Online Appendix
"Where Are the Missing Dead? How Metrics Management Mitigates Official Data Misreporting in China"*

Guoer Liu†

May 10, 2023

Contents

*I thank Iza Ding, Christian Fong, Mary Gallagher, Benjamin Goehring, Mai Hassan, Nahomi Ichino, Kenneth Lowande, Kevin Quinn, Fiona Shen-Bayh, Chuck Shipan, Yuki Shiraito, participants in the 2018 APSA annual conference presentation and comparative politics workshop at the University of Michigan, the members of the Ichino Research Lab, and two anonymous reviewers for the helpful comments and discussions on earlier drafts.

†Ph.D. Candidate, Department of Political Science, University of Michigan, Ann Arbor, Michigan, USA. Email: guoerliu@umich.edu.

A Data

A.1 Sources

For the outcome variables, I scraped an event-based workplace accidents dataset from OrientTrans@IT's work safety accidents section between July 11, 2000, and December 31, 2005. OrientTrans@IT is an Internet company that owns, publishes, and manages the largest work safety-related dataset in China. Information contained in this dataset is identical to that published by the State Administration of Work Safety's (SAWS) website after comparison. However, the dataset published by the government website is not consistently available to the public. Sometimes, data only shows in the Internet Explorer browser in the Windows system but not in other browsers and operating systems.[1] The first accident in this dataset took place on July 11, 2000. Data after December 31, 2005, is available, but it is not useful for the difference-in-differences (DD) design.

There were 13,382 entries in total. Each entry recorded the time, location, technical cause, and number of deaths but did not contain any identifiers. Here are two examples translated into English:

Example 1. *"At 11:10 p.m. on September 8, 2002, in Fengxian county, Xi'an, Shaanxi province, an oil tank (registered plate number A20493) overturned in the Jiudiangou section of National Route 316 and resulted in 3 deaths."*

Example 2. *"At 9:50 a.m. on October 5, 2004, in Zhongkai Industrial Park in Huizhou, Guangdong province, a fire accident occurred in a three-floored storage building owned by LG Electronics Ltd. The accident resulted in four minor injuries, one death, and one person missing."*

Leaving out the cases where the number of death cannot be determined,[2] there are 13,356 reported accidents. Figure A1 shows the distribution of the death toll in all work safety accidents. As expected, the data was right-skewed: accidents with one death accounted for about 50% of entries. The maximum death toll in an accident was 300, the minimum was zero, and the average number of deaths per accident was 3.25. There were six reported cases with zero death. The accident description clearly indicated that no one died in these accidents (rather than "missing" or "corpse not found"). I dropped the six cases from the

[1] The website address is `http://media.chinasafety.gov.cn:8090/iSystem/shigumain.jsp`. When I tried to access it on January 4, 2019, via Chrome in macOS Sierra (Version 10.12.6), the pop-up message read "This page isn't working," and "media.chinasafety.gov.cn didn't send any data.". I compared the dataset for this research period from both sources, which are identical.

[2] There are 26 cases where casualty information is missing. For example, in case 14861, it reads, *"At 7:30 p.m. on May 16, 2001, an explosion happened in Tiebei Tiexing gasoline station located in Meihekou, Jilin province. One of the seven injured staff members was in critical condition. All of them were sent to the hospital for treatment. However, the cause remained unknown."*

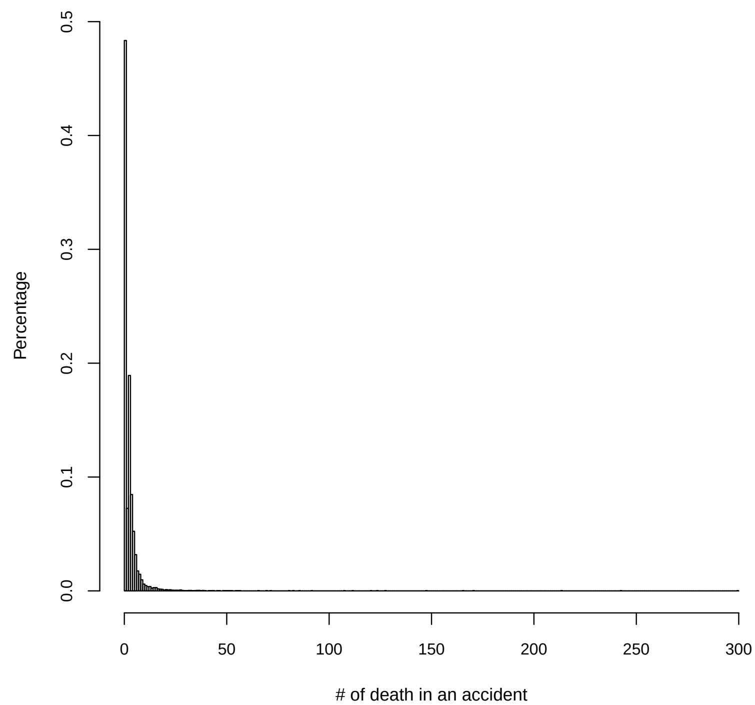

Figure A1: Histogram of Workplace Accident Casualties

analysis because this research focuses on deaths in work safety accidents. This step resulted in 13,350 accidents in total.

This data set has several advantages compared with other available sources. First, the data covered the before and after the death cap incentive change for 31 province-unit in China. Such systematic documentation of work safety accidents since the early 2000s was rare. While data before 2000 was unavailable, there are enough observations for meaningful statistical analysis. Second, the event-based data contained a wealth of information on each accident compared with aggregated official records. This structure affords more flexibility in data aggregation. For instance, statistical yearbooks in China, such as *China Coal Industry Yearbook*, record only the number of deaths in some key sectors annually (Jia, 2017). Fisman and Wang (2017) gathered a more granular data set. They hand-collected province-level quarterly reported work safety deaths from the *People's Daily* for 2005-2012 and requested non-publicly available data from SAWS. Comparing Fisman and Wang's data and my data in 2005, the only overlap period, I found that the aggregated death toll in each province was smaller in my data set than that in theirs. Because their data set was aggregated, I could not compare the event-based entries in two datasets and determine how they overlap. However, for the DD design, the missing data may not pose a severe threat to inference, especially when covariates are included in the estimator. If I invoke the assumption of "missing at random," this data set could be considered a random sample of Fisman and Wang's data. Further, because Fisman and Wang's data was aggregated at the province level, it is not useful for comparisons across sectors. And because only post-treatment periods were recorded in the dataset, it is not suitable for the DD design.

A.2 Coding Procedures

Coding the outcome variables took three steps. First, I used a dictionary method to parse the timing and location information from text chunks in each entry and code all the data. I used the county-level administrative code from the National Bureau of Statistics 2017 version to code the data for the location.[3] The first two digits in the administrative code represent the province, the next two represent the prefecture, and the last two represent the county. One problem with the dictionary method is that it overwrites codes. If the accident description contains different location names, this approach is error-prone. For instance, if the text chunk says, "in province A, a truck owned by a commercial transportation company in province B, overturned." This coding procedure would match the location of this accident first with province A and then overwrite it as province B. It is incorrect. To address this concern, I manually coded all the locations again and corrected the remaining mistakes.

Second, I coded sectoral information for each entry. In the official documents published in 2006, there were six categories: coal mining, IMCT, fireworks, transportation (including road, railway, navigation, and agricultural machinery), fire, and fishery, where the IMCT category includes workplace deaths in industrials, non-coal mining, commercial, and trade (Wang, 2006). Because I were dealing with data before 2006, I slightly modified the categorization for this research. I classified accidents into seven categories: coal mining, other mining, into seven different categories: coal mining, other mining, fire, IMCT, railway, road traffic, and unclassified. Unclassified includes fishery, agriculture, navigation, and aviation; IMCT includes industries, machinery production, commerce, and trade; and other death categories are straightforward. The difference in categorization is unlikely to bias the estimates. Additionally, I used a dictionary method to classify each entry into different categories to cross-validate the manual coding (see Table A1). I compared the inconsistent coding and corrected the remaining mistakes.

Lastly, because the treatment occurred at the province level, I collapsed the data by province-year. This step should generate a panel data set for 31 province-level units for six years, from 2000 to 2005, hence 186 provinces $\times$ year observations. However, because

[3] I downloaded the information on July 15, 2017, from the government website, which is no longer valid. The new address with similar information can be found here, last accessed on January 5, 2019.

the earliest recorded accident in the dataset occurred on July 11, 2000, the reported work safety accidents covered 21 provincial units only in 2000. This resulted in 176 province-year observations. Note that not all sector-level information is available for all province-year during the research period.

Table A1: Accident Classification Cross-Validation: Dictionary Method vs. Manual Coding

Dict. method	**Manual coding**							
	Coal mine	Other mine	Fire	IMCT	Railway	Road	Others	Total
Mine	7,651	1,424	14	639	3	104	13	9,845
Fire	0	0	136	6	0	0	1	143
Railway	0	0	0	0	1	0	0	1
Road	5	4	2	17	2	2,520	9	2,559
IMCT	0	6	0	193	0	0	3	202
Others	5	3	8	86	0	29	469	600
Total	7,661	1,434	160	941	6	2,653	495	13,350

A.3 Summary Statistics

The summary statistics of all variables in the main specification and robustness checks are presented in Table A2.

Table A2: Summary Statistics

Description	Obs	Mean	Std. Dev.	Min	Max
Outcomes variables					
Deaths (all)	176	244.10	230.40	3.00	1246.00
Deaths (road only)	140	94.10	98.50	1.00	426.00
Deaths (other sectors)[a]	175	170.10	168.70	3.00	865.00
Accident counts (all)	176	75.60	98.40	1.00	517.00
Accident counts (road only)	140	18.80	22.70	1.00	105.00
Accident counts (other sectors)	175	61.00	87.90	1.00	493.00
Accident counts (for death=2)	125	7.73	9.70	1.00	48.00
Accident counts (for death=3)	155	16.20	17.60	1.00	93.00
ln(Deaths (all))	176	4.93	1.24	1.10	7.13
ln(Deaths (road only))	140	3.92	1.28	0.00	6.05
ln(Deaths (Other sectors))	175	4.52	1.28	1.10	6.76
ln(Accident counts (all))	176	3.50	1.48	0.00	6.25
ln(Accident counts (road only))	140	2.20	1.28	0.00	4.65
ln(Accident counts (other sectors))	175	3.21	1.51	0.00	6.20
ln(Accident counts (for death=2))	125	1.48	1.04	0.00	3.87
ln(Accident counts (for death=3))	155	2.22	1.14	0.00	4.53
Occupational injury insurance claimants (% of total Insured)	170	0.75	1.30	0.00	15.1[b]
Other variables					
Avg. death per accident (road only)	140	6.14	3.07	1.00	24.30
Avg. death per accident (other sectors)	175	4.74	5.06	1.10	55.30
Covariates					
Party secretary: tenure	186	5.64	2.84	1.00	14.00
Party secretary: age	186	57.90	3.83	46.00	66.00
Party secretary: promotion[c]	186	0.04	0.20	0.00	1.00
Governor: tenure	186	4.85	1.94	1.00	10.00
Governor: age	186	57.60	4.12	43.00	65.00
Governor: promotion	186	0.10	0.30	0.00	1.00
ln(Population)	186	17.20	0.91	14.70	18.40
% of rural population	186	67.80	15.50	15.50	86.00
ln(DMSP-OLS satellite nighttime light)	186	0.32	1.78	-4.63	3.50
ln(Tax revenue per capita)	186	7.04	0.91	5.71	9.83
ln(1+ Coal output)	186	6.89	2.95	0.00	10.90
ln(1+ Mine wage)	186	8.19	3.18	0.00	10.20
ln(Transportation wage)	186	9.66	0.31	8.91	10.40

[a] *Deaths (other sectors)* where "other sectors" include all categories except for "Road". Through out this article I use "other sectors" the same way.

[b] This is Liaoning province in 2002, which is an outlier in the dataset.

[c] *Party secretary: Promotion* is coded as a dummy variable. It equals to 1 if the party secretary has been promoted at the beginning of the following year, and 0 otherwise. The same coding principal applies to provincial governors.

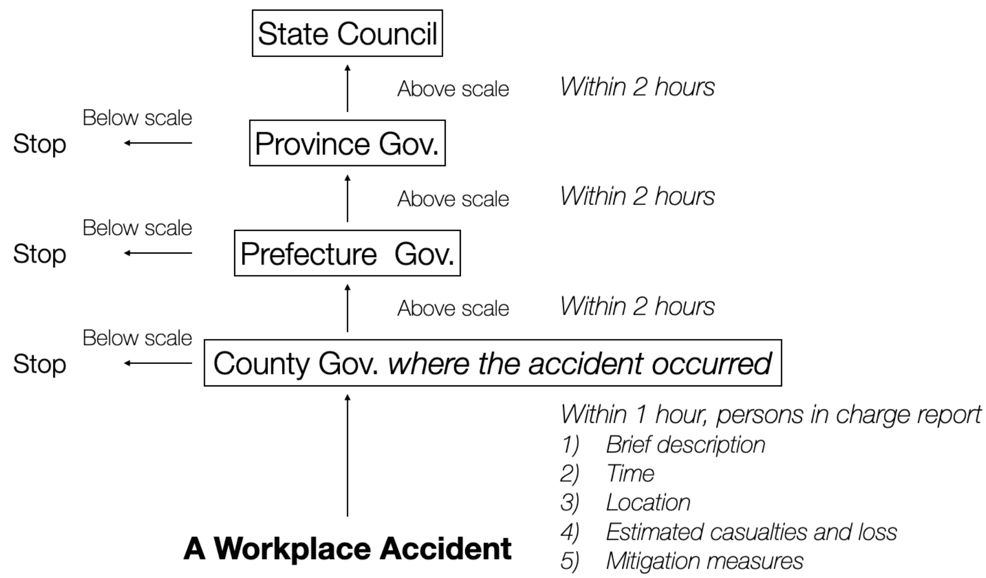

Figure A2: Workplace Accident Reporting Protocols. According to the "Regulations on the Reporting, Investigation, and Disposition of Work Safety Accidents" (2007), county-level governments must be notified of all workplace accidents regardless of the scale within one hour of occurrence. As the severity level increases, more high-level authorities *should* be informed.

A.4 Workplace Accidents Reporting Protocols

Figure A2 illustrates the workplace accidents reporting protocols.

A.5 Interviews

I conducted 40 semi-structured interviews with government officials, scholars, and factory managers in Shanghai and Sichuan. I relied on snowball sampling to get the contact information of interviewees, starting with a few scholars and local officials. Among the 40 interviews, five interviews were cited in this paper. Two interviewees were former department heads in the prefecture-level Administration of Work Safety (i.e., LAsh90507102, YXsh90507102). One interviewee was a deputy party secretary of a county-level Development and Reform Bureau and resigned from the post to join a private insurance company as a manager (i.e., XGsh22507102). An interviewee was a professor at a local party school (i.e., DYsh12307102). Another interviewee was a factory manager (i.e., MMsh90507102). Interviews with others also inform this paper, but the content was not cited.

Before each interview, I did a short self-introduction and informed the interviewees about the research project. I also explicitly told them they could end the interview if they did not want to continue. I did not ask interviewees to sign the consent form on paper, because researchers with more field experience advised me that such documents may put contacts in an awkward position, given the tightly controlled political environment. To put interviewees at ease, I neither recorded the interview nor took notes during the meeting. But I wrote detailed notes immediately after the meeting, according to memory. The interviews were conducted as part of the fieldwork approved by the research ethics committee.

B Main Analysis Regression Tables

Table A3: The DD Estimates of the Overall Effect of the Death Cap

Outcome Variables	(1) Death (all)	(2) Death (all)	(3) Death (all)	(4) Death (all)	(5) Acc. count (all)	(6) Acc. count (all)	(7) Acc. count (all)	(8) Acc. count (all)
Treated group × Treated time	−0.281	−0.652	−0.539	−0.412	−0.303	−0.781	−0.619	−0.348
	(0.236)	(0.190)	(0.169)	(0.157)	(0.278)	(0.233)	(0.197)	(0.165)
Treated time	0.736	0.243	2.349	2.757	0.882	0.198	3.272	5.240
	(0.103)	(0.218)	(0.362)	(0.931)	(0.129)	(0.291)	(0.383)	(1.054)
Treated group	0.893	0.041	0.086	−1.229	0.635	−0.257	−0.195	−7.218
	(0.387)	(0.380)	(0.410)	(4.033)	(0.428)	(0.413)	(0.456)	(3.663)
Control	×	✓	✓	✓	×	✓	✓	✓
Year FE	×	×	✓	✓	×	×	✓	✓
Province FE	×	×	×	✓	×	×	×	✓
R2	0.121	0.553	0.779	0.893	0.089	0.428	0.794	0.926
R2 Adj.	0.105	0.508	0.750	0.851	0.073	0.371	0.768	0.897
Cluster Std.Errors	by province	by province	by province	by province	by province	by province	by province	by province
Observations	176	176	176	176	176	176	176	176

[a] The outcome variables are log transformed.

Table A4: The DD Estimates of the Effect of the Death Cap: Traffic Accident Only

Outcome Variables	(1) Death (Road)	(2) Death (Road)	(3) Death (Road)	(4) Acc. count (Road)	(5) Acc. count (Road)	(6) Acc. count (Road)
Treated group × Treated time	−0.052	−0.018	0.475	−0.295	−0.242	0.073
	(0.449)	(0.453)	(0.566)	(0.421)	(0.424)	(0.512)
Treated time	0.385	0.664	−0.744	0.323	0.772	−0.875
	(0.254)	(0.357)	(1.250)	(0.238)	(0.350)	(1.331)
Treated group	−0.329	−0.320	6.622	−0.288	−0.289	6.290
	(0.754)	(0.781)	(9.629)	(0.650)	(0.696)	(9.729)
Control	✓	✓	✓	✓	✓	✓
Year FE	×	✓	✓	×	✓	✓
Province FE	×	×	✓	×	×	✓
R2	0.488	0.505	0.755	0.510	0.529	0.804
R2 Adj.	0.421	0.427	0.626	0.447	0.454	0.701
Cluster Std.Errors	by province	by province	by province	by province	by province	by province
Observations	140	140	140	140	140	140

[a] The outcome variables are log transformed.

Table A5: The DD Estimates of the Effect of the Death Cap: Other Sectors

Outcome Variables	(1) Death (others)	(2) Death (others)	(3) Death (others)	(4) Acc. count (others)	(5) Acc. count (others)	(6) Acc. count (others)
Treated group × Treated time	−0.773	−0.680	−0.552	−0.954	−0.803	−0.532
	(0.151)	(0.135)	(0.145)	(0.256)	(0.232)	(0.213)
Treated time	0.067	1.811	2.388	0.088	3.062	4.961
	(0.231)	(0.352)	(0.698)	(0.296)	(0.329)	(0.820)
Treated group	0.264	0.304	−1.379	−0.098	−0.034	−8.614
	(0.445)	(0.465)	(4.024)	(0.466)	(0.503)	(3.571)
Control	✓	✓	✓	✓	✓	✓
Year FE	×	✓	✓	×	✓	✓
Province FE	×	×	✓	×	×	✓
R2	0.567	0.731	0.875	0.431	0.760	0.919
R2 Adj.	0.523	0.696	0.826	0.373	0.729	0.887
Cluster Std.Errors	by province	by province	by province	by province	by province	by province
Observations	175	175	175	175	175	175

[a] The outcome variables are log transformed.

Table A6: The DD Estimates of the Effect of the Death Cap on Borderline Cases

Outcome Variables	(1) Acc. count with 2 Deaths	(2) Acc. count with 2 Deaths	(3) Acc. count with 2 Deaths	(4) Acc. count with 2 Deaths	(5) Acc. count with 3 Deaths	(6) Acc. count with 3 Deaths	(7) Acc. count with 3 Deaths	(8) Acc. count with 3 Deaths
Treated group × Treated time	−0.865	−0.943	−0.934	−0.435	−0.358	−0.589	−0.480	−0.576
	(0.247)	(0.324)	(0.315)	(0.396)	(0.295)	(0.238)	(0.241)	(0.243)
Treated time	0.293	0.030	0.509	0.001	0.618	0.228	1.552	2.276
	(0.135)	(0.240)	(0.316)	(0.584)	(0.140)	(0.160)	(0.628)	(1.327)
Treated group	0.279	−0.156	−0.134	0.939	1.055	0.194	0.167	2.443
	(0.441)	(0.436)	(0.447)	(6.306)	(0.442)	(0.465)	(0.454)	(7.736)
Control	✕	✓	✓	✓	✕	✓	✓	✓
Year FE	✕	✕	✓	✓	✕	✓	✓	✓
Province FE	✕	✕	✕	✓	✕	✕	✓	✓
R2	0.025	0.431	0.467	0.870	0.136	0.562	0.618	0.837
R2 Adj.	0.001	0.347	0.371	0.788	0.119	0.511	0.561	0.760
Cluster Std.Errors	by province	by province	by province	by province	by province	by province	by province	by province
Observations	125	125	125	125	155	155	155	155

[a] The outcome variables are log transformed.

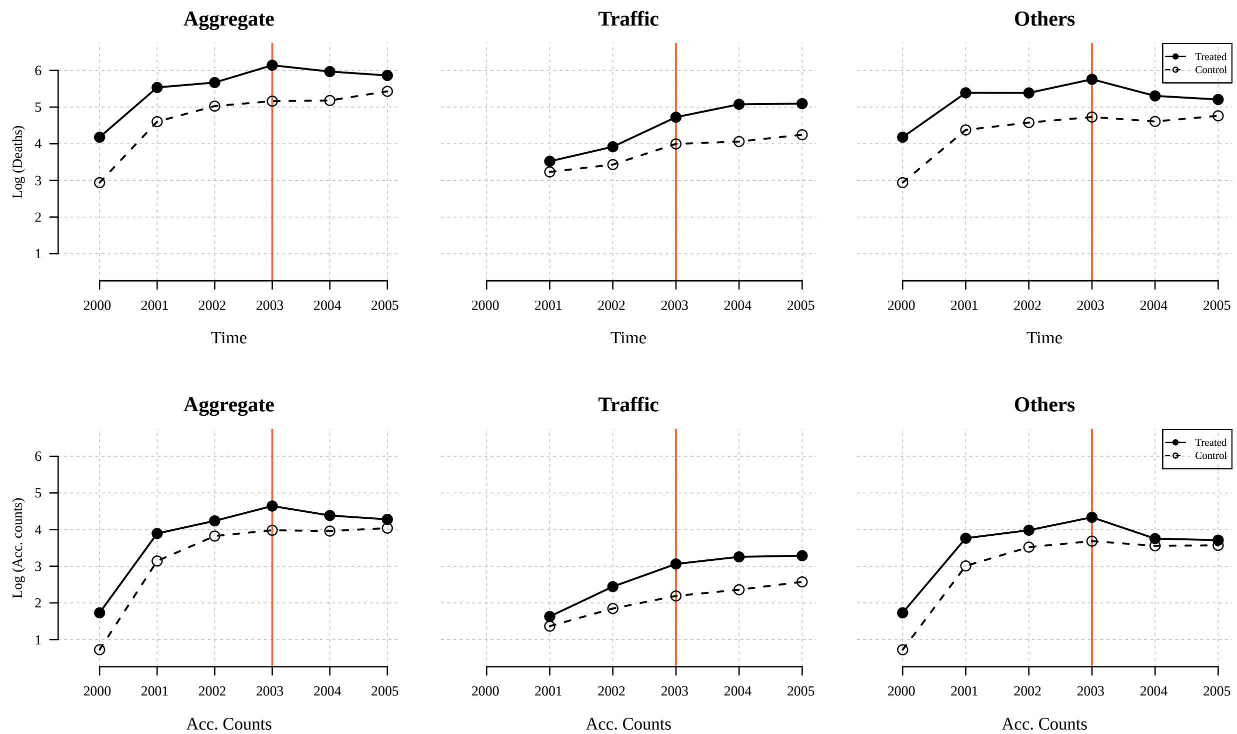

Figure A3: Parallel-Trend Visualization in the Treated and Control Group. The outcome variables for the three figures at the top are workplace accident deaths. The outcome variables for the figures at the bottom are workplace accident counts. Missing data points indicates no such accident was recorded in the dataset. The vertical line separates the pre-and post-treatment period by marking 2003, which is one year before the treatment period.

C Additional Robustness Checks

C.1 Parallel Trend Plot

Figure A3 plots the treated and control group average for all outcome variables. The trends between the treated and control groups were heading in the same direction before 2004.

C.2 Prefecture-Year and Province-Month Analysis

Table A7: The DD Estimates of the Effect of Death Cap (Unit: Prefecture-Year)

Outcome Variables	Death (all)	Death (road)	Death (others)
Treated group × Treated time	−0.205	0.139	−0.391
	(0.116)	(0.126)	(0.088)
Treated time	0.775	0.325	0.372
	(0.231)	(0.149)	(0.201)
Treated group	1.487	0.394	1.237
	(0.042)	(0.078)	(0.032)
Year FE	✓	✓	✓
Province FE	✓	✓	✓
R2	0.082	−0.158	0.239
R2 Adj.	0.232	0.231	0.219
Std.Errors	by province	by province	by province
Observations	1610	891	1411

[a] The outcome variables are log transformed.

Table A8: The DD Estimates of the Effect of Death Cap (Unit: Province-Month)

Outcome Variables	Death (all)	Death (road)	Death (others)
Treated group × Treated time	−0.014	0.134	−0.177
	(0.181)	(0.141)	(0.093)
Treated time	1.341	1.704	0.856
	(0.215)	(0.221)	(0.247)
Treated group	1.579	0.160	1.175
	(0.103)	(0.146)	(0.054)
Month FE	✓	✓	✓
Province FE	✓	✓	✓
R2	0.261	−0.151	0.475
R2 Adj.	0.463	0.232	0.439
Std.Errors	by province	by province	by province
Observations	1616	864	1510

[a] The outcome variables are log transformed.

Note that the province-month results are mixed. The direction of the coefficients is consistent with the hypotheses, but some estimates are not statistically significant. Missing value may be a major underlying factor. A lower level of aggregation—represented by a narrower time frame—leads to a higher number of missing data points. For all accidents, 21% out of 2046 province-month observations are missing; for traffic accidents, 52% out of 1798 province-month observations are missing; for other accidents, 26% out of 2046 province-month accidents are missing.

Doing analysis at the prefecture-month level meets two challenges. First, missing observation is a more severe problem at a lower level of aggregation, such as prefecture-month. For all accidents, 72% out of 24,420 prefecture-month observations are missing; for traffic deaths, 89% out of 18,734 prefecture-month observations are missing; for other sectors, 77% out of 23,892 prefecture-month observations are missing. For example, Figure A4 presents the missing values across prefecture-month during the research period when the aggregated death is the outcome variable. This figure details the data from our research period, with each row representing a different prefecture and every column representing one month.[4] Second, severe missingness means methods like multiple imputation is not applicable because there is little observed data to condition on. Additionally, much of the data is time-varying, which will be difficult to impute accurately.

[4]The figure is created with the `R` package `PanelView`, version 1.1.16.

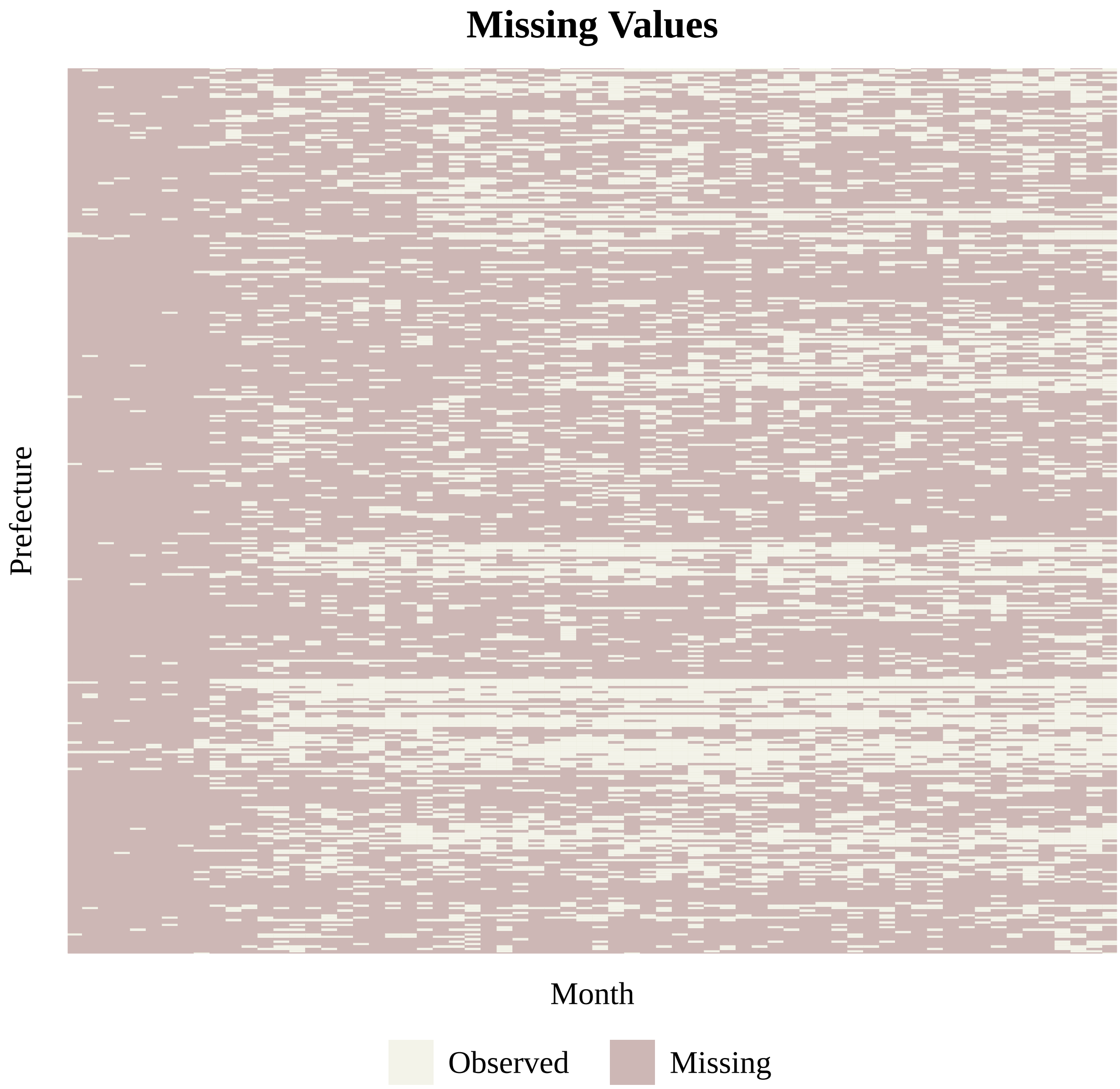

Figure A4: The Distribution of Missing Value by Prefecture-Month. The color ivory denotes that the prefecture-month accident deaths (in aggregate) is observed, and the color pink represents a missing observation.

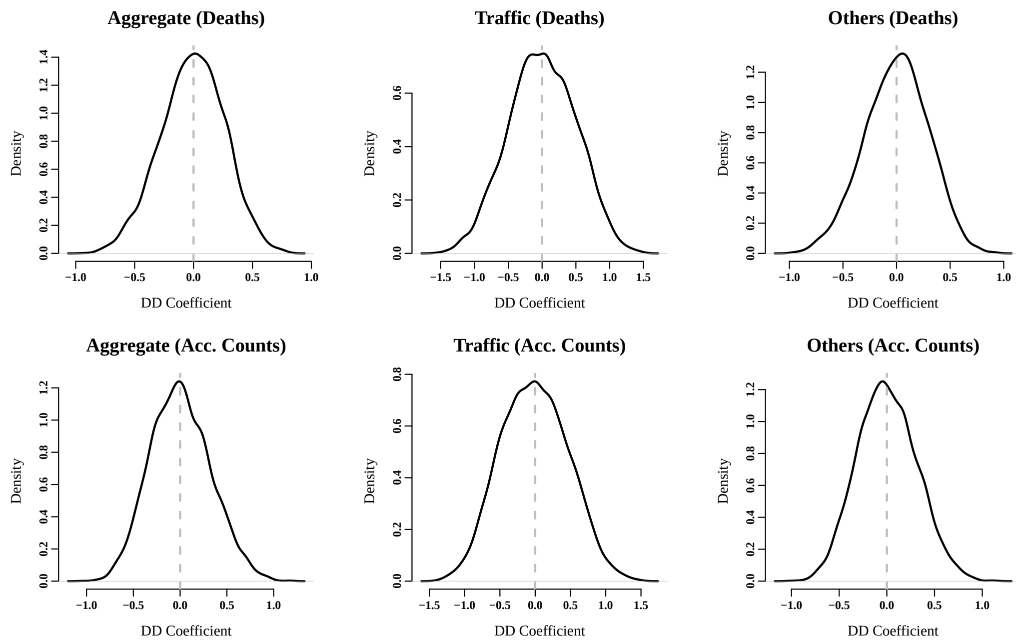

Figure A5: The Coefficient Density of the Placebo Tests. The outcome variables for the three figures at the top are workplace accident deaths. The outcome variables for the figures at the bottom are workplace accident counts. 5,000 Monte Carlo simulations generate the coefficient distribution in each plot with the control group data. In each simulation, I randomly picked four provinces as the "placebo-treated," and the remaining provinces were the "placebo-control." Then I calculated the DD estimate with the most stringent specification. The vertical line marks zero.

C.3 Placebo Test

Another empirical concern is that the main result is a "lucky" draw from all possible combinations of treated and control group assignments. The death cap affects the outcome, not through any meaningful mechanism I considered earlier, but merely as a result of one particular realization of the assignment. If this is true, the result should be vulnerable to a placebo test. Following this intuition, I performed 5,000 Monte Carlo simulations on the *control* provinces. In each trial, I randomly sampled four provinces and labelled them as a "placebo-treated group" and the rest as "placebo-control." Then I computed the DD coefficients with the most stringent specification. Since none of the control provinces had implemented the death cap prior to 2006, the coefficients density plot for each outcome variable was expected to be centered around zero. If this was not the case, it indicates that the placebo test was unsuccessful. Figure A5 shows the result. The distribution is consistent with the expectation.

C.4 Outcome Variables: Raw Data

In the main analysis, I did the log transformation for all outcome variables. One concern over this practice is that data transformation should be avoided when there is no good theoretical reason to do so. I used the logarithm to discount extreme values in work safety accidents by province-year. In the event-based raw data, about 50% entries with precisely one death, and the maximum death toll in an accident was 300. In the aggregated province-year dataset, the minimum death toll is three, and the maximum is 1,246. The mean is 244.10. The accident counts range from 1 to 517, with the mean at 75.6 (see Table A2 Summary Statistics). The large variation may bias the analysis. To address the concern, I redid all the analyses using raw data. Figure A6 summarizes the regression results.[5] The directionality remains consistent with the main hypotheses. Figure A7 visualize the trend of the outcome variables in the treated and control group.

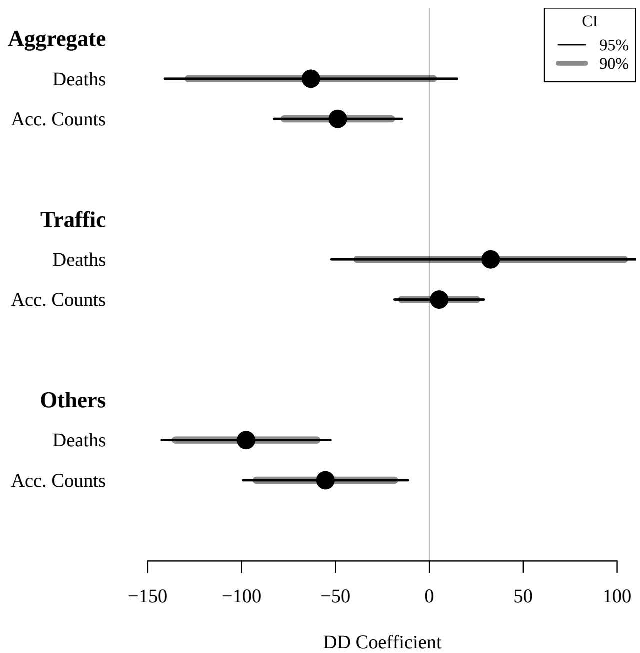

Figure A6: The DD Estimates of the Effect of the Death Cap on Reported Workplace Accident Deaths and Accident Counts (Raw Data). The wide horizontal bars represent 90% asymptotic confidence intervals, and narrow horizontal bars represent 95% confidence intervals. The vertical bar marks zero.

[5]The graph is plotted according to results in Table A9, Table A10, and Table A11.

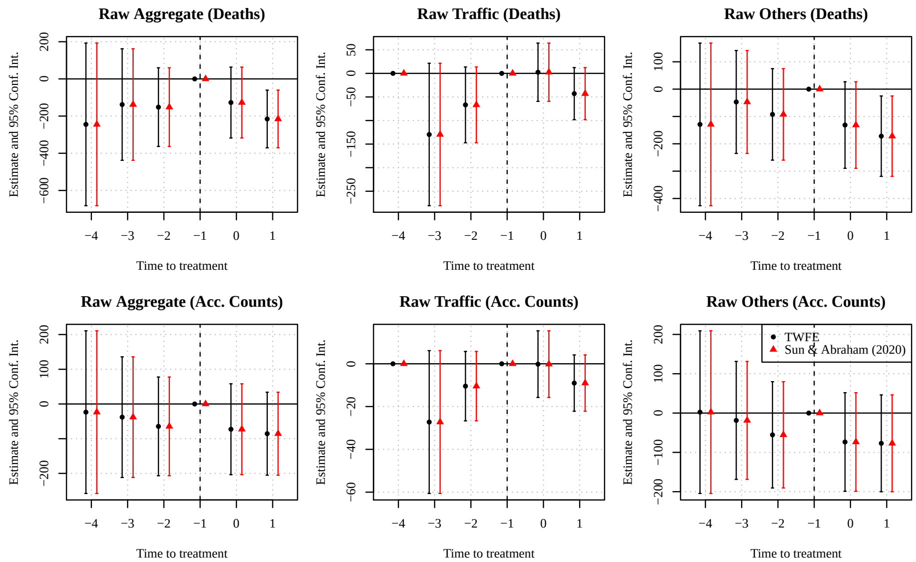

Figure A7: Parallel-Trend Visualization in the Treated and Control Group. The outcome variables for the three figures at the top are workplace accident deaths. The outcome variables for the figures at the bottom are workplace accident counts. Missing data points indicates no such accident was recorded in the dataset. The solid lines connect the annual average for the treated group, represented by solid circles. The dash lines connect the annual average for the control group, represented by empty circles. The vertical line separates the pre-and post-treatment period by marking 2003, which is one year before the treatment period.

Table A9: **Raw data**: The DD Estimates of the Overall Effect of the Death Cap

Outcome Variables	(1) Death (all)	(2) Death (all)	(3) Death (all)	(4) Death (all)	(5) Acc. count (all)	(6) Acc. count (all)	(7) Acc. count (all)	(8) Acc. count (all)
Treated group × Treated time	−10.133	−84.902	−70.848	−63.101	−24.189	−51.050	−44.355	−48.754
	(66.114)	(52.614)	(48.389)	(39.715)	(16.679)	(21.734)	(20.486)	(17.386)
Treated time	109.758	10.500	224.832	341.459	35.189	2.030	94.904	209.986
	(24.404)	(37.142)	(62.775)	(200.884)	(9.022)	(20.039)	(34.852)	(90.492)
Treated Group	161.202	21.324	26.788	−879.144	23.818	−31.815	−29.339	−580.890
	(124.102)	(127.082)	(132.391)	(957.950)	(41.148)	(46.744)	(48.823)	(388.254)
Controls	×	✓	✓	✓	×	✓	✓	✓
Year FE	×	×	✓	✓	×	×	✓	✓
Province FE	×	×	×	✓	×	×	×	✓
R2	0.104	0.481	0.560	0.822	0.029	0.332	0.423	0.822
R2 Adj.	0.088	0.429	0.503	0.753	0.012	0.264	0.349	0.753
Cluster Std.Errors	by province	by province	by province	by province	by province	by province	by province	by province
Observations	176	176	176	176	176	176	176	176

[a] The outcome variables are raw data.

Table A10: **Raw Data**: The DD Estimates of the Effect of Death Cap: Traffic Accident Only

Outcome Variables	(1) Death (road)	(2) Death (road)	(3) Death (road)	(4) Acc. count (road)	(5) Acc. count (road)	(6) Acc. count (road)
Treated group × Treated time	24.143	25.980	32.668	5.211	5.929	5.243
	(42.930)	(43.002)	(43.277)	(13.053)	(12.970)	(12.155)
Treated time	22.705	38.206	−77.568	5.714	11.708	−21.648
	(19.164)	(23.434)	(109.021)	(4.460)	(6.013)	(28.988)
Treated group	3.851	3.754	918.578	−0.399	−0.368	272.040
	(47.663)	(49.543)	(702.825)	(10.197)	(10.654)	(178.206)
Control	✓	✓	✓	✓	✓	✓
Year FE	×	✓	✓	×	✓	✓
Province FE	×	×	✓	×	×	✓
R2	0.458	0.464	0.759	0.427	0.433	0.747
R2 Adj.	0.387	0.379	0.632	0.352	0.344	0.614
Cluster Std.Errors	by province	by province	by province	by province	by province	by province
Observations	140	140	140	140	140	140

[a] The outcome variables are raw data.

Table A11: **Raw data**: The DD Estimates of the Effect of Death Cap: Other Sectors

Outcome Variables	(1) Death (others)	(2) Death (others)	(3) Death (others)	(4) Acc. count (others)	(5) Acc. count (others)	(6) Acc. count (others)
Treated group × Treated time	−128.457	−118.239	−97.613	−60.334	−54.714	−55.335
	(34.948)	(32.582)	(22.946)	(24.827)	(24.463)	(22.416)
Treated time	−15.145	141.616	294.966	−4.589	81.808	193.855
	(31.550)	(44.903)	(111.573)	(19.313)	(26.151)	(68.108)
Treated group	27.086	31.933	−960.715	−30.133	−27.420	−641.238
	(96.425)	(100.198)	(617.373)	(41.331)	(43.138)	(321.701)
Control	✓	✓	✓	✓	✓	✓
Year FE	×	✓	✓	×	✓	✓
Province FE	×	×	✓	×	×	✓
R2	0.421	0.502	0.813	0.286	0.379	0.813
R2 Adj.	0.363	0.437	0.740	0.214	0.298	0.739
Std.Errors	by province	by province	by province	by province	by province	by province
Observations	175	175	175	175	175	175

[a] The outcome variables are raw data.

References

Fisman, R. and Y. Wang (2017, April). The Distortionary Effects of Incentives in Government: Evidence from China's "Death Ceiling" Program. *American Economic Journal: Applied Economics 9*(2), 202–218.

Jia, Ruixue Nie, H. (2017, March). Decentralization, Collusion, and Coal Mine Deaths. *Review of Economics and Statistics 99*(1), 105–118.

The State Council (2007). Regulations on the Reporting, Investigation and Disposition of Work Safety Accidents [Shengchan anquan shigu baogao he diaocha chuli tiaoli].

Wang, X. (2006, January). Notice on the 2006 National Work Safety Performance Indicators [Guanyu 2006 nian quanguo anquan shengchan kongzhi kaohe zhibiao de shuoming].