Environment and Development Economics (2019), **24**, 1–22
doi:10.1017/S1355770X18000414

EDE

RESEARCH ARTICLE

A dragon eating its own tail: public control of air pollution information in China

Chiara Ravetti[1]*, Tim Swanson[2], Yana Jin[3], Quan Mu[3] and Shiqiu Zhang[4]

[1]Polytechnic of Turin, Turin, Italy; [2]Graduate Institute of Geneva (IHEID), Geneva, Switzerland; [3]The College of William and Mary, Williamsburg, Virginia, USA and [4]Peking University, College of Environmental Sciences and Engineering, Beijing, China
*Corresponding author. E-mail: chiara.ravetti@polito.it

(Submitted 4 August 2017; revised 30 May 2018; accepted 27 August 2018; first published online 9 November 2018)

Abstract

This paper analyses the implications of government control over public information about air pollution. First, we model the incentives for a local government with control over the media to affect popular perception concerning pollution. We argue that biased announcements can influence the inflows of labour force in a municipality beyond economic factors. Then, we examine some evidence on information misreporting in the context of Beijing, China. We show that official air pollution announcements diverge systematically from an alternative source of information, provided by the US Embassy. The results point at a manipulation of popular perception consistent with the motives indicated in our model. Furthermore, using an original household survey, we examine whether the distorted public signal affects agents' behaviour. We find that households that depend upon government-controlled media are significantly less responsive to pollution peaks.

Keywords: Air pollution; averting behaviour; government policy; information; labour

JEL Classification: H41; Q53; Q56; Q58

1. Introduction

Information is power, and knowledge asymmetries can benefit those who control information. Often individual economic agents cannot afford to acquire complete data about the state of the world. Hence the government can reduce knowledge asymmetries through public sector information, on issues ranging from meteorological forecasts, to inflation targets, to official socio-economic statistics (Morris and Shin, 2002). Such information is a valuable public good that can influence the expectations and behaviour of private agents. However, if a government can exercise some degree of control over the media, it may have an incentive to limit or distort information to redirect public opinion and economic choices (Williams, 2009). Several studies identify the positive effects of independent public information and a free press (Brunetti and Weder, 2003; Besley and Prat, 2006), but what happens instead when the media is not free?

In this article, we investigate whether a government with control over public information would misrepresent environmental issues to its public, and to what effect. We explore these questions in the context of air pollution in Beijing, China. This country is an emblematic example of a nation where information is strictly controlled by the government.[1] We focus on urban air pollution to analyse information control because data about air quality is hard to acquire by individual citizens: gathering precise data is costly and requires specialized technology (monitors) and knowledge of atmospheric and epidemiological research to understand the consequences for human health. Visibility, which can be used as a private proxy for pollution, is not a strong indicator for air quality, as it varies with wind and humidity, confounding with fog.[2] Therefore, air pollution information is provided in most countries by public state agencies.

In this article, we argue that a local government with control over the media has an economic incentive to manipulate the public perception of environmental hazards. We build a model of imperfect information and local governments competing for labour to define the driving forces behind pollution biases. Next, we exploit data from the capital of China to illustrate the mechanism and consequences of the distorted information signal. We find evidence of misreporting around specific thresholds of air quality in Beijing, which points at a manipulation of public perception. Moreover, looking at the the impacts of distorted information on household behaviour, we find that the agents who rely on government-controlled media are the most vulnerable to pollution.

This article contributes to the literature on how public information shapes expectations and economic outcomes. Much research has focused on macroeconomic variables or government policies,[3] with fewer articles on environmental effects (Kennedy *et al.*, 1994). Most of these works assume that the state agency only provides truthful information, or that the production of information by the media is separate from policy-makers (Dur and Swank, 2005; Angeletos and Pavan, 2007). Few articles highlight that governments might have an incentive to misreport data strategically (Michalski and Stoltz, 2013). In the case of China, a growing body of literature finds evidence of misreporting in pollution data (Andrews, 2008; Chen *et al.*, 2012; Ghanem and Zhang, 2014; Jia, 2014; Stoerk, 2016). These works highlight the *political* incentives for local government officials to present optimistic data to the central government. Our approach complements this literature from a different viewpoint, analysing an important and relatively unexplored issue: the *economic* rationale for distorted public information and its potential consequences on households. We do not rule out the political motives described in the existing literature, since our results are compatible with its findings, but we propose a further explanation that considers also the economic benefits for a government distorting information.

We proceed in three steps. First, we develop a model of public control over information in a municipality competing for labour with other regions. The government faces a trade-off between output and public health costs, both increasing with pollution levels and with the number of workers in that area. The local government has some control

[1]China ranks 176/180 countries in Press Freedom Index (Reporters Without Borders, 2016) and 100 per cent of TV, radio and major newspapers are owned by the Communist Party (Djankov *et al.*, 2003).

[2]It is difficult to distinguish if the hazy sky is caused by particulate matter or fog (Liu *et al.*, 2016). Pollution has different effects on visibility, depending on humidity and temperature (Chang *et al.*, 2009).

[3]For examples see Besley and Burgess (2002) and Gavazza and Lizzeri (2009) for policy transparency and responsiveness.

over public perception of pollution, and can thereby manipulate air quality announcements to retain workers, without the need to increase wages to compensate for pollution damages. This model predicts that a negative bias – declaring that pollution is lower than in reality – ensues whenever the government benefits more from the contribution of (cheap) labour than it pays in public health expenditures per worker.

We then illustrate our theoretical hypothesis with a study of public information and air pollution in Beijing. This city offers a unique setting to analyse information about air pollution, because of an alternative source of data besides the official Chinese one: readings reported by the US Embassy. Comparing the two sources, we find that the Chinese government downplays emissions' information as pollution increases, especially around emission levels where it can strongly affect public opinion. For instance, once pollution crosses the threshold of 100 points, the bias becomes a further 34 per cent more negative. This evidence suggests that the Chinese signal is distorted with the purpose of improving public perception of air quality. Finally, we examine the implications of these distorted signals at the household level. We employ an original household survey to show that agents who rely primarily on public information sources (publicly owned media such as television, radio and newspaper) are directly influenced by public signals, and thus less capable of responding to pollution peaks.

Our contribution is twofold: firstly, we consider theoretically the economic motives for misreporting air pollution information by a local government. This analysis complements the literature on the political incentives of Chinese officials, offering a long-term economic explanation for pollution distortions. Secondly, we find support for this mechanism of information manipulation in the case of Beijing. As a consequence, we argue that centralized control over public information can hinder decentralized decision-making in the case of self-protective health behaviours against pollution. The article proceeds as follows: section 2 develops the model about information distortion; section 3 presents the empirical analysis of information about Beijing's air pollution; section 4 examines households' behaviour with different sources of information; and finally section 5 concludes.

2. Model

We consider a country uniformly populated with two distinct groups of citizens, those who are relatively critical and well-informed, and those who are more passive with respect to information gathering and hence relatively uninformed. Only the critical group updates expectations about pollution, correcting for a potential information bias, while the passive group wholly relies on any official source of information. Individual agents cannot evaluate air quality, but they have some prior beliefs about its characteristics, namely

$$\hat{p}_t = p^n + p_t \quad \text{with} \quad p_t \sim N(0, \sigma_p^2), \tag{1}$$

where p^n is some 'natural' level of pollution, given by the geographic conformation and location of a city, and p_t captures emissions shocks that vary with meteorological conditions, traffic, construction work, etc. We assume without loss of generality that p^n is zero and focus on the variations in emissions shocks. The pollution level on a given day is unknown to individual agents, because they are unable to measure the shocks p_t. A government agency can measure actual emissions in the local economy, and thus knows the true p_t. The agency releases announcements about the quality of the air. However this government announcement A_t can include a potential bias B:

$$A_t = p_t + B \quad \text{with} \quad B \sim N(\overline{B}, \sigma_B^2). \tag{2}$$

Citizens do not know the actual bias of the government, but critical households have a prior over its distribution. This group forms expectations by solving a signal extraction problem: using their prior beliefs about the distribution of pollution and the bias, plus observations of the announcement over time $t = \{1, 2 \ldots T\}$, they form their current beliefs about pollution. The passive fraction of the population, $\lambda \in [0, 1]$, is incapable of updating expectations concerning the bias. For them, expected pollution is just the announcements, $E(p_t) = A_t$, and the expectation about the bias is zero, $E(B) = 0$. This means that, on average, only $(1 - \lambda)$ of the bias is factored out of the announcement in the whole population, yielding the following aggregate expression (derivation in the online appendix):

$$\lim_{T \to \infty} E(p) = \frac{A - (1 - \lambda)B}{z}, \tag{3}$$

where z is a weight that captures how precisely the corrected announcement translates into expectations.[4] In other words, z is a measure of how much the government can affect expectations with its announcements. Over time, aggregate expected pollution converges to the above expression, because $\underset{T \to \infty}{\text{Lim}} E(B) = B$ for those who update expectations (see online appendix). As long as there are people who cannot fully update their beliefs ($\lambda > 0$), expectations would never converge to the true value of pollution. We can then simplify this expression to $E(p) = (p + \lambda B)/z$. From this signal extraction problem, we can derive the following proposition.

Proposition 1. *An announcement A regarding emissions p that includes a bias B affects expectations about air pollution directly, as it reaches the whole population, and indirectly, entering the bias-updating process of the critical group. The marginal effect of the announcement is stronger the larger the fraction of the population that does not update expectations,* $\partial E(p)^2/\partial B \partial \lambda > 0$.

Proof : It follows straightforwardly from differentiating equation (3) with respect to B and λ. □

2.1 Competition among local governments

Next, we model how local governments release announcements about pollution, given the signal extraction process described above. We assume that each municipality competes for workers with other jurisdictions within the country, similarly to the classic literature on tax competition (Tiebout, 1956; Bucovetsky, 1991; Wilson, 1995; Janeba and Osterloh, 2013). In this literature, each government chooses the optimal policy (e.g., taxes), given mobile capital and/or labour (people 'voting with their feet'). In our case, local governments affect the expectations of mobile labour about economic conditions and expected pollution costs. Workers maximize their individual utility function, $U(c_i)$, where c_i is consumption in a location i. The workers all have identical preferences and

[4] Precisely, this factor captures a combination of the variance from the announcement and the one from the pollution emissions $z \equiv W^2(1/\sigma_p^2 + 1/W^2)$ where W^2 is the variance of $A - E(B)$.

https://doi.org/10.1017/S1355770X18000414 Published online by Cambridge University Press

the utility function is well behaved: $\partial U/\partial c_i > 0$ and $\partial^2 U/\partial c_i^2 < 0$. Perfect mobility of the workforce implies utility equalization across locations: $U(c_i) = U(c_j) \forall i; j$.[5]

Workers' consumption depends first and foremost on the real wage, but also on the expected damages from pollution. This is because pollution can hinder workers' productivity, resulting in absence from work and medical costs. The amount of damages associated with a given level of pollution is represented here by θ. So, workers' consumption in a location i is equal to the local real wage w_i less the damages from pollution in locality i., that is $c_i = w_i - \theta E(p_i)$. Assuming linear utility functions, workers' migration equalizes consumption in two locations, $w_i - \theta E(p_i) = w_j - \theta E(p_j)$.

Hence the first driver of migration is the real wage gap between two regions, a proxy for economic conditions, such as job opportunities, expected salaries, likelihood of unemployment and so on. This follows the extensive literature on the role of wage gaps in the economics of migration (Chiquiar and Hanson, 2005; McKenzie *et al.*, 2014; Munshi and Rosenzweig, 2016). The second and more novel factor is the environmental quality differential, which captures the gap between the two locations in expected pollution costs. For simplicity, we assume that production is quadratic, $\Pi(L) = aL - b/2L^2$, and labour is paid its marginal productivity, $w \equiv \Pi'(L) = a - bL$. Total labour is $\sum_{j=1}^{N} L_j = \overline{L}$. Labour in a location is then (proof in the online appendix):

$$L_i = \frac{\overline{L}}{N} - \frac{\theta}{bN}\left[(N-1)E(p_i) - \sum_{j=1}^{N-1} E(p_j)\right]. \tag{4}$$

If region i has higher expected pollution levels, it will lose workers to other regions, *ceteris paribus*.

2.2 Optimal bias in pollution announcements

A local government can achieve higher output, and thus tax revenues, by attracting labour. However it must also pay the costs of providing public goods and services to the local population. We assume that these costs derive chiefly from health care services, increasing with the size of the labour force. In this way, the government faces a trade-off between the number of workers it wishes to attract for production, and the number of individuals to which it must provide health care. The objective function of each governance unit is to maximise tax revenues less the costs of local public goods (e.g., medical/health care).[6] The government chooses the pollution announcements (and tax revenues)[7] subject to the competition for the supply of labour described above:

$$\max_{A_i, \tau_i} V_i = \tau_i R(L_i) - \phi L_i \quad \text{s. t.} \quad L_i(w_i; E(p_i)), \tag{5}$$

where V_i is the value function for a government, τ_i is the tax rate on revenues R, which derives directly from the supply of workers, and ϕ is the health care costs of local workers

[5]This model abstracts from migration costs and socio-economic factors specific to China, such as the permanent residence registration system (hukou), land-sale policies, infrastructure, migrants' networks, etc.

[6]The government's revenues are proportional to the value of production, and hence a function of labour as well, $R(L) = \varrho \Pi(L)$. We normalize the price of output ϱ to 1 to simplify exposure.

[7]See Naso and Swanson (2017) for a discussion of the optimal selection of taxes τ_i.

that is borne by the government. The objective function is subject to the labour supply function $L_i(w_i; E(p_i))$ from equation (4), with expected pollution deriving from the signal extraction process from equation (3). We then calculate the optimal bias for a local government[8]

$$B^* = \frac{1}{\lambda}\left[\frac{z}{(N-1)}\left(\frac{b\overline{L}}{\theta} + \sum_{j=1}^{N-1} E(p_j)\right) - p - \frac{zN}{(N-1)}\left(\frac{\tau a - \phi}{\tau\theta}\right)\right]. \tag{6}$$

This result shows that a local government has an incentive to distort information on account of the net benefit it receives from labour's production. If provided with accurate information on the environment, there will come a point when the effective wage in a given locality $(w_i - E(p_i))$ is no longer attractive. The local government can counteract the real wage gap, via its control over $E(p_i)$. In this way, the information distortion becomes an instrument for marginally mobilising labour beyond existing economic conditions. The optimal distortion in equation (6) depends on several factors: the country's overall labour force and pollution in all localities (first term), the local pollution (second term), and the net value of attracting workers (last term). We summarise the effect of each term of equation (6) in Proposition 2, focusing on the case in which B becomes more negative, announcing air pollution to be lower than in reality:

Proposition 2. *The optimal bias B^* for the government of locality i becomes increasingly negative to the extent that: i) the overall labour stock $\overline{L}$ gets smaller; ii) pollution in other municipalities falls, creating further competition; iii) local labour pollution p rises; iv) the net benefits from attracting labour are greater – tax revenues per labour unit of production τa exceed public labour cost ϕ. The relative weight of these factors determines the sign of the bias and its magnitude. The absolute value of the bias increases with the proportion of the informed populace (1-λ). Finally, the noise in the informational environment z impacts upon optimal distortion with an ambiguous sign.*

Proof : From inspection of the various terms of equation (6). □

In sum, the competition for labour between municipalities can offer an explanation for long-term distortions in information concerning pollution levels. We hypothesize that information biases can be used to attract at the margin mobile labour to polluted localities, even when it is sub-optimal from the worker's perspective to go there, given the real wage. Of course there are multiple other political factors that create an incentive for pollution misreporting, as already well documented in the case of China by the literature (Andrews, 2008; Chen *et al.*, 2012; Ghanem and Zhang, 2014). The number of *Blue Sky Days*, for instance, is one of the indicators that qualifies a city as a 'national environmental protection model' (Chen *et al.*, 2012) and enters the performance assessment of local officials (Stoerk, 2016). Given the media salience of Blue Sky Days, these political motives can give a strong short-term incentive to distort the official data. However, we want to highlight in this model that the local government would also achieve long-term economic benefits from convincing its local workers that pollution is not a major problem. We now turn to the empirical evidence in support of this theory of informational interventions.

[8]We suppress subscripts for ease of exposition, since the problem is identical for each of the governance units *i*.

3. Empirical analysis

In the previous section, we argue that a local government with control over the media may have an incentive to distort environmental information to retain labour at a given real wage. In order to find some suggestive evidence for this hypothesis, we need to prove that data manipulations actually occur, and that the behaviour of the local population is affected by these distortions. First of all, we test whether the bias relates to the local pollution level and the characteristics of the local informational environment. Secondly, we examine the response to pollution peaks in the share of the population λ that does not 'update' expectations. For the first part, we rely on the insights on the optimal bias B^* in equation (6) and Proposition 2. We focus on a reduced form expression:

$$B = \alpha_1 p + \alpha_2 z_1 - \alpha_3 z_2, \tag{7}$$

where B is the bias, p is pollution and z a proxy for how effectively the government can manipulate public perception. We are not able to fully test the model in the context of Beijing, because it captures the case of only one municipality, but we can test whether the bias seems to target public perceptions of air pollution, as hypothesized in our model. Next, we can test if this mechanism affects households that fully rely on announcement without updating for potential biases.

3.1 Air quality signals and health risk perceptions in Beijing

As a micro-illustration of the phenomenon of data manipulation, we now examine public information about air pollution from the point of view of one locality, namely the capital of China, Beijing. This is only one case in the heterogeneous landscape of Chinese urban centres, and while other studies have analysed the challenges of China's urban pollution across the country (Zheng and Kahn, 2013), we focus on Beijing because it offers a useful context in which to examine data manipulation. This analysis provides an illustration of the mechanism presented above, and is not intended as a universal proof that could be generalized to other cities (Beijing is quite unique in its institutional, geographical and cultural characteristics).

In China, the Ministry of Ecology and Environment, formerly known as the Ministry of Environmental Protection of the People's Republic of China (MEP), communicates to the public the state of air quality through an Air Pollution Index (API), with a format analogous to indexes used in the USA, Canada and the European Union. These indexes reflect international standards and health risks defined by the World Health Organization (WHO, 2005), and convey information about the local data on pollution risk through a simple rating of air quality. The signal needs to be understandable to the general public, thus it uses color-codes ranging from green (lowest pollution) to dark red (highest pollution). Each colour corresponds the potential health damages associated with that pollution range.

Moreover, in Beijing, there exists a second source of information about air quality: the hourly Twit provided by the US Embassy. Table A2 in the online appendix compares the two measures up until 2013, showing the differences in the construction of the two indexes. By construction, the single values of the two indexes can differ substantially: the US index is based on real time data from the Embassy district of Chaoyang, while the Chinese signal is the average of monitors all over the city; the single pollutants considered differ, as the US includes only PM2.5, particulate matter of fine diameter,

https://doi.org/10.1017/S1355770X18000414 Published online by Cambridge University Press

while the Chinese index measures PM_{10}, SO_2, NO_x, but in our time frame not $PM_{2.5}$.[9] Lastly, the US index is an hourly measurement over the whole day, while the Chinese index provides the average pollution in an 11-hour window. However, interestingly for our study, both indexes convey alternative *information signals* about health risks that can influence the public perception of pollution. Crucially, the colour coding used to communicate information is the same, following WHO recommendations. An index below 100 implies little risk of health damages (green); then, as the index rises between 100–200 (yellow) and 200–300 (orange), more people can be affected by pollution; and a signal above 300 is defined as a health alert (dark red), with all the population risking severe health consequences.

The two indexes are not directly comparable in terms of their pollution measurements, because of spatial differences, time of measurement and pollutants content. Nevertheless, for the purpose of our analysis what matters is the end-point information signal seen by the population. A Beijing dweller only observes the final information delivered by the government and possibly compares it to the information from the US Twit. In line with our model, not all Chinese people might have easy access to the latter, due to internet restrictions, and hence we have a source of variation in announcements and access to these announcements that can illustrate our mechanism of information control.

3.2 Air pollution information

We analyse four and a half years of daily pollution announcements in Beijing, both from the official government and from the US Embassy, starting from the first available date, 25 August 2008, up to January 2013. We do not consider further dates because recently the Chinese index has come under revision (Stoerk, 2016).[10] Since the US Embassy data is more frequent, as it is reported hourly, while the Chinese data is only daily, we can then experiment with different aggregation strategies over the day to compare the US and the Chinese index (see the Robustness section). To be conservative, we use as our baseline the daily *minimum* of air pollution signal communicated by the US embassy. Figure 1 shows a snapshot (not statistically representative) of the two indexes during a short time period within our dataset. During this period, in August 2008, the Beijing Summer Olympic games took place, and economic activity and construction work were restricted to improve air quality in the city. Yet, despite these precautionary measures, we still see that the two sources of information often convey a different message in terms of health hazards.

We must examine a longer time period to see if there is a systematic difference between the message conveyed by the two sources. The difference on a given day could be partly attributed to random noise or measurement error in one (or both) signals. However we can test if there is a recurring pattern in the gap between the indexes. To confirm our hypothesis that information control is used to influence public perception, we should not observe just a constant difference between the two time series, which could be due to monitors' sensitivity, or even to their orientation towards a polluted street. We must

[9] For this reason, we compare only those days when the main pollutant in the Chinese index was PM_{10}, which has been shown to be highly correlated with $PM_{2.5}$ (Liu *et al.*, 2015).

[10] In February 2012 (regulation HJ 633–2012), China defined a new air quality index that includes $PM_{2.5}$ and ozone. This change does not take effect nationwide until 2016, but Beijing already began piloting it in January 2013 (Ministry of Environmental Protection – source: http://www.mep.gov.cn).

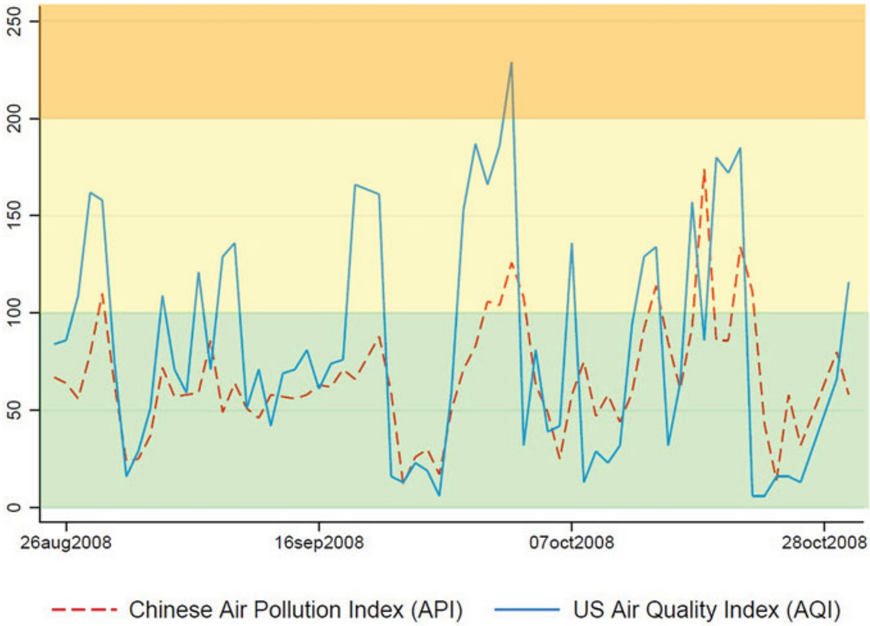

Figure 1. Mismatch between Chinese and US index (daily minimum) in some months of 2008.

examine if there are any systematic shifts that would suggest an intentional manipulation of public perception of air pollution.

3.3 Empirical model

The key issue for our analysis is how pollution indexes translate into information for the population. This loosely corresponds to the informational environment discussed in the model. Most people in a country would not pay systematic attention to the value of a pollution index and its minor variations. Most likely, the population would notice only the colour on the health risks scale. The colour coding system is artificially overlapped on a continuous variable, the air quality measure. It groups together ranges of values and imposes thresholds and discontinuities that do not exist intrinsically in the atmospheric concentration of pollution.[11] This discontinuous signal is useful to identify any potential manipulation of popular perception. All the factors that can influence a pollution measure (monitor sensitivity, measurement error, spatial variability, and so on) are unrelated to the colour coding scale, unless there is some bias introduced in the announcements explicitly targeting popular perception. Following the reduced form model in equation (7), we exploit the variations in colour coding to estimate an

[11]For instance, an index value of 99 represents a pollution level similar to 101, but the first would be perceived as clean air (green), and the second as mildly polluted (yellow). Conversely, an index of 101 seems virtually identical to 149, because both fall in the yellow region.

https://doi.org/10.1017/S1355770X18000414 Published online by Cambridge University Press

auto-regressive moving-average model by unconditional maximum likelihood:

$$G_t = \alpha + \beta \ln(\mathrm{AQI})_t^{\mathrm{US}} + \sum_{i=1}^{p} \gamma_i T_i + \sum_{i=1}^{p} \delta_i T_i \ln(\mathrm{AQI})_t^{\mathrm{US}}$$

$$+ \sum_{i=1}^{q} \phi_i G_{t-j} + \sum_{l=1}^{r} p_l \epsilon_{t-l} + \eta_m + \sigma_y + \epsilon_t,$$

where the dependent variable G_t is the gap between the information signal provided by the Chinese government and the US one, namely $G_t \equiv ln(API)_t^{China} - ln(AQI)_t^{US}$, as a proxy for the bias.[12] Its average value is -0.1, indicating that the Chinese signal is slightly lower than the US one.[13] As an alternative dependent variable, we construct a categorical variable taking the value of zero whenever the two indexes indicate the same colour (even if their specific value is not the same), and a different value if instead they indicate different risks. More precisely, we construct the categorical gap to take a positive value whenever the Chinese signal indicates a higher risk level than the US signal, and negative otherwise. Then the absolute value of the gap takes the value of 1, 2 or 3 depending on how many categories of distance differentiate the two indexes (for instance, if the US index signals Orange, while the Chinese one Yellow, the categorical gap would take the value of -1). In most cases the distance in information signals is only of 1 degree of risk. A full description of all cases of this categorical dependent variable is provided in table A4 in the online appendix. T_i is a dummy variable equal to zero when the US measure of air pollution is below an information threshold, and equal to 1 above it, with $i \in [100, 200, 300]$. It captures the different 'regions' of information about pollution damage: for example a value of the index above 300 means that the air is 'Heavy Polluted' (dark red). Any significant action around these thresholds suggests some manipulation of the qualitative message that the index conveys.

The thresholds loosely capture the effect of z, the noise (non-informative part of the signal) in the announcements from the theoretical model. The 'imprecision' with which agents translate a change in government signal into their expectations depends on the proximity to a threshold. Changes around a threshold convey a strong difference in signals (e.g., Red means something clearly different than Orange). However, farther away from the thresholds, the announcements become more noisy: for example, when in the middle of the Yellow category, a change in the government announcement imprecisely affects public expectations, because of the way the signals are structured. Thus, to model the proximity to a threshold, we interact the T-thresholds with the US pollution index, $T_i \times ln(AQI^{US})$. Intuitively, the incentive to introduce a bias should be stronger near the threshold, where the informational environment changes significantly, but then would diminish as pollution gets away from the crossing point. Then, to capture persistence in shocks and stock of pollution, and to correct for serial correlation over time, we include an autoregressive term G_{t-j}, with lags of the dependent variable, and lags of the error

[12] Or, equivalently, the natural log of the ratio of the two indexes, with coefficients interpreted as percentage changes.

[13] See the online appendix for summary statistics (table A3) and autocorrelation and partial autocorrelation functions (figure A5), showing significant autocorrelation in the dependent variable, which requires modelling of the time-dependent components to ensure that the error term is white noise.

term, ϵ_{t-l}.[14] We include month and year fixed effects, η_m and σ_y, to capture anomalous events in the dataset, like the Beijing Olympic games.[15]

3.4 Air pollution results

The results of different model specifications are presented in table 1 for the gap in the (continuous) indexes and, for robustness, in table 2 for the categorical dependent variable.[16]

The negative and significant coefficient of the US index derives automatically from our definition of the gap, such that, whenever pollution captured by the US measurement rises, the gap between the Chinese and US announcements (the empirical counterpart for our bias) becomes more negative: a 10 per cent increase in the minimum daily pollution measured by the US index makes the ratio of Chinese/US signals more negative by 9 per cent. More interestingly, we can observe an effect around information thresholds. Crossing a threshold for the US monitors always has a negative effect, although in this empirical specification the only significant one is the 100 threshold. Crossing the 100 threshold, we find that: i) it directly reduces the ratio of indexes (intercept) by 0.44, and ii) this negative effect is hampered as pollution gets higher, as shown by the positive coefficient of the interaction term (slope). The combination of the two effects shows that the strongest influence on the bias is when immediately surpassing the crossing point of 100, but it gets weaker as we move away from the thresholds and pollution increases. This second effect reflects the hypothesis about 'noise' z mentioned above.[17] In terms of the percentage impact of the threshold dummy in this semi-logarithmic specification, we have a decline of around 34 per cent in the ratio API/AQI.[18] The different columns of table 1 show three different specifications changing the number of autoregressive and moving average terms.[19]

[14]Our time series for pollution are long and, according to Dickey-Fuller and Phillips-Perron unit root tests, stationary (see online appendix). The underlying data-generating process may have some persistence (since pollution can last a few days), but no extremely long memory. Figures A3 and A4 in the online appendix show that most of the autocorrelation takes place in the first period.

[15]We also checked for seasonal components, but found no evidence for them, and the log of the time series takes care of heteroskedasticity in the variance.

[16]Specifications with more lags yield similar results, but we keep the most parsimonious models with white noise error terms, according to a Portmanteau (Q) test that all autocorrelation coefficients are jointly equal to zero.

[17]The magnitude of the second effect, however, is quite small: overall, the combined effect of crossing the 100-point threshold is $-0.44 + 0.004 \times \ln(\text{AQI})$, which ranges from -0.42 when the AQI is exactly 100 to -0.41 for higher pollution levels. A simple Wald test shows that both coefficients are jointly significant in the model estimated.

[18]This is calculated as $100[\exp(-0.4) - 1]$, following Halvorsen and Palmquist (1980). If we assumed a normal distribution of the errors, a consistent and almost unbiased estimator of the effect would be one that corrects for the variance of the estimated coefficient $100[\exp(-0.4 - 1/2\nu * (-0.4)) - 1]$, where $\nu*$ is the estimated variance. This however yields an almost identical result, a 33 per cent fall.

[19]These models can be compared with the Akaike and the Bayesian Information Criterion for model selection through the relative goodness of fit. Keeping the models with lowest information criteria, we compute in-sample forecasts to see which one performs best in terms of predictive power. Comparing the mean squared errors of our forecasts, the most suitable lag structure is the moving average one period lag, MA(1). Forecasts for the best fitting model are plotted in figure A6 in the online appendix. The model closely follows the fluctuations in the bias.

https://doi.org/10.1017/S1355770X18000414 Published online by Cambridge University Press

Table 1. Discrepancy between the Chinese and US index (daily minimum)

	Dependent variable: Gap China – US minimum signal		
	(1)	(2)	(3)
US AQI (min)	−0.952*** (0.02)	−0.950*** (0.02)	−0.951*** (0.02)
Min. AQI above T100	−0.441*** (0.09)	−0.449*** (0.09)	−0.442*** (0.09)
Min. AQI above T200	−0.066 (0.39)	−0.037 (0.39)	−0.062 (0.39)
Min. AQI above T300	0.333 (0.93)	0.308 (0.95)	0.326 (0.93)
T100 * Min. AQI	0.004*** (0.00)	0.004*** (0.00)	0.004*** (0.00)
T200 * Min. AQI	−0.000 (0.00)	−0.000 (0.00)	−0.000 (0.00)
T300 * Min. AQI	−0.001 (0.00)	−0.001 (0.00)	−0.001 (0.00)
Constant	4.018*** (0.14)	4.010*** (0.14)	4.016*** (0.14)
Month FE	YES	YES	YES
Year FE	YES	YES	YES
ARMA			
L.ar	0.235*** (0.03)		0.187 (0.12)
L.ma		0.234*** (0.04)	0.052 (0.13)
sigma			
Constant	0.276*** (0.01)	0.276*** (0.01)	0.276*** (0.01)
Observations	876	876	876
AIC	290.799	291.657	292.729
BIC	414.958	415.817	421.664

Notes: Standard errors in parentheses. ***$p < 0.01$.
The lower panel shows the autoregressive moving average lags (ARMA) components.

If we are concerned that the values of the two indexes could include significant structural differences or measurement error, we can analyse the difference between the informational content of the two signals, looking at the categorical dependent variable. Results are shown in table 2. All three thresholds are now significantly negative. The interpretation of the interaction coefficient is not the same as in the previous continuous gap, because in this case there are no values at the center of a category (like a 149 and a 101 both being Yellow in the previous case). The significant interaction of T100 × AQI only indicates that the more downward bias is present when AQI is closer to 100, as one

https://doi.org/10.1017/S1355770X18000414 Published online by Cambridge University Press

Table 2. Categorical discrepancy between the Chinese and US index (daily minimum)

	Dependent variable: *Categorical* Gap China – US minimum signal		
	(1)	(2)	(3)
US AQI (daily minimum)	0.157*** (0.04)	0.159*** (0.04)	0.150*** (0.04)
Min. AQI above T100	−0.973*** (0.06)	−0.969*** (0.06)	−0.967*** (0.06)
Min. AQI above T200	−2.528*** (0.53)	−2.513*** (0.53)	−2.514*** (0.54)
Min. AQI above T300	−4.071*** (1.58)	−4.068** (1.68)	−4.006** (1.57)
T100 * Min. AQI	−0.002*** (0.00)	−0.002*** (0.00)	−0.002*** (0.00)
T200 * Min. AQI	0.003* (0.00)	0.003 (0.00)	0.004 (0.00)
T300 * Min. AQI	0.005 (0.00)	0.005 (0.00)	0.005 (0.00)
Constant	−0.283 (0.17)	−0.294* (0.17)	−0.253 (0.17)
ARMA			
L.ar	0.162*** (0.03)		0.602*** (0.13)
L.ma		0.143*** (0.03)	−0.459*** (0.15)
sigma			
Constant	0.428*** (0.01)	0.429*** (0.01)	0.427*** (0.01)
Observations	876	876	876
AIC	1044.812	1047.200	1042.141
BIC	1145.095	1147.483	1147.199

Notes: Standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.
The lower panel shows the autoregressive moving average lags (ARMA) components.

would expect. The same specification without interaction terms would yield the same results.

3.5 Discussion

In the previous analysis we performed the most conservative comparison possible, confronting the Chinese signal with the *minimum* daily record of the US embassy. We find that the information thresholds significantly increase the downward bias, confirming the predictions of our model and indicating some manipulation of popular perception around salient discontinuities. The result for the 100-point threshold is also in accordance with the literature on strategic political manipulations of air pollution data: in particular, China applies a 'national environmental protection model city' award, based

https://doi.org/10.1017/S1355770X18000414 Published online by Cambridge University Press

on various environmental measures, including a certain number of Blue Sky days with API below 100 points (Chen *et al.*, 2012). This award could also create political incentives to manipulate specifically the 100 threshold. Up to 2012 Beijing never won the award, so it is not clear how strong this political incentive was there, but we cannot rule it out. However, when considering a categorical dependent variable, all three thresholds are significant, and the negative bias actually gets larger for higher risk levels. We argue that this is suggestive evidence that the distortion is also targeting popular perception, rather than exclusively political objectives. To see if the colour-coded information ranges have a consistent impact, we consider for robustness also the *average* and *maximum* of all US observations during the day, and report the results in the online appendix (tables A5 and A6 for the continuous dependent variable and tables A7 and A8 for the categorical dependent variable).

Moreover, we consider more flexible specifications, including a quadratic pollution term (tables A9–A11) and we try restricting our sample to a narrower bandwidth around the T100 threshold, one of ± 50 points (tables A12–A14) and one of only ± 20 points (tables A15–A17). Even with these specifications, the threshold effects are still present and in most cases significant.

This analysis shows that the Chinese public signal about air quality systematically diverges from the announcements of the US Embassy. In combination with the theoretical model of an optimal government bias, we take this as evidence that the local government is introducing a downward bias, especially around significant information levels, to misguide popular perception about pollution. Part of the story could be driven by political incentives, but these would not explain the significant changes around the higher threshold. The next section is dedicated to the analysis of household responses to pollution information given this distorted air pollution signal.

4. Households

Households living in a polluted city like Beijing can respond to the environmental hazard presented by air pollution, by incurring an ex-ante cost (monetary or in terms of time) to protect themselves, or ex-post, for instance through medical expenditures. This cost was generically captured in the theoretical model as a loss in consumption of θ. In order to test if people incur these costs and adopt self-protective behaviours, we collected individual household data through a survey in urban Beijing. From this data, we elicit the expenditure and time allocation regarding self-protective activities against pollution risks. A set of questions was dedicated specifically to sources of information, allowing us to identify which groups access specific signals during peak pollution days, and which ones are most likely to update their expectations over time. For a sample questionnaire, see online appendix E. For a detailed analysis of the responses to each question, see Ravetti *et al.* (2014).

4.1 Data

The survey was administered in August 2012 in three districts of Beijing (Haidian, Chaoyan and Dongcheng), to a total of 1672 individuals in 578 households. The sample selection was designed to represent the total population: we applied probability proportional to size (PPS) to select families at the district and street level and random selection at the community and household level, so that all households in Beijing had equal chances of selection.[20] The questionnaire inquired in detail about four categories of questions, namely: i) the socio-economic characteristics of the household; ii) habits and self-protective behaviours against pollution hazards: wearing masks over mouth and nose, reducing time outdoor, changing means of transportation, doing preventive health checks, and using air purifiers; iii) health of family members and particularly airborne diseases, cost of illness and insurance; and iv) how the family gathered information about air pollution. The respondents (one per household) could only answer for themselves and for close family members who spent most of the time in the household. Comparing various demographic features of the sample with the Statistics Bureau of Beijing, the survey is in line with the characteristics of the total population, so the sample can be considered representative.

The data from the household survey varies in three dimensions: across individuals, within households and somewhat over time. We introduce a degree of time variation, asking respondents to recall their averting behaviour choices in periods of extreme pollution peaks as compared to the rest of the year. This distinction provides some variation between 'normal' times and extreme pollution events.[21] The use of recall data to introduce this time dimension is not free from limitations, but it gives a sense of people's variation in behaviour vis-a-vis pollution peaks. In the following section we test whether the change in self-protective behaviours during extremely polluted circumstances relates to the source of information used.

4.2 Stylized facts

A simple analysis of the ex-post medical expenditures incurred by households reveals that there should be a strong incentive for families to protect themselves from the damages of air pollution. The average private cost of illness from airborne diseases in our sample is quite high: the mean annual expenditure including medical costs, medicines and foregone wage is more than 3000 yuan, almost a month of average salary, and this is just a lower bound (cost of illness is a conservative measure for how much a person is truly willing to pay to avoid diseases (Alberini and Krupnick, 2000)). At the same time, though, the damage to the workforce of these airborne diseases is not large: on average workers suffered from 12 days per year of hindered activities due to sickness, but lost less than one day a year of paid sick leave. So, according to our survey, the population bears substantial health costs, but without significant effects on labour force availability for production.

The survey captures in detail weekly exposure to outdoor pollution and several self-protective behaviours. *Reducing time outdoor* captures the decision to spend less time

[20]The sampling probability for a given household was

$$p_0 \frac{[N_H]_{D1}}{[N_h]_{\text{TOT}}} * p_1 \frac{[[N_H]_{S1}]_{D1}}{[N_h]_{D1}} * p_2 \frac{1}{[[N_{C1}]_{S1}]_{D1}} * \frac{x}{[[[N_H]_{C1}]_{S1}]_{D1}} = c,$$

where each term captures respectively the probability of a given district, street, community and household being chosen. Overall, the sampling design yielded a constant probability c for a household in any district, street or community to be selected.

[21]To distinguish between extreme and normal times, the respondents needed to recall the two worst episodes of air pollution in Beijing in the previous year, and to locate them in time. In the year before the survey, in fact, there were two major pollution alerts during hazardous pollution days. Only 66 per cent of respondents had noticed the extremely polluted days in Beijing, indicating that even in those cases there was no widespread information about the pollution risks.

https://doi.org/10.1017/S1355770X18000414 Published online by Cambridge University Press

outside for leisure and exercise purposes. *Transport change* implies a change in means of transportation, from those with high exposure to pollution (such as walking or biking) to relatively less-exposed forms of transport, such as using a car. This is not a strategy that many people in Beijing can afford, as less than 6 per cent of our sample adopts it normally. *Masks* are also a relatively infrequent behaviour, adopted by less than 20 per cent of the sample. Finally, the questionnaire includes more expensive, long-term ex-ante strategies: *preventive medical checks* and *buying an air purifier*. These capture medical check-ups of the respiratory system for which the person had to pay some medical costs personally. Air purifiers are like capital investments, and typically are installed at home. These strategies are quite different in nature from the previous behaviours, because they do not respond immediately to pollution peaks, and relate to long-term perception of air pollution, rather than daily signals of an alert, so we keep them separate from the main analysis.

The survey also reports on the various modes of accessing public pollution information in Beijing. Personal monitoring devices were too expensive in 2012 to be relevant for our sample, so we have data on external sources from third parties, or self-perception (usually an approximate measurement from visibility, rheumatisms, smell of the air, etc.). In our sample, internet use for the purpose of collecting information about air pollution is limited. The majority of people interviewed relied on government controlled sources of information, such as television, radio or newspapers. The US Embassy measurements are freely available via Twitter every hour, and they can even be downloaded on a mobile device, however, since the internet in China is restricted, typically only young people declared accessing this sort of information indirectly or using virtual private networks. For a majority of the population in Beijing, the government is the sole mean of accessing information on pollution.

4.3 Averting behaviour

Next, we analyse self-protective behaviours in response to high pollution depending on the source of information. We use a treatment-effect model, where the treatment is the government signal. Do people who rely on government information act differently during pollution peaks? Here we attempt to separate out the group of people that does not update beliefs from more critical individuals, and see if indeed the effect of public information is stronger for the former, as modelled theoretically for the fraction of the population λ.

We apply a two-step procedure, with a bi-probit model – see Greene (2012: 738–752) and Pindyck and Rubinfeld (1998) – to examine how people respond to pollution peaks depending on their information sources.[22] We include in the first stage a dummy variable, Λ, which takes the value of 1 for those people who consider the information they have sufficient to understand the quality of the air, and zero otherwise. This variable captures the fraction of the population that, as in our theoretical model, is not likely to look for further information about pollution to adjust their expectations for any possible

[22] We cannot observe how a person behaves both with and without government information, thus we consider an average treatment effect (ATE), but given the non-random assignment of treatment we need to account for selection bias, as people choose what signal they want to listen to. We only take within-individual changes, namely how the same person responds to different air pollution levels. Since many of the unobserved characteristics of a person remain constant under different pollution situations, this reduces the problem of omitted variables.

information bias.[23] The two-step empirical model is:

$$P_h = \beta_0 + \mathbf{X_i}\beta_{i1} + \mathbf{X_h}\beta_{h2} + \beta_3\Lambda_h + \epsilon_{it} \tag{8}$$

$$A_{ij} = \alpha_0 + \alpha_{h1}P_h + \mathbf{X_i}\alpha_{h2} + \mathbf{X_h}\alpha_3 + \eta_{it}. \tag{9}$$

Averting behaviours, A, vary over individuals, i and over three possible activities, $j \in$ {masks, transport, time outdoor}. The dependent variable is measured in changes, taking the value of 1 if a person switches to more averting behaviour on extremely polluted days, compared to normal days. We leave the results of the air purifiers and the preventive health checks for the online appendix (table A26), with a caveat for the interested reader: those behaviours are not measured in changes, since they do not rapidly respond to pollution peaks. Therefore, omitted variables relative to the profession, degree of awareness or other sources of social protection (health insurance) could be relevant. For instance, public servants in Beijing tend to have access to better health insurance and thus might do more preventive medical controls in all medical realms, even if they receive systematically biased air quality news.

P is the use of public information controlled by the government: it takes the value of 1 when a person uses as principal source of information government-controlled public media (TV, radio, newspapers), zero otherwise. We add some further controls at the individual and household level, $\mathbf{X_i}$ and $\mathbf{X_h}$: age, gender, education level, and a dummy for smokers to capture health-risks aversion and one for workers, to distinguish individuals with different degrees of time flexibility; household income, and a dummy for households with children, which could possibly be more careful about the health damages of pollution; and for the transport specification a control for car ownership, which may be particularly important as a sunk investment in averting choices. We estimate a separate equation for each averting behaviour, rather than combining them in a multinomial logit, since these are not mutually exclusive behaviours.

The first stage with Λ is to identify those agents who not only use government information, but also believe completely in that signal and do not update expectations of a bias. This way, we can directly relate the averting behaviour to the distorted government signal. In our theoretical model, this corresponds to the non-critical fraction of the population λ. In addition, we could argue that the first stage isolates the variation in the choice of public information that does not have to do with pollution peaks, but rather with how people consider information overall. In the theoretical model, we even assumed that the critical and non-critical groups of agents are exogenous. In practice, however, there could be a long-term attitude towards media sources that determines if people search for extra information or are satisfied.

4.4 Household results

Table 3 shows the results from the bi-probit estimation. In the first stage we note that Λ, the dummy capturing non-updating individuals, has a positive and significant correlation with the use of government-controlled media. Those people who consider sufficient the information they have, are also more likely to choose government media, controlling for other factors. Then the second stage shows that this group of people who fully relies

[23]In the survey, these are the respondents who answered that information was enough to the question '*Do you think [your current choice of] information is enough for you or would you like more of it?*'. Not all were users of government media for air quality news.

Table 3. Self-protective behaviours and information (Bi-Probit)

	Outdoor Δ	Mask Δ	Transport Δ
Government media	−1.58*** (0.47)	−2.25*** (0.43)	−0.76 (0.54)
Respondent	0.14** (0.07)	0.02 (0.10)	0.37*** (0.13)
Age	0.00 (0.00)	−0.01 (0.00)	−0.01 (0.01)
Male	−0.07 (0.07)	−0.12 (0.08)	−0.09 (0.11)
Education	−0.07** (0.04)	−0.01 (0.05)	0.01 (0.05)
Smoker	−0.18 (0.12)	−0.47** (0.21)	0.12 (0.15)
Worker	−0.03 (0.14)	0.36* (0.20)	0.29 (0.24)
Children	0.11 (0.16)	0.25 (0.18)	−0.21 (0.24)
Household Income	0.32 (1.04)	−0.00 (0.93)	0.10 (1.39)
Migrant	0.06 (0.18)	0.09 (0.21)	−0.21 (0.28)
Car			0.50*** (0.20)
Constant	0.73 (0.49)	0.78* (0.46)	−0.94 (0.58)
Government media			
Sufficient info	0.37** (0.17)	0.39*** (0.15)	0.38** (0.17)
Respondent	−0.25*** (0.08)	−0.23** (0.09)	−0.28*** (0.08)
Age	0.01** (0.00)	0.01** (0.00)	0.01* (0.00)
Male	0.03 (0.07)	0.04 (0.06)	0.03 (0.07)
Education	−0.04 (0.05)	−0.01 (0.05)	−0.03 (0.05)
Smoker	−0.34** (0.13)	−0.32** (0.13)	−0.28** (0.13)
Worker	0.03 (0.16)	0.06 (0.15)	0.05 (0.16)
Children	−0.18 (0.19)	−0.15 (0.20)	−0.15 (0.20)
Household Income	−0.18 (1.20)	0.06 (1.10)	−0.45 (1.10)

https://doi.org/10.1017/S1355770X18000414 Published online by Cambridge University Press

Table 3. Continued

Migrant	−0.16 (0.27)	−0.15 (0.28)	−0.12 (0.27)
Constant	0.40* (0.23)	0.28 (0.23)	0.40* (0.24)
Athrho Constant	0.89** (0.44)	1.55** (0.65)	0.36 (0.26)
Observations	1103	1103	1093

Notes: Clustered standard errors (household) in brackets. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.
District dummies for Chaoyang and Dongchen omitted.

on media controlled by the Chinese Communist Party is less likely to switch to more averting behaviours during peak pollution days. This result is valid both in terms of time spent outdoors and for wearing masks. In the case of transport switch (third column), the sign of the coefficient is still negative, but insignificant in the Bi-Probit specification: information does not play such a strong role, since car ownership is a strong determinant of this behaviour.

These results are robust to an IV-probit specification that considers the use of government media an endogenous regressor (table A25 in the online appendix). In summary, in our sample of the population of Beijing, the individuals who rely uncritically on the government as a main source of information about air pollution are also less likely to adopt short-term self-protective behaviours during pollution peaks, like staying indoors or wearing a mask. Combining this result with the previous evidence that the government signal might contain a systematic bias, this evidence suggests that the Chinese government is effectively distorting popular perceptions about air pollution risks in Beijing. Overall these results fit well with the theoretical analysis, which suggested an economic mechanism for why and how a government should report optimistically low values for pollution. This evidence opens up many further questions about the joint provision of public information and other public goods, such as pollution abatement, when information can be manipulated.

5. Conclusion

Many economic studies have examined how firms or individuals use information strategically. In this article, we argue that this analysis should extend to governments. We focus on Chinese air pollution, showing that a local government with control over public media may introduce a bias in information signals, possibly with the intention of attracting labour and maximizing its taxable revenue from production. Public information control provides the local government with an instrument to influence popular responses to pollution, without adjusting real variables such as wages or air quality. We illustrate this mechanism empirically, comparing the official public announcements about air pollution with the ones coming from the US Embassy, finding that the public signal is significantly downward biased around critical thresholds, suggesting that the government manipulates information in order to affect the popular perception of pollution. As a consequence, we find in an original household survey that those urban dwellers who rely on government-controlled media adopt fewer short-term measures to protect themselves during pollution peaks. Our analysis does not rule out other alternative explanations for the manipulation of the media in autocratic regimes: in particular, political incentives (such as the well documented Blue Sky Day policy) are also relevant in this context. We should also be cautious in extrapolating from these results about the capital of China to other contexts: our analysis provides an illustration of this type of distortions in environmental information, but is not conclusive evidence that this is the only or most important mechanism everywhere.

A potential implication of this analysis is that, whenever a government can control public information, there is an incentive to distort popular perceptions, and reduced motivation to resolve the problem itself. In the case of pollution, the government could have fewer incentives for abatement or environmental regulation, because it can count on information control as a policy tool. Media control gives the government the power to shift the costs of pollution to the population, so that they translate into individual health damages rather than production losses. Therefore, in some countries the problem with public goods' provision might not be the lack of capacity of the state, but rather the opposite, the excessive control of the government on public information. We leave this issue to be explored in future research. Tackling pollution requires a richer understanding of this system, in which a cycle of distorted information, reduced public responsiveness and possibly reduced provision of public goods is a self-reinforcing reality, similar to a snake, or in this case a dragon, eating its own tail. What could break the cycle? This issue would call for a variety of checks-and-balances that could increase transparency and media freedom. These could come from popular initiatives (e.g., Tang *et al.* (2018) finds that petitions against air pollution have been mildly effective) and from environmental watchdogs, both at the international level (Greenpeace is a prominent example active in China) and from local NGOs or networks, like the All-China Environment Federation.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S1355770X18000414.

Acknowledgements. The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007–2013) under the grant agreement no 26 6992 (GLOBAL IQ). The funding source had no involvement in the study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. We benefited from the comments of participants in the International Economics Association World Congress, in the Swiss Program of Environmental and Energy Economics, in the Italian Society for Climate Sciences and in the Italian Association of Resource Economics and in the Young Swiss Economist Meeting. We thank Richard Baldwin, Nicolas Berman, Anthony Venables, Christopher Meissner, Ugo Panizza, Wessel Vermeulen, Michael Stimmelmayr, Martina Bozzola, Arun Jacob and Rahul Mehrotra for their insightful comments and suggestions. We are particularly grateful to Pedro Naso for his research support. The usual disclaimer applies.

References

Alberini A and Krupnick A (2000) Cost-of-illness and willingness-to-pay estimates of the benefits of improved air quality: Evidence from Taiwan. *Land Economics* **76**, 37–53.

Andrews SQ (2008) Inconsistencies in air quality metrics: 'Blue Sky' days and PM10 concentrations in Beijing. *Environmental Research Letters* **3**, 1–14.

Angeletos GM and Pavan A (2007) Efficient use of information and social value of information. *Econometrica* **75**, 1103–1142.

Besley T and Burgess R (2002) The political economy of government responsiveness: theory and evidence from India. *Quarterly Journal of Economics* **117**, 1415–1451.

Besley T and Prat A (2006) Handcuffs for the grabbing hand? Media capture and government accountability. *American Economic Review* **96**, 720–736.

Brunetti A and Weder B (2003) A free press is bad news for corruption. *Journal of Public Economics* **87**, 1801–1824.

Bucovetsky S. (1991) Asymmetric tax competition. *Journal of Urban Economics* **30**, 167–181.

Chang D, Song Y and Liu B (2009) Visibility trends in six megacities in China 1973–2007. *Atmospheric Research* **94**, 161–167.

Chen Y, Zhe JG, Naresh K and Guang S (2012) Gaming in air pollution data? Lessons from China. *The B.E. Journal of Economic Analysis and Policy* **13**, 1–43.

Chiquiar D and Hanson GH (2005) International migration, self-selection, and the distribution of wages: evidence from Mexico and the United States. *Journal of Political Economy* **113**, 239–281.

Djankov S, McLeish C, Nenova T and Shleifer A (2003) Who owns the media? *Journal of Law and Economics* **46**, 341–81.

Dur R and Swank OH (2005) Producing and manipulating information. *Economic Journal* **115**, 185–199.

Gavazza A and Lizzeri A (2009) Transparency and economic policy. *Review of Economic Studies* **76**, 1023–1048.

Ghanem D and Zhang J (2014) Effortless perfection: do Chinese cities manipulate air pollution data? *Journal of Environmental Economics and Management* **68**, 203–225.

Greene WH (2012) Econometric Analysis, 7th ed. Upper Saddle River, NJ: Prentice Hall.

Halvorsen R and Palmquist R (1980) The interpretation of dummy variables in semilogarithmic equations. *American Economic Review* **70**, 474–475.

Janeba E and Osterloh S (2013) Tax and the city – a theory of local tax competition. *Journal of Public Economics* **106**, 89–100.

Jia R (2014) Pollution for promotion. Mimeo San Diego: University of California.

Kennedy PW, Laplante B and Maxwell J (1994) Pollution policy: the role for publicly provided information. *Journal of Environmental Economics and Management* **26**, 31–43.

Liu Z, Hu B, Wang L, Wu F, Gao W and Wang Y (2015) Seasonal and diurnal variation in particulate matter (PM10 and PM2.5) at an urban site of Beijing: analyses from a 9-year study. *Environmental Science and Pollution Research* **22**, 627–42.

Liu C, Tsow F, Zou Y and Tao N (2016) Particle pollution estimation based on image analysis. *PLoS ONE* **11** (2), e0145955.

McKenzie D, Theoharides C and Yang D (2014) Distortions in the international migrant labor market: evidence from Filipino migration and wage responses to destination country economic shocks. *American Economic Journal: Applied Economics* **6**, 49–75.

Michalski T and Stoltz G (2013) Do countries falsify economic data strategically? Some evidence that they might. *Review of Economics and Statistics* **95**, 591–616.

Morris S and Shin HS (2002) Social value of public information. *American Economic Review* **92**, 1521–1534.

Munshi K and Rosenzweig M (2016) Networks and misallocation: insurance, migration, and the rural-urban wage gap. *American Economic Review* **106**, 46–98.

Naso P and Swanson T (2017) Competition between governance units within a federal structure: impacts on migration and pollution. *CIES Research Paper*, forthcoming.

Pindyck RS and Rubinfeld DL (1998) Econometric Models and Economic Forecasts, 4th ed. New York: McGraw-Hill.

Ravetti C, Swanson T, Jin YP, Quan M and Shiquiu Z (2014) A household survey of the cost of illness due to air pollution in Beijing, China. *CIES Research Paper*, No. 28.

Reporters Without Borders (2016), World press freedom index. Paris: RSF.

Stoerk T (2016) Statistical corruption in Beijing's air quality data has likely ended in 2012. *Atmospheric Environment* **127**, 365–371.

Tang X, Chen W and Wu T (2018) Do authoritarian governments respond to public opinion on the environment? evidence from china. *International Journal of Environmental Research and Public Health* **15**, 266–281.

Tiebout C (1956) A pure theory of local expenditures. *Journal of Political Economy* **64**, 416–424.

WHO (2005) *Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide. Global update 2005. Summary of risk assessment*, World Health Organization.

Williams A (2009) On the release of information by governments: causes and consequences. *Journal of Development Economics* **89**, 124–138.

Wilson J (1995) Mobile labor, multiple tax instruments, and tax competition. *Journal of Urban Economics* **38**, 333–356.

Zheng S and Kahn ME (2013) Understanding China's urban pollution dynamics. *Journal of Economic Literature* **51**, 731–72.

Cite this article: Ravetti C, Swanson T, Jin Y, Mu Q, Zhang S (2019). A dragon eating its own tail: public control of air pollution information in China. *Environment and Development Economics* **24**, 1–22. https://doi.org/10.1017/S1355770X18000414