Online Appendix: Ambiguous Air Pollution Effects of China's COVID-19 Lockdown*

Douglas Almond,[†‡] Xinming Du,[§] Valerie Karplus,[¶] and Shuang Zhang[‖]

January 8, 2021

Contents

*We thank the National Science Foundation for support through Award SES-1658888: "Collaborative Research: Market Based Emissions Policies" in China. We thank Janet Currie, Anouch Missirian, Wolfram Schlenker and Eric Zou for helpful comments and Solveig Asplund for proofreading.

[†]Columbia University and NBER

[‡]To whom correspondence should be addressed: da2152@columbia.edu

[§]Columbia University: xd2197@columbia.edu

[¶]Carnegie Mellon University: vkarplus@andrew.cmu.edu

[‖]University of Colorado Boulder and CEIBS: Shuang.Zhang@colorado.edu

1 Data description

Air quality monitor data

We use monitor-specific hourly data from air quality monitor stations across China. With the increasing concern of air pollution, the Chinese government built the National Urban Air Quality Real-time Publishing Platform and mandated detailed quality assurance and quality control programs at each monitoring station. The Platform is required to report six primary pollutants and air quality index since 2013. By the end of our study period, the reporting system covers 367 prefecture-level cities and 1642 monitors across China.

We collected data from 1642 monitors. To construct a balanced panel, monitors are required to report at least one non-missing data each day over 150 days (day -21 to 28 around the Lunar New Year in 2018-2020). In our final sample, the number of monitors for NO_2, SO_2, $PM_{2.5}$, O_3, CO is 1213, 1213, 1202, 1208, 1208 respectively. Detailed number for each province is shown in Supplementary Material Table S2. For each monitor, we collapse hourly data into daily average. We add monitor fixed effects in our regression to control for possible unobserved monitor-specific factors.

Weather station data

We obtain data on weather conditions including temperature, wind speed and precipitation from NCDC Global Summary of the Day. This dataset is derived from The Integrated Surface Hourly dataset and includes data from over 9000 weather stations. We use all active weather stations in China over our study period. To match weather data with air pollution measure, we average weather indicators for each province-month.

CEMS data

We collect CEMS hourly pollutant emissions data from each province's public platform. The national CEMS network covers most thermal power plants and large industrial pollution sources. Monitors installed on the stacks of emitting units measure the emission concentrations of diverse air pollutants. In this study, we focus on two primary pollutants, SO_2 and NO_x.

We use CEMS data in eleven provinces where data are consistently reported in 2019 and 2020: Anhui, Heilongjiang, Henan, Hubei, Jiangsu, Jiangxi, Liaoning, Inner Mongolia, Shaanxi, Shandong and Zhejiang. Four of them are Hubei's neighboring provinces.

We require firms with non-missing data at least 10 days over day -21 to -1 and 14 days over day 0 to 28 in both 2019 and 2020. We average hourly emissions to firm-pollutant-day level data and add firm fixed effects in the regression. When using the number of firms as dependent variable, we construct our sample at province-day level and control for province fixed effects.

Table S1: Infection Rate on February 21, 2020

Province	Confirmed cases	Deaths	Cases per million people	Deaths per million people
Hubei	63,454	2,250	1,072.4	38.03
Neighbors	5,023	37	14.5	0.107
Non-neighbors	7,810	58	7.9	0.058

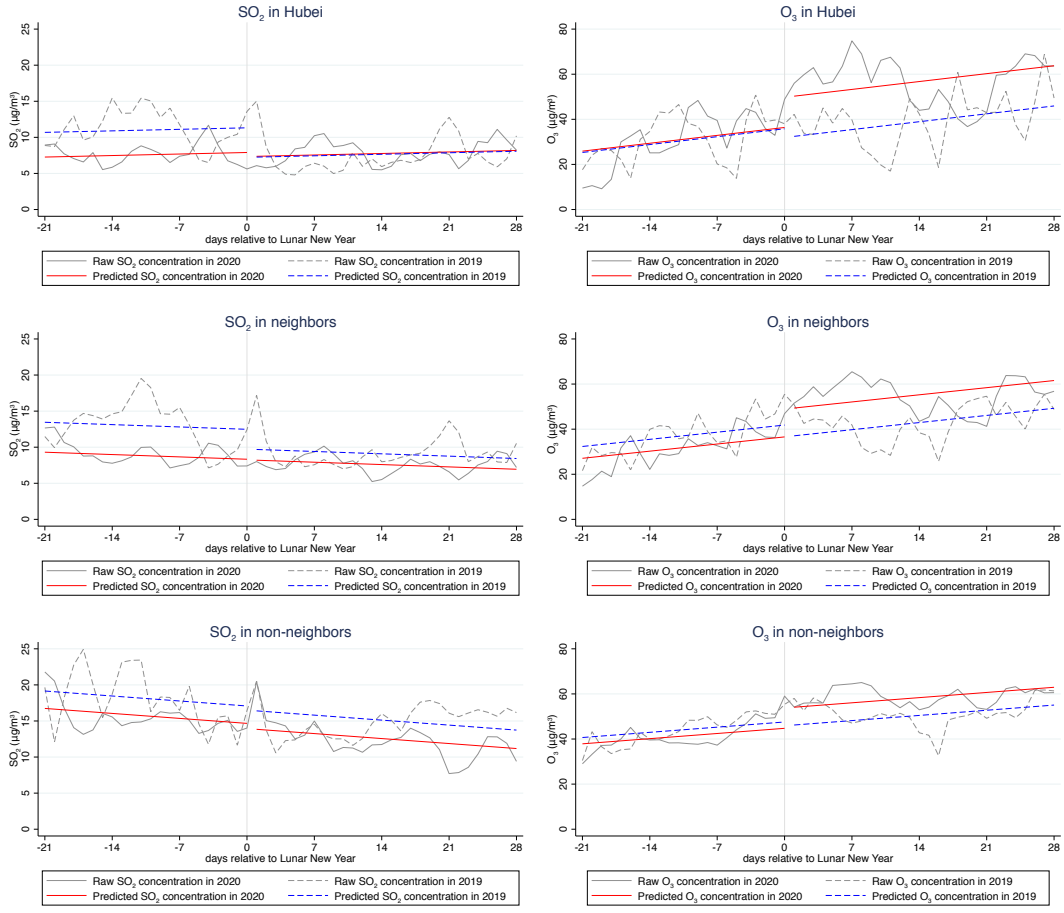
Table S2: Number of monitors

Province	Hubei's neighbor	NO_2	SO_2	$PM_{2.5}$	O_3	CO
Anhui	Y	61	61	61	61	61
Beijing	N	12	12	12	12	12
Chongqing	Y	11	11	11	11	11
Fujian	N	33	33	31	33	33
Gansu	N	27	27	27	27	27
Guangdong	N	87	86	86	86	88
Guangxi	N	44	44	44	41	44
Guizhou	N	31	31	31	31	31
Hainan	N	7	7	7	7	7
Hebei	N	46	46	47	46	47
Heilongjiang	N	42	43	42	42	42
Henan	Y	63	63	61	61	63
Hubei		37	38	37	37	37
Hunan	Y	64	65	63	63	65
Inner Mongolia	N	37	37	36	37	37
Jiangsu	N	62	62	62	62	58
Jiangxi	Y	46	46	46	45	45
Jilin	N	29	29	27	29	29
Liaoning	N	74	74	74	74	72
Ningxia	N	17	17	17	17	17
Qinghai	N	9	10	10	10	10
Shaanxi	Y	42	42	42	42	42
Shandong	N	62	60	61	62	61
Shanghai	N	10	10	10	10	10
Shanxi	N	51	51	51	51	51
Sichuan	N	84	84	85	85	84
Tianjin	N	15	15	15	15	15
Tibet	N	8	8	7	9	9
Xinjiang	N	34	34	32	33	34
Yunnan	N	29	28	29	29	28
Zhejiang	N	39	39	38	40	38
Neighbors		287	288	284	283	287
Non-neighbors		889	887	881	888	884
All China		1213	1213	1202	1208	1208

2 Figures with regression lines

Figure S1 provides raw and predicted SO_2 and O_3. We run a simplified difference-in-difference model and plot predicted values after regression. Control variables include `Post`, `Y2020`, `Post*Y2020` and `time`. `time` is the same as X-axis in Figure S1, defined as day number relative to LNY. The parallel lines are to smooth the raw data and to show the trend and level change before and after LNY in year 2019 and 2020. They are consistent with our difference-in-difference level estimates with full set of controls reported in Table S4, which indicates our results are robust with and without weather controls, day of week and monitor fixed effects.

Figure S1: SO_2 and O_3 concentration by day before and after the Lunar New Year in 2019 (dash) and 2020 (solid). We normalize days around the Lunar New Year (grey vertical line). Trend lines predicted from a simplified difference-in-difference model are in blue and red, raw concentrations are in grey.



3 Double difference results

Table S3: Double difference results

	$\ln(NO_2)$	$\ln(SO_2)$	$\ln(PM_{2.5})$	$\ln(O_3)$	$\ln(CO)$
Panel A: All China					
Post	-0.314***	-0.174***	-0.062***	0.143***	-0.040***
	(0.006)	(0.008)	(0.010)	(0.008)	(0.004)
Post×Y2020	-0.494***	0.010	-0.252***	0.334***	-0.245***
	(0.007)	(0.009)	(0.010)	(0.008)	(0.007)
Y2020	-0.059***	-0.219***	-0.063***	-0.081***	0.032***
	(0.007)	(0.011)	(0.009)	(0.007)	(0.006)
Observations	181950	181950	180300	181200	181200
R-squared	0.584	0.633	0.409	0.374	0.423
Panel B: Hubei					
Post	-0.436***	-0.431***	-0.024	0.059**	-0.072***
	(0.024)	(0.032)	(0.029)	(0.027)	(0.023)
Post×Y2020	-0.551***	0.254***	-0.132***	0.732***	-0.130***
	(0.042)	(0.041)	(0.028)	(0.037)	(0.034)
Y2020	-0.315***	-0.187***	-0.495***	-0.111**	-0.119***
	(0.028)	(0.060)	(0.024)	(0.045)	(0.033)
Observations	5550	5700	5550	5550	5550
R-squared	0.612	0.412	0.356	0.351	0.230
Panel C: Neighbors					
Post	-0.447***	-0.331***	-0.069***	0.116***	-0.029***
	(0.012)	(0.015)	(0.023)	(0.015)	(0.009)
Post×Y2020	-0.498***	0.147***	-0.262***	0.544***	-0.290***
	(0.013)	(0.018)	(0.021)	(0.017)	(0.012)
Y2020	-0.181***	-0.366***	-0.142***	-0.205***	0.006
	(0.014)	(0.023)	(0.019)	(0.015)	(0.013)
Observations	43050	43200	42600	42450	43050
R-squared	0.616	0.511	0.387	0.378	0.393
Panel D: Non-neighbors					
Post	-0.270***	-0.110***	-0.055***	0.157***	-0.041***
	(0.007)	(0.008)	(0.012)	(0.009)	(0.005)
Post×Y2020	-0.500***	-0.061***	-0.242***	0.257***	-0.235***
	(0.009)	(0.010)	(0.012)	(0.009)	(0.008)
Y2020	-0.006	-0.158***	-0.027**	-0.037***	0.046***
	(0.008)	(0.013)	(0.011)	(0.008)	(0.008)
Observations	133350	133050	132150	133200	132600
R-squared	0.581	0.661	0.402	0.369	0.431
Monitor FEs	Y	Y	Y	Y	Y
Weather	Y	Y	Y	Y	Y
Y2018	Y	Y	Y	Y	Y
DOW FEs	Y	Y	Y	Y	Y

Notes: Standard errors are clustered at the monitor level. * significant 10% level; ** significant at 5% level; *** significant at 1% level.

4 Level estimates

Table S4: Double difference results

	NO_2	SO_2	$PM_{2.5}$	O_3	CO
Panel A: All China					
Post	-9.155***	-2.139***	1.109*	5.873***	-0.036***
	(0.199)	(0.118)	(0.635)	(0.339)	(0.005)
Post×Y2020	-10.969***	0.240*	-16.829***	11.286***	-0.235***
	(0.246)	(0.144)	(0.826)	(0.304)	(0.008)
Y2020	-2.854***	-2.733***	-0.964	-1.583***	0.039***
	(0.246)	(0.160)	(0.639)	(0.245)	(0.007)
Observations	181950	181950	180300	181200	181200
R-squared	0.514	0.554	0.347	0.410	0.431
Panel B: Hubei					
Post	-15.249***	-4.757***	-3.623*	1.870*	-0.099***
	(1.012)	(0.408)	(2.026)	(1.052)	(0.024)
Post×Y2020	-6.872***	2.927***	-6.071***	26.192***	-0.083**
	(0.663)	(0.368)	(1.694)	(0.994)	(0.033)
Y2020	-11.269***	-2.056***	-28.614***	-5.408***	-0.143***
	(0.908)	(0.542)	(1.636)	(1.123)	(0.035)
Observations	5550	5700	5550	5550	5550
R-squared	0.533	0.396	0.367	0.402	0.218
Panel C: Neighbors					
Post	-13.600***	-3.815***	3.909**	3.790***	-0.029***
	(0.323)	(0.173)	(1.634)	(0.624)	(0.010)
Post×Y2020	-8.645***	2.362***	-20.569***	18.495***	-0.258***
	(0.419)	(0.203)	(1.844)	(0.564)	(0.013)
Y2020	-8.125***	-4.192***	-8.341***	-4.970***	-0.001
	(0.401)	(0.250)	(1.430)	(0.428)	(0.014)
Observations	43050	43200	42600	42450	43050
R-squared	0.540	0.409	0.353	0.388	0.400
Panel D: Non-neighbors					
Post	-7.425***	-1.450***	1.162*	6.775***	-0.034***
	(0.224)	(0.142)	(0.687)	(0.414)	(0.006)
Post×Y2020	-12.213***	-0.831***	-15.488***	8.774***	-0.236***
	(0.301)	(0.179)	(0.923)	(0.334)	(0.010)
Y2020	-0.658**	-2.188***	1.948***	-0.351	0.061***
	(0.276)	(0.201)	(0.693)	(0.291)	(0.008)
Observations	133350	133050	132150	133200	132600
R-squared	0.515	0.558	0.342	0.408	0.441
Monitor FEs	Y	Y	Y	Y	Y
Weather	Y	Y	Y	Y	Y
Y2018	Y	Y	Y	Y	Y
DOW FEs	Y	Y	Y	Y	Y

Notes: Standard errors are clustered at the monitor level. * significant 10% level; ** significant at 5% level; *** significant at 1% level.

Table S5: Triple difference results

	NO_2	SO_2	$PM_{2.5}$	O_3	CO
Panel A: Hubei and Neighbors					
Post	-13.704***	-3.903***	3.658**	3.851***	-0.034***
	(0.320)	(0.170)	(1.619)	(0.626)	(0.010)
Post×Y2020	-8.606***	2.358***	-20.409***	18.541***	-0.256***
	(0.418)	(0.203)	(1.835)	(0.564)	(0.013)
Post×Y2020×Hubei	2.919***	1.087**	12.925***	4.258***	0.185***
	(0.783)	(0.422)	(2.706)	(1.110)	(0.039)
Y2020	-8.176***	-4.268***	-8.533***	-4.775***	-0.002
	(0.397)	(0.253)	(1.414)	(0.427)	(0.014)
Post×Hubei	-1.92***	.17	-8.62***	-1.91**	-.064***
	(.709)	(.39)	(2.06)	(.89)	(.0178)
Y2020×Hubei	-3.4***	1.43*	-12.9***	2.13**	-.122***
	(1.03)	(.776)	(2.1)	(.913)	(.0269)
Observations	48600	48900	48150	48000	48600
R-squared	0.539	0.410	0.353	0.389	0.387
Panel B: Neighbors and Non-neighbors					
Post	-7.727***	-1.823***	3.411***	7.061***	-0.030***
	(0.214)	(0.143)	(0.676)	(0.408)	(0.006)
Post×Y2020	-12.018***	-0.619***	-16.340***	8.597***	-0.237***
	(0.299)	(0.174)	(0.911)	(0.334)	(0.010)
Post×Y2020×Neighbor	3.285***	2.905***	-3.300	9.491***	-0.018
	(0.538)	(0.271)	(2.218)	(0.642)	(0.016)
Y2020	-0.764***	-2.194***	1.187*	-0.125	0.061***
	(0.274)	(0.208)	(0.705)	(0.289)	(0.008)
Post×Neighbor	-5.05***	-.921***	-8.64***	-4.07***	-.0152
	(.327)	(.236)	(1.54)	(.665)	(.00942)
Y2020×Neighbor	-6.89***	-1.94***	-5.68***	-5.69***	-.0634***
	(.479)	(.411)	(1.83)	(.505)	(.015)
Observations	176400	176250	174750	175650	175650
R-squared	0.520	0.554	0.350	0.411	0.435
Monitor FEs	Y	Y	Y	Y	Y
Weather	Y	Y	Y	Y	Y
Y2018	Y	Y	Y	Y	Y
DOW FEs	Y	Y	Y	Y	Y

Notes: Variable `Hubei` in Panel A and `Neighbor` in Panel B are absorbed by pollution monitor fixed effects. Standard errors are clustered at the monitor level. * significant 10% level; ** significant at 5% level; *** significant at 1% level.

5 Robustness checks

There is a tradition for many Chinese families to set off fireworks and commemorate their ancestors during the Lunar New Year, particularly in suburban and rural areas, which leads to heavy bout of pollution. Air quality concerns have caused hundreds of cities to ban fireworks. To our knowledge, there is no ban due to COVID-19, but the quarantine and heavy sadness could result in different fireworks patterns this year. We drop the Lunar New Year's Eve, the second day and the Lantern Festival Day to address this concern. Results in Table S6 and S7 Panel A confirm the robustness of estimates.

The shock of COVID-19 on the economy and the late back-to-work time this year provide incentives for China's speeding up on economic recovery. If pollution increase during recovery is larger this year than that in the back-to-work in previous years especially for Hubei and neighboring provinces, our observed pollution increase is driven by recovery rather than COVID-19. We use a shorter post period to rule out the recovery story. Results in Panel B still show less reduction in SO_2 and more increase in O_3 in Hubei and neighboring provinces.

As mentioned in Background Point 1., China's air quality gets improved in recent years under great efforts of the central and provincial governments. Hubei, neighboring and non-neighboring provinces are in some ways different in industrial structures and provincial pollution control. Different baseline pollution levels do not threaten our results but do suggest the importance of allowing for different air quality trajectories. We address this by including province specific linear day trends. Results in Panel C show our estimates remain robust.

Table S6: Triple difference results, Hubei and neighbors

	$\ln(NO_2)$	$\ln(SO_2)$	$\ln(PM_{2.5})$	$\ln(O_3)$	$\ln(CO)$
Panel A: Drop firework days					
Post	-0.439***	-0.370***	-0.129***	0.101***	-0.056***
	(0.012)	(0.015)	(0.024)	(0.016)	(0.009)
Post×Y2020	-0.509***	0.181***	-0.225***	0.564***	-0.293***
	(0.014)	(0.019)	(0.025)	(0.018)	(0.013)
Post×Y2020×Hubei	-0.022	0.149***	0.086*	0.101**	0.205***
	(0.049)	(0.047)	(0.047)	(0.041)	(0.042)
Observations	45684	45966	45261	45120	45684
R-squared	0.608	0.500	0.381	0.373	0.385
Panel B: 14 days as post period					
Post	-0.560***	-0.397***	-0.162***	0.088***	-0.046***
	(0.012)	(0.017)	(0.022)	(0.016)	(0.008)
Post×Y2020	-0.363***	0.175***	0.011	0.640***	-0.202***
	(0.014)	(0.018)	(0.018)	(0.019)	(0.011)
Post×Y2020×Hubei	0.061	0.184***	0.107***	0.158***	0.111***
	(0.045)	(0.048)	(0.031)	(0.045)	(0.034)
Observations	34992	35208	34668	34560	34992
R-squared	0.637	0.524	0.341	0.388	0.364
Panel C: Add province by day trend					
Post	-0.593***	-0.370***	0.013	-0.142***	0.087***
	(0.015)	(0.020)	(0.024)	(0.018)	(0.012)
Post×Y2020	-0.508***	0.149***	-0.254***	0.530***	-0.281***
	(0.013)	(0.018)	(0.022)	(0.017)	(0.012)
Post×Y2020×Hubei	-0.041	0.144***	0.071*	0.130***	0.167***
	(0.046)	(0.046)	(0.042)	(0.039)	(0.040)
Observations	48600	48900	48150	48000	48600
R-squared	0.621	0.506	0.390	0.399	0.394
Monitor FEs	Y	Y	Y	Y	Y
Weather	Y	Y	Y	Y	Y
Y2018	Y	Y	Y	Y	Y
DOW FEs	Y	Y	Y	Y	Y

Notes: Standard errors are clustered at the monitor level. * significant 10% level; ** significant at 5% level; *** significant at 1% level.

Table S7: Triple difference results, neighbors and non-neighbors

	$\ln(NO_2)$	$\ln(SO_2)$	$\ln(PM_{2.5})$	$\ln(O_3)$	$\ln(CO)$
Panel A: Drop firework days					
Post	-0.259***	-0.141***	-0.054***	0.149***	-0.042***
	(0.007)	(0.008)	(0.012)	(0.009)	(0.006)
Post×Y2020	-0.520***	-0.047***	-0.249***	0.264***	-0.249***
	(0.009)	(0.010)	(0.014)	(0.009)	(0.008)
Post×Y2020×Neighbor	0.017	0.226***	0.026	0.284***	-0.043***
	(0.017)	(0.021)	(0.030)	(0.020)	(0.015)
Observations	165816	165675	164265	165111	165111
R-squared	0.587	0.636	0.406	0.379	0.424
Panel B: 14 days as post period					
Post	-0.371***	-0.151***	-0.096***	0.178***	-0.067***
	(0.008)	(0.009)	(0.011)	(0.009)	(0.006)
Post×Y2020	-0.374***	0.015	-0.098***	0.251***	-0.132***
	(0.009)	(0.010)	(0.010)	(0.009)	(0.008)
Post×Y2020×Neighbor	0.033**	0.179***	0.109***	0.364***	-0.062***
	(0.016)	(0.021)	(0.021)	(0.020)	(0.014)
Observations	127008	126900	125820	126468	126468
R-squared	0.619	0.654	0.432	0.395	0.442
Panel C: Add province by day trend					
Post	-0.434***	-0.100***	0.001	0.026***	0.024***
	(0.009)	(0.010)	(0.013)	(0.008)	(0.006)
Post×Y2020	-0.529***	-0.041***	-0.264***	0.224***	-0.228***
	(0.009)	(0.010)	(0.012)	(0.009)	(0.008)
Post×Y2020×Neighbor	0.018	0.184***	0.010	0.289***	-0.055***
	(0.016)	(0.021)	(0.027)	(0.019)	(0.014)
Observations	176400	176250	174750	175650	175650
R-squared	0.603	0.649	0.422	0.394	0.431
Monitor FEs	Y	Y	Y	Y	Y
Weather	Y	Y	Y	Y	Y
Y2018	Y	Y	Y	Y	Y
DOW FEs	Y	Y	Y	Y	Y

Notes: Standard errors are clustered at the monitor level. * significant 10% level; ** significant at 5% level; *** significant at 1% level.

6 Results using CEMS data

Using firm-level, hourly emission data from the CEMS, we do similar difference-in-difference estimations in Hubei and non-Hubei provinces in Table S8. First of all, we examine the change in the number of firms reporting emission data before and after the New Year day. In all Chinese firms in the CEMS, the number of firms that report emission data went down more after the New Year day in 2020 than that in 2019. Hubei has more firms that do not report data after the New Year day, compared to that in other provinces. This pattern is consistent with more firms shut down in Hubei given the strictest lockdown policy. Therefore, our analyses on the changes in emissions are limited to firms that consistently report data before and after the New Year day. Within this subsample, we find that a similar 18% decrease in NO_x concentration among firms in Hubei and other provinces. In contrast, for SO_2, there is little change in concentration in Hubei, while a 18% decrease is observed in other provinces. These findings provide suggestive evidence for explaining the less improvement in ambient SO_2 in Hubei during the COVID-19.

Table S8: Double difference results using CEMS data

	Number of firms reporting NO_x	NO_x	$\ln(NO_x)$	Number of firms reporting SO_2	SO_2	$\ln(SO_2)$
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All China						
Post	-7.002*	-3.714***	-0.078***	-5.980	-0.683**	-0.023**
	(3.735)	(0.414)	(0.010)	(3.727)	(0.284)	(0.010)
Post×Y2020	-17.755***	-5.335***	-0.201***	-19.072***	-0.606	-0.189***
	(5.283)	(0.552)	(0.014)	(5.271)	(0.377)	(0.014)
Y2020	30.823***	-7.759***	-0.141***	31.658***	-2.954***	-0.079***
	(4.021)	(0.450)	(0.011)	(4.012)	(0.307)	(0.011)
Observations	1100	60719	60719	1100	61445	61445
R-squared	0.833	0.819	0.662	0.833	0.835	0.781
Panel B: Hubei						
Post	0.016	4.118	0.008	-0.141	-0.274	-0.194**
	(1.760)	(2.925)	(0.055)	(1.793)	(2.661)	(0.078)
Post×Y2020	-34.969***	-10.737**	-0.203**	-35.813***	3.843	-0.038
	(2.489)	(4.795)	(0.091)	(2.535)	(4.180)	(0.123)
Y2020	27.952***	-59.543***	-0.627***	29.857***	-31.868***	-0.962***
	(1.894)	(4.249)	(0.080)	(1.930)	(3.831)	(0.113)
Observations	100	985	985	100	1013	1013
R-squared	0.829	0.792	0.642	0.836	0.643	0.838
Panel C: Non-Hubei						
Post	-7.704*	-3.945***	-0.081***	-6.564	-0.727**	-0.021**
	(4.100)	(0.417)	(0.010)	(4.091)	(0.284)	(0.010)
Post×Y2020	-16.034***	-5.420***	-0.202***	-17.398***	-0.761**	-0.195***
	(5.798)	(0.555)	(0.014)	(5.786)	(0.377)	(0.014)
Y2020	31.110***	-7.283***	-0.137***	31.838***	-2.652***	-0.069***
	(4.414)	(0.452)	(0.011)	(4.404)	(0.307)	(0.011)
Observations	1000	59734	59734	1000	60432	60432
R-squared	0.828	0.819	0.661	0.829	0.838	0.780
Province FEs	Y			Y		
Firm FEs		Y	Y		Y	Y
DOW FEs	Y	Y	Y	Y	Y	Y

Notes: Standard errors are clustered at the province level in Column (1) and (4), and at the firm level in other columns. * significant 10% level; ** significant at 5% level; *** significant at 1% level.