SUPPORTING INFORMATION

The Impact of Exposure to Air Pollution on Cognitive Performance

PART 1: Description of Data

A. Cognitive Test Scores Data

We utilize cognitive test scores from the China Family Panel Studies (CFPS), a nationally representative survey of Chinese families and individuals conducted in 2010 and 2014. The CFPS is funded by Peking University and carried out by the university's Institute of Social Science Survey (1). The survey uses multistage probability proportional to size sampling with implicit stratification to better represent Chinese society. The 2010 CFPS baseline sample is drawn through three stages (i.e. county, village, and household) from 25 provinces. The 162 randomly chosen counties largely represent Chinese society (2). The CFPS includes questions on a wide range of topics for families and individuals, including their economic activities, education outcomes, family dynamics and relationships, health, and cognitive abilities.

The CFPS is suitable for our study for several reasons. First, the survey includes several standardized cognitive tests. Second, exact information about the geographic locations and dates of interviews is available to us for all respondents, enabling us to precisely match individual test scores in the survey with local air-quality data. Third, the longitudinal data allow us to remove unobserved individual factors that may bias estimates. Further, the survey embodies rich information at multiple levels, allowing us to control for a wide range of covariates. Finally, because the cognitive tests are administered to all age cohorts older than 10, we can study the effects of air pollution on different age groups.

CFPS 2010 and CFPS 2014 contain the same cognitive ability module, i.e., 24 standardized mathematics questions and 34 word-recognition questions. All these

www.pnas.org/cgi/doi/10.1073/pnas.1809474115

questions are obtained from standard textbooks and are sorted in ascending order of difficulty. The starting question depends on the respondent's education level. Specifically, those whose education level is primary school or below start with the 1st question; those who attended middle school begin with the 9th question in the verbal test and the 5th question in the math test; and those who finished high school or above start with the 21st question in the verbal test and the 13th question in the math test. The test ends when the individual incorrectly answers three questions in succession. The final test score is defined as the rank of the hardest question a respondent is able to answer correctly. If the respondent fails to answer any questions during the test, his or her test score is assigned as the rank of the starting question minus one. For example, a respondent with middle school education begins with the 9th question in the verbal test. If the hardest question he can correctly answer is the 14th question, then his verbal test scores would be 14. However, if he fails the 9th, 10th, and 11th questions consecutively, his verbal test scores would be 8. The respondents did not know the rules before they were interviewed, so they did not have the incentive to fail the tests on purpose.

B. Air Pollution and Weather Data

We measure air quality using the air pollution index (API), which is aggregated based on daily readings for three atmospheric pollutants, namely sulfur dioxide (SO_2), nitrogen dioxide (NO_2), and particulate matter smaller than 10 micrometers (PM10).[1] Carbon monoxide (CO), ozone, and particulate matter smaller than 2.5 micrometers (PM2.5) were not added to the basket of the index until 2014. Because all the cognitive tests were administered between 2010 and 2014, we transform the air quality index (AQI), which includes six pollutants, to the API in 2014 and use the API in our paper. The API ranges

[1] We use the Chinese Ministry of Environmental Protection's (MEP's) breakpoints table (see **Table S5**) and the following formula to generate the API measurement: $I_P = ((I_{HI} - I_{LO}) / (BP_{HI} - BP_{LO})) * (C_P - BP_{LO}) + I_{LO}$, where I_P is the index for pollutant P, C_P is the rounded concentration of pollutant P, BP_{HI} is the breakpoint that is greater than or equal to C_P, BP_{LO} is the breakpoint that is less than or equal to C_P, I_{HI} is the API value corresponding to BP_{HI}, and I_{LO} is the API value corresponding to BP_{LO}. The API represents the highest index value calculated for each pollutant.

from 0 to 500, with larger values indicating worse air quality. Daily API observations are obtained from the city-level air-quality report published by the Chinese Ministry of Environmental Protection (MEP) (3). The report includes 86 major cities in 2000 and covers most of the cities in China in 2014. If the government indeed manipulates the API data as suggested by Chen et al. (4) and Ghanem and Zhang (5), using the official API data would underestimate the true impact of air pollution. In this case, our estimates would represent a lower bound. **Figure S2** plots the daily API in China from 2010 to 2014, showing large temporal and regional variations.

The weather data are derived from the National Centers for Environmental Information of the US National Oceanic and Atmospheric Administration (6). The dataset contains daily records of rich weather conditions from 402 monitoring stations in China. Graff Zivin, Hsiang, and Neidell (7) find that high temperature is associated with significant decreases in cognitive performance on math in the short run. Hence, we control for a set of temperature bins (that is, <25°F, 25–45°F, 45–65°F, 65–85°F, and >85°F), total precipitation, mean wind speed, and a dummy for bad weather to capture the effects. Bad weather is defined as fog, rain/drizzle, snow/ice pellets, hail, thunder, and tornadoes/funnel clouds in the data.

C. Other Covariates

As a complement to the CFPS, we also use the China county census data to control for confounders that might vary across locations and years in China. For example, in the empirical analysis, we include a vector of county-level characteristics, including GDP per capita (deflated to 2010 yuan), population density, and industrial value share. These additional covariates are taken from the China Data Center of the University of Michigan and linked to the CFPS counties by year (8). Besides, the CFPS data also provide us with rich demographic characteristics at the individual level. We control gender, age and its square and cubic terms, log form of household per capita income and years of education in the main regressions. **Table S6** describes all the key variables and their summary statistics.

PART 2: Robustness of Results to Alternative Matching Radiuses and Strategies

We employ two matching methods. In each method, we test a wide range of matching radiuses. The findings are robust to different matching radiuses and strategies. Section A and Section B below describe them in detail.

A. Matching CFPS Counties to the Boundaries of API Reporting Cities

As the first alternative matching strategy, we calculate the distance between the centroid of each CFPS county and the boundary of API reporting cities. The MEP of China reports APIs at the city level. **Figure S3** plots the boundaries of API cities.[2] In specific, if a CFPS county is located within an API reporting city, we set the matching distance as zero and treat the city's API readings as the CFPS county's readings. If a CFPS county is not located in any cities with API readings, we match it to the nearest API reporting city within 40 kilometers of radius between the CFPS county centroid and city boundaries. We use the radius of ***40 km*** (i.e. ***25 miles***) in our analyses to keep consistency with the convention of the literature (9).

There is a tradeoff between the precision of matching and sample attenuation. When the radius becomes smaller, the matched counties becomes more precise at the cost of losing samples. In order to evaluate the tradeoff, **Table S7** and **Table S8** report estimation results based on nine different cutoff radiuses ranging from zero km to 80 km for the verbal test scores and math test scores, respectively. For example, Panel A of **Table S7** reports regression results on a subsample including only respondents who live in API reporting cities (matching distance equals to zero) for verbal test scores. While in Panel I of **Table S7**, the sample is expanded to include residents living within 80 km to the nearest API cities. The main findings hold no matter which radius is used, suggesting that measurement errors and attenuation bias do not affect our key findings. In particular, as shown in **Figure S5**

---

[2] Due to the confidential agreement, we are not allowed to plot the locations of CFPS counties on the map.

and **Figure S6**, the results for age-cohort effects are robust to the choice of radius.

B. Matching CFPS Counties to the Monitoring Stations

The city-level APIs are computed based on readings in multiple monitoring stations. While we have access to the exact latitude and longitude information of all the monitoring stations in each API reporting city in 2014, unfortunately the readings at the monitoring station level are not available in 2010. **Figure S4** displays the spatial distribution of air quality monitoring stations.[3] To improve the matching precision, our second alternative matching strategy assigns the reported city-level APIs in 2010 to all stations within the city boundary and match CFPS counties to the nearest air quality monitoring station within a specific radius. This alternative strategy assumes that the geographic locations of air quality monitoring stations do not change much between 2010 and 2014, and variation in daily air pollution across monitoring stations within each city is small. All the regressions are weighted by the inverse distance to monitoring stations. As presented in **Table S9** and **Table S10**, our results using this new matching strategy are robust to various radiuses between 40 km and 90 km. Besides, **Figure S7** and **Figure S8** plot the results for age-cohort effects across various radiuses. Once again, the pattern is consistent: while there is a negative effect on verbal tests for older men, the effect is much more muted on math tests.

C. Matching CFPS with Weather Data

The weather conditions are obtained as the inverse distance-weighted average of all monitoring stations within a radius of 100 km of the county centroid. The matching radius is comparable to those used in Deschenes, Greenstone and Guryan (10) and Deschenes and Greenstone (11) for the weather data. The binary indicator for bad weather comes from the nearest monitoring station.

D. Number of Observations in the Final Data

[3] Still, due to the confidential agreement, we are not allowed to plot the locations of CFPS counties on the map.

The CFPS surveyed a balanced panel of 25,486 individual respondents over age 10 in 2010 and 2014, for a total of 50,972 observations.[4] Of the individuals surveyed in both waves, 282 are missing values for cognitive test scores. Among the remaining 50,789 observations, 37,918 observations could be matched to API and weather data.[5] Due to some missing values for household demographics, the final dataset used in this study includes 31,955 observations. **Figure S9** displays the 24-hour time distribution of respondents who took the cognitive tests and the hourly pollutant concentration. Most of the cognition tests were conducted in the afternoon and evening. Among the three pollutants, PM10 is a dominant one throughout the day.

PART 3: Falsification Tests

In the falsification test, we employ a strategy similar to Bensnes (12), which tests the effects of API readings on the days after the interviews on cognitive test scores. **Figure S10** presents the estimated coefficients with their 95 and 99 percent confidence intervals from a regression of test scores on API readings one to six days into the future by gender. For both the whole sample and the subsamples, all the coefficients are statistically indifferent from zero, largely dismissing the concern about potential omitted variables.

PART 4: Robustness of Results to Alternative Specifications and Samples

In this section, we examine the robustness of our results according to several different

---

[4] The attrition rates for consecutive waves, that is, 2010–2012 and 2012–2014, are 19.3 percent and 13.9 percent, respectively. We compare the attrition rate of the CFPS with that of the UK Household Longitudinal Survey (UKHLS). The two surveys were conducted during the same period and followed similar interview methods, so the UKHLS serves as a good benchmark for the CFPS. Compared to the UKHLS, the CFPS's attrition rate is reasonable. The key reason for using the 2010 and 2014 waves is that the two waves included exactly the same test modules, whereas the short memory and logic tests employed in the 2012 wave are not comparable with the tests used in the other two waves.

[5] Counties unmatched to any API report cities within 40 km or weather stations within 100 km are dropped. The matching rate of 74.7 percent (37,918 out of 50,789) is within a reasonable range compared with other studies. For example, Levinson (2012) was able to maintain 52.3 percent of the observations when matching the US General Social Survey with PM10 readings from the Environmental Protection Agency's Air Quality System.

tests to confirm whether our results were qualitatively affected by the decisions made in our paper along several dimensions, such as model specification, sample selection, mechanism tests, and weighted regressions.

A. Alternative Specifications of Table S1

In **Table S11**, we further display alternative specifications of **Table S1** with and without demographic controls and individual fixed effects. Panel A is for verbal test scores, while Panel B is for math test scores. In each panel, there are three parts, which correspond to 7-day, 90-day and 1-year windows, respectively. The last column in each part just keeps the original column in **Table S1** for ease of comparison. The first column in each part addresses the concern that household income per capita and years of education may be bad controls, i.e., they are endogenous. We re-estimate our models with income per capita and years of education excluded. The results in the first column are qualitatively identical to those in the last columns, suggesting they are not bad controls. The middle column in each part further explores the difference between longitudinal and cross-sectional estimations. The comparison between the middle column and the last column indicates that the statistical significance is basically the same between these two specifications, but the size of the effect of air pollution without individual fixed effects is only two thirds of that with individual fixed effects controlled.

B. Giving Interviews in Winter Months Greater Weights

Most of the survey interviews were conducted in the summer months. There is a concern that the results are driven by the overwhelmingly large sample in the summer. We divide the sample into two groups. Respondents in Group 1 were interviewed at least once in winter months (November, December and January), while respondents in Group 2 were only interviewed in non-winter months (from February to October). Observations in Group 1 are reweighted by the ratio of the number of observations in Group 2 divided by the number of observations in Group 1. In doing so, we give observations in the winter months greater weights. As revealed in **Table S12**, the weighted regression indicates that the results are robust, i.e., the size of the effects is similar to that estimated in the baseline results. Besides, **Figure S11** further shows that the pattern of age-cohort effects still holds for verbal tests. Hence, the underestimation of the effect of air pollution due to over representation of the summer months, if any, is small.

C. Excluding the Channels of Impatience and Noncooperation

The negative effect of air pollution on cognitive performance may be driven by behavior change. People may become more impatient or uncooperative when exposed to more polluted air, thereby hampering their cognitive tests. The CFPS includes evaluation on interviewees' degrees of impatience and cooperation rated by the interviewers. The ratings for impatience and cooperation are both scaled from 1 (low) to 7 (high). We explore the effects of exposure to air pollution on impatience and cooperation using a similar specification in Equation (1). **Table S13** displays the results. Panels A and B are for impatience cooperation, respectively. The results indicate that there is no significant association between air pollution and interviewees' impatience and cooperation, ruling out the behavioral channel.

D. Restricting the Sample to Non-Migrants Only

The CFPS tracked and interviewed individuals who moved. For respondents who moved between waves 2010 and 2014, our analysis matches API measures according to their places of residence by the 2014 survey. A measurement issue arises that individuals may not stay in their counties of residence for the whole period of cumulative measure of exposure used. To more precisely match air pollution exposure in wave 2014, as a robust check we exclude 1.3% of respondents who migrated across counties between 2010 and 2014. Similarly, to ensure precise matching of air pollution exposure in wave 2010, we further use information on the time of moving into the latest addresses by the 2010 survey to exclude those respondents who moved within each time window of pollution exposure.

For example, we exclude respondents who moved into the latest addresses in 2010 for the 1-year time window, 2009-2010 for the 2-year time window, and 2008-2010 for the 3-year time window. As reported in **Table S14**, the majority of respondents in CFPS were non-migrants across all time windows. In addition, **Figure S12** also reveals a larger effect on verbal test for the older male cohorts. Therefore, our key findings are robust to using non-migrants only.

E. Excluding Polluted Occupations

In **Table S15** and **Figure S13**, we perform robustness checks by excluding polluted occupations. Polluted occupations include "Geology and mineral industry workers", "Workers in metal smelting and refining industry", "Chemical product manufacturing personnel", "Textile workers", "Production workers (wood processing, artificial board, wood products, pulp and paper industry)", and "Production and processing worker (construction materials)". Our baseline results are qualitatively unchanged after excluding polluted occupations.

F. Controlling for Province-by-year Fixed Effects and Clustering Standard Errors at the Province Level

As revealed in **Table S16**, our baseline results are robust to controlling for province-by-year fixed effects and clustering standard errors at the province level.

PART 5: Estimating Movement in the Test Distribution Using the Coefficient Estimates

Figure S14 plots the percentile of scores for verbal and math tests, respectively. As revealed in **Figure S14**, a one-point increase in verbal test scores corresponds to moving people from the median (i.e. the 50th percentile) to the 55th percentile in the verbal test distribution, while a one-point increase in math test scores is equivalent to moving people from the median to the 68th percentile in the math test distribution.

The population-weighted annual mean concentration of PM10 over 2014 in China is 112 μg/m^3, much higher than the National Ambient Air Quality Standards (NAAQS) published by the U.S. Environmental Protection Agency (EPA).[6] Reducing the annual mean PM10 to levels below the standard, which amounts to 56 units in one-year-mean API derived from **Table S5**, will lead to a sizable increase in verbal test scores by 2.41 points (or moving people from the median to the 63rd percentile in the verbal test distribution) and math test scores by 0.39 point (or moving people from the median to the 58th percentile in the math test distribution) calculated from **Table S1**.

Besides, we also evaluate the effect of exposure to air pollution on the older cohort using estimated coefficients in Panel A and Panel B of **TableS4a** separately by educational attainment. A one standard deviation decrease in 3-year-mean API leads to an increase in verbal test scores by 9.18 points (or moving people from the median to the 87th percentile in the verbal test distribution) for less educated men above age 65 relative to their counterparts below age 25, while the same decline in API is associated with an increase in verbal test scores by 1.88 points (or moving people from the median to the 69th percentile in the verbal test distribution) for more educated older men relative to their younger counterparts. The negative effect on the less (more) educated older men provides an upper (lower) bound of the detrimental impact of air pollution exposure on older persons.

PART 6: Scientific Backgrounds and Potential Mechanisms

Air pollution may affect cognition through both physiological and psychological pathways. In this Appendix, we hypothesize that differences in brain composition may help explain why men appear more sensitive to air pollution. It is beyond the scope of this paper to formally test this mechanism. We leave it as a future research topic.

[6] The annual mean PM10 data at the city level are obtained from the *China Environmental Statistical Yearbook 2015*, and the population data (for the weighting purpose) come from *China City Statistical Yearbook 2015*. The standard of annual mean PM10 published by the EPA is 50 μg/m^3. Source: https://www3.epa.gov/ttn/naaqs/standards/pm/s_pm_history.html.

A few of these physiological pathways have been documented in the literature (13). First, multiple pollutants (or toxic compounds bonded to the pollutants) may directly affect brain chemistry. For example, ozone in the air can react with body molecules to create toxins, causing asthma and respiratory problems (14).[7] Particulate matter (PM), especially fine particles, can carry toxins through small passageways and directly enter the brain. Braniš, Řezáčová, and Domasová (15) show that exposure to high PM concentrations compromises cognitive performance even for people working indoors.[8]

Second, people breathing polluted air are more likely to be subject to oxygen deficiency, which in turn impairs their cognitive abilities (16, 17). Carbon monoxide (CO), one important element of air pollution, prevents the body from releasing adequate oxygen to vital organs, in particular to the brain, which consume a large fraction of total oxygen intake. Third, air pollution could also damage the immune system, hinder neurological development, and impair neuron behavior, all of which contribute to long-term memory formation (18). Fourth, long-term exposure to pollution leads to the growth of white-matter lesions, potentially inhibiting cognition (19). Further, exposure to highly concentrated air pollution can be linked to markers of neuroinflammation and neuropathology that are associated with neurodegenerative conditions, such as Alzheimer's disease (20, 21). Finally, a recent study on healthy children living in polluted environment with APOE Ɛ4 allele (known to increase risk of developing Alzheimer's) demonstrates compromised cognitive responses compared with those carrying APOE gene with Ɛ3 allele (22). However, this gene environment interaction is only verified for children, while our main findings are towards elder persons.

In addition to physiological pathways, air pollution could also disrupt cognitive functioning through some psychological pathways. For example, high concentrations of CO and nitrogen dioxide (NO_2) are significantly associated with headache, eye irritation,

[7] Ozone is formed through a chemical reaction between nitrogen oxides, sunlight, and various gaseous pollutants.

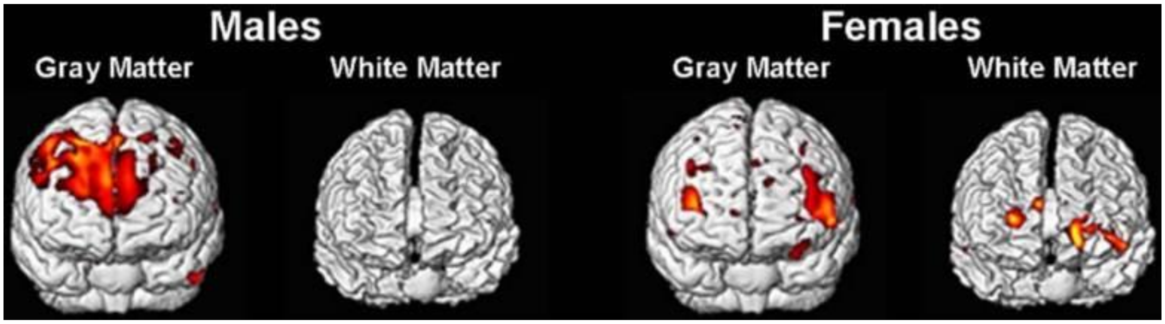
[8] PM is generated by power plants, factories, vehicles, dust, pollen and forest fires.

and respiratory problems (23).[9] High levels of ozone and sulfur dioxide (SO_2) have also been found to cause psychiatric distress (24).[10] Exposure to high concentrations of CO, NO_2, SO_2, ozone, and PM may also increase the risk of depression (25).

Our central nervous system has two important tissues: gray matter and white matter. Gray matter represents information processing centers, and white matter represents the networking of – or connections between – these processing centers. Mathematics abilities, which require more local processing, mainly depend on gray matter. While language skills, which require integrating and assimilating information from distributed gray-matter regions in the brain, mainly rely on white matter.[11]

A brain scanning study conducted by Haier et al. (26) reveals that men have approximately 6.5 times the amount of gray matter *activated* during general intelligence tests than women do, but women have nearly 10 times the amount of white matter *activated* during general intelligence tests than men do. See the figure below for a front view of grey and white matter *activation* during IQ tests. This finding may help explain why men tend to excel in math tests, while women tend to excel in verbal tests.

Figure: Front view of grey and white matter *activation* during IQ tests



Source: Haier et al. (2005).

A large body of literature has proven that air pollution can reduce the density of white matter in the brain (19, 27, 28), which may directly explain why air pollution appears to

---

[9] NO_2 and CO are emitted by coal-burning power plants and the burning of fossil fuels.

[10] SO_2 is mainly emitted by coal-burning power plants.

[11] University of California, Irvine. "Intelligence in Men and Women Is a Gray and White Matter." Science Daily. www.sciencedaily.com/releases/2005/01/050121100142.htm [accessed January 25, 2017].

have a larger effect on verbal test than on math test scores. Besides, since men have a much smaller amount of white matter *activated* during intelligence tests, their cognitive performance, especially in the verbal domain, tends to be more affected by exposure to air pollution.

References

1. *China Family Panel Studies (CFPS): 2010-2014*. Institute of Social Science Survey of Peking University, China. http://www.isss.pku.edu.cn/cfps/EN/Data/.
2. Xie, Y., and J. Hu. 2014. "An Introduction to the China Family Panel Studies (CFPS)." *Chinese Sociological Review* 47 (1): 3–29.
3. *Air Quality Daily Reports: 2007-2014*. Chinese Ministry of Environmental Protection. http://data.epmap.org/eia/air.
4. Chen, Y., G. Z. Jin, N. Kumar, and G. Shi. 2012. "Gaming in Air Pollution Data? Lessons from China." *B.E. Journal of Economic Analysis and Policy* 13 (3): 1–43.
5. Ghanem, D. and J. Zhang. 2014. "Effortless Perfection": Do Chinese Cities Manipulate Air Pollution Data? *Journal of Environmental Economics and Management* 68: 203–225.
6. *Global Summary of the Day (GSOD): 2010-2014*. National Centers for Environmental Information of the US National Oceanic and Atmospheric Administration, ftp://ftp.ncdc.noaa.gov/pub/data/gsod.
7. Graff Zivin, J. S., S. M. Hsiang, and M. J. Neidell. 2015. *Temperature and Human Capital in the Short- and Long-Run*. NBER Working Paper 21157. Cambridge, MA, US: National Bureau of Economic Research.
8. *China County Statistics: 2010-2014*. China Data Center of University of Michigan. http://chinadataonline.org/member/county/.
9. Levinson, A. 2012. "Valuing Public Goods Using Happiness Data: The Case of Air Quality." *Journal of Public Economics* 96: 869–880.
10. Deschenes, O., M. Greenstone, and J. Guryan. 2009. "Climate Change and Birth Weight." *American Economic Association* 99 (2): 211–217.
11. Deschenes, O., and M. Greenstone. 2011. "Climate Change, Mortality, and Adaptation: Evidence from Annual Fluctuations in Weather in the US." *American Economic Journal: Applied Economics* 3 (October): 152–185.
12. Bensnes, S. S. 2016. "You Sneeze, You Lose: The Impact of Pollen Exposure on Cognitive Performance During High-Stakes High School Exams." *Journal of Health Economics* 49: 1–13.
13. Block, M. L., and L. Calderón-Garcidueñas. 2009. "Air Pollution: Mechanisms of Neuroinflammation and CNS Disease." *Trends in Neurosciences* 32 (9): 506–516.
14. Sanders, N. J. 2012. "What Doesn't Kill You Makes You Weaker: Prenatal Pollution Exposure and Educational Outcomes." *Journal of Human Resources* 47: 826–850.
15. Braniš, M., P. Řezáčová, and M. Domasová. 2005. "The Effect of Outdoor Air and Indoor Human Activity on Mass Concentrations of PM10, PM2.5, and PM1 in a Classroom." *Environmental Research* 99 (2): 143–149.

16. Amitai, Y., Z. Zlotogorski, V. Golan-Katzav, A. Wexler, and D. Gross. 1998. “Neuropsychological Impairment from Acute Low-Level Exposure to Carbon Monoxide.” *Archives of Neurology* 55 (6): 845–848.

17. Kampa, M., and E. Castanas. 2007. “Human Health Effects of Air Pollution.” *Environmental Pollution* 151: 362–367.

18. Perera, F. P., Z. Li, R. Whyatt, L. Hoepner, S. Wang, D. Camann, and V. Rauh. 2009. “Prenatal Airborne Polycyclic Aromatic Hydrocarbon Exposure and Child IQ at Age 5 Years.” *Pediatrics* 124 (2): e195–e202.

19. Calderón-Garcidueñas, L., A. Mora-Tiscareño, E. Ontiveros, G. Gómez-Garza, G. Barragán-Mejía, J. Broadway, S. Chapman, G. Valencia-Salazar, V. Jewells, R. R. Maronpot, C. Henríquez-Roldán, B. Pérez-Guillé, R. Torres-Jardón, L. Herrit, D. Brooks, N. Osnaya-Brizuela, M. Monroy, A. González-Maciel, R. Reynoso-Robles, R. Villarreal-Calderon, A. Solt, and R. Engle. 2008. “Air Pollution, Cognitive Deficits and Brain Abnormalities: A Pilot Study with Children and Dogs.” *Brain and Cognition* 68 (2): 117–127.

20. Calderón-Garcidueñas, L., Reed, W., Maronpot, RR., Henriquez-Roldán, C., Delgado-Chavez, R., Calderón-Garcidueñas, A., Swenberg, J.A. 2004. “Brain inflammation and Alzheimer’s-like pathology inindividuals exposed to severe air pollution.” *Toxicologic Pathology* 32: 650–658.

21. Levesque, S., Surace, MJ., McDonald, J., and Block, ML. 2011. Air pollution and the brain: Subchronic diesel exhaust exposure causes neuroinflammation and elevates early markers of neurodegenerative disease. *Journal of Neuroinflammation* 8: 105.

22. Calderón-Garcidueñas, L. et al. 2015. Mexico City normal weight children exposed to high concentrations of ambient PM2.5 show high blood leptin and endothelin-1, vitamin D deficiency, and food reward hormone dysregulation versus low pollution controls. Relevance for obesity and Alzheimer disease. *Environmental Research* 140: 579-592.

23. Nattero, G., and A. Enrico. 1996. “Outdoor Pollution and Headache.” *Headache* 36: 243–245.

24. Rotton, J., and J. Frey. 1984. “Psychological Costs of Air Pollution: Atmospheric Conditions, Seasonal Trends, and Psychiatric Emergencies.” *Population and Environment* 7 (1): 3–16.

25. Szyszkowicz, M. 2007. “Air Pollution and Emergency Department Visits for Depression in Edmonton, Canada.” *International Journal of Occupational Medicine and Environmental Health* 20 (3): 241–245.

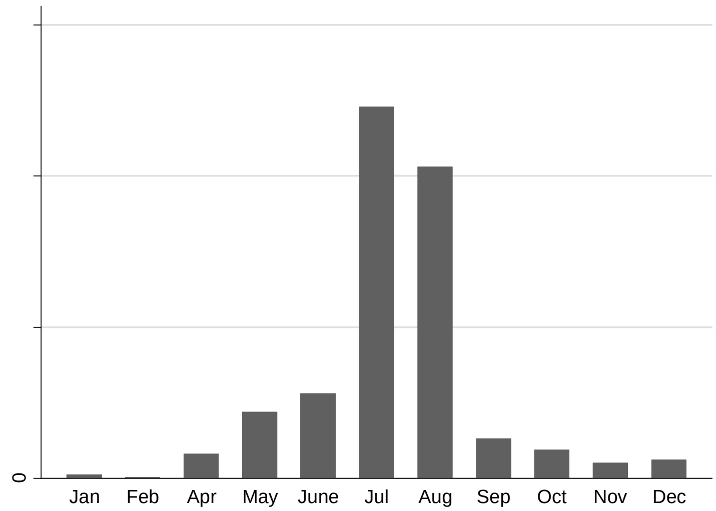
26. Haier, Richard J., Rex E. Jung, Ronald A. Yeo, Kevin Head, and Michael T. Alkire. 2005. “The Neuroanatomy of General Intelligence: Sex Matters.” *NeuroImage* 25: 320-332.

27. Calderón-Garcidueñas, L., R. Engle, A. Mora-Tiscareño, M. Styner, G. Gómez-Garza, H. Zhu, V. Jewells, R. Torres-Jardón, L. Romero, M. Monroy-Acosta, C. Bryant, L. González-González, H. Medina-Cortina, A. D’Angiulli. 2011. “Exposure to Severe Urban Air Pollution Influences Cognitive Outcomes, Brain Volume and Systemic

Inflammation in Clinically Healthy Children.” *Brain and Cognition* 77: 345-355.

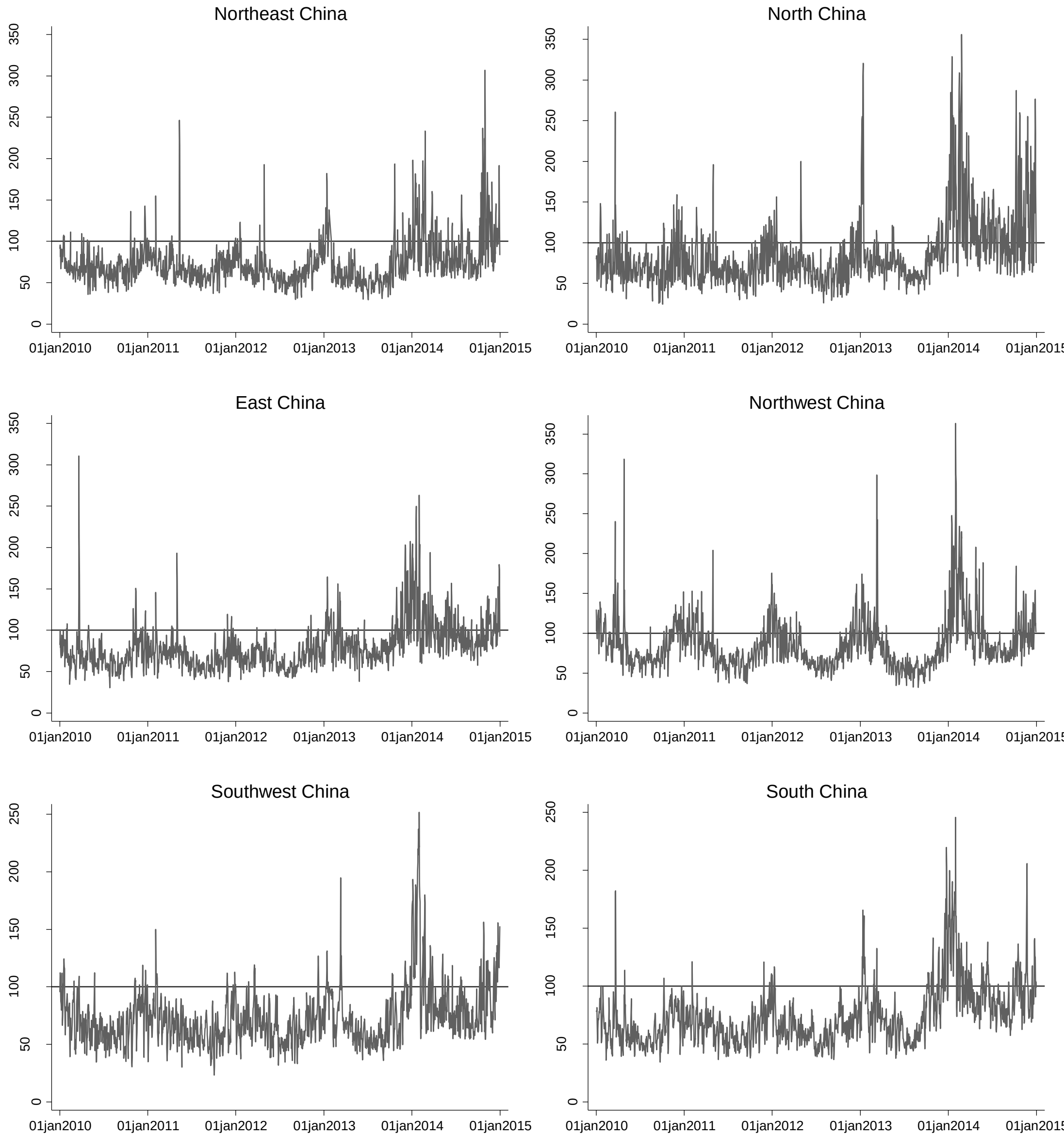
28. Wilker, Elissa H., Sarah R. Preis, Alexa S. Beiser, Philip A. Wolf, Rhoda Au, Itai Kloog, Wenyuan Li, Joel Schwartz, Petros Koutrakis, Charles DeCarli, Sudha Seshadri, Murray A. Mittleman. 2015. “Long-Term Exposure to Fine Particulate Matter, Residential Proximity to Major Roads and Measures of Brain Structure.” *Stroke* 46:1161-1166.

Figure S1: Distribution of interview months in 2010 and 2014



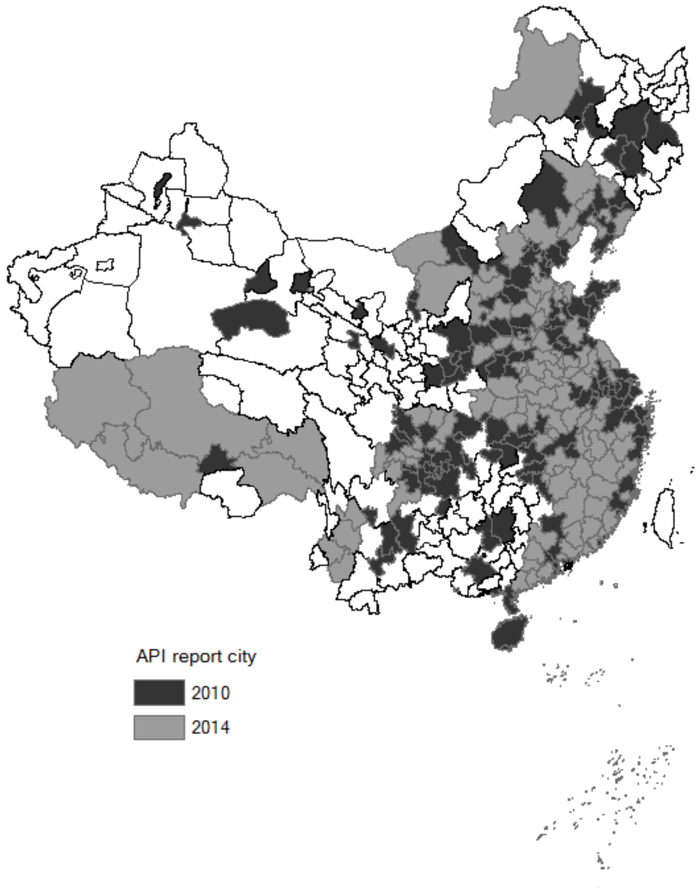
Source: CFPS survey 2010 and 2014.

Figure S2: Daily air pollution index (API) in China, 2010–2014



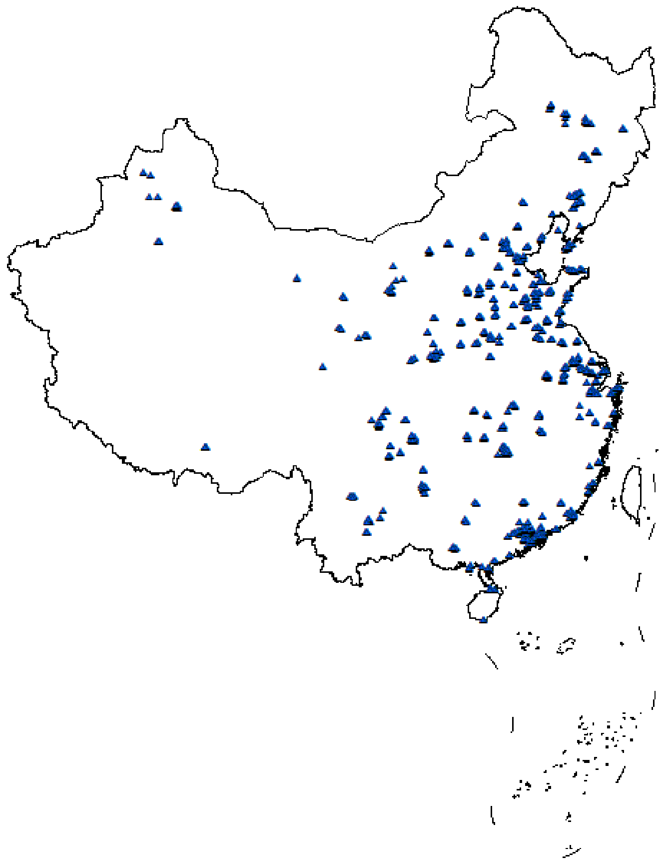
Source: Daily air-quality report, Ministry of Environmental Protection of the People's Republic of China. Note: The daily mean API is calculated by finding the weighted average of all the API report cities within the region, where the weights are the yearly population in each city. The US National Ambient Air Quality Standard for fine particulate matter smaller than 10 micrometers is 0.15 mg/m^3, which corresponds to API = 100 in China. Northeast China includes Heilongjiang, Jilin, and Liaoning. North China includes Beijing, Hebei, Inner Mongolia, Shanxi, and Tianjin. East China includes Anhui, Fujian, Jiangsu, Jiangxi, Shandong, Shanghai, and Zhejiang. Northwest China includes Gansu, Ningxia, Qinghai, Shanxi, and Xinjiang. Southwest China includes Guizhou, Sichuan, Tibet, Yunnan, and Chongqing. South China includes Guangdong, Guangxi, Hainan, Henan, Hubei, and Hunan.

Figure S3: The distribution of API reporting cities



Source: The Ministry of Environmental Protection of China.
Note: The legend 2010 represents all 86 API reporting cities in 2010, and the legend 2014 indicates newly added API reporting cities in 2014, which cover most of the cities in China by 2014. This figure is plotted using ArcMap 10.3.1. API = air pollution index.

Figure S4: The distribution of monitoring stations



Source: The Ministry of Environmental Protection of China.
Note: This figure is plotted using ArcMap 10.3.1.

Figure S5: Robustness checks — matching distance to the boundaries of API reporting cities (verbal test scores)

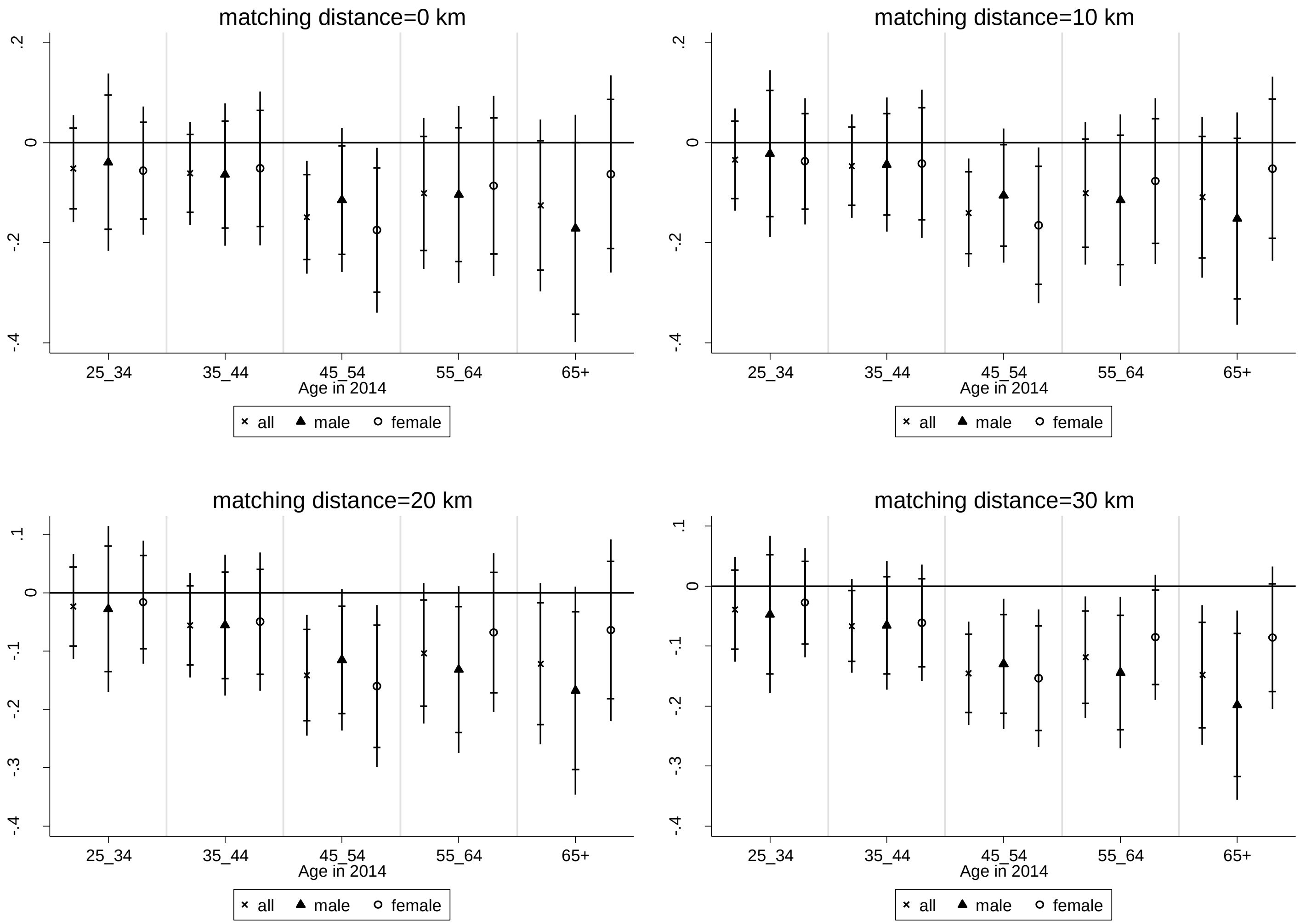


Figure S5 (continued): Robustness checks — matching distance to the boundaries of API reporting cities (verbal test scores)

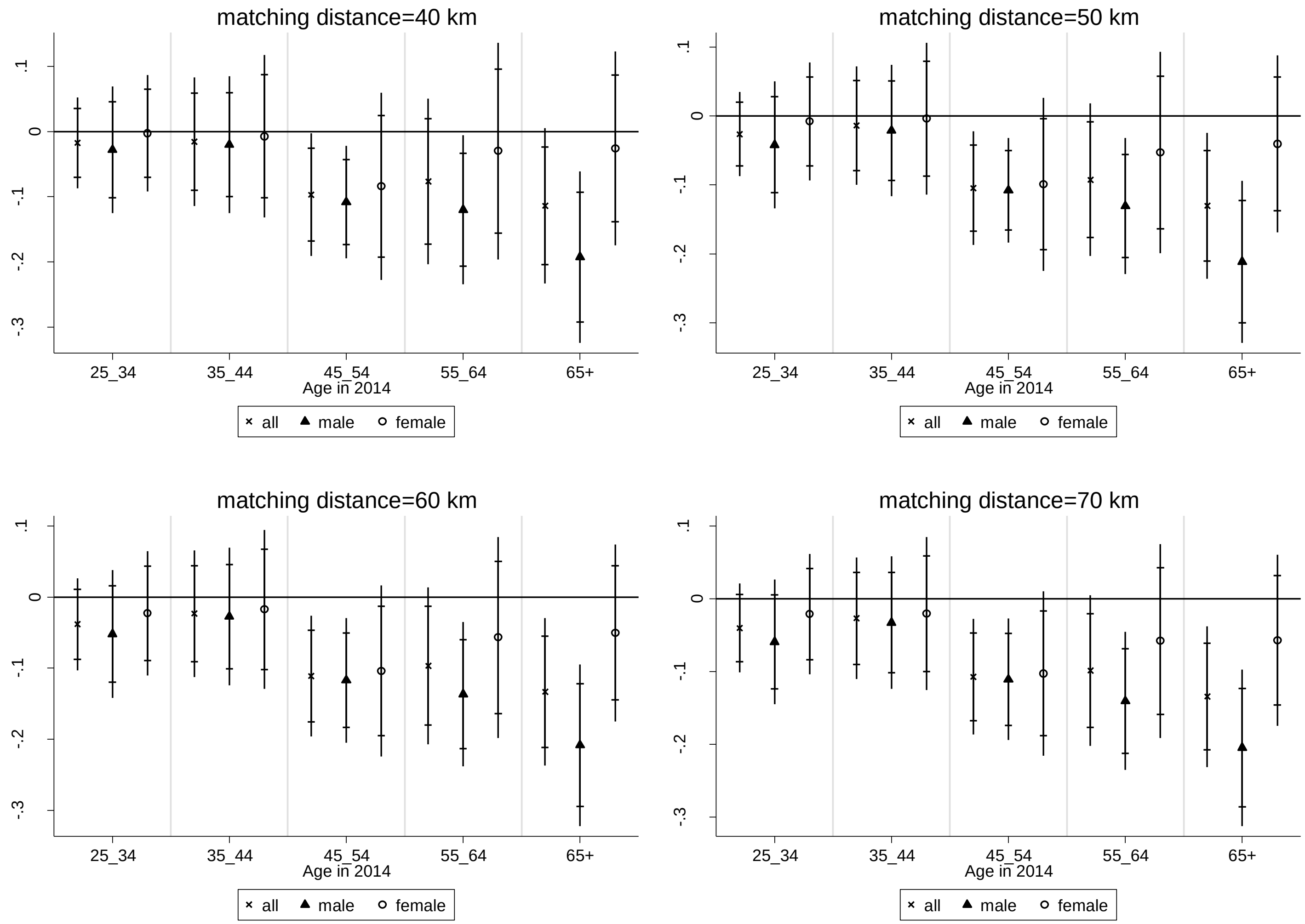
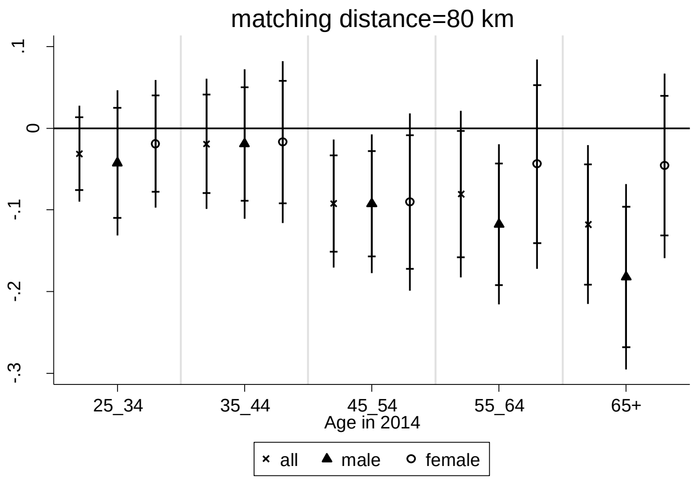


Figure S5 (continued): Robustness checks — matching distance to the boundaries of API reporting cities (verbal test scores)



Note: The age-cohort effects include interaction terms between 3-year-mean API and age cohort dummies 25-34, 35-44, 45-54, 55-64, and 65+ in 2014. The age band 10-24 is the reference category. The figures plot the estimated coefficients on the interaction terms for the whole sample as well as the male and female subsamples with 95% and 99% confidence intervals. As APIs are reported at the city level, we calculate the distance between the centroid of each CFPS county and the boundary of API reporting cities. If a CFPS county is located within an API reporting city, we set the matching distance as zero and treat the city's API readings as the CFPS county's readings. If a CFPS county is not located in any cities with API readings, we match it to the nearest API reporting city within a specific radius between the CFPS county centroids and the city boundaries. The figure reports estimation results based on nine cutoff radiuses ranging from zero km to 80 km.

Figure S6: Robustness checks — matching distance to the boundaries of API reporting cities (math test scores)

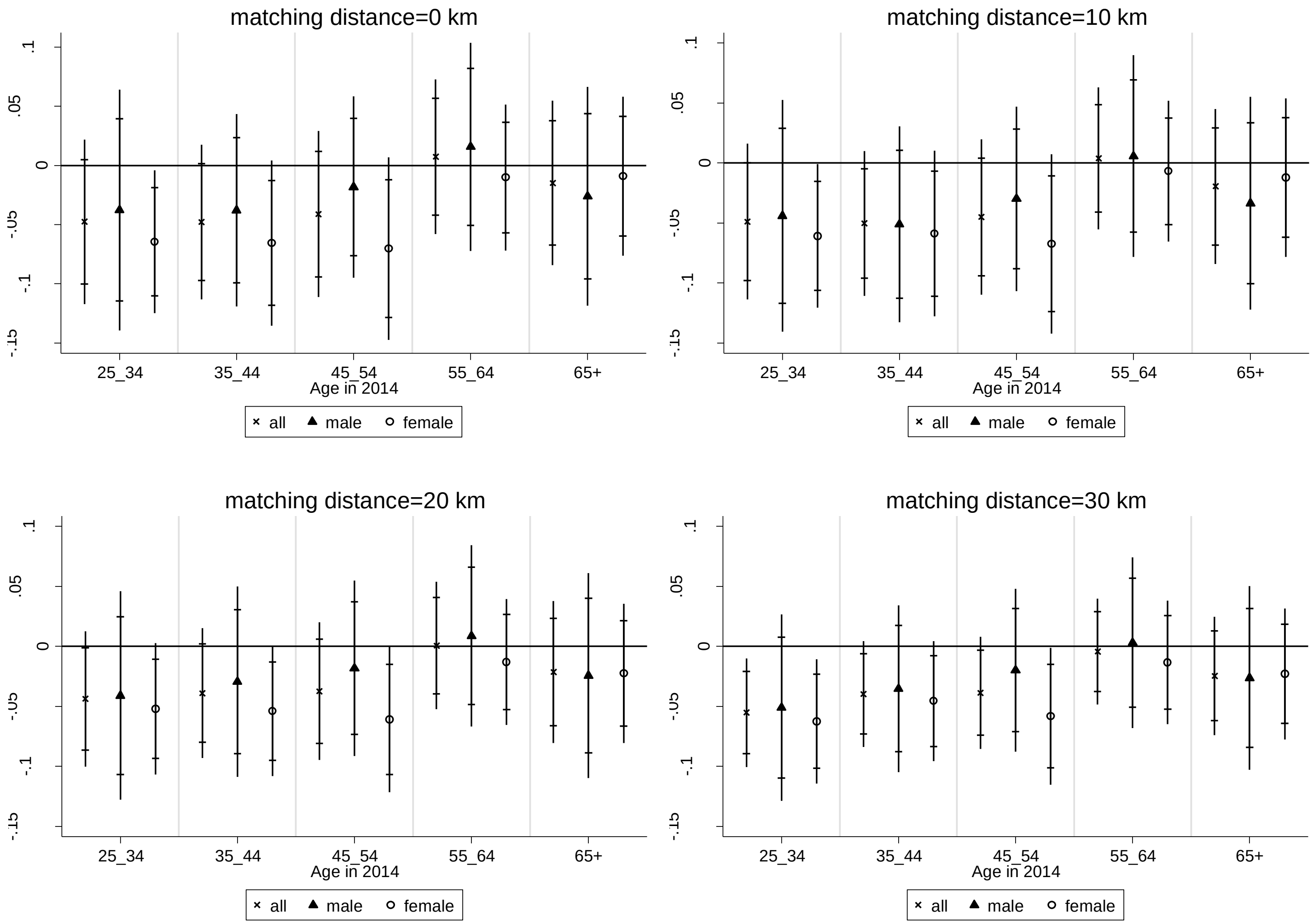


Figure S6 (continued): Robustness checks — matching distance to the boundaries of API reporting cities (math test scores)

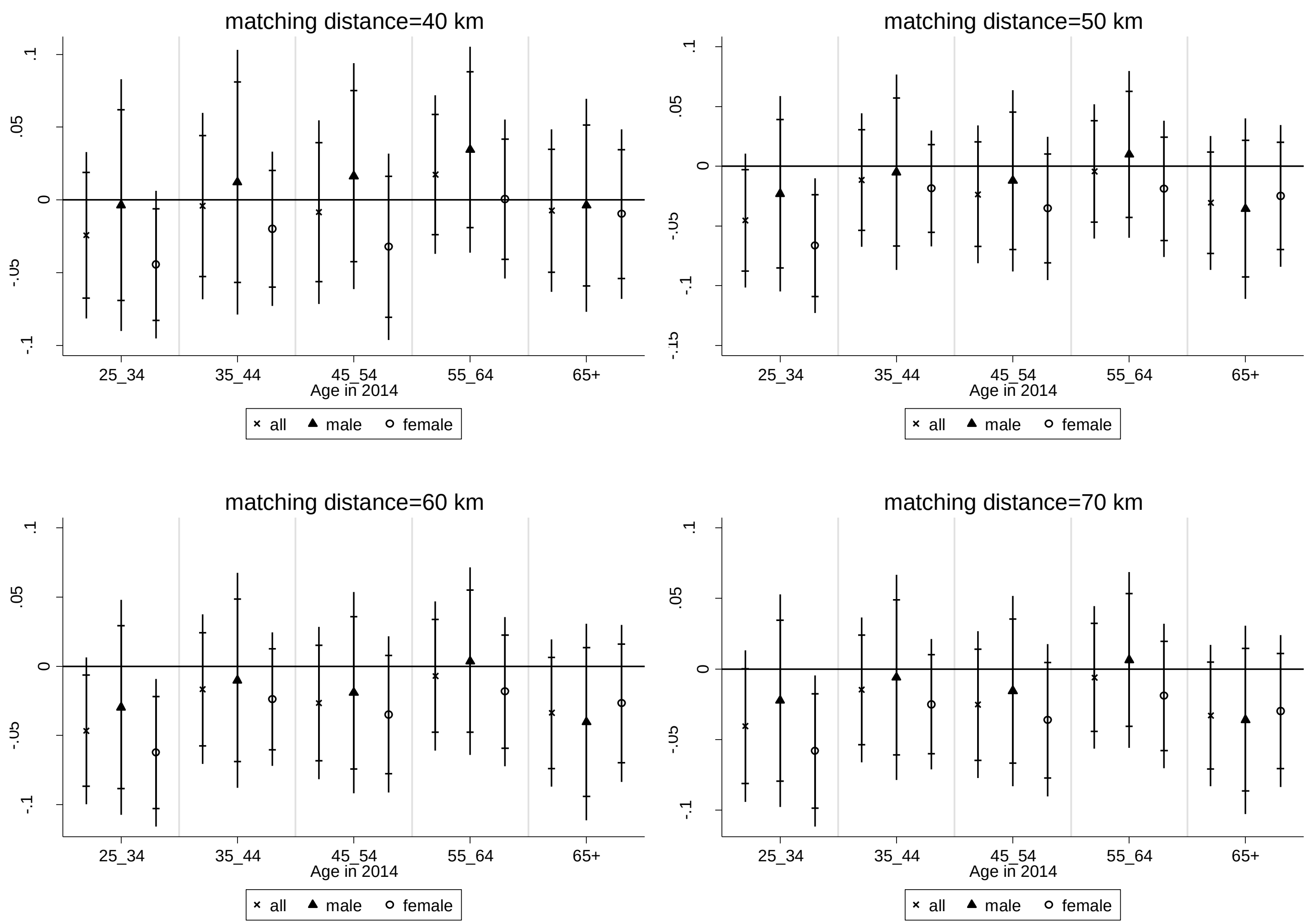
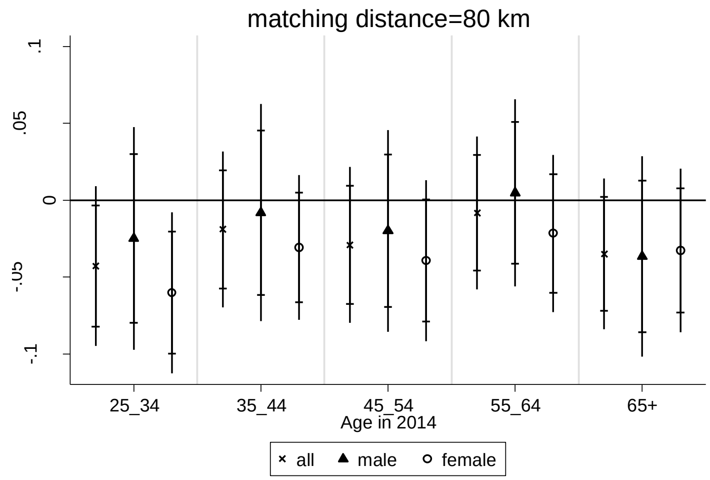


Figure S6 (continued): Robustness checks — matching distance to the boundaries of API reporting cities (math test scores)



Note: The age-cohort effects include interaction terms between 3-year-mean API and age cohort dummies 25-34, 35-44, 45-54, 55-64, and 65+ in 2014. The age band 10-24 is the reference category. The figures plot the estimated coefficients on the interaction terms for the whole sample as well as the male and female subsamples with 95% and 99% confidence intervals. As APIs are reported at the city level, we calculate the distance between the centroid of each CFPS county and the boundary of API reporting cities. If a CFPS county is located within an API reporting city, we set the matching distance as zero and treat the city's API readings as the CFPS county's readings. If a CFPS county is not located in any cities with API readings, we match it to the nearest API reporting city within a specific radius between the CFPS county centroids and the city boundaries. The figure reports estimation results based on nine cutoff radiuses ranging from zero km to 80 km.

Figure S7: Robustness checks — matching distance to monitoring stations (verbal test scores)

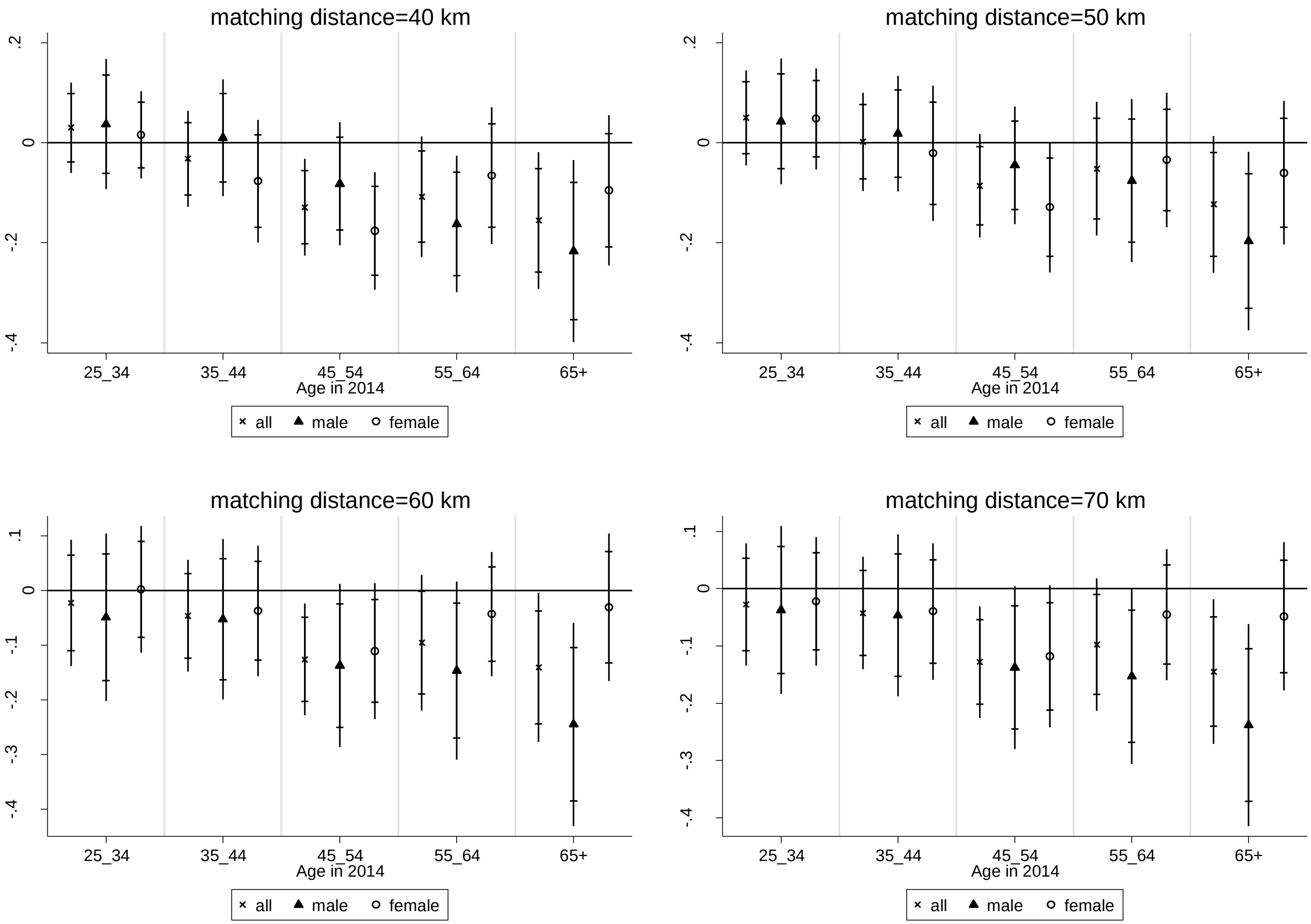
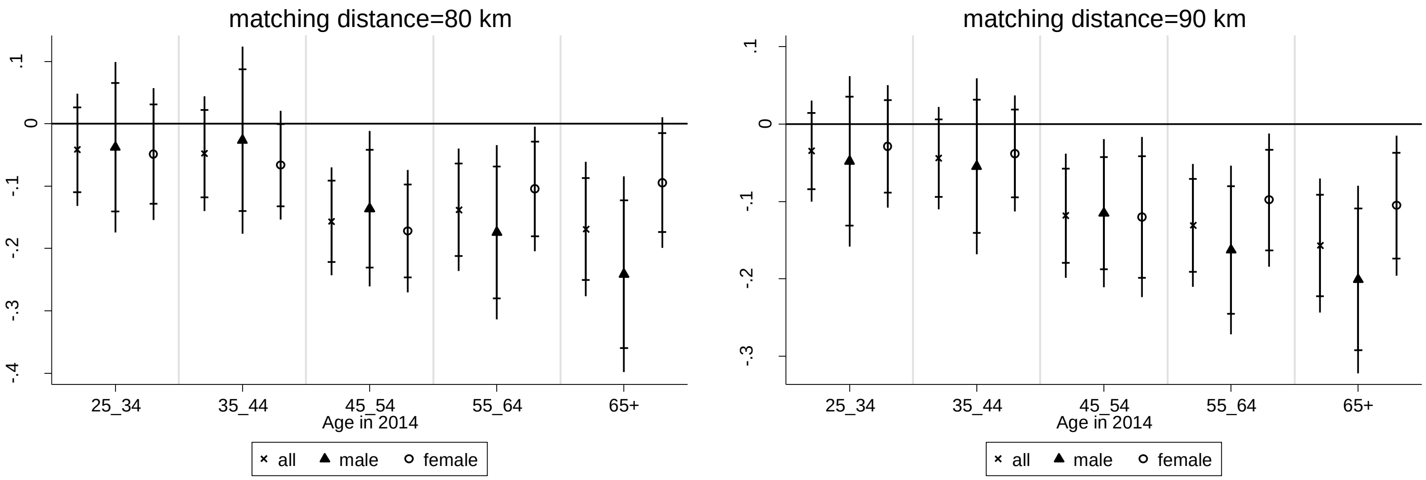


Figure S7 (continued): Robustness checks — matching distance to monitoring stations (verbal test scores)



Note: The age-cohort effects include interaction terms between 3-year-mean API and age cohort dummies 25-34, 35-44, 45-54, 55-64, and 65+ in 2014. The age band 10-24 is the reference category. The figures plot the estimated coefficients on the interaction terms for the whole sample as well as the male and female subsamples with 95% and 99% confidence intervals. The city-level APIs are computed based on readings from multiple monitoring stations in each API reporting city, but the readings from the monitoring stations along with their latitude and longitude information have been available since 2014. We assign the reported city-level APIs in 2010 to all stations within the city boundary under the assumption that that the geographic locations of air quality monitoring stations do not change much between 2010 and 2014 and variations in daily air pollution across monitoring stations within each city are small. We match CFPS counties to the nearest air quality monitoring station within a specific radius between the CFPS county centroids and the monitoring stations. **All the results are weighted by the inverse distance to monitoring stations.** The table reports estimation results based on six cutoff radiuses ranging from 40 km to 90 km.

Figure S8: Robustness checks — matching distance to monitoring stations (math test scores)

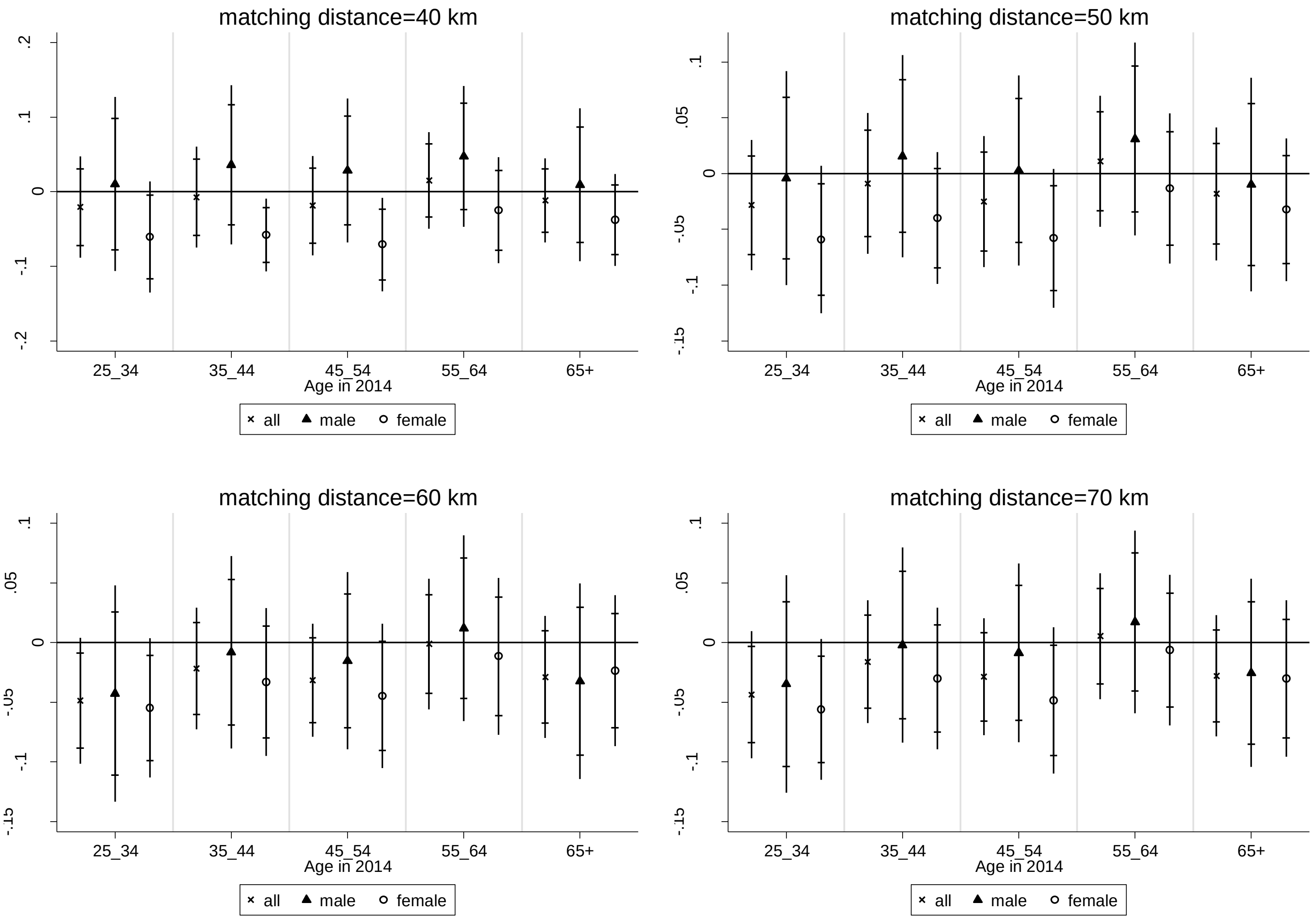
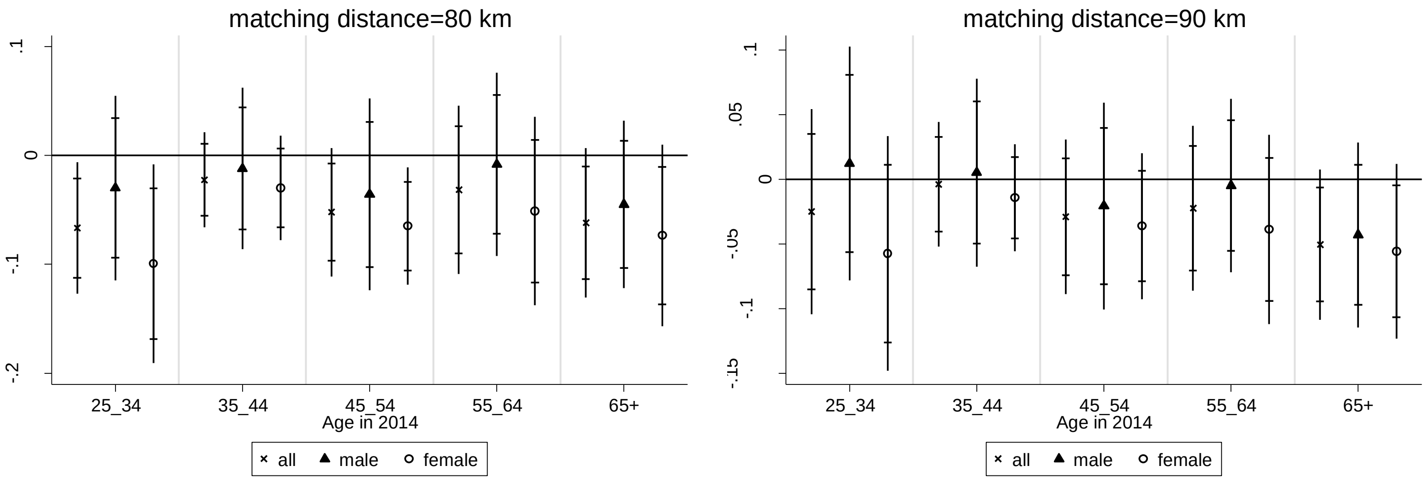
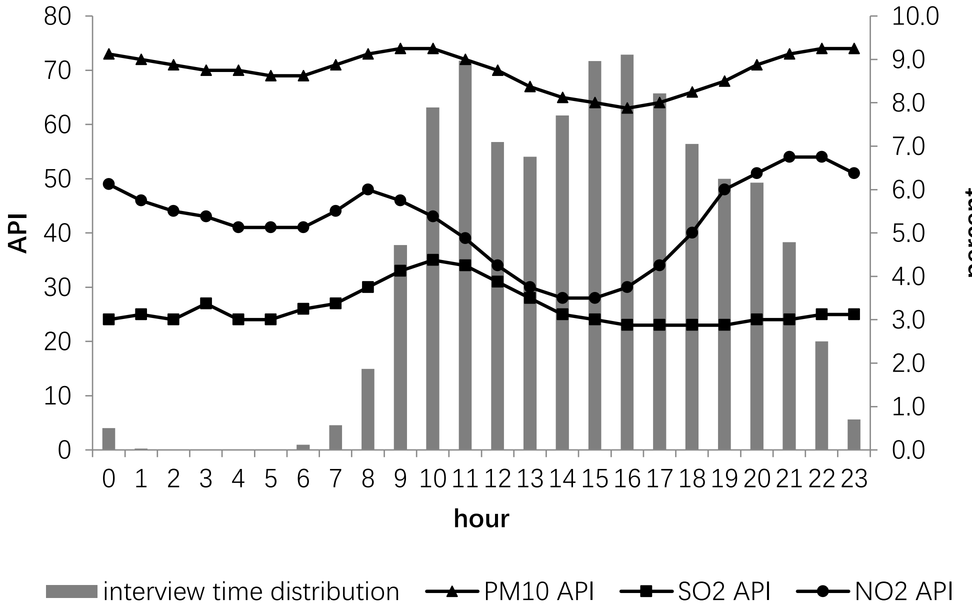


Figure S8 (continued): Robustness checks — matching distance to monitoring stations (math test scores)



Note: The age-cohort effects include interaction terms between 3-year-mean API and age cohort dummies 25-34, 35-44, 45-54, 55-64, and 65+ in 2014. The age band 10-24 is the reference category. The figures plot the estimated coefficients on the interaction terms for the whole sample as well as the male and female subsamples with 95% and 99% confidence intervals. The city-level APIs are computed based on readings from multiple monitoring stations in each API reporting city, but the readings from the monitoring stations along with their latitude and longitude information have been available since 2014. We assign the reported city-level APIs in 2010 to all stations within the city boundary under the assumption that that the geographic locations of air quality monitoring stations do not change much between 2010 and 2014 and variations in daily air pollution across monitoring stations within each city are small. We match CFPS counties to the nearest air quality monitoring station within a specific radius between the CFPS county centroids and the monitoring stations. **All the results are weighted by the inverse distance to monitoring stations.** The table reports estimation results based on six cutoff radiuses ranging from 40 km to 90 km.

Figure S9: PM10 API, SO_2 API, and NO_2 API during the day, 2014

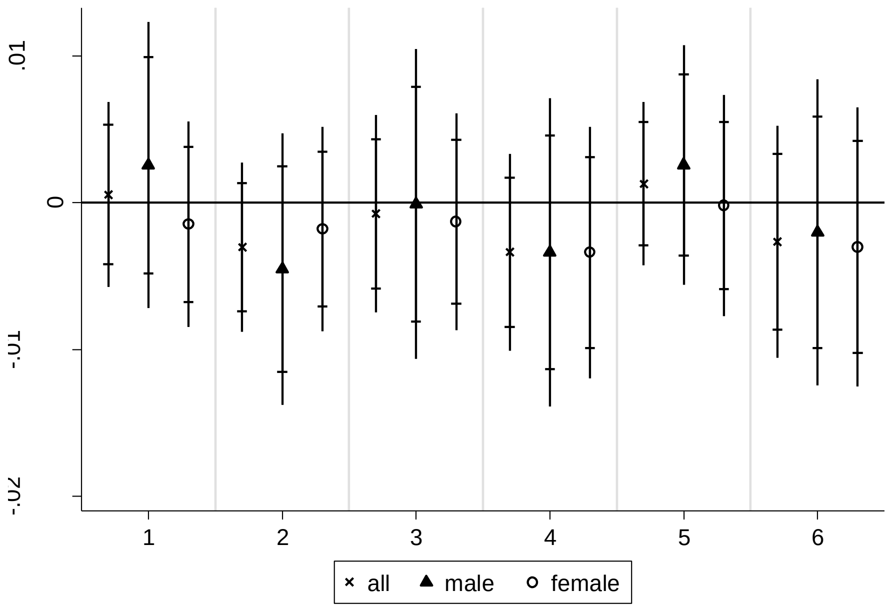


Source: Hourly air-quality report, Ministry of Environmental Protection (MEP) of the People's Republic of China.

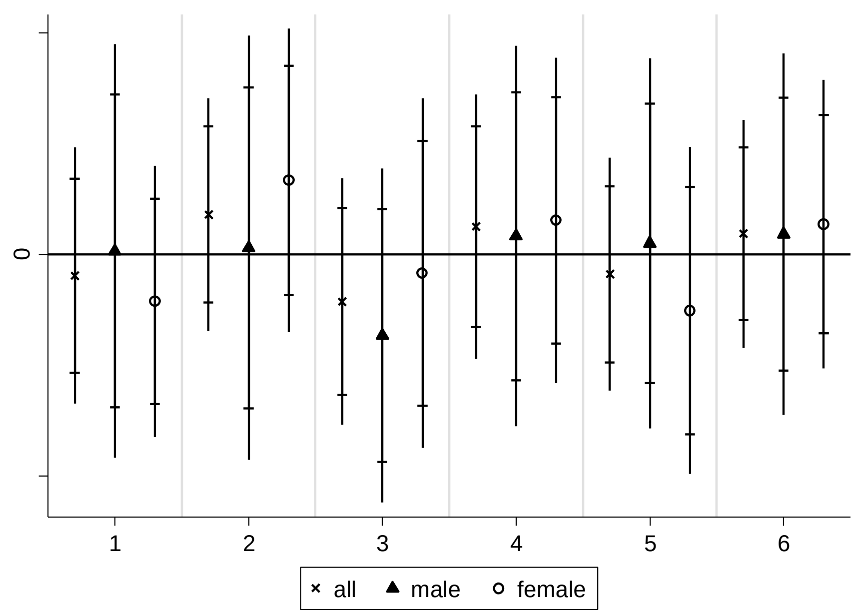
Note: The hourly mean pollution concentrations are calculated using the average values from all the monitoring stations in China. The left axis indicates the pollutant API that converts the corresponding pollutant measure in micrograms per cubic meter ($\mu g/m^3$) into an API score ranging from 0 to 500 using a formula devised by the MEP. The right axis indicates the distribution of interview time (percent). As this detailed air pollution component dataset has only been available since 2014, we cannot use it in our main empirical analysis. API = air pollution index; NO_2 = nitrogen dioxide; PM10 = particulate matter 10 micrometers or less in diameter; SO_2 = sulfur dioxide.

Figure S10: Falsification tests - Effects of air pollution on test scores in the days after the interview

Panel A: Verbal test scores



Panel B: Math test scores

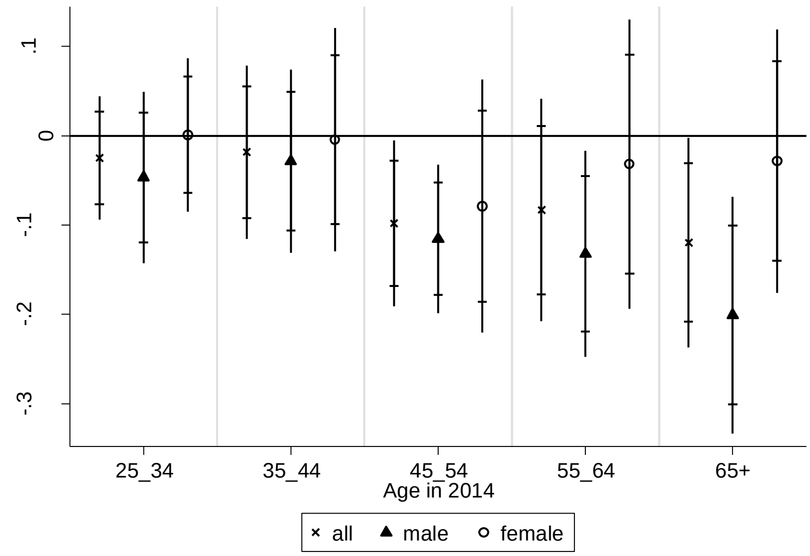


Source: Authors' estimations using CFPS survey 2010 and 2014.

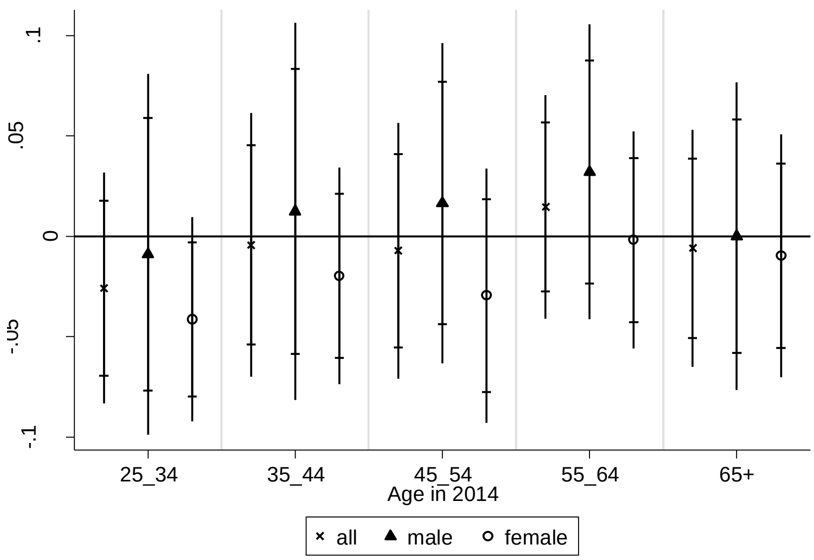
Note: The figure plots the coefficients with 95% and 99% confidence intervals from a regression of test scores on air pollution index (API) readings in the days after the interview. Other controls and fixed effects are the same as those presented in Table S1.

Figure S11: Robustness checks – giving interviews in winter months greater weights

A. Verbal test scores



B. Math test scores

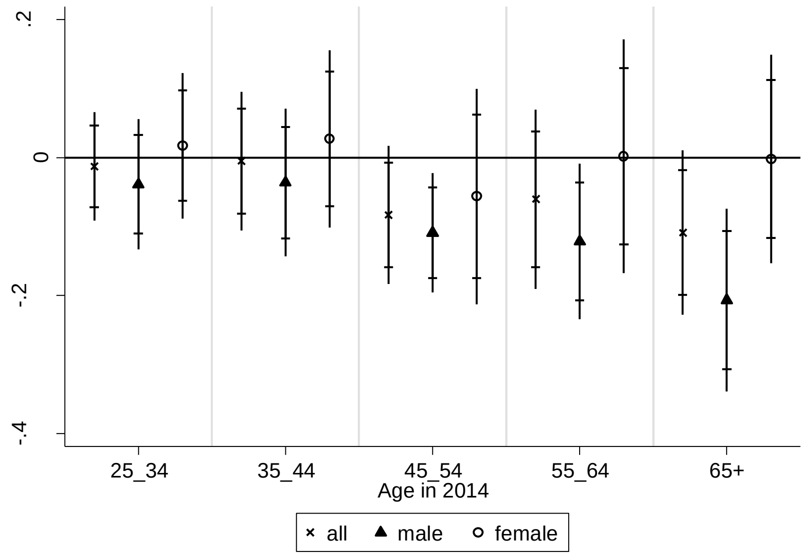


Source: Authors' estimations using CFPS survey 2010 and 2014.

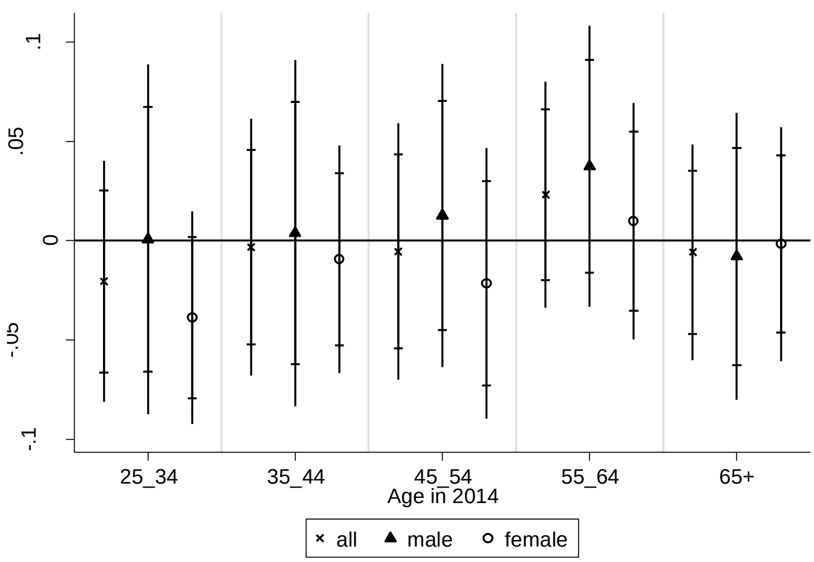
Note: The age-cohort effects include interaction terms between 3-year-mean API and age cohort dummies 25-34, 35-44, 45-54, 55-64, and 65+ in 2014. The age band 10-24 is the reference category. The figures plot the estimated coefficients on the interaction terms for the whole sample as well as the male and female subsamples with 95% and 99% confidence intervals. Air pollution data are matched between each CFPS county centroid and its nearest API reporting city boundary within a radius of 40km (i.e. 25miles). We divide the sample into two groups. Respondents in Group 1 were interviewed at least once in winter months (November, December and January), while respondents in Group 2 were only interviewed in non-winter months (from February to October). Observations in Group 1 are reweighted by the ratio of the number of observations in Group 2 divided by the number of observations in Group 1.

Figure S12: Robustness checks – using non-migrants

A. Verbal test scores



B. Math test scores

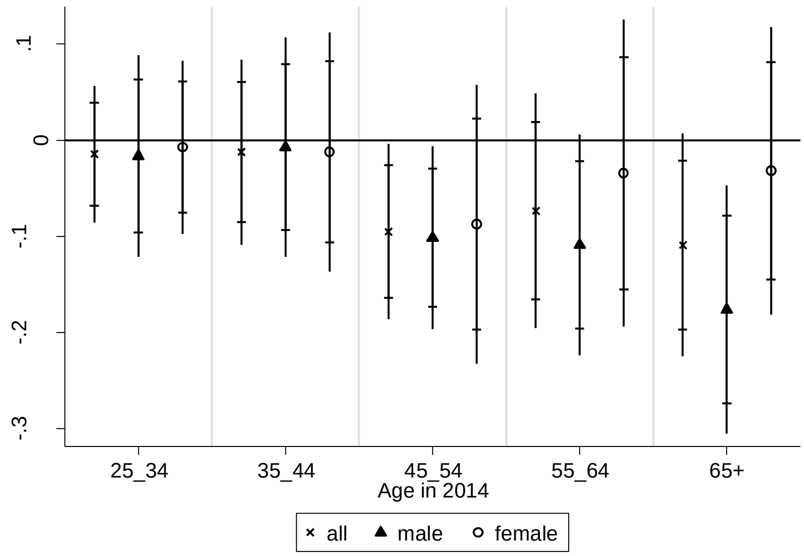


Source: Authors' estimations using CFPS survey 2010 and 2014.

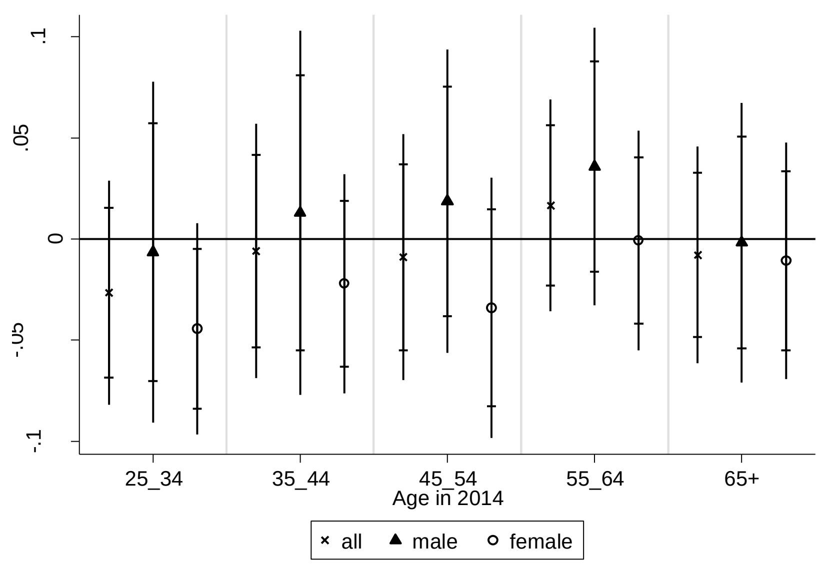
Note: The age-cohort effects include interaction terms between 3-year-mean API and age cohort dummies 25-34, 35-44, 45-54, 55-64, and 65+ in 2014. The age band 10-24 is the reference category. The figures plot the estimated coefficients on the interaction terms for the whole sample as well as the male and female subsamples with 95% and 99% confidence intervals. Air pollution data are matched between each CFPS county centroid and its nearest API reporting city boundary within a radius of 40km (i.e. 25miles). In all the regressions, we exclude the respondents who moved across counties between 2010 and 2014. In addition, we exclude respondents who moved into the latest addresses in 2008-2010 for the 3-year time window.

Figure S13: Robustness checks – polluted occupations excluded

A. Verbal test scores



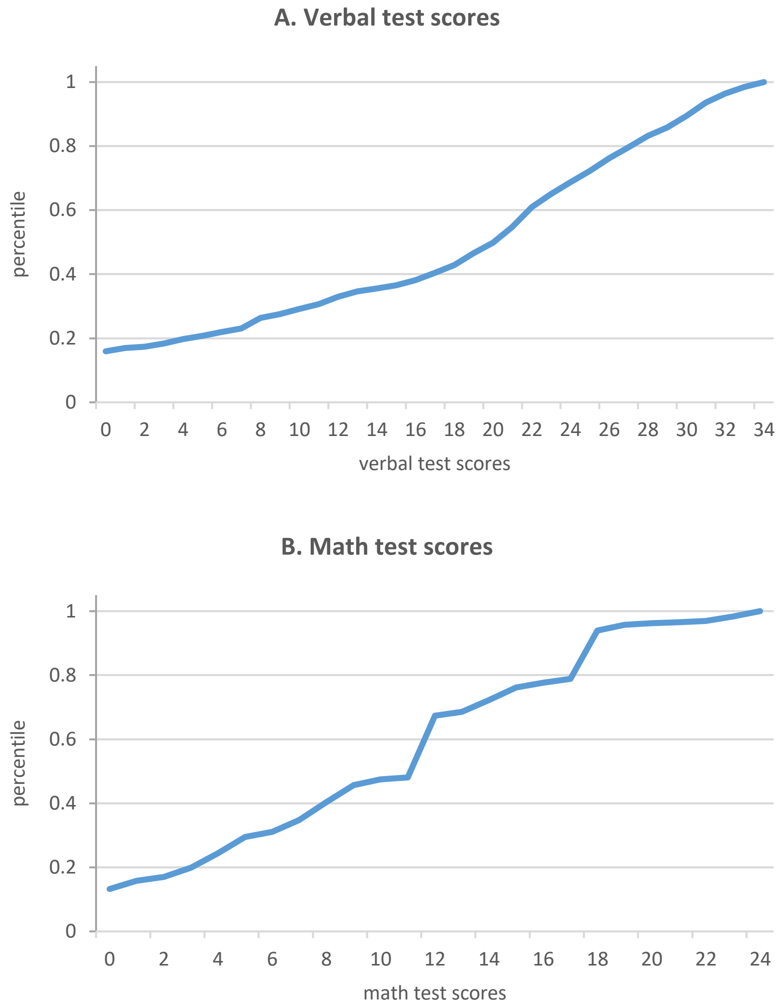
B. Math test scores



Source: Authors' estimations using CFPS survey 2010 and 2014.

Note: The age-cohort effects include interaction terms between 3-year-mean API and age cohort dummies 25-34, 35-44, 45-54, 55-64, and 65+ in 2014. The age band 10-24 is the reference category. The figures plot the estimated coefficients on the interaction terms for the whole sample as well as the male and female subsamples with 95% and 99% confidence intervals. Air pollution data are matched between each CFPS county centroid and its nearest API reporting city boundary within a radius of 40km (i.e. 25miles). The regressions exclude respondents with polluted jobs. Polluted occupations include "Geology and mineral industry workers", "Workers in metal smelting and refining industry", "Chemical product manufacturing personnel", "Textile workers", "Production workers (wood processing, artificial board, wood products, pulp and paper industry)", and "Production and processing worker (construction materials)".

Figure S14: Percentiles of cognitive test scores



Source: CFPS survey 2010 and 2014.
Note: The figure plots the percentiles of scores for verbal and math tests, respectively.

Table S1: Effects of air pollution on cognitive test scores

	1-day	7-day	30-day	90-day	1-year	2-year	3-year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Verbal test scores							
API_t	-0.004*	-0.001	0.000	-0.002	-0.002	-0.003	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.013**	-0.035***	-0.044***	-0.043***	-0.060***	-0.086***
		(0.005)	(0.008)	(0.011)	(0.012)	(0.016)	(0.021)
Observations	31,955	31,955	31,955	31,955	31,955	31,955	31,955
Overall *R*-squared	0.285	0.279	0.288	0.291	0.281	0.279	0.278
Impact of a one SD reduction in mean API on test scores (SDs of test scores)	0.131 (0.012)	0.278 (0.026)	0.599 (0.057)	0.712 (0.068)	0.895 (0.085)	0.942 (0.090)	1.132 (0.108)
B. Math test scores							
API_t	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.003	-0.004*	-0.009**	-0.007**	-0.010**	-0.016**
		(0.002)	(0.002)	(0.003)	(0.003)	(0.005)	(0.007)
Observations	31,955	31,955	31,955	31,955	31,955	31,955	31,955
Overall *R*-squared	0.449	0.440	0.455	0.447	0.451	0.450	0.447
Impact of a one SD reduction in mean API on test scores (SDs of test scores)	0.033 (0.005)	0.064 (0.010)	0.068 (0.011)	0.146 (0.023)	0.146 (0.023)	0.157 (0.025)	0.211 (0.033)

Source: Authors' estimations using CFPS survey 2010 and 2014.

Note: $\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$ indicates the mean of API readings in the past *k* days, where *k* equals 1, 7, 30, 90, 365, 730, and 1,095, respectively. All the regressions include individual fixed effects; county fixed effects; year, month, day of week, and post meridiem hour fixed effects; and a quadratic monthly time trend. Demographic controls include gender, age and its square and cubic terms, household per capita income, and years of education. Weather controls include 20°F indicators for temperature bins (that is, <25°F, 25–45°F, 45–65°F, 65–85°F, and >85°F), total precipitation, mean wind speed, and a dummy for bad weather. County-level characteristics include gross domestic product (GDP) per capita, population density, and industrial value share. Robust standard errors, clustered at the county level, are presented in parentheses. Air pollution data are matched between each CFPS county centroid and its nearest API reporting city boundary within a radius of 40km (i.e. 25miles). API = air pollution index; SD = standard deviation. *10% significance level; **5% significance level; ***1% significance level.

Table S2a: Effects of air pollution on verbal test scores, by gender

Dependent variable verbal test scores	1-day (1)	7-day (2)	30-day (3)	90-day (4)	1-year (5)	2-year (6)	3-year (7)
A. Male subsample							
API_t	-0.007**	-0.003	-0.001	-0.003	-0.004	-0.005	-0.005
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.014**	-0.047***	-0.053***	-0.052***	-0.072***	-0.103***
		(0.007)	(0.010)	(0.014)	(0.014)	(0.020)	(0.026)
Observations	15,315	15,315	15,315	15,315	15,315	15,315	15,315
Overall *R*-squared	0.195	0.193	0.193	0.194	0.187	0.187	0.186
B. Female subsample							
API_t	-0.002	0.001	0.000	-0.000	-0.001	-0.001	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.011**	-0.024***	-0.035***	-0.035***	-0.049***	-0.069***
		(0.005)	(0.007)	(0.010)	(0.011)	(0.014)	(0.018)
Observations	16,640	16,640	16,640	16,640	16,640	16,640	16,640
Overall *R*-squared	0.392	0.390	0.392	0.395	0.383	0.381	0.381
C. Gender differences							
Gender differences	-0.005	-0.003	-0.023***	-0.018*	-0.017**	-0.023*	-0.034**
Wald test chi-square	2.38	0.22	11.25	3.64	4.67	3.50	4.37

Source: Authors' estimations using CFPS survey 2010 and 2014.

Note: $\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$ indicates the mean of API readings in the past *k* days, where *k* equals 1, 7, 30, 90, 365, 730, and 1,095, respectively. All the regressions include individual fixed effects; county fixed effects; year, month, day of week, and post meridiem hour fixed effects; and a quadratic monthly time trend. Demographic controls include gender, age and its square and cubic terms, household per capita income, and years of education. Weather controls include 20°F indicators for temperature bins (that is, <25°F, 25–45°F, 45–65°F, 65–85°F, and >85°F), total precipitation, mean wind speed, and a dummy for bad weather. County-level characteristics include gross domestic product (GDP) per capita, population density, and industrial value share. Robust standard errors, clustered at the county level, are presented in parentheses. Air pollution data are matched between each CFPS county centroid and its nearest API reporting city boundary within a radius of 40km (i.e. 25miles). The significance of the male-female differences is derived from Wald tests. API = air pollution index; SD = standard deviation. *10% significance level; **5% significance level; ***1% significance level.

Table S2b: Effects of air pollution on math test scores, by gender

Dependent variable math test scores	1-day (1)	7-day (2)	30-day (3)	90-day (4)	1-year (5)	2-year (6)	3-year (7)
A. Male subsample							
API_t	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002
	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.001	-0.004	-0.007*	-0.005	-0.006	-0.012
		(0.003)	(0.003)	(0.004)	(0.004)	(0.006)	(0.008)
Observations	15,315	15,315	15,315	15,315	15,315	15,315	15,315
Overall *R*-squared	0.444	0.445	0.429	0.442	0.422	0.442	0.428
B. Female subsample							
API_t	-0.001	0.000	-0.001	-0.000	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.004	-0.004	-0.010**	-0.008**	-0.012*	-0.019**
		(0.003)	(0.003)	(0.005)	(0.004)	(0.006)	(0.009)
Observations	16,640	16,640	16,640	16,640	16,640	16,640	16,640
Overall *R*-squared	0.482	0.487	0.489	0.485	0.482	0.480	0.480
C. Gender differences							
Gender differences	-0.001	0.003	0.000	0.003	0.003	0.006	0.007
Wald test chi-square	1.15	0.75	0.01	0.17	0.52	0.65	0.54

Source: Authors' estimations using CFPS survey 2010 and 2014.

Note: $\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$ indicates the mean of API readings in the past *k* days, where *k* equals 1, 7, 30, 90, 365, 730, and 1,095, respectively. All the regressions include individual fixed effects; county fixed effects; year, month, day of week, and post meridiem hour fixed effects; and a quadratic monthly time trend. Demographic controls include gender, age and its square and cubic terms, household per capita income, and years of education. Weather controls include 20°F indicators for temperature bins (that is, <25°F, 25–45°F, 45–65°F, 65–85°F, and >85°F), total precipitation, mean wind speed, and a dummy for bad weather. County-level characteristics include gross domestic product (GDP) per capita, population density, and industrial value share. Robust standard errors, clustered at the county level, are presented in parentheses. Air pollution data are matched between each CFPS county centroid and its nearest API reporting city boundary within a radius of 40km (i.e. 25miles). The significance of the male-female differences is derived from Wald tests. API = air pollution index; SD = standard deviation. *10% significance level; **5% significance level; ***1% significance level.

Table S3: Age-cohort effects of air pollution on cognitive test scores, by gender

Dependent variable	Verbal test scores				Math test scores			
	All	Male	Female	Gender differences (chi-square)	All	Male	Female	Gender differences (chi-square)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
3-year-mean API×age25-34	-0.017	-0.028	-0.003	-0.025	-0.024	-0.004	-0.044**	0.040
	(0.027)	(0.037)	(0.034)	(0.29)	(0.022)	(0.033)	(0.019)	(1.69)
3-year-mean API×age35-44	-0.015	-0.020	-0.007	-0.013	-0.004	0.012	-0.020	0.032
	(0.038)	(0.040)	(0.048)	(0.08)	(0.024)	(0.035)	(0.020)	(1.22)
3-year-mean API×age45-54	-0.097***	-0.108***	-0.084	-0.024	-0.008	0.016	-0.032	0.048*
	(0.036)	(0.033)	(0.055)	(0.18)	(0.024)	(0.030)	(0.024)	(3.48)
3-year-mean API×age55-64	-0.077	-0.120***	-0.030	-0.090*	0.017	0.035	0.001	0.034
	(0.049)	(0.044)	(0.064)	(3.09)	(0.021)	(0.027)	(0.021)	(1.80)
3-year-mean API×age65+	-0.114**	-0.192***	-0.026	-0.166***	-0.007	-0.004	-0.010	0.006
	(0.046)	(0.050)	(0.057)	(8.06)	(0.021)	(0.028)	(0.022)	(0.04)
Observations	31,955	15,315	16,640		31,955	15,315	16,640	
Overall *R*-squared	0.644	0.644	0.644		0.644	0.644	0.644	

Source: Authors' estimations using CFPS survey 2010 and 2014.

Note: Other covariates include API on the interview date, 3-year-mean API and age cohort dummies 25-34, 35-44, 45-54, 55-64, and 65+ in 2014. The age band 10-24 is the reference category. All the regressions include individual fixed effects; county fixed effects; year, month, day of week, and post meridiem hour fixed effects; and a quadratic monthly time trend. Demographic controls include gender, household per capita income, and years of education. Weather controls include 20°F indicators for temperature bins (that is, <25°F, 25–45°F, 45–65°F, 65–85°F, and >85°F), total precipitation, mean wind speed, and a dummy for bad weather. County-level characteristics include gross domestic product (GDP) per capita, population density, and industrial value share. Robust standard errors, clustered at the county level, are presented in parentheses. Air pollution data are matched between each CFPS county centroid and its nearest API reporting city boundary within a radius of 40km (i.e. 25miles). The significance of the male-female differences is derived from Wald tests. API = air pollution index; SD = standard deviation. *10% significance level; **5% significance level; ***1% significance level.

Table S4a: Age-cohort effects of air pollution on verbal test scores, by gender and education level

Dependent variable verbal test scores	A. Primary school or below				B. Middle school or above			
	All	Male	Female	Gender differences (chi-square)	All	Male	Female	Gender differences (chi-square)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
3-year-mean API×age25-34	-0.319**	-0.545***	-0.008	-0.537**	0.006	0.003	0.015	-0.012
	(0.154)	(0.195)	(0.176)	(6.06)	(0.027)	(0.039)	(0.037)	(0.05)
3-year-mean API×age35-44	-0.248	-0.437**	0.037	-0.474**	0.006	-0.004	0.030	-0.034
	(0.151)	(0.199)	(0.144)	(5.37)	(0.033)	(0.035)	(0.049)	(0.44)
3-year-mean API×age45-54	-0.336**	-0.649***	0.018	-0.667***	-0.070**	-0.058*	-0.071	0.013
	(0.136)	(0.146)	(0.146)	(14.61)	(0.033)	(0.035)	(0.052)	(0.05)
3-year-mean API×age55-64	-0.272*	-0.650***	0.114	-0.764***	-0.083**	-0.063	-0.098	0.035
	(0.161)	(0.178)	(0.158)	(19.61)	(0.039)	(0.042)	(0.064)	(0.27)
3-year-mean API×age65+	-0.320**	-0.694***	0.089	-0.783***	-0.126***	-0.142***	-0.038	-0.104
	(0.145)	(0.163)	(0.154)	(19.29)	(0.043)	(0.050)	(0.082)	(1.36)
Observations	12,515	4,927	7,588		19,436	10,385	9,051	
Overall *R*-squared	0.166	0.166	0.166		0.166	0.166	0.166	

Source: Authors' estimations using CFPS survey 2010 and 2014.

Note: Other covariates include API on the interview date, 3-year-mean API and age cohort dummies 25-34, 35-44, 45-54, 55-64, and 65+ in 2014. The age band 10-24 is the reference category. All the regressions include individual fixed effects; county fixed effects; year, month, day of week, and post meridiem hour fixed effects; and a quadratic monthly time trend. Demographic controls include gender, household per capita income, and years of education. Weather controls include 20°F indicators for temperature bins (that is, <25°F, 25–45°F, 45–65°F, 65–85°F, and >85°F), total precipitation, mean wind speed, and a dummy for bad weather. County-level characteristics include gross domestic product (GDP) per capita, population density, and industrial value share. Robust standard errors, clustered at the county level, are presented in parentheses. Air pollution data are matched between each CFPS county centroid and its nearest API reporting city boundary within a radius of 40km (i.e. 25miles). The significance of the male-female difference is derived from Wald tests. API = air pollution index; SD = standard deviation. *10% significance level; **5% significance level; ***1% significance level.

Table S4b: Age-cohort effects of air pollution on math test scores, by gender and education level

Dependent variable math test scores	A. Primary school or below				B. Middle school or above			
	All	Male	Female	Gender differences (chi-square)	All	Male	Female	Gender differences (chi-square)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
3-year-mean API×age25-34	-0.110	-0.111	-0.093	-0.018	-0.011	0.005	-0.027	0.032
	(0.068)	(0.120)	(0.093)	(0.01)	(0.022)	(0.032)	(0.021)	(1.04)
3-year-mean API×age35-44	-0.107	-0.104	-0.092	-0.012	0.019	0.024	0.018	0.006
	(0.074)	(0.128)	(0.071)	(0.01)	(0.025)	(0.032)	(0.024)	(0.05)
3-year-mean API×age45-54	-0.100	-0.120	-0.073	-0.047	0.015	0.037	-0.007	0.044
	(0.073)	(0.113)	(0.074)	(0.12)	(0.023)	(0.029)	(0.025)	(2.22)
3-year-mean API×age55-64	-0.066	-0.096	-0.033	-0.063	0.046**	0.065**	0.027	0.038
	(0.072)	(0.113)	(0.078)	(0.19)	(0.020)	(0.025)	(0.027)	(1.43)
3-year-mean API×age65+	-0.082	-0.114	-0.041	-0.073	0.005	0.018	-0.002	0.020
	(0.069)	(0.110)	(0.079)	(0.26)	(0.024)	(0.026)	(0.043)	(0.22)
Observations	12,515	4,927	7,588		19,436	10,385	9,051	
Overall *R*-squared	0.268	0.268	0.268		0.268	0.268	0.268	

Source: Authors' estimations using CFPS survey 2010 and 2014.

Note: Other covariates include API on the interview date, 3-year-mean API and age cohort dummies 25-34, 35-44, 45-54, 55-64, and 65+ in 2014. The age band 10-24 is the reference category. All the regressions include individual fixed effects; county fixed effects; year, month, day of week, and post meridiem hour fixed effects; and a quadratic monthly time trend. Demographic controls include gender, household per capita income, and years of education. Weather controls include 20°F indicators for temperature bins (that is, <25°F, 25–45°F, 45–65°F, 65–85°F, and >85°F), total precipitation, mean wind speed, and a dummy for bad weather. County-level characteristics include gross domestic product (GDP) per capita, population density, and industrial value share. Robust standard errors, clustered at the county level, are presented in parentheses. Air pollution data are matched between each CFPS county centroid and its nearest API reporting city boundary within a radius of 40km (i.e. 25miles). The significance of the male-female difference is derived from Wald tests. API = air pollution index; SD = standard deviation. *10% significance level; **5% significance level; ***1% significance level.

Table S5: Breakpoints for API value calculation

API index value	PM10 (μg/m^3)	SO_2 (μg/m^3)	NO_2 (μg/m^3)
0	0	0	0
50	50	50	40
100	150	150	80
150	250	475	180
200	350	800	280
300	420	1600	565
400	500	2100	750
500	600	2620	940

Note: API = air pollution index; NO_2 = nitrogen dioxide; PM10 = particulate matter 10 micrometers or less in diameter; SO_2 = sulfur dioxide.

Table S6: Summary statistics

Variable	All		Male		Female	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**
verbal test scores	18.115	10.489	19.729	9.430	16.629	11.172
math test scores	10.438	6.403	11.496	5.924	9.464	6.667
API	73.516	32.684	73.197	31.714	73.810	33.549
7-day mean API	72.885	21.360	72.619	21.108	73.130	21.587
30-day mean API	72.992	17.118	72.801	17.078	73.168	17.153
90-day mean API	75.516	16.184	75.342	16.133	75.676	16.231
1-year mean API	84.002	20.806	83.822	20.863	84.167	20.753
2-year mean API	77.738	15.706	77.572	15.782	77.891	15.634
3-year mean API	74.882	13.166	74.705	13.227	75.044	13.108
log form of household per capita income (Chinese *yuan*)	8.874	1.154	8.891	1.153	8.858	1.155
age	44.742	17.892	44.925	18.158	44.573	17.642
years of education	7.475	4.451	8.220	4.058	6.789	4.681

Source: Authors' estimations using CFPS survey 2010 and 2014.
Note: API = air pollution index; SD = standard deviation.

Table S7: Robustness checks — matching distance to the boundaries of API reporting cities (verbal test scores)

Dependent variable	1-day	7-day	30-day	90-day	1-year	2-year	3-year
verbal test scores	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. matching distance = 0 km							
API_t	-0.009***	-0.007**	-0.006**	-0.006**	-0.007**	-0.007***	-0.008***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.004	-0.021**	-0.040***	-0.031***	-0.042**	-0.060**
		(0.006)	(0.009)	(0.013)	(0.011)	(0.016)	(0.024)
Observations	17,160	17,160	17,160	17,160	17,160	17,160	17,160
Overall *R*-squared	0.297	0.295	0.294	0.295	0.294	0.295	0.290
B. matching distance = 10 km							
API_t	-0.008***	-0.005**	-0.005**	-0.006**	-0.007***	-0.007***	-0.007***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.010*	-0.030***	-0.045***	-0.027**	-0.036**	-0.060**
		(0.005)	(0.008)	(0.013)	(0.010)	(0.016)	(0.024)
Observations	20,049	20,049	20,049	20,049	20,049	20,049	20,049
Overall *R*-squared	0.278	0.278	0.273	0.276	0.270	0.274	0.269
C. matching distance = 20 km							
API_t	-0.007***	-0.004*	-0.003*	-0.005**	-0.005**	-0.006***	-0.006***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.009*	-0.033***	-0.044***	-0.032**	-0.040**	-0.068***
		(0.005)	(0.009)	(0.015)	(0.014)	(0.017)	(0.023)
Observations	23,423	23,423	23,423	23,423	23,423	23,423	23,423
Overall *R*-squared	0.278	0.277	0.275	0.275	0.274	0.277	0.276

Table S7 (continued): Robustness checks — matching distance to the boundaries of API reporting cities (verbal test scores)

Dependent variable verbal test scores	1-day (1)	7-day (2)	30-day (3)	90-day (4)	1-year (5)	2-year (6)	3-year (7)
D. matching distance = 30 km							
API_t	-0.006***	-0.003	-0.002	-0.004	-0.005**	-0.005**	-0.005**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.011**	-0.035***	-0.043***	-0.033***	-0.041**	-0.062***
		(0.005)	(0.008)	(0.012)	(0.012)	(0.016)	(0.022)
Observations	28,000	28,000	28,000	28,000	28,000	28,000	28,000
Overall R-squared	0.262	0.262	0.263	0.255	0.254	0.259	0.253
E. matching distance = 40 km							
API_t	-0.004*	-0.001	0.000	-0.002	-0.002	-0.003	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.013**	-0.035***	-0.044***	-0.043***	-0.060***	-0.086***
		(0.005)	(0.008)	(0.011)	(0.012)	(0.016)	(0.021)
Observations	31,955	31,955	31,955	31,955	31,955	31,955	31,955
Overall R-squared	0.285	0.279	0.288	0.291	0.281	0.279	0.278
F. matching distance = 50 km							
API_t	-0.004*	-0.000	0.001	-0.001	-0.002	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.015***	-0.038***	-0.047***	-0.039***	-0.055***	-0.082***
		(0.005)	(0.008)	(0.010)	(0.010)	(0.014)	(0.019)
Observations	33,953	33,953	33,953	33,953	33,953	33,953	33,953
Overall R-squared	0.271	0.277	0.270	0.270	0.268	0.272	0.274

Table S7 (continued): Robustness checks — matching distance to the boundaries of API reporting cities (verbal test scores)

Dependent variable verbal test scores	1-day (1)	7-day (2)	30-day (3)	90-day (4)	1-year (5)	2-year (6)	3-year (7)
G. matching distance = 60 km							
API_t	-0.004*	0.000	0.001	-0.001	-0.002	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.015***	-0.037***	-0.045***	-0.036***	-0.050***	-0.072***
		(0.005)	(0.008)	(0.010)	(0.009)	(0.013)	(0.018)
Observations	35,265	35,265	35,265	35,265	35,265	35,265	35,265
Overall *R*-squared	0.278	0.269	0.272	0.270	0.269	0.273	0.269
H. matching distance = 70 km							
API_t	-0.004*	0.000	0.001	-0.001	-0.002	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.015***	-0.035***	-0.042***	-0.034***	-0.045***	-0.066***
		(0.005)	(0.008)	(0.010)	(0.009)	(0.012)	(0.017)
Observations	36,388	36,388	36,388	36,388	36,388	36,388	36,388
Overall *R*-squared	0.266	0.274	0.279	0.270	0.264	0.265	0.269
I. matching distance = 80 km							
API_t	-0.003	0.001	0.001	-0.000	-0.001	-0.001	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.012**	-0.032***	-0.039***	-0.034***	-0.045***	-0.060***
		(0.005)	(0.008)	(0.009)	(0.008)	(0.012)	(0.016)
Observations	38,142	38,142	38,142	38,142	38,142	38,142	38,142
Overall *R*-squared	0.273	0.278	0.277	0.281	0.268	0.273	0.278

See the notes to Table S1. As APIs are reported at the city level, we calculate the distance between the centroid of each CFPS county and the boundary of API reporting cities. If a CFPS county is located within an API reporting city, we set the matching distance as zero and treat the city's API readings as the CFPS county's readings. If a CFPS county is not located in any cities with API readings, we match it to the nearest API reporting city within a specific radius between the CFPS county centroids and the city boundaries. The table reports estimation results based on nine cutoff radiuses ranging from zero km to 80 km.

Table S8: Robustness checks — matching distance to the boundaries of API reporting cities (math test scores)

Dependent variable	1-day	7-day	30-day	90-day	1-year	2-year	3-year
math test scores	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. matching distance = 0 km							
API_t	-0.003**	-0.002	-0.003*	-0.002	-0.003*	-0.003**	-0.003**
	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.004	-0.006	-0.018***	-0.010***	-0.013**	-0.020***
		(0.003)	(0.005)	(0.005)	(0.003)	(0.005)	(0.007)
Observations	17,160	17,160	17,160	17,160	17,160	17,160	17,160
Overall *R*-squared	0.388	0.382	0.389	0.385	0.388	0.391	0.387
B. matching distance = 10 km							
API_t	-0.002	-0.001	-0.001	-0.001	-0.002	-0.002	-0.002
	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.003	-0.006*	-0.016***	-0.007**	-0.008*	-0.015**
		(0.002)	(0.004)	(0.005)	(0.003)	(0.005)	(0.007)
Observations	20,049	20,049	20,049	20,049	20,049	20,049	20,049
Overall *R*-squared	0.425	0.407	0.421	0.428	0.426	0.432	0.424
C. matching distance = 20 km							
API_t	-0.002	-0.001	-0.001	-0.001	-0.001	-0.002	-0.002
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.004*	-0.009***	-0.016***	-0.006*	-0.005	-0.012**
		(0.002)	(0.003)	(0.005)	(0.003)	(0.004)	(0.006)
Observations	23,423	23,423	23,423	23,423	23,423	23,423	23,423
Overall *R*-squared	0.447	0.445	0.444	0.450	0.453	0.452	0.452

Table S8 (continued): Robustness checks — matching distance to the boundaries of API reporting cities (math test scores)

Dependent variable math test scores	1-day (1)	7-day (2)	30-day (3)	90-day (4)	1-year (5)	2-year (6)	3-year (7)
D. matching distance = 30 km							
API_t	-0.002	-0.001	-0.001	-0.001	-0.001	-0.002	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.004*	-0.005*	-0.010**	-0.006*	-0.006	-0.010*
		(0.002)	(0.003)	(0.004)	(0.003)	(0.004)	(0.006)
Observations	28,000	28,000	28,000	28,000	28,000	28,000	28,000
Overall *R*-squared	0.394	0.395	0.395	0.391	0.388	0.392	0.387
E. matching distance = 40 km							
API_t	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.003	-0.004*	-0.009**	-0.007**	-0.010**	-0.016**
		(0.002)	(0.002)	(0.003)	(0.003)	(0.005)	(0.007)
Observations	31,955	31,955	31,955	31,955	31,955	31,955	31,955
Overall *R*-squared	0.449	0.440	0.455	0.447	0.451	0.450	0.447
F. matching distance = 50 km							
API_t	-0.002	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.003	-0.005**	-0.012***	-0.011***	-0.013**	-0.020**
		(0.002)	(0.002)	(0.004)	(0.004)	(0.006)	(0.008)
Observations	33,953	33,953	33,953	33,953	33,953	33,953	33,953
Overall *R*-squared	0.433	0.436	0.433	0.433	0.427	0.431	0.436

Table S8 (continued): Robustness checks — matching distance to the boundaries of API reporting cities (math test scores)

Dependent variable math test scores	1-day (1)	7-day (2)	30-day (3)	90-day (4)	1-year (5)	2-year (6)	3-year (7)
G. matching distance = 60 km							
API_t	-0.002*	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.003	-0.005**	-0.010***	-0.010***	-0.013**	-0.019**
		(0.002)	(0.002)	(0.004)	(0.003)	(0.005)	(0.008)
Observations	35,265	35,265	35,265	35,265	35,265	35,265	35,265
Overall *R*-squared	0.410	0.405	0.404	0.405	0.403	0.401	0.402
H. matching distance = 70 km							
API_t	-0.002*	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.003	-0.005**	-0.012***	-0.009**	-0.011*	-0.017**
		(0.002)	(0.002)	(0.004)	(0.004)	(0.006)	(0.007)
Observations	36,388	36,388	36,388	36,388	36,388	36,388	36,388
Overall *R*-squared	0.393	0.397	0.398	0.391	0.384	0.395	0.390
I. matching distance = 80 km							
API_t	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.002	-0.004*	-0.011***	-0.009***	-0.010**	-0.016**
		(0.002)	(0.002)	(0.004)	(0.003)	(0.005)	(0.007)
Observations	38,142	38,142	38,142	38,142	38,142	38,142	38,142
Overall *R*-squared	0.400	0.405	0.399	0.404	0.389	0.399	0.394

See the notes to Table S1. As APIs are reported at the city level, we calculate the distance between the centroid of each CFPS county and the boundary of API reporting cities. If a CFPS county is located within an API reporting city, we set the matching distance as zero and treat the city's API readings as the CFPS county's readings. If a CFPS county is not located in any cities with API readings, we match it to the nearest API reporting city within a specific radius between the CFPS county centroids and the city boundaries. The table reports estimation results based on nine cutoff radiuses ranging from zero km to 90 km.

Table S9: Robustness checks — matching distance to monitoring stations (verbal test scores)

Dependent variable	1-day	7-day	30-day	90-day	1-year	2-year	3-year
verbal test scores	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. matching distance = 40 km							
API_t	-0.008**	-0.001	-0.004	-0.004	-0.005	-0.006*	-0.006*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.023***	-0.035***	-0.047***	-0.040***	-0.057**	-0.083**
		(0.007)	(0.007)	(0.011)	(0.015)	(0.023)	(0.033)
Observations	18,996	18,996	18,996	18,996	18,996	18,996	18,996
Overall *R*-squared	0.246	0.250	0.256	0.249	0.247	0.246	0.246
B. matching distance = 50 km							
API_t	-0.007**	-0.000	-0.001	-0.003	-0.004	-0.005*	-0.005*
	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.024***	-0.044***	-0.051***	-0.033**	-0.044*	-0.058
		(0.007)	(0.009)	(0.012)	(0.015)	(0.025)	(0.038)
Observations	21,770	21,770	21,770	21,770	21,770	21,770	21,770
Overall *R*-squared	0.260	0.255	0.259	0.258	0.254	0.258	0.256
C. matching distance = 60 km							
API_t	-0.007**	-0.001	-0.000	-0.003	-0.004	-0.004	-0.004
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.022***	-0.049***	-0.047***	-0.041***	-0.055**	-0.078***
		(0.007)	(0.010)	(0.013)	(0.015)	(0.021)	(0.029)
Observations	24,879	24,879	24,879	24,879	24,879	24,879	24,879
Overall *R*-squared	0.245	0.252	0.251	0.250	0.243	0.247	0.251

Table S9 (continued): Robustness checks — matching distance to monitoring stations (verbal test scores)

Dependent variable	1-day	7-day	30-day	90-day	1-year	2-year	3-year
verbal test scores	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			D. matching distance = 70 km				
API_t	-0.007***	-0.002	-0.001	-0.004*	-0.005*	-0.005*	-0.005*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.019**	-0.054***	-0.052***	-0.046***	-0.069***	-0.096***
		(0.008)	(0.010)	(0.013)	(0.016)	(0.023)	(0.033)
Observations	26,723	26,723	26,723	26,723	26,723	26,723	26,723
Overall *R*-squared	0.262	0.255	0.261	0.256	0.252	0.247	0.249
			E. matching distance = 80 km				
API_t	-0.005*	-0.002	0.001	-0.002	-0.002	-0.003	-0.003
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.013**	-0.050***	-0.046***	-0.041***	-0.050**	-0.067**
		(0.006)	(0.010)	(0.015)	(0.015)	(0.023)	(0.031)
Observations	30,086	30,086	30,086	30,086	30,086	30,086	30,086
Overall *R*-squared	0.237	0.247	0.243	0.237	0.241	0.242	0.242
			F. matching distance = 90 km				
API_t	-0.005*	-0.001	0.000	-0.002	-0.003	-0.003	-0.004
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.014**	-0.048***	-0.045***	-0.037***	-0.043**	-0.053**
		(0.006)	(0.009)	(0.014)	(0.012)	(0.019)	(0.024)
Observations	31,624	31,624	31,624	31,624	31,624	31,624	31,624
Overall *R*-squared	0.241	0.244	0.246	0.246	0.242	0.249	0.247

See the notes to Table S1. The city-level APIs are computed based on readings from multiple monitoring stations in each API reporting city, but the readings from the monitoring stations along with their latitude and longitude information have been available since 2014. We assign the reported city-level APIs in 2010 to all stations within the city boundary under the assumption that that the geographic locations of air quality monitoring stations do not change much between 2010 and 2014 and variations in daily air pollution across monitoring stations within each city are small. We match CFPS counties to the nearest air quality monitoring station within a specific radius between the CFPS county centroids and the monitoring stations. **All the results are weighted by the inverse distance to monitoring stations.** The table reports estimation results based on six cutoff radiuses ranging from 40 km to 90 km.

Table S10: Robustness checks — matching distance to monitoring stations (math test scores)

Dependent variable math test scores	1-day (1)	7-day (2)	30-day (3)	90-day (4)	1-year (5)	2-year (6)	3-year (7)
A. matching distance = 40 km							
API_t	-0.005***	-0.003**	-0.004***	-0.004***	-0.004***	-0.004***	-0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.006*	-0.003	-0.010*	-0.009**	-0.011**	-0.014*
		(0.003)	(0.004)	(0.005)	(0.004)	(0.005)	(0.007)
Observations	18,996	18,996	18,996	18,996	18,996	18,996	18,996
Overall *R*-squared	0.371	0.366	0.368	0.370	0.361	0.372	0.370
B. matching distance = 50 km							
API_t	-0.005***	-0.004***	-0.004***	-0.004***	-0.004***	-0.005***	-0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.005*	-0.007**	-0.012***	-0.008**	-0.011*	-0.011
		(0.003)	(0.003)	(0.005)	(0.004)	(0.006)	(0.009)
Observations	21,770	21,770	21,770	21,770	21,770	21,770	21,770
Overall *R*-squared	0.361	0.354	0.360	0.361	0.360	0.360	0.363
C. matching distance = 60 km							
API_t	-0.003*	-0.001	-0.002	-0.002	-0.002	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.006**	-0.009***	-0.011**	-0.009**	-0.012**	-0.016**
		(0.003)	(0.003)	(0.004)	(0.004)	(0.006)	(0.008)
Observations	24,879	24,879	24,879	24,879	24,879	24,879	24,879
Overall *R*-squared	0.335	0.336	0.332	0.332	0.328	0.332	0.332

Table S10 (continued): Robustness checks — matching distance to monitoring stations (math test scores)

Dependent variable math test scores	1-day (1)	7-day (2)	30-day (3)	90-day (4)	1-year (5)	2-year (6)	3-year (7)
D. matching distance = 70 km							
API_t	-0.002*	-0.001	-0.001	-0.002	-0.002	-0.002	-0.002
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.004	-0.009***	-0.010**	-0.007*	-0.008	-0.014*
		(0.003)	(0.003)	(0.004)	(0.004)	(0.006)	(0.008)
Observations	26,723	26,723	26,723	26,723	26,723	26,723	26,723
Overall R-squared	0.341	0.338	0.341	0.338	0.342	0.340	0.336
E. matching distance = 80 km							
API_t	-0.000	0.001	0.000	0.000	-0.000	-0.000	-0.000
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.004	-0.008**	-0.008	-0.003	0.000	-0.002
		(0.003)	(0.004)	(0.005)	(0.004)	(0.007)	(0.009)
Observations	30,086	30,086	30,086	30,086	30,086	30,086	30,086
Overall R-squared	0.346	0.350	0.348	0.348	0.350	0.357	0.355
F. matching distance = 90 km							
API_t	-0.001	0.000	0.000	-0.000	-0.000	-0.001	-0.001
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.004	-0.008**	-0.007	-0.002	-0.000	-0.002
		(0.002)	(0.003)	(0.005)	(0.004)	(0.005)	(0.007)
Observations	31,624	31,624	31,624	31,624	31,624	31,624	31,624
Overall R-squared	0.339	0.342	0.340	0.342	0.347	0.346	0.349

See the notes to Table S1. The city-level APIs are computed based on readings from multiple monitoring stations in each API reporting city, but the readings from the monitoring stations along with their latitude and longitude information have been available since 2014. We assign the reported city-level APIs in 2010 to all stations within the city boundary under the assumption that that the geographic locations of air quality monitoring stations do not change much between 2010 and 2014 and variations in daily air pollution across monitoring stations within each city are small. We match CFPS counties to the nearest air quality monitoring station within a specific radius between the CFPS county centroids and the monitoring stations. **All the results are weighted by the inverse distance to monitoring stations.** The table reports estimation results based on six cutoff radiuses ranging from 40 km to 90 km.

Table S11: Robustness checks - alternative specifications of Table S1

	7-day mean			90-day mean			1-year mean		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A: Verbal test scores									
API_t	-0.001	-0.000	-0.001	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$	-0.012**	-0.011**	-0.013**	-0.039***	-0.027**	-0.044***	-0.040***	-0.035***	-0.043***
	(0.005)	(0.004)	(0.005)	(0.011)	(0.011)	(0.011)	(0.012)	(0.011)	(0.012)
Income per capita		0.395***	0.152*		0.400***	0.165*		0.397***	0.163**
		(0.044)	(0.086)		(0.044)	(0.084)		(0.044)	(0.081)
Years of education		1.444***	0.447***		1.443***	0.456***		1.443***	0.435***
		(0.021)	(0.128)		(0.021)	(0.128)		(0.021)	(0.125)
Observations	33,803	31,955	31,955	33,803	31,955	31,955	33,803	31,955	31,955
Overall *R*-squared	0.279	0.279	0.279	0.291	0.291	0.291	0.281	0.281	0.281
B: Math test scores									
API_t	-0.000	-0.000	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$	-0.003*	-0.002	-0.003	-0.007**	-0.005	-0.009**	-0.007**	-0.008**	-0.007**
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Income per capita		0.137***	0.022		0.138***	0.024		0.138***	0.023
		(0.028)	(0.042)		(0.028)	(0.041)		(0.028)	(0.041)
Years of education		1.136***	0.703***		1.136***	0.705***		1.136***	0.701***
		(0.010)	(0.060)		(0.010)	(0.060)		(0.010)	(0.060)
Observations	33,803	31,955	31,955	33,803	31,955	31,955	33,803	31,955	31,955
Overall *R*-squared	0.440	0.440	0.440	0.447	0.447	0.447	0.451	0.451	0.451
Individual fixed effects	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
County-level characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Source: Authors' estimations using CFPS survey 2010 and 2014.

Note: $\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$ indicates the mean of the air pollution index (API) in the past *k* days, where *k* equals 1, 7, and 365, respectively. All the regressions include county fixed effects; year, month, day of week, and post meridiem hour fixed effects; and a monthly quadratic time trend. Other demographic controls include gender, age and its square and cubic terms. Weather controls include 20°F indicators for temperature bins (that is, <25°F, 25–45°F, 45–65°F, 65–85°F, and >85°F), total precipitation, mean wind speed, and a dummy for bad weather. County-level characteristics include gross domestic product (GDP) per capita, population density, and industrial value share. Robust standard errors, clustered at the county level, are presented in parentheses. Air pollution data are matched between each CFPS county centroid and its nearest API reporting city boundary within a radius of 40km (i.e. 25miles). *10% significance level; **5% significance level; ***1% significance level.

Table S12: Robustness checks – giving interviews in winter months greater weights

	1-day	7-day	30-day	90-day	1-year	2-year	3-year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Verbal test scores							
API_t	-0.005*	-0.001	-0.000	-0.002	-0.002	-0.003	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.012**	-0.036***	-0.044***	-0.042***	-0.058***	-0.083***
		(0.005)	(0.008)	(0.011)	(0.012)	(0.016)	(0.021)
Observations	31,955	31,955	31,955	31,955	31,955	31,955	31,955
Overall *R*-squared	0.279	0.279	0.279	0.279	0.279	0.279	0.279
B. Math test scores							
API_t	-0.002	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.002	-0.004*	-0.009**	-0.007**	-0.009*	-0.015**
		(0.002)	(0.002)	(0.004)	(0.003)	(0.005)	(0.007)
Observations	31,955	31,955	31,955	31,955	31,955	31,955	31,955
Overall *R*-squared	0.432	0.432	0.432	0.432	0.432	0.432	0.432

See the notes to Table S1. We divide the sample into two groups. Respondents in Group 1 were interviewed at least once in winter months (November, December and January), while respondents in Group 2 were only interviewed in non-winter months (from February to October). Observations in Group 1 are reweighted by the ratio of the number of observations in Group 2 divided by the number of observations in Group 1.

Table S13: Mechanism tests – interviewees' impatience and cooperation

	1-day	7-day	30-day	90-day	1-year	2-year	3-year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Interviewees' impatience on a scale from 1 (*low*) to 7 (*high*)							
API_t	0.005	-0.015	0.023	-0.004	-0.003	0.004	0.005
	(0.091)	(0.074)	(0.076)	(0.091)	(0.090)	(0.090)	(0.090)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		0.093	-0.351	1.070	1.554	0.739	0.982
		(0.253)	(0.622)	(0.879)	(1.594)	(4.276)	(7.657)
Observations	17,903	17,903	17,903	17,903	17,903	17,903	17,903
Overall *R*-squared	0.249	0.249	0.249	0.250	0.249	0.249	0.249
B. Interviewees' cooperation on a scale from 1 (*low*) to 7 (*high*)							
API_t	-0.117	-0.113	-0.106	-0.098	-0.101	-0.114	-0.121
	(0.146)	(0.160)	(0.151)	(0.141)	(0.137)	(0.140)	(0.142)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.014	-0.092	-0.310	-0.322	-0.092	0.197
		(0.166)	(0.321)	(0.457)	(0.399)	(0.608)	(0.854)
Observations	33,285	33,285	33,285	33,285	33,285	33,285	33,285
Overall *R*-squared	0.001	0.002	0.001	0.001	0.001	0.000	0.001

Source: Authors' estimations using CFPS survey 2010 and 2014.

Note: Interviewees' impatience rated by interviewers is only available in the CFPS2014 wave. In Panel A, other covariates and fixed effects are the same as those in column (2) of Table S11. In Panel B, other covariates and fixed effects are the same as those in column (3) of Table S11. **All the coefficients are scaled by 100 to make them more readable.** Robust standard errors, clustered at the county level, are presented in parentheses. Air pollution data are matched between each CFPS county centroid and its nearest API reporting city boundary within a radius of 40km (i.e. 25miles). *10% significance level.

Table S14: Robustness checks – using non-migrants

	1-day	7-day	30-day	90-day	1-year	2-year	3-year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Verbal test scores							
API_t	-0.004*	-0.001	0.000	-0.002	-0.002	-0.003	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.013**	-0.036***	-0.044***	-0.042***	-0.054***	-0.075***
		(0.005)	(0.008)	(0.011)	(0.012)	(0.015)	(0.021)
Observations	31,529	31,529	31,529	31,529	30,874	29,086	27,768
Overall *R*-squared	0.281	0.281	0.281	0.281	0.281	0.281	0.281
B. Math test scores							
API_t	-0.001	-0.001	-0.001	-0.001	-0.001	-0.002	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.003	-0.004*	-0.009**	-0.007**	-0.009*	-0.015**
		(0.002)	(0.002)	(0.003)	(0.003)	(0.005)	(0.007)
Observations	31,529	31,529	31,529	31,529	30,874	29,086	27,768
Overall *R*-squared	0.470	0.470	0.470	0.470	0.470	0.470	0.470
Follow-up rate (removing respondents who may not stay in their counties of residence for as long as the cumulative measure of exposure applies)	98.7%				96.6%	91.0%	86.6%

See the note to Table S1. In all the regressions, we exclude the respondents who moved across counties between 2010 and 2014. In addition, we exclude respondents who moved into the latest addresses in 2010 for the 1-year time window; we exclude respondents who moved into the latest addresses in 2009-2010 for the 2-year time window; we exclude respondents who moved into the latest addresses in 2008-2010 for the 3-year time window.

Table S15: Robustness checks – polluted occupations excluded

	1-day	7-day	30-day	90-day	1-year	2-year	3-year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Verbal test scores							
API_t	-0.004*	-0.001	-0.000	-0.001	-0.002	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.012**	-0.034***	-0.044***	-0.044***	-0.063***	-0.090***
		(0.005)	(0.008)	(0.011)	(0.012)	(0.016)	(0.021)
Observations	31,414	31,414	31,414	31,414	31,414	31,414	31,414
Overall *R*-squared	0.274	0.274	0.274	0.274	0.274	0.274	0.274
B. Math test scores							
API_t	-0.002	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.003	-0.004*	-0.009**	-0.007**	-0.009*	-0.016**
		(0.002)	(0.002)	(0.004)	(0.003)	(0.005)	(0.007)
Observations	31,414	31,414	31,414	31,414	31,414	31,414	31,414
Overall *R*-squared	0.424	0.424	0.424	0.424	0.424	0.424	0.424

See the notes to Table S1. The regressions exclude respondents with polluted jobs. Polluted occupations include “Geology and mineral industry workers”, “Workers in metal smelting and refining industry”, “Chemical product manufacturing personnel”, “Textile workers”, “Production workers (wood processing, artificial board, wood products, pulp and paper industry)”, and “Production and processing worker (construction materials)”.

Table S16: Robustness checks – including province-by-year fixed effects and standard errors clustered at the province level

	1-day	7-day	30-day	90-day	1-year	2-year	3-year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Verbal test scores							
API_t	-0.003*	-0.001	-0.000	-0.001	-0.001	-0.002	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.011**	-0.034***	-0.046***	-0.056***	-0.105***	-0.144***
		(0.005)	(0.007)	(0.009)	(0.015)	(0.020)	(0.028)
Observations	31,953	31,953	31,953	31,953	31,953	31,953	31,953
Overall *R*-squared	0.263	0.263	0.263	0.263	0.263	0.263	0.263
B. Math test scores							
API_t	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$		-0.002	-0.004	-0.011**	-0.010**	-0.019**	-0.027**
		(0.001)	(0.003)	(0.004)	(0.004)	(0.008)	(0.010)
Observations	31,953	31,953	31,953	31,953	31,953	31,953	31,953
Overall *R*-squared	0.405	0.405	0.405	0.405	0.405	0.405	0.405

Source: Authors' estimations using CFPS survey 2010 and 2014.

Note: $\frac{1}{k}\sum_{i=0}^{k-1} API_{t-i}$ indicates the mean of API readings in the past *k* days, where *k* equals 1, 7, 30, 90, 365, 730, and 1,095, respectively. All the regressions include individual fixed effects, county fixed effects, province-by-year fixed effects, year, month, day of week, and post meridiem hour fixed effects. Demographic controls include gender, age and its square and cubic terms, household per capita income, and years of education. Weather controls include 20°F indicators for temperature bins (that is, <25°F, 25–45°F, 45–65°F, 65–85°F, and >85°F), total precipitation, mean wind speed, and a dummy for bad weather. County-level characteristics include gross domestic product (GDP) per capita, population density, and industrial value share. Robust standard errors, clustered at the province level, are presented in parentheses. Air pollution data are matched between each CFPS county centroid and its nearest API reporting city boundary within a radius of 40km (i.e. 25miles). API = air pollution index; SD = standard deviation. *10% significance level; **5% significance level; ***1% significance level.