Online Appendix

Air Pollution and the Labor Market: Evidence from Wildfire Smoke

Mark Borschulte, David Molitor, & Eric Yongchen Zou

Appendix A: Mortality Effects Estimation

In Section 5.2, we calculate mortality costs of air pollution using established estimates from Deryugina et al. (2019). Here we consider a complementary approach to benchmark the costs of lost earnings to those of premature deaths by directly estimating the effect of smoke (and the resulting pollution increases) on mortality. A conceptual appeal of this approach is that we rely on the same source of variation in deriving the mortality damages as we did in deriving the earnings losses, facilitating a direct comparison of labor market and mortality costs of pollution.

We measure mortality outcomes using micro-data provided by the National Vital Statistics System. The underlying data are taken from death certificates which contain age of death. We use the restricted data files containing month of death and covering all counties in the United States to measuring monthly mortality at the county level. The data are available from 2007 to 2015.

We begin by estimating the mortality effect of smoke exposure at the monthly level, the temporal level of our mortality data, using a regression specification that mirrors that from our earnings analysis. The outcome, $Mortality_{cm}$, is measured as deaths per million in county c and month m. We regress this outcome on the number of days $SmokeDay_{cm}$ in which the county was exposed to wildfire smoke that month:

$$Mortality_{cm} = \beta \cdot SmokeDay_{cm} + \alpha_{c \times month\text{-}of\text{-}year} + \alpha_{state \times year} + \varepsilon_{cm}. \tag{A–1}$$

The primary coefficient of interest is β, which describes the effect of an additional day of smoke on mortality in the month of exposure. Analogous to equation (4) we include county-by-month-of-year fixed effects and state-by-year fixed effects to control for county-specific seasonality as well as common shocks at the state-year level. Standard errors are two-way clustered at the county and state by month levels. Like in the labor market analysis, we also report OLS regressions in which we use monthly average $PM_{2.5}$ concentration as the key independent variable in equation (A–1). We further implement IV models instrumenting for monthly $PM_{2.5}$ using $SmokeDay_{cm}$.

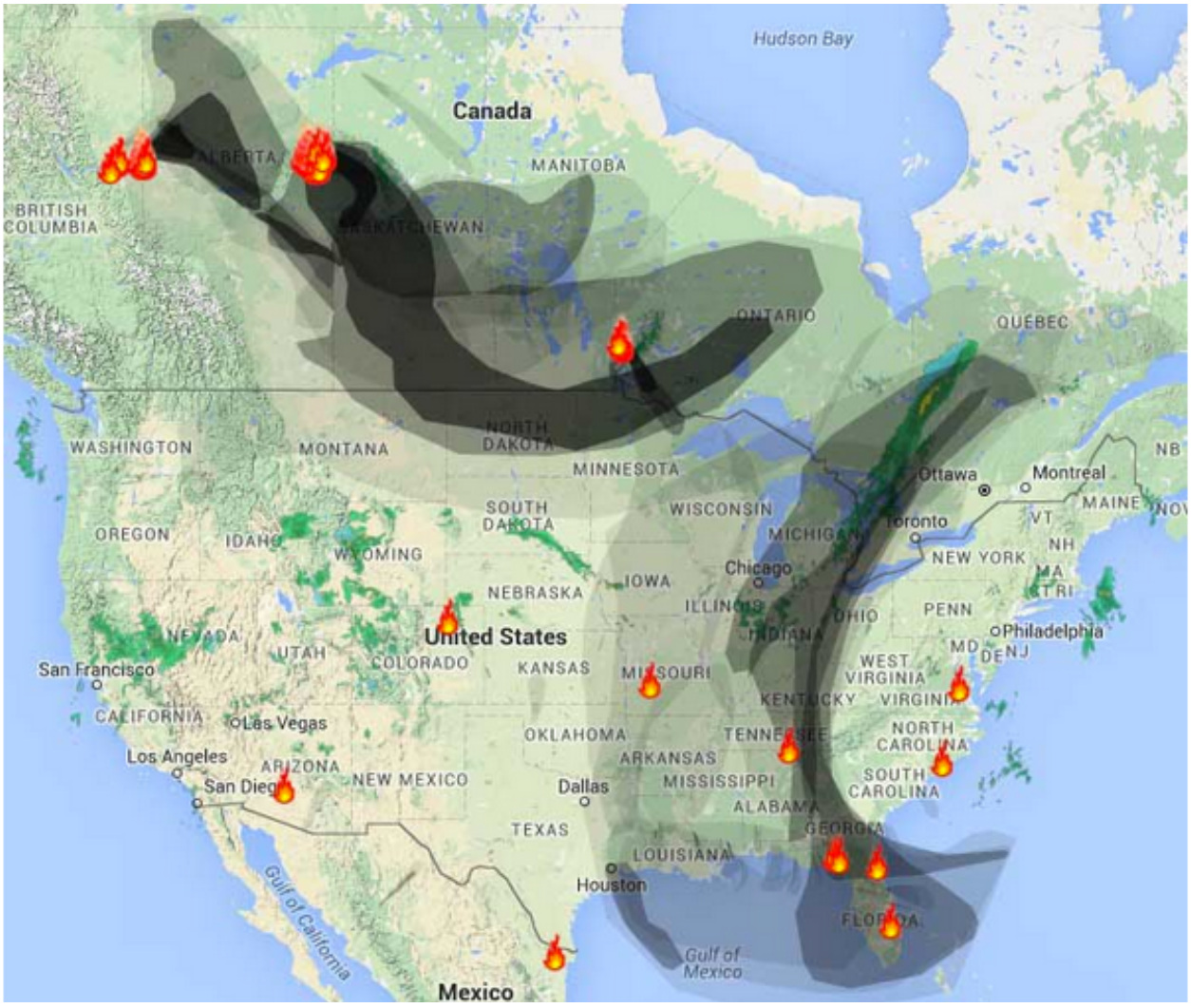
Panel A of Appendix Table A.6 reports the estimates for all age groups (column (1)), the non-elderly group (column (2)), and the elderly population (column (3)). The mortality estimates allow us to compare the costs of mortality to social welfare losses of a 1 unit increase in $PM_{2.5}$, now using IV estimates that come directly out of the smoke quasi-experiment. Column (3), Panel A

shows that each day of smoke in a month increases the elderly mortality rate by 0.972 deaths per million people. This estimate has a standard error of 0.522 and is marginally significant at the 10 percent level. The corresponding IV estimate is noisily estimated, and suggests a unit increase in monthly $PM_{2.5}$ increases elderly mortality by 4.45 deaths per million (SE=3.34, p-value=0.183), or an annual effect of 53.4 additional deaths per million people. The magnitude of our monthly $PM_{2.5}$-mortality IV estimate is comparable to the three-day mortality effects reported in Deryugina et al. (2019) who also study the elderly population during a similar period. The implied monthly effect of a 1 $\mu g/m^3$ $PM_{2.5}$ on elderly mortality is $(0.69 \div 3) \times 30 = 6.9$ deaths per million. Using the two VSL approaches mentioned in Section 5.2, we conclude that the annual mortality cost among the elderly is \$5.2 billion to \$19.9 billion. If we instead use the all-age IV estimate of 0.379 deaths per million people for the calculation (Table A.6, Panel A, column (1)), the implied mortality cost based on the EPA's VSL estimate of \$9.25 million per life lost is \$13 billion annually.

In Panel B of Appendix Table A.6, we repeat reduced-form, OLS, and IV models but aggregating data at the monthly level to the quarterly level, the temporal level of our labor market analysis. Unfortunately, standard errors in the quarterly estimates are too large to draw conclusions. These noisy results may partially reflect well-known challenges of estimating the effect of transient pollution changes on longer-run mortality outcome due to issues such as harvesting and behavioral responses. Most of the existing literature we are aware of focuses on either the pollution's effect in the short run (e.g., Knittel, Miller and Sanders, 2016; Schlenker and Walker, 2016; Deryugina et al., 2019), or on the longer-run mortality effects of sustained pollution exposure (e.g., Deschênes, Greenstone and Shapiro, 2017; Anderson, 2020; Ebenstein, Lavy and Roth, 2016).

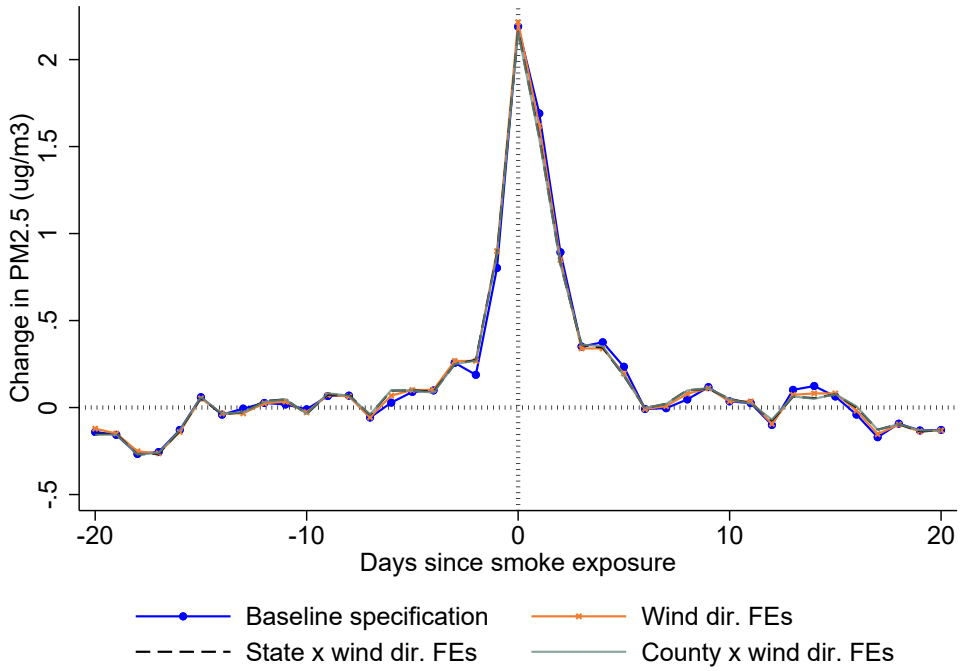
Appendix Figures and Tables

Figure A.1: Fire and Smoke on May 7, 2016



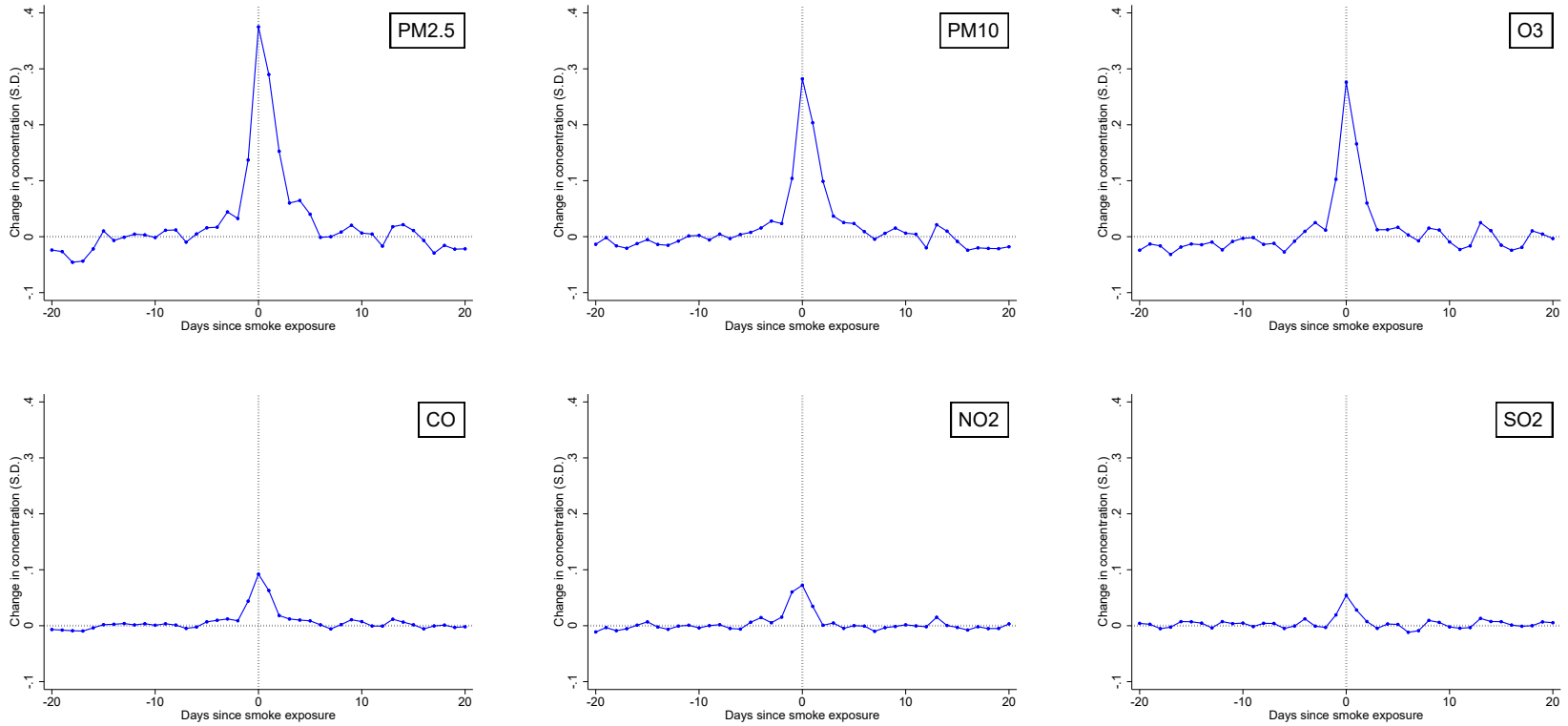
Notes: This map depicts smoke patterns on May 7, 2016, at 9:20 a.m. The Fort McMurray fires in Northern Canada can be seen north of Alberta. This large wildfire produces a smoke plume that reaches the upper Midwest of the United States. Wildfires in the U.S. Southeast produce plumes reaching Canada. Source: WeatherUnderground.com via WildfireToday.com.

Figure A.2: Wildfire Smoke and Ground-level $PM_{2.5}$: Robustness to Flexible Wind Controls



Notes: This figure shows coefficients from a regression of daily $PM_{2.5}$ on indicators of daily smoke exposure up to 20 days before and after the day of observation. Three specifications show varying degrees of controls of wind direction: no controls, 60-degree angle bins of daily wind direction, 60-degree angle bins of daily wind direction fully interacted with state dummies, and 60-degree angle bins of daily wind direction fully interacted with county dummies. The regression incorporates 41 smoke indicators and controls for county-by-day-of-year fixed effects and state-by-year fixed effects.

Figure A.3: Wildfire Smoke Shock to Ground-level Pollution



Notes: Each panel shows coefficients from a regression of daily standardized (mean of zero, standard deviation of one) pollutant concentration on indicators of daily smoke exposure up to 20 days before and after the day of observation. The regression incorporates 41 smoke indicators and controls for county-by-day-of-year fixed effects and state-by-year fixed effects.

Table A.1: Robustness Checks: Reduced-Form Regressions

	(1)	(2)	(3)	(4)
	$PM_{2.5}$	earnings	employment	lfp
A. Smoke measurement				
$\sum$ %county smoked	0.053***	-4.102***	-75.0***	-25.3***
	(0.006)	(0.779)	(20.0)	(9.6)
B. Weather controls				
Temp. Ppt. Wdir. Wspd.	0.043***	-4.950***	-76.3***	-32.8***
	(0.006)	(0.791)	(18.9)	(9.6)
Wdir. $\times$ state	0.048***	-5.019***	-59.6***	-32.0***
	(0.006)	(0.793)	(19.0)	(9.3)
Wdir. $\times$ county	0.048***	-5.312***	-63.1***	-32.0***
	(0.007)	(0.864)	(20.3)	(9.9)
C. Fixed effects controls				
state-by-year FEs (baseline)	0.056***	-5.217***	-79.6***	-38.7***
	(0.007)	(0.776)	(21.9)	(9.2)
division-by-year FEs	0.053***	-5.399***	-100.7***	-27.9
	(0.007)	(0.759)	(26.2)	(17.5)
region-by-year FEs	0.054***	-5.212***	-109.2***	-26.0
	(0.007)	(0.773)	(29.6)	(20.5)
year FEs	0.057***	-4.421***	-132.9***	13.5
	(0.008)	(0.797)	(33.2)	(22.3)
D. Annual data				
state-by-year FEs	0.036***	1.083	-3.5	-30.2
	(0.005)	(4.252)	(37.1)	(25.0)
division-by-year FEs	0.016***	-5.902***	-80.0***	15.8
	(0.004)	(2.075)	(26.1)	(16.9)
region-by-year FEs	0.017***	-4.139***	-79.1***	11.2
	(0.003)	(1.454)	(21.8)	(13.7)
year FEs	0.019***	0.045	-82.8***	51.2***
	(0.003)	(1.377)	(18.3)	(10.4)
E. Outcome specification				
Lvl spec. w/ trends	0.060***	-2.454***	-42.7***	-24.7***
	(0.006)	(0.596)	(11.1)	(7.4)
Lvl spec. wo/ trends	0.056***	-1.581	10.8	-25.3***
	(0.007)	(0.965)	(8.6)	(2.1)
First-diff spec	0.063***	-3.663***	-51.6***	-26.9***
	(0.006)	(0.654)	(16.1)	(9.0)
F. Standard errors clustering				
	0.056	-5.217	-79.6	-38.7
County + division-by-quarter	(0.009)***	(1.062)***	(37.7)**	(13.1)***
County + region-by-quarter	(0.011)***	(1.432)***	(51.9)	(16.4)**
County + quarter	(0.011)***	(2.077)**	(81.6)	(26.6)
County	(0.004)***	(0.470)***	(10.0)***	(5.7)***
State	(0.013)***	(0.949)***	(23.0)**	(9.5)***

Notes: Each cell is a separate regression. Row names indicate the type of robustness checks performed. All regressions are weighted by county population (columns 1, 2, and 4) and population aged over 16 (column 3), and include county-by-quarter-of-year fixed effects and state-by-year fixed effects. Unless otherwise noted, standard errors are two-way clustered at the county and state-by-quarter levels. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Table A.2: Robustness Checks: IV Regressions

		(1)	(2)	(3)
		earnings	employment	lfp
A. Smoke measurement				
∑ %county smoked	F=89.9	-83.409***	-1738.9***	-506.3**
		(19.517)	(394.1)	(205.1)
B. Weather controls				
Temp. Ppt. Wdir. Wspd.	F=45.9	-124.074***	-2065.7***	-814.8***
		(28.967)	(544.2)	(231.5)
Wdirx × state	F=56.0	-113.794***	-1643.9***	-758.3***
		(25.824)	(453.2)	(199.7)
Wdirx × county	F=50.0	-123.416***	-1768.0***	-778.0***
		(29.247)	(513.3)	(226.9)
B. Fixed effects controls				
state-by-year FEs (baseline)	F=71.8	-103.1***	-1750.1***	-790.9***
		(20.4)	(434.8)	(182.1)
division-by-year FEs	F=62.5	-110.3***	-2183.7***	-637.7*
		(21.8)	(561.8)	(348.9)
region-by-year FEs	F=61.7	-105.040***	-2191.8***	-715.3*
		(20.492)	(596.1)	(400.9)
year FEs	F=47.8	-85.368***	-2607.6***	-25.7
		(19.977)	(710.0)	(418.9)
C. Annual data				
state-by-year FEs	F=50.2	41.562	-99.6	-658.9
		(144.336)	(1172.0)	(753.6)
division-by-year FEs	F=20.1	-357.457*	-4851.0**	1048.7
		(185.348)	(2350.1)	(1233.9)
region-by-year FEs	F=25.7	-219.809*	-3929.2**	-171.9
		(115.260)	(1849.6)	(935.4)
year FEs	F=42.1	14.799	-4585.5***	2093.1***
		(89.288)	(1479.2)	(769.0)
D. Outcome specification				
Lvl spec. w/ trends	F=115.8	-43.572***	-848.0***	-492.6***
		(12.989)	(201.0)	(157.3)
Lvl spec. wo/ trends	F=71.9	-27.721	206.9***	-451.9***
		(22.200)	(68.8)	(73.1)
First-diff spec.	F=97.9	-65.9***	-1069.6***	-462.8***
		(14.8)	(291.5)	(153.3)
E. Standard errors clustering				
		-103.077	-1750.1	-790.9
County + division-by-quarter	F=35.9	(27.610)***	(761.5)**	(262.9)***
County + region-by-quarter	F=26.3	(34.802)***	(1012.3)*	(333.9)**
County + quarter	F=25.0	(43.372)**	(1579.7)	(511.1)
County	F=190.5	(13.611)***	(233.4)***	(123.9)***
State	F=18.2	(39.845)**	(503.0)***	(285.1)***

Notes: Each cell is a separate 2SLS regression. Row names indicate the type of robustness checks performed. All regressions are weighted by county population (columns 1, 2, and 4) and population aged over 16 (column 3), and include county-by-quarter-of-year fixed effects and state-by-year fixed effects. Unless otherwise noted, standard errors are two-way clustered at the county and state-by-quarter levels. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Table A.3: Air Pollution and Earnings: OLS Regressions with Multiple Pollutants

	(1)	(2)	(3)	(4)	(5)	(6)
$PM_{2.5}$	-10.6***	-13.6***	-11.1***	-14.1***	-10.0**	-15.4***
	(3.1)	(3.5)	(3.4)	(4.4)	(4.0)	(5.1)
PM_{10}	-	1.9*	-	-	-	2.6*
	-	(1.1)	-	-	-	(1.4)
O_3	-	-	-1.2	-	-	-2.5
	-	-	(1.1)	-	-	(1.7)
SO_2	-	-	-	-3.6	-	-7.6*
	-	-	-	(2.3)	-	(4.5)
NO_2	-	-	-	-	-5.5	-4.2
	-	-	-	-	(3.9)	(3.6)
Outcome mean	5,687.6	5,975.2	5,763.9	6,114.5	6,211.8	6,390.4
Observations	74,725	42,616	64,248	40,363	31,534	23,373

Notes: Each column is a separate regression. Pollutants are measured in $\mu g/m^3$ ($PM_{2.5}$ and PM_{10}), ppb (O_3), and ppm (SO_2 and NO_2). All regressions are weighted by county population, and include county-by-quarter-of-year fixed effects and state-by-year fixed effects. Standard errors are two-way clustered at the county and state-by-quarter levels. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Table A.4: Population Flow Responses

	(1)	(2)	(3)
	in-migration (log)	out-migration (log)	tax-exemptions (log)
Smoke	-0.003	0.0001	-0.010
	(0.015)	(0.012)	(0.007)
Observations	37,254	37,256	37,284

Notes: The table reports estimated effects of an additional day of wildfire smoke exposure on annual IRS SOI migration outcomes. Each column corresponds to a separate regression using county-year observations and relevant county population weights. The focal independent variables capture the number of days in a year on which a county was exposed to wildfire smoke. All regressions include county fixed effects, and state-by-year fixed effects. Standard errors are clustered at both the county and the state-by-year levels.

Table A.5: Sub-industry IV Estimates for the Agricultural Sector

	(1)	(2)	(3)	(4)	(5)	(6)
NAICS code:	11	111	112	113	114	115
	ag total	crop production	animal production	forestry logging	fishing hunting	support activities
$\hat{PM}_{2.5}$	-54.2*	-44.2***	0.7	1.6	-0.7	-26.9
	(29.1)	(15.3)	(1.6)	(1.9)	(1.6)	(23.7)
Outcome mean	5,147.3	2,464.0	735.8	437.5	99.0	2,462.8
Kleibergen-Paap F	186.2	160.0	134.3	138.5	15.1	140.5
Observations	68,846	50,816	43,921	19,841	3,582	38,634

Notes: Each cell is a separate regression following the IV estimation equations (5) and (6). The dependent variable is QWI employment for the corresponding sector indicated by the column title. The smoke variable is used as an instrument for county's quarterly average $PM_{2.5}$. All regressions are weighted by county population aged over 16 and include county-by-quarter-of-year fixed effects and state-by-year fixed effects. Standard errors are two-way clustered at the county and state-by-quarter levels. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Table A.6: Wildfire Smoke and Mortality

	(1)	(2)	(3)
	all ages	ages 60-	ages 60+
A. Monthly mortality			
Smoke (reduced form)	0.132	0.081*	0.972*
	(0.110)	(0.043)	(0.522)
$PM_{2.5}$ (OLS)	0.480***	0.148**	1.273
	(0.167)	(0.061)	(0.820)
$\hat{PM}_{2.5}$ (IV)	0.379	0.270	4.446
	(0.648)	(0.259)	(3.335)
	[F=103.5]	[F=104.2]	[F=97.9]
B. Quarterly mortality			
Smoke (reduced form)	-0.118	0.045	1.162
	(0.186)	(0.049)	(0.934)
$PM_{2.5}$ (OLS)	0.628	0.247	-2.131
	(1.032)	(0.310)	(4.997)
$\hat{PM}_{2.5}$ (IV)	-5.908	1.468	-8.889
	(5.415)	(1.480)	(28.206)
	[F=42.6]	[F=43.3]	[F=39.1]
Mean monthly mortality	678.762	168.445	2845.111
Mean quarterly mortality	2033.419	506.850	8600.272
Observations (monthly)	330,442	330,442	330,442
Observations (quarterly)	123,422	123,422	123,422

Notes: Each cell represents a separate regression. Outcome variables are all-age mortality (column 1), mortality among age below 60 (column 2), and mortality among age above 60 (column 3). "Smoke" counts the number of days a county is fully covered by wildfire smoke plumes. In IV estimation, the smoke variable is used as an instrument for county's quarterly average $PM_{2.5}$. All regressions are weighted by county population in the relevant age groups, and include county×month-of-year fixed effects and census state×year fixed effects. Standard errors are two-way clustered at the county and state-by-month levels (panel A) and county and state-by-quarter levels (panel B). *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.