Pollution? Evidence From China](#)[Show details](#) 

Available access

[The Rewards to Running: Prize Structure and Performance in Professional Road Racing](#)[Show details](#)

PDF

Help

Privacy

[View more](#)

SAGE recommends:

SAGE Knowledge

Book chapter

[Air Pollution: History of Actions and Effectiveness of Change](#)

Show details 

SAGE Knowledge

Entry

[Ambient Air Quality](#)

Show details 

SAGE Knowledge

Entry

[Air Pollution](#)

Show details 

[View more](#)

PDF

Help

Also from SAGE Publishing

CQ Library

American political resources

Data Planet

A universe of data

Lean Library

SAGE Business Cases

Privacy

Increase the visibility of your library

Real-world cases at your fingertips

SAGE Campus

Online skills and methods courses

SAGE Knowledge

The ultimate social science library

SAGE Research Methods

The ultimate methods library

SAGE Video

Streaming video collections

PDF

Help