"Four Years After Declaring War on Pollution, China is Winning"
Technical Support Document

This document describes the data cleaning and construction process used in this article's analysis.

Calculating Monthly Average Monitor Data, 2013 and 2017

The data referred to in this article are based upon daily $PM_{2.5}$ measurements from 750 – 1,600 ground monitors across China between 2013 and 2017, collected by the China National Environmental Monitoring Center. We construct monthly $PM_{2.5}$ estimates for each monitor by taking a simple average of all daily readings within each month. To ensure temporal coverage, only monitor months with at least 25 non-missing daily readings are considered valid. Next, we construct a balanced panel of monitor-month averages by only keeping monitors which appear in both 2013 and 2017 and have 8 or more months of available data. This ensures constant spatial coverage in our dataset for the two years, and that the selected monitors provide data throughout the year. Table 2 shows the sensitivity of pollution values to these restrictions.[1] In both 2013 and 2017, pollution values are typically missing in the months of June and August. This is a limitation of the data, although air quality is typically better in summer months.

Applying Monitor Values to Prefectures

Using the dataset described above, we estimate $PM_{2.5}$ concentrations in Chinese prefectures as the inverse distanced-weighted average of all monitors within 150 kilometers of the prefecture's area centroid.[2] That is, for prefecture i in month t, $PM_{2.5}$ is estimated using monitor values x_j for all monitors j that are within 150 kilometers of the prefecture centroid as:

$$PM_{2.5_{i,t}} = \frac{\sum_j x_j w_j}{\sum_j w_j}$$

where $w_j = \frac{1}{d_j}$ and d_j is the distance from monitor j to the prefecture centroid. This process yields PM2.5 Yearly $PM_{2.5}$ estimates for each prefecture are calculated as the simple average of these monthly estimates, yielding our final dataset. In Table 2, we show that pollution values are robust to several choices of the distance cutoff.

Prefecture Population Estimates

Population counts are calculated by aggregating LandScan gridded population datasets from 2015 to prefecture administrative boundaries. These population counts are merged with the prefecture level pollution data.

Table 1 summarizes prefecture pollution values using the completed dataset generated by the process above.

[1] Table 3 shows of the sensitivity of these values to using an unbalanced approach that uses all monitors with available data in 2013 and 2017. The full set of monitors for the two years gives lower $PM_{2.5}$ values for 2013 and higher values for 2017, than the balanced panel. We prefer the balanced panel in this analysis to ensure comparability of readings across the two years when assessing improvements.

[2] Prefecture administrative boundaries are consistent with the 2010 Chinese Census.

Estimating Years of Life Saved by Pollution Reductions

Our estimated impact of pollution on life expectancy is based on the finding in Ebenstein et al. (2017) that a 10 $\mu g/m^3$ increase in sustained exposure to PM_{10} reduces life expectancy by 0.64 years. To apply this PM_{10} estimate to our measures of $PM_{2.5}$ concentrations, we estimate the ratio of $PM_{2.5}$-to-PM_{10} in China using monitor data from the national air quality monitoring network in years where both PM_{10} and $PM_{2.5}$ data is available (2013 – 2015). The data provided to us contains annualized monitor-level measurements of the ratio of $PM_{2.5}$-to-PM_{10} from 1,612 monitors across China, for a total of 3,233 unique monitor readings. These monitor values are aggregated to a national $PM_{2.5}$-to-PM_{10} ratio by averaging monitor readings within the county they are located, and then calculating a population weighted average ratio of $PM_{2.5}$-to-PM_{10} across all counties. This procedure yields a national $PM_{2.5}$-to-PM_{10} ratio of 0.62; applying this to the Ebenstein et al. (2017) estimate of 0.64 life years lost per 10 $\mu g/m^3$ of PM_{10} yields an estimated effect 1.03 life years lost per 10 $\mu g/m^3$ of $PM_{2.5}$.

Table 1
Pollution and Life Expectancy Gains in China's Most Populated Prefectures

		$PM_{2.5}$ Concentrations ($\mu g/m^3$)			Potential
	Population	2013	2017	Difference	Life Year Gain
	(1)	(2)	(3)	(4)	(5)
All Prefectures	-	72.9	49.9	-23.0	2.4
Shanghai (Districts)	22,810,750	62.5	40.5	-21.9	2.3
Beijing (Districts)	19,323,943	90.6	58.8	-31.8	3.3
Chonqing (Districts)	16,078,945	58.6	42.7	-15.9	1.6
Chengdu	14,418,717	94.2	54.3	-39.8	4.1
Guangzhou	13,052,193	51.7	37.7	-14.0	1.4
Tianjin	11,297,312	99.1	63.6	-35.6	3.7
Baoding	11,210,792	115.5	71.7	-43.8	4.5
Haerbin	10,901,443	67.1	52.8	-14.3	1.5
Suzhou	10,707,893	70.6	44.7	-25.9	2.7
Shijiazhuang	10,454,812	133.4	81.7	-51.7	5.3

Notes: This table lists average PM2.5 concentrations for years 2013 and 2017 in 10 of China's most populated prefectures. The "All Prefectures" row refers to the 204 prefectures for which we have pollution data, and is weighted by prefecture population. The body of the technical support document describes the data generating process underlying each column. Pollution units are in micrograms per cubic meter.

Table 2
Sensitivity of Estimates to Choice of Acceptable Distance from Prefecture Centroids

	2013			2017		
	50km	100km	150km	50km	100km	150km
	(1)	(2)	(3)	(4)	(5)	(6)
All Prefectures	72.3	73.5	72.9	49.3	50.1	49.9
Shanghai (Districts)	59.9	61.3	62.5	38.7	39.4	40.5
Beijing (Districts)	89.2	90.6	90.6	58.5	58.8	58.8
Chonqing (Districts)	49.7	58.6	58.6	42.0	42.7	42.7
Chengdu	94.2	94.2	94.2	54.3	54.3	54.3
Guangzhou	52.4	52.9	51.7	37.7	38.6	37.7
Tianjin	-	103.0	99.1	-	63.5	63.6
Baoding	-	-	115.5	-	-	71.7
Haerbin	-	71.2	67.1	-	53.7	52.8
Suzhou	71.6	70.8	70.6	45.5	44.5	44.7
Shijiazhuang	135.1	135.1	133.4	83.1	83.1	81.7

Notes: This table presents prefecture-level PM2.5 estimates using three different radii (50km, 100km, and 150km) around prefecture centroids to compute inverse distance weighted averages of monitor readings. The "All Prefectures" row refers to the set prefectures for which we have pollution data, and is weighted by prefecture population. Pollution units are in micrograms per cubic meter.

Table 3
Sensitivity of PM2.5 Estimates to Monitor Selection Criteria

	2013			2017		
	Balanced Approach	All Monitors (Unbalanced)	Difference	Balanced Approach	All Monitors (Unbalanced)	Difference
	(1)	(2)	(3)	(4)	(5)	(6)
All Prefectures	72.9	69.6	3.3	49.9	50.5	-0.6
Shanghai (Districts)	62.5	60.0	2.5	40.5	42.5	-2.0
Beijing (Districts)	90.6	90.0	0.6	58.8	60.6	-1.8
Chonqing (Districts)	58.6	60.6	-2.0	42.7	46.1	-3.4
Chengdu	94.2	85.5	8.7	54.3	56.2	-1.9
Guangzhou	51.7	47.4	4.3	37.7	34.5	3.1
Tianjin	99.1	98.9	0.2	63.6	66.1	-2.5
Baoding	115.5	118.7	-3.2	71.7	81.7	-9.9
Haerbin	67.1	61.7	5.4	52.8	56.6	-3.8
Suzhou	70.6	66.1	4.5	44.7	45.5	-0.8
Shijiazhuang	133.4	126.3	7.1	81.7	87.8	-6.1

Notes: This table compares PM2.5 estimates between the balanced approach described in the body of the technical support document with an unbalanced approach that uses all monitors with available data in 2013 and 2017. The "All Prefectures" row refers to the 204 prefectures for which we have pollution data in the balanced approach. Pollution units are in micrograms per cubic meter.