Social cost of air pollution:

The Impact of Outdoor Air Pollution on Chinese Expressed Happiness through Social Media

JEL Classifications: Q5, H2

Yiren Wang
Duke Univesity
international Master of Environmental Policy 2020

Duke

Outline:

1. Problem identification
2. Literature review
3. Data & methods
4. Results & discussion
5. Conclusion

Duke

Literature review

1. **Air pollution and emotions:** Zhang and Wang (2019) found out that the control of $PM_{2.5}$ concentrations can help decrease happiness inequality and increase the individuals' **levels of happiness** in the long term.

2. **Social media and emotion sharing:** The statistical 0.718 ($n=74$, $p<.001$) result from Wang et al. (2015) shows a significant correlation between particle pollution levels and pollution-related **messages** in 74 cities in Sina Weibo.

3. **Sentiment Analysis Methods:** On average, 91.48% of language model-based sentiment classifications can automatically capture writers' opinions and match human perception (Siebert et al 2019).

Problem identification

1.

Air Pollution Impacts on Expressed Happiness

Concentrations of $PM_{2.5}$ or Air Quality Index

10 thousands of posted Weibo contents data

Baidu sentiment analysis Service

2. Air pollution: Ministry of Ecology and Environment, **national** government;
3. Baidu sentiment analysis: http://ai.baidu.com/tech/nlp/sentiment_classify;
4. Weather and holiday data**: regional** governments.

Duke

Problem identification

5. Weibo contents: obtain by the web crawler
   - Time selection: from **January to March**;
   - Search items: location (3) + key words (3) = **9 combinations**;

E.g.: Beijing + haze ← 北京+雾霾

Shanghai + emotion; Guangzhou + weather 上海+心情；广州+天气

   - Random selected: however, retrieve full text documents for the top 50 pages of main posts, **not a quasi-collection** process.

- Stimulation: define, direct mentioned ~ indirect behaviors.

Haze > weather> emotion

Duke

Problem identification

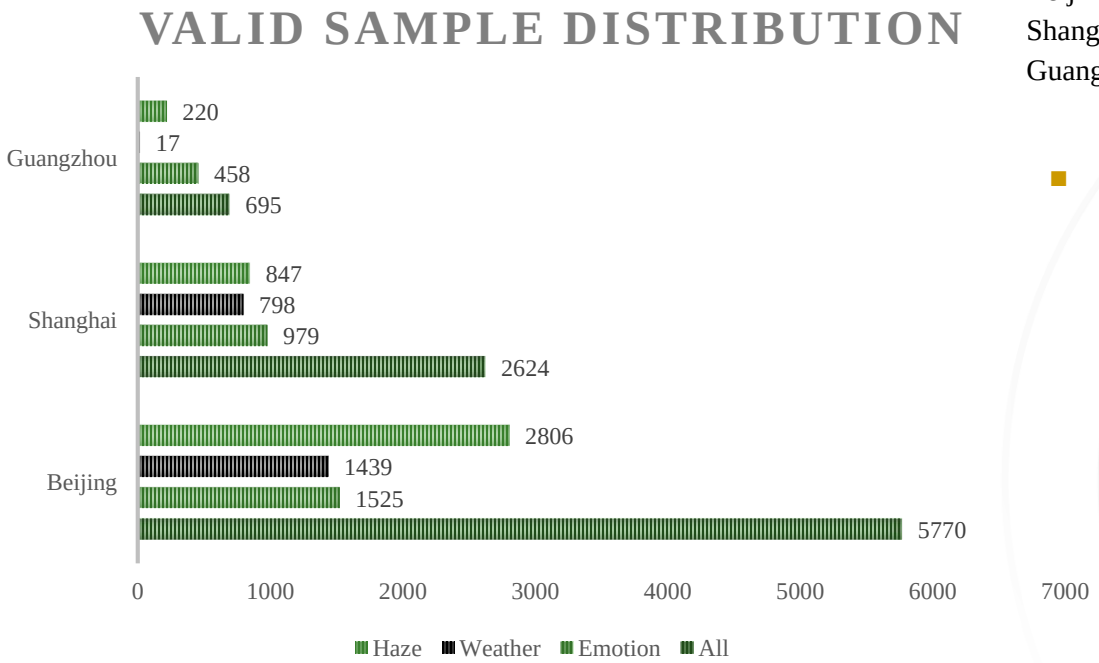
- Hypothesis:

    1. Air quality worse (AQI increase) ↑
    2. People's expressed happiness ↓

Duke

Date and Methods

1. Sample distribution



Valid samples	Haze	Weather	Emotion	All
Beijing	2806	1439	1525	5770
Shanghai	847	798	979	2624
Guangzhou	220	17	458	695

- 9087 valid data;
  - Locations:
    - **63.5% Beijing**;
    - 28.9% Shanghai;
  - Keywords
    - 42.6% Haze;
    - 24.8% Weather;
    - 32.6% Emotion.

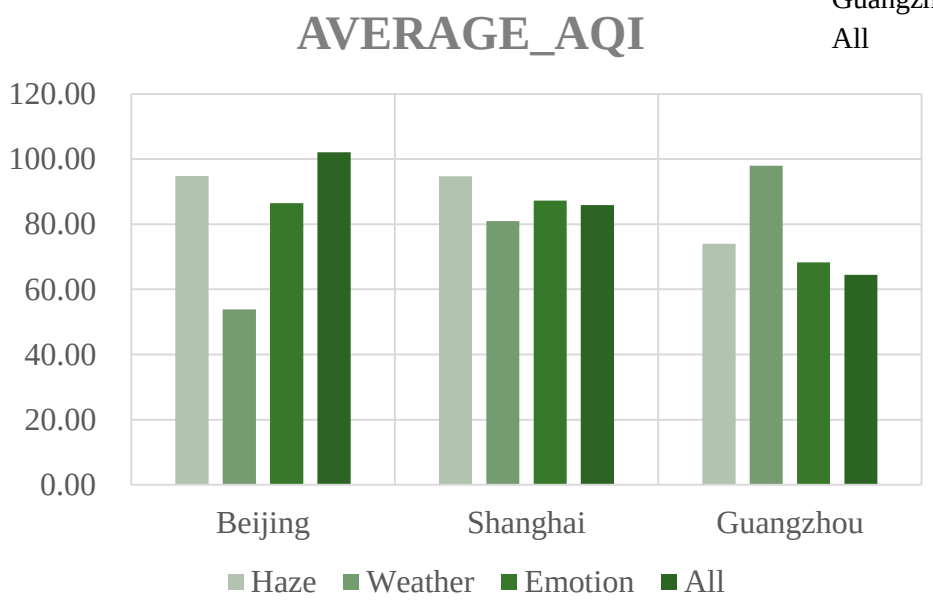
Duke

Date and Methods

2. Air Quality Index

Average_AQI	Haze	Weather	Emotion	All
Beijing	**94.79**	53.82	**102.13**	86.52
Shanghai	94.69	80.95	85.92	87.24
Guangzhou	74.00	97.94	64.47	68.31
All	93.59	63.76	90.95	85.33

AQI classification standard

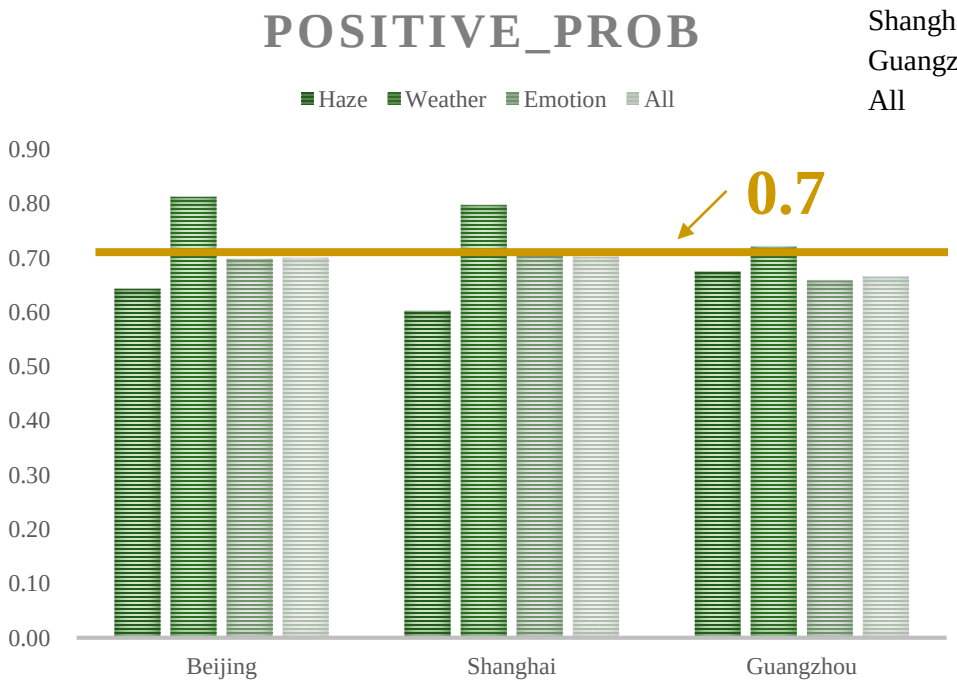


- Heavily polluted:
  - Beijing>Shanghai>Guangzhou
- Trigger or stimulation:
  - Keywords;
  - Complain more;
  - Guangzhou: lack of data.
  - …

Duke

Date and Methods

3. Expressed happiness



Positive_prob	Haze	Weather	Emotion	All
Beijing	0.64	0.81	0.70	0.70
Shanghai	0.60	0.80	0.71	0.70
Guangzhou	0.68	0.72	0.66	0.67
All	0.64	0.81	0.70	0.70

- **Stable**: 0.67~0.7
- Expressed Happiness:
  - Happier: weather, **0.81**;
  - Middle: emotion, 0.70;
  - Less happy: haze, **0.64.**

Duke

Results

1. Sentiment & Positive probability

Table 4: The effect of using sentiment and positive probability

	(1)	(2)
VARIABLES	Sentiment	Positive prob
AQI	-0.00205***	-0.00170***
		-
	(0.000499)	(0.000526)
Constant	1.197***	1.370***
	(0.0497)	(0.0523)
Observations	9,087	9,087

*** p < 0.01, ** p < 0.05, * p < 0.1

Significant

- **Confidence degree (0~1)**
- **Neutral** (!!)
  Unknown: identification rules
- **"Positive probability"**
  With six significant digits
  Certainty?

Duke

Results

2. Locations & Keywords

Table 2 | The effect of locations and keywords

VARIABLES	(1) positive prob	(2) positive prob	(3) positive prob	(4) positive prob
AQI	-0.0382***	-0.0408***	-0.00149	-0.00263
	(0.00635)	(0.00638)	(0.00641)	(0.00648)
Constant	73.02***	73.50***	63.74***	64.27***
	(0.617)	(0.665)	(0.745)	(0.773)
Observations	9,087	9,087	9,087	9,087
R-squared	0.004	0.006	0.057	0.058
Locations	NO	Yes	NO	YES
Keywords	NO	NO	YES	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Significant

- **Locations** (!!)
  - Beijing
  - Shanghai
  - Guangzhou
- **Keywords** (?)
  - Haze
  - Weather
  - Emotion

Duke

Results

Table 6: The effect of various factors

	(1)	(2)	(3)	(4)
VARIABLES	positive prob	positive prob	positive prob	positive prob
AQI	-0.0336***	-0.0456***	-0.00168	-0.00974
	(0.00646)	(0.00655)	(0.00660)	(0.00690)
Max temperature	0.663***	0.960***	0.233***	0.402***
	(0.0571)	(0.0647)	(0.0612)	(0.0730)
Precipitation	-0.634	0.826	-0.140	0.884
	(1.989)	(1.997)	(1.955)	(1.970)
Holiday	-4.273***	-2.478***	-2.724***	-2.052***
	(0.770)	(0.788)	(0.762)	(0.778)
Constant	65.10***	63.19***	62.07***	61.55***
	(0.937)	(0.963)	(0.942)	(0.963)
Observations	9,087	9,087	9,087	9,087
R-squared	0.026	0.036	0.060	0.062
Locations	NO	Yes	NO	YES
Keywords	NO	NO	YES	YES

*** p < 0.01, ** p < 0.05, * p < 0.1

3. Parameters:

- **How to understand**
- ✓ Negative:
  AQI
  Holiday
- ✓ Positive:
  Maximum temperature
  Precipitation
- **Location matters**
- **“Concern” keyword?**

Table 7: The effect of a potential delay

VARIABLES	(1) instant	(2) 1-day delay	(3) 3-day delay	(4) 7-day delay
AQI	-0.0456***			
	(0.00655)			
1-day delay		-0.0514***		
		(0.00718)		
3-day delay			-0.0696***	
			(0.00869)	
7-day delay				-0.0973***
				(0.0131)
Max temperature	0.960***	0.942***	0.932***	0.993***
	(0.0647)	(0.0643)	(0.0640)	(0.0654)
Precipitation	0.826	-0.0541	-1.035	-1.899
	(1.997)	(1.992)	(1.993)	(2.005)
Holiday	-2.478***	-3.093***	-3.830***	-3.089***
	(0.788)	(0.769)	(0.760)	(0.768)
Constant	63.19***	62.58***	64.31***	65.47***
	(0.963)	(0.927)	(1.002)	(1.110)
Observations	9,087	9,087	9,087	9,087
R-squared	0.036	0.037	0.038	0.037
Locations	YES	Yes	YES	YES

*** p < 0.01, ** p < 0.05, * p < 0.1

5. Potential delay:

Time lag effect?

- ✓ Instant: at the same day:

1_day: the average (instant + 1 day before);

3_day: the average (instant + 3 days before);

7_day: the average (instant + 7 days before).

Yes

Table 8-1: The effect of expressed happiness in three cities

	(1)	(2)	(3)
VARIABLES	Beijing	Shanghai	Guangzhou
AQI	-0.0355***	-0.0337***	-0.0455***
	(0.00647)	(0.00646)	(0.00655)
Max temperature	0.731***	0.662***	0.949***
	(0.0595)	(0.0571)	(0.0642)
Precipitation	0.423	-0.793	0.517
	(2.005)	(2.000)	(1.983)
Holiday	-3.954***	-4.262***	-2.512***
	(0.773)	(0.770)	(0.787)
Constant	62.68***	64.98***	63.05***
	(1.110)	(0.949)	(0.957)
Observations	9,087	9,087	9,087
R-squared	0.028	0.027	0.036
Keywords	NO	NO	NO

*** p < 0.01, ** p < 0.05, * p < 0.1

Table 8-2: The effect of three different "concern" keywords

	(1)	(2)	(3))
VARIABLES	Haze	Weather	Emotion
AQI	-0.0303***	-0.0125*	-0.0466***
	(0.00661)	(0.00692)	(0.00658)
Max temp	0.654***	0.505***	0.963***
	(0.0687)	(0.0721)	(0.0647)
Precipitation	1.082	0.674	0.868
	(1.980)	(1.977)	(1.997)
Holiday	-2.303***	-2.079***	-2.491***
	(0.781)	(0.780)	(0.788)
Constant	69.30***	62.78***	63.00***
	(1.074)	(0.953)	(0.972)
Observations	9,087	9,087	9,087
R-squared	0.053	0.056	0.037
Locations	YES	YES	YES

*** p < 0.01, ** p < 0.05, * p < 0.1

Limitation

- Data collection
  - Only provide 50 pages of results;
  - The priority for rankings.

- Sentiment analysis
  - Black box: Baidu sentiment analysis service;
  - Picture sharing behaviors.

Overestimate or underestimate?

Not so sure

Duke

Limitation

- Time span
  - 3 months;
  - Close the access permissions.

- Heavily polluted:
  - AQI (151~235) only composed of 7.51%;
  - Provide limited support.

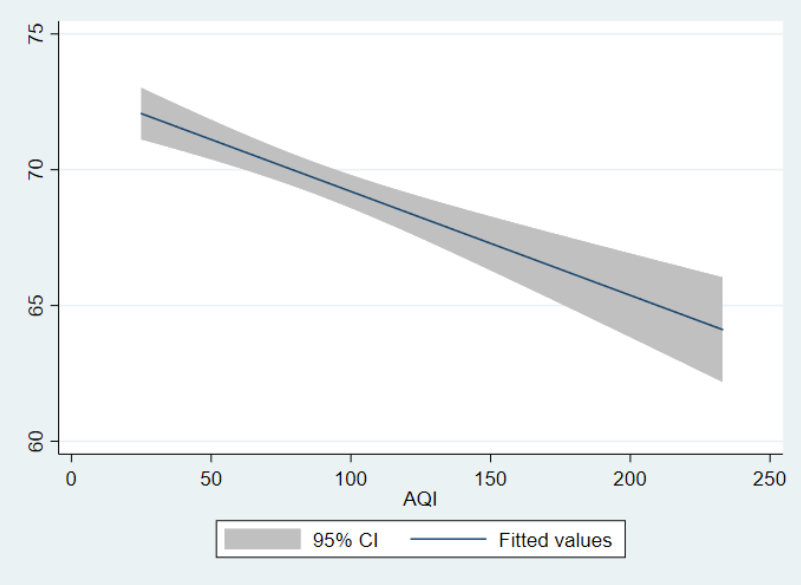
Overestimate or underestimate?

Not so sure

Duke

Conclusion

- AQI: **significant** effect of the expressed happiness
- 0 (air quality pretty good) to 150 (middle level polluted), the average expressed happiness decreased by **5.73%**.



Duke

References

Sha, Y., Yan, J., & Cai, G. (2014, January). Detecting public sentiment over PM2. 5 pollution hazards through analysis of Chinese microblog. In *ISCRAM*.

Cui, C., Wang, Z., He, P., Yuan, S., Niu, B., Kang, P., & Kang, C. (2019). Escaping from pollution: the effect of air quality on inter-city population mobility in China. *Environmental Research Letters*, *14*(12), 124025.

He, K., Yang, F., Ma, Y., Zhang, Q., Yao, X., Chan, C. K., Cadle, S., Chan, T. & Mulawa, P. (2001). The characteristics of PM2. 5 in Beijing, China. *Atmospheric Environment*, *35*(29), 4959-4970.

Ebenstein, A., Fan, M., Greenstone, M., He, G., Yin, P., & Zhou, M. (2015). Growth, pollution, and life expectancy: China from 1991-2012. *American Economic Review*, *105*(5), 226-31.

Zheng, S., & Kahn, M. E. (2013). Understanding China's urban pollution dynamics. *Journal of Economic Literature*, *51*(3), 731-72.

Zhang, X., Zhang, X., & Chen, X. (2017). Valuing air quality using happiness data: the case of China. *Ecological economics*, *137*, 29-36.

Zhang, X., Zhang, X., & Chen, X. (2017). Happiness in the air: how does a dirty sky affect mental health and subjective well-being?. *Journal of environmental economics and management*, *85*, 81-94.

Franklin, M., Zeka, A., & Schwartz, J. (2007). Association between $PM_{2.5}$ and all-cause and specific-cause mortality in 27 US communities. *Journal of Exposure Science and Environmental Epidemiology*, *17*(3), 279.

Darçın, M. (2014). Association between air quality and quality of life. *Environmental Science and Pollution Research*, *21*(3), 1954-1959.

Duke

References

Cuñado, J., & De Gracia, F. P. (2013). Environment and happiness: New evidence for Spain. *Social Indicators Research*, *112*(3), 549-567.

Li, Y., Guan, D., Tao, S., Wang, X., & He, K. (2018). A review of air pollution impact on subjective well-being: Survey versus visual psychophysics. *Journal of cleaner production*, *184*, 959-968.

Welsch, H. (2002). Preferences over prosperity and pollution: environmental valuation based on happiness surveys. *Kyklos*, *55*(4), 473-494.

Zhang, P., & Wang, Z. (2019). PM2. 5 Concentrations and Subjective Well-Being: Longitudinal Evidence from Aggregated Panel Data from Chinese Provinces. *International journal of environmental research and public health*, *16*(7), 1129.

Davis, J. J., & O'Flaherty, S. (2012). Assessing the accuracy of automated Twitter sentiment coding. *Academy of marketing studies journal*, *16*, 35.

Zheng, S., Wang, J., Sun, C., Zhang, X., & Kahn, M. E. (2019). Air pollution lowers Chinese urbanites' expressed happiness on social media. *Nature Human Behaviour*, *3*(3), 237.

He, K., Yang, F., Ma, Y., Zhang, Q., Yao, X., Chan, C. K., Chan, T. & Mulawa, P. (2001). The characteristics of PM2. 5 in Beijing, China. *Atmospheric Environment*, *35*(29), 4959-4970.

Ebenstein, A., Fan, M., Greenstone, M., He, G., Yin, P., & Zhou, M. (2015). Growth, pollution, and life expectancy: China from 1991-2012. *American Economic Review*, *105*(5), 226-31.

Zheng, S., & Kahn, M. E. (2013). Understanding China's urban pollution dynamics. *Journal of Economic Literature*, *51*(3), 731-72.

Duke

Thanks for listening!

Duke