Gaming in Air Pollution Data? Lessons from China

Yuyu Chen
Peking University

Ginger Zhe Jin
University of Maryland & NBER

Naresh Kumar
University of Miami

Guang Shi
Peking University

Abstract

Protecting the environment during economic growth is a challenge facing every country. This paper focuses on two regulatory measures that China has adopted to incentivize air quality improvement: one is publishing a daily air pollution index (API) for major cities since 2000; the other is linking the API to performance evaluation of local governments. In particular, China defines a day with the API at or below 100 as a blue sky day. Starting 2003, a city with at least 80% blue sky days in a calendar year was qualified (among other criteria) for the award of "national environmental protection model city." This cutoff was increased to 85% in 2007.

Using officially reported API data from 37 large cities during 2000-2009, we find a significant discontinuity at the threshold of 100 and this discontinuity is more pronounced after 2003. Moreover, model cities were found to be less likely to report the API right above 100 when they were close to the targeted blue sky days in the fourth quarter of the year before they won the model city award. That being said, we also find significant correlation of the API with two alternative measures of air pollution – namely visibility reported by the China Meteorological Administration and aerosol optical depth (corrected for meteorological conditions) from NASA satellites. The discontinuity around 100 suggests that some improvement on the count of blue sky days could have been subject to data manipulation, but the API does contain useful information about air pollution.

Keywords: Gaming, Incentive, Air Pollution, China.
JEL code: D8, H7, I18, L3, L5.

Acknowledgement: Yuyu Chen and Guang Shi are affiliated with the Guanghua School of Management, Peking University. Ginger Zhe Jin (contact author) is affiliated with NBER, and Department of Economics at the University of Maryland. Naresh Kumar is affiliated with Department of Epidemiology and Public Health, Miller School of Medicine, University of Miami. Email: chenyuyu@pku.edu.cn, jin@econ.umd.edu, nkumar@med.miami.edu, shiguang1@gsm.pku.edu.cn. We thank Douglas Almond, Hanming Fang and participants at the 2011 NBER China Group meeting for constructive comments on a previous draft that contains part of the materials used in this study. Financial support from Peking University is gratefully acknowledged. All errors are ours.

1. Introduction

Protecting the environment during economic growth is a challenge facing every country. Because private sectors may not fully internalize the environmental consequence of their activity, environmental protectionists have called for government regulation. Existing research has examined the effect of environmental regulations on firm behavior[1], environmental measures[2], and health outcomes[3], but fewer studies look at the policies that motivate *local governments* to reduce pollution.[4] Nevertheless, a great number of international treaties were established, assuming that each targeted government, facing the incentives specified in an international treaty, can effectively reduce pollution in or near its territory. Unfortunately, any policy that motivates local governments to reduce pollution could also motivate them to report better environmental outcomes on paper, especially if it is more costly to make actual improvement, gaming is difficult to detect, and disaggregated objectively measured data are unavailable.

China provides a unique opportunity to study local government incentives in environmental protection. While China has enjoyed a steady GDP growth for the past 30

---

[1] Among others, Henderson (1996), Becker and Henderson (2000), Greenstone (2002) and List et al. (2003) examine the effect of environmental regulations on firm entry, exit, and size change in the US. See the review of Millimet, Roy and Senqupta (2009) for more details. A number of other studies focus on the effect of environmental regulations on trade flow, capital flow and international pollution havens, for example, Dean, Lovely and Wang (2009), Ederington, Levinson and Minier (2005), Keller and Levinson (2002), Wheeler (2004) and Zeng and Zhao (2009).

[2] For example, Greenstone (2004) studies the impact the US Clean Air Act on Sulfur Dioxide, Davis (2008) studies the effect of driving restrictions on air quality in Mexico City, and Chen et al. (2011) studies the effect of environmental measures adopted in the name of the 2008 Olympic Games on Beijing's air quality.

[3] For example, Chay and Greenstone (2003), Currie and Neidell (2005), Currie, Neidell and Schmieder (2009) study the impact of air pollution on infant health and mortality, and Chay, Dobkin and Greenstone (2003) examine the effect of the 1970 Clean Air Act on adult mortality.

[4] At the country-level, Congleton (1992) and Murdoch and Sandler (1997) show that the democratic countries are more likely to support and enforce chlorofluorocarbon emissions control under the Montreal Protocol. See Oats and Portney (2007) for a review on the political economy of environmental policy.

years, 16 of the world's top 20 most polluted cities are located in China as of 2007.[5] Given the regional decentralized authoritarian (RDA) regime in China, Xu (2011) argues that local government officials have an incentive to sacrifice environmental protection in order to boost local GDP growth. This is because the local government leaders are appointed by the central government based on the local performance and GDP growth is much easier to measure than environmental protection. To be fair, the central government is aware of the problem and has incorporated some environmental measures in performance evaluation. Lessons learned from these environmental incentives could be valuable for other domestic and international policies that target local governments for environmental protection.

This paper focuses on two regulatory measures that China has adopted to incentivize local governments to improve air quality: one is publishing daily air pollution index (API) for 86 cities since 2000; the other is linking the API to performance evaluation of local governments. In particular, China defines a day with the API at or below 100 as a blue sky day. Starting 2003, a city with at least 80% blue sky days in a calendar year was qualified (among other criteria) for the award of "national environmental protection model city." The cutoff for model city increased further to 85% in 2007. While these incentive policies were adopted to reduce air pollution, they also created incentives to game the API data, as the API data are reported by local governments and misreporting is less costly than actual improvement of air quality.

Using officially reported API data from 37 large cities during 2000-2009, we show that there is a significant discontinuity at the threshold of 100, despite the fact that

[5] http://www.cbsnews.com/stories/2007/06/06/eveningnews/main2895653.shtml.

API is calculated as a city-day average of multiple pollutants across multiple monitoring stations. This discontinuity is more pronounced after 2003, and model cities are less likely to report right above 100 when they face more pressure to reach the cutoff by the end of the year before they can win the award. That being said, we also find significant correlation of the API with two alternative measures of air pollution – namely visibility reported by the China Meteorological Administration (CMA) and aerosol optical depth (AOD) derived from NASA satellites. These findings suggest that some improvement on the count of "blue sky days" could be subject to data manipulation, but the API does contain useful information about air pollution.

We also show that, controlling for weather, national trend, and city-specific factors, there is no statistically significant improvement in API, visibility or AOD immediately before or after a city won the model city award. This implies that model city is not awarded to acknowledge significant air quality improvement within a city, which in turn casts doubt about the role that the model city policy has played in providing incentives for a city to improve air quality.

The rest of the paper is organized as follows: Section 2 describes data and policies on blue sky days and model city evaluation. Section 3 reviews the related literature. Section 4 presents evidence on the discontinuity of the API, in comparison with other objective measures of air quality: satellite based AOD and visibility. Section 5 examines the extent to which the pressure to achieve the model-city goal of blue sky days has affected the reported API. Section 6 examines the correlation between API, visibility and AOD more systematically in light of the API discontinuity. Section 7 tests whether the three measures of air quality have improved immediately before or after a city wins the model city award. Section 8 concludes.

2 Data and Background

China has been known for poor air quality since the 1990s. The 1996 national standards on Sulfur Dioxide (SO_2), Nitrogen Dioxide (NO_2), total suspended particles (TSP), and particulate matter with an aerodynamic diameter of 10 microns or smaller (PM_{10}) were 2-7 times higher than the standards established by the World Health Organization (UNEP 2009). An amendment in 2000 further weakened the Chinese standards for NO_2 and Ozone. Even so, the relatively generous standards are hard to enforce in China, partly because each local environment protection agency, although a branch of the Ministry of Environmental Protection (MEP), is also part of the local government and subject to local governance.

API Among other environmental protection efforts, the MEP has started to publish the daily air pollution index (API) for 86 cities since June 2000. These cities cover most median- and large-size cities of China, including all the provincial-level municipalities and all provincial capitals. For each city, the MEP aggregates the measured intensities of NO_2, SO_2 and TSP into a daily API ranging from 0 to 500, where TSP includes any particulate matter with an aerodynamic diameter of 100 microns or smaller.[6] Specifically, suppose a city has M stations and each station monitors NO_2, SO_2 and TSP

[6] MEP monitors the intensity of CO, but does not include it in the current API calculation because the calculation formula was set ten years ago and at that time the vehicle volume in China was very low. MEP is considering adding CO and other pollutants for future API. Source: http://news.163.com/09/0312/11/5470SBA9000120GU.html

for N times each day,[7] the MEP first computes the daily average of all the MxN measures for each pollutant and then translates the daily mean intensity into pollutant-specific API according to linear spines with the cutoff points defined in Table 1.[8] The overall API is the maximum of all the pollutant-specific APIs. If that maximum is above 500, the overall API is capped at 500. An API of 0-50 is defined as “excellent” air quality, 51-100 as “good”, 101-200 as “slightly polluted”, 201-300 as “moderately polluted” and 301-500 as “heavily polluted”. The MEP also reports the category of the dominant pollutant(s) if the API is above 50. Over our analysis sample (from 06/06/2000 to 10/31/2009 for 37 cities), 72.9% of API observations report TSP as the main pollutant, 0.35% report NO_2, and 6.85% for SO_2. The remaining 19.9% has the API below 50.

Although the API data are disclosed on the MEP website, they are collected and reported by local MEP branches. At the frequency of city-day, it is virtually impossible for the central MEP to verify every number reported by a local branch. To the extent that local MEP officials are subject to local governance, the reliability of API data may depend on the data collection method (defined by the MEP), as well as local political incentive to report a good number to the central government.

Blue Sky Day A cruder categorization of air pollution refers to a day with the API at or below 100 as a “blue sky” day. Both national and local environmental authorities have used the number of blue sky days in a year as a public measure of air pollution. For example, the city-level environmental bureau of Beijing claims steady air quality

[7] The MEP stipulates the number of monitoring stations according to city population and the size of the established area. For a large city like Beijing, one monitoring station is required for every 25-30 km^2 and the total number of stations must be at least 8.

[8] For example, if the daily mean of TSP is 370 $\mu g/m^3$, the corresponding API of TSP is (370-300)/(500-300)*(200-100)+100 = 135.

improvement because the number of blue sky days has increased from 274 in 2008[9], to 285 in 2009[10] and 286 days in 2010.[11] However, in our analysis sample, the average API of all blue sky days has increased from 65.62 in 2008 to 71.11 in 2009 (up to 10/31/2009). This implies that continuous API and binary count of blue sky days can paint different pictures of air quality. Nevertheless, the number of blue sky days is more visible in mass media and its over-time improvement is often cited by local governments as a political goal at the beginning of a calendar year and an achievement at the end of the year.[12] The phrase of "blue sky day" has also been challenged by a local resident of Beijing who snapped a picture of sky each day and found that the number of days with real blue sky is 180 instead of 285 in 2009.[13] While naked eyes and statistics differ in the definition of blue sky, it reflects intensive public attention on blue sky days.

Model city policy In as early as 1997, the central government of China started to evaluate whether a city was qualified for a "national environmental protection model city" award based on environmental quality and economic measures. While the API is always the only measure for air quality, the model-city criteria for the API was vague until 2003. In particular, the 2003 regulation specified that a model city must have over 80% of days in a calendar year with the API below 100. The 80% cutoff was increased further to 85% in 2007.[14]

---

[9] See news report at http://news.qq.com/a/20081231/001928.html (reported on 12/31/ 2008, accessed on 02/10/2012).

[10] See news report at http://news.qq.com/a/20091231/001311.html (reported on 12/31/2009, accessed on 02/10/2012).

[11] See news report at http://news.xinhuanet.com/fortune/2011-01/01/c_13672919.htm (reported on 1/1/2011, accessed on 02/10/2012).

[12] See news report at http://news.163.com/11/1107/01/7IyI1FB100014AED.html (reported on 07/11/2011, accessed on 02/10/2012).

[13] See news report at http://news.xinhuanet.com/society/2010-10/27/c_12704483.htm (reported on 10/27/2010, accessed on 02/10/2012).

[14] The 1997 regulation was a pilot program which specified the criterion on air quality as "API<100." The

It is worth noting that participation in model city evaluation is voluntary. We do not observe who has applied for the award in a given year, and we assume that every city that satisfies the explicit criteria will apply. Upon application, the central government's evaluation committee will visit the city in person and typically announce the winner(s) sometime in the year. Among the 37 big cities for which we have complete API and visibility data from 6/5/2000 to 10/31/2009, nine of them have won the model city award in our sample period. They are Qingdao (2000), Hangzhou (2001), Changchun (2002), Nanjing (2003), Fuzhou (2004), Shenyang (2004), Nantong (2006), Tianjin (2006) and Guangzhou (2007). Another six cities won the award before the start of our sample: Shenzhen (1997), Dalian (1997), Xiamen (1997), Haikou (1999), Shantou (1999) and Suzhou (1999). Because the award is announced in the middle of the year, we assume the evaluation is conducted based on the data of the previous year.[15]

The model city award is semi-permanent. According to the MEP regulation, a city that has won the model city award is subject to reexamination every three years; if it fails in the reexamination, it has two years to correct the problem; if underperformance remains after the correction period, the model city award will be revoked. In reality, some reexaminations were conducted more than three years after the initial model city award,

2003 regulation clarified the air quality criterion as the percent of days of API<100 bigger than 80%. It was issued on 11/19/2002 and effective on 7/1/2003 (http://www.mep.gov.cn/gkml/zj/bgt/200910/t20091022_173806.htm). The 2007 regulation (effective on 1/1/2007) stipulated that the percent of days of API<=100 is bigger than 85%. Although the 1997 and 2003 regulations specified API<100 instead of API<=100, we believe the actual implementation was always API<=100 because both definitions of blue sky days and the good category of API includes 100 as the upper bound. This assumption is also confirmed in the below discontinuity study and a local MEP branch website in Xiaomen.

[15] In one particular application, we observe the applicant citing environmental and economic measures in the past two years.

and some cities were even exempted from reexamination.[16] To our best knowledge no model city award has ever been revoked. This suggests that it is more difficult to earn a model city award than it is to keep it. It may also introduce the incentive to slack after winning the model city award, a hypothesis that we will examine in Section 7.

Visibility In addition to the API, we resort to two additional proxies of air pollution. The first proxy is visibility, defined as the greatest distance at which an observer with normal eyesight can discern a dark object from the horizontal sky. We obtain daily visibility data from the China Meteorological Administration (CMA), along with local temperature, precipitation, barometric pressure, sunshine, humidity and wind velocity as reported at 2pm each day at a fixed point in each city. Researchers have shown that the API and visibility are negatively correlated (Che et al. 2006, Fan and Li 2008) and visibility is considered to be an important predictor of fine particulates (Ozkaynak et al. 1985, Huang et al. 2009). [17] Like the API, visibility is reported by government officials, but it is not disclosed to the public (we purchase the data from CMA), not used in the evaluation of government officials, and therefore subject to fewer gaming incentives.

AOD The second proxy for air pollution is the daily 10km AOD data (Level 2, collection 5.0) retrieved from Moderate Resolution Imaging Spectroradiometer (MODIS) aboard Terra and Aqua satellites (NASA 2010). The extraction procedure is available

---

[16] For example, Yangzhou earned the award in 2002 but was reexamined in 2006 (http://www.mep.gov.cn/gkml/zj/bgth/200910/t20091022_174312.htm); Changchun earned the award in 2002, but was reexamined in 2008 (http://www.mep.gov.cn/gkml/zj/bgth/200910/t20091022_174443.htm). According to http://wfs.mep.gov.cn/mfcs/mfcsxx/gldt/200503/t20050327_65797.htm, some model cities can be exempted from reexamination.

[17] Fine particulates ($PM_{2.5}$), are defined as particulates with less than 2.5μm in aerodynamic diameter.

elsewhere (Chu, Kaufman et al. 2003; Levy, Remer et al. 2007a, 2007b). By definition, the AOD captures the amount of radiation absorbed, reflected and scattered due to the presence of solid and liquid particulates suspended in the atmospheric column (Kaufman, Gobron et al. 2002; Kaufman, Tanre et al. 2002). While AOD is potentially available everywhere at the satellite crossing time (~10:30am and ~1:30pm of Beijing time), it is sensitive to the point and time specific weather and only available for days with minimal cloud cover, such as less than 10% cloud cover, otherwise the quality of AOD retrieval can be influenced by cloud contamination. Despite this, researchers have shown that the AOD, corrected for meteorological conditions, can predict air quality (Gupta, Christopher et al. 2006; Kumar, Chu et al. 2011). Focusing on Delhi and Kanpur in India and Cleveland in US, Kumar et al. (2009; 2011) demonstrate how the AOD can be converted to PM_{10} estimates. They develop an empirical relationship between *in situ* measurements of PM_{10} and the AOD. They conclude that the AOD captured 70% of the variations in the PM_{10} (monitored on the surface) after controlling for meteorological conditions and seasonality. Since the *in situ* PM_{10} data were not available in China during our sample period, this paper utilizes AOD directly (corrected for meteorological conditions and spatiotemporal structure).

The comparison of API, visibility and AOD is far from perfect. Even if the reported dominant pollutant of API is TSP, it covers all particulate matter with diameter up to 100 microns, where visibility and AOD are more related to finer particulate matter with diameters smaller than 10 and 2.5 microns. According to Brook, Dann and Burnett (1997), Canada data suggest that $PM_{2.5}$ accounted for 49% of the PM_{10}, and PM_{10} accounted for 44% of the TSP. The composition of suspended particles is likely coarser in developing countries: for example, Kumar and Foster (2009) and Kumar, Chu and Foster (2008) show that $PM_{2.5}$ accounts for only 24% of PM_{10} in Delhi, India. Moreover, all particulate matter in the atmosphere could affect the AOD, whereas visibility and API are more related to particulate matter on the ground. The third difference is due to the mismatch in the spatial resolution of AOD and API. Although we know the centroid latitude and longitude of each AOD observation (that represents ~10km at nadir i.e. the satellite crossing path), we do not know the exact location of each API or visibility monitoring station.

Analysis Sample Conditional on having non-break API and visibility data, our analysis consists of 37 cities, which includes major provincial-level municipalities such as Beijing, Tianjin, Shanghai and Chongqing, as well as 24 provincial capitals.[18]

The API, visibility and other meteorological data from CMA are reported by city-day, covering 126,688 observations from 6/5/2000 to 10/31/2009.

For the 37 cities in our sample, we retrieved 2,614,734 valid 10km AOD observations from 2/25/2000 to 12/31/2009. Of all the 3,598 calendar days in the sample period, only 49.9% had valid AOD observations due to gaps in the data[19]. On average, we had 39.36 data points of AOD per city-day.

To control for time-specific meteorological conditions at the observation time of AOD, we acquired hourly global surface meteorological data from the monitoring stations in and around the selected cities. The details on these data are available

---

[18] Although the MEP reports API for 86 cities and the CMA visibility data cover 69 cities, only 42 cities has API data in 2000 and the visibility data are incomplete for some cities between 1993 and 2009. For an unknown reason, the API data are missing on June 4, 2008 for all cities. So the “non-break” criterion ignores the missing data on June 4, 2008.

[19] For 37 cities from February 25, 2000 to December 31, 2009, there are totally 133126 city/day cells. In our data, there are actually 66427 city/day cells with valid AOD observations, which amounts to 49.9%.

elsewhere (NCDC 2007). These data were collocated with the AOD data within one hour time interval of AOD time on a given day. This means we assigned the same value of meteorological conditions (from the closest station) to all AOD values in a given city on a same day. Since there were subtle gaps in the meteorological and AOD data, it resulted in missing values in 6% of the sample. Therefore, meteorological conditions were imputed for missing days when AOD was available. The procedure impute was employed to estimate missing values with the aid of continuous time and other city specific meteorological conditions in STATA (StataCorp 2010).

To facilitate the comparison between API, visibility and AOD data, our AOD analysis focuses on the city-day average of AOD conditional on AOD availability. This leads to 63,948 city-day observations of AOD, of which 50,672 city-days report TSP as the dominant pollutant.

3 Literature and potential ways to game the API

We are not the first one to question the reliability of API data. Andrews (2008) expressed the concern that Beijing may have manipulated the official API report because (1) Beijing seems to have relocated monitoring stations over time; (2) the 2000 MEP standard for air quality has weakened the limits of nitrogen oxides and ozone; and (3) the number of days with the API between 96 and 100 is significantly higher than the number of days with the API between 101 and 105. However, Guinot (2008) suggests that it is not uncommon to add monitoring stations with economic and urban growth and the uncertainty in the API metrics may range from 15% to 25% due to measurement errors in pollutant intensities. This casual debate motivates us to examine the discontinuity of API in a more scientific way.

A few other studies have used alternative measures of air pollution in addition to the official API. In February 2009, United Nations published a summary report on the 2008 Beijing Olympic Games, with a complete chapter devoted to air quality (UNEP 2009). This report uses the API and pollution intensity data from the Beijing Environmental Protection Bureau (EPB), plus a brief discussion of coarse resolution AOD data (i.e.100km instead of 10km AOD used in our study). They conclude that Beijing's air quality has improved from before to during and shortly after the 2008 Games. Chen et al. (2011) investigate the impact of the 2008 Games on air quality using the same API and AOD data as in this paper. After controlling for city-specific attributes and a nationwide trend towards better air, they find that the air-cleaning actions adopted in the name of the 2008 Olympic Games led to real but temporary improvement of air quality in Beijing. This result is supported by both API and AOD data, suggesting that the API contains useful information about air pollution. However, this conclusion does not rule out gaming of the API given the imperfect comparison between API and AOD data.

Wang et al. (2009) collected their own PM_{10} and $PM_{2.5}$ data in Peking University between 7/28/2008 and 10/7/2008. They found a significant correlation between the self-measured and published PM_{10}, but the absolute level of their self-measure is 30% higher. This finding triggered some concerns that the official API may be subject to manipulation, but the discrepancy may be attributed to sampling and methodological differences (Tang et al., 2009, Yao et al. 2009, Simorich 2009). Wang et al. (2009) also find that meteorological conditions such as wind, precipitation and humidity account for

40% of the total variation in PM_{10}. This is why we need daily meteorological data for every city in our sample.

Any systematic study of gaming needs to ask why and how. The incentive to improve the reported API is rooted in the unique structure of the Chinese political system. As described in Xu (2011), China is characterized by a combination of political centralization and economic regional decentralization: the central government controls the appointment, promotion, and demotion of local political leaders, while leaving subnational governments (provinces, municipalities and counties) the responsibility for initiating and coordinating reforms, providing public services, and making and enforcing laws within their jurisdictions. The central control of personnel is a powerful instrument to induce regional officials to follow the central government's policies. This so-called regionally decentralized authoritarian (RDA) regime is a clear distinction from federalism (where governors or mayors are elected from the bottom) and central-planning.

Researchers have shown that the central government stipulates performance criteria for local leaders, and local leaders negotiate narrower and more precisely defined performance contracts with its sub-levels. For example, Tsui and Wang (2004) show that 60 percent of provincial leaders were assigned to targets related to economic construction. More generally, work achievement accounts for 60 to 70 percent of the evaluation of regional officials, while political integrity, competence, diligence and other aspects account for the rest (Edin 2003). Similar personnel control is documented between county governments and township and village officials (Whiting 2000). Within this structure, every level of government may use absolute and/or performance in the political contract for the next level. Maskin, Qian and Xu (2000) provide evidence that officials from relatively better-performing regions have a higher chance of being promoted. Similarly, Chen, Li and Zhou (2005) find that provincial officials' performances relative to the national average and to their immediate predecessors had significant impacts on their promotions. All these evidence suggests that the central personnel control on local governments is effective and the model city award policy is likely one of the many performance criteria that the central government uses to evaluate local officials.

For gaming of the API to exist, two conditions must hold. First, there is enough noise in the true API that one cannot precisely target a particular number (say the upper bound of blue sky days) via actual improvement. This is easy to satisfy given the fact that meteorological conditions such as wind, precipitation and humidity account for 40% of the total variation in PM_{10} (Wang et al. 2009).

The second necessary condition for gaming is that it is difficult to detect gaming. To the extent that the MEP uses the reported API without verification, a local MEP branch could report any number in theory. However, the reported API will be disclosed to the public, and common people (including local media) will have their own judgment on how precise the reported API is relative to their personal experience in that city-day. The recent smog of Beijing demonstrates high public awareness of air quality and the power of public outcry if the official API is not consistent with personal experience.[20] This suggests that any misreport cannot be too far away from the truth.

One can underreport an API slightly above 100 to be slightly below 100. More sophisticated gaming may spread the underreporting if the public cannot tell the

[20] The heavy smog of Beijing on December 4-6, 2011 has caused massive flight delays and cancellation. See BBC December 5, 2011 report "China morning round-up: Beijing heavy smog scrutinised" accessed at December 7, 2011 at http://www.bbc.co.uk/news/world-asia-china-16044098.

difference between small changes in the API (say 99 and 95). Data manipulation can also be achieved by relocating monitoring stations or computing the aggregated index from a selective sample of existing stations, both of which are difficult to detect since the reported API is not station specific and *in situ* density of air pollutants is not publicly available in our sample period.[21]

As summarized in Zitzewitz (forthcoming), "forensic economics" relies on several techniques to detect gaming: one is to compare the reported data with other data sources. For example, Fisman and Wei (2004, 2007) and Mishra, Subramanian, and Topalova (2007) compare custom data in both origin and destination countries in order to identify missing imports or missing exports. Zinman and Zitzewitz (2009) compare resort-reported snowfall with official weather data and find that resorts are more likely to over-report snowfalls when they can benefit more from such over-reports. Snyder and Zidar (2009) compare economists' self-reported publications (in vitas) with journals' tables of contents and identify subtle forms of inflation. This method of collation is limited to the availability of other reliable data sources.

A second method to detect gaming is searching for data patterns that are consistent with gaming, for example bunching around a threshold (Slemrod 1985 and Saez 2010 on income tax, Degeorge, Patel and Zeckhauser 1999 on earnings management, and Forbes, Lederman and Tombe 2011 on airline delays), patterns that should not exist without cheating (Jacob and Levitt 2003 on school test scores), or a correlation between the reported data and the situations that present strong incentives to

[21] Beijing Municipal Environmental Monitoring Center started to publish pollutant densities by hour and monitoring stations since January 12, 2012. There are not enough historical data to study its reliability.

game (Michaely and Womack 1999 on stock recommendation, Levitt and Syverson 2007 on real estate sales).

We use both methods. First, we focus on discontinuity of the raw data. Given the clear threshold definition for blue sky days, we expect the density of the API to be lower immediately right of 100 than immediately left of 100 if gaming exists. This implies discontinuity around 100. Moreover, this discontinuity should only exist after the central government introduces a quantitative measure of blue sky days in the evaluation of model city. With the specified cutoff (80% from 2003 to 2006 and 85% after 2007), we expect cities that are close enough to the cutoff in the fourth quarter of a year have incentive to underreport the API at or below 100. Given the permanency of the model city award, we also expect cities that have won the model city award before be less eager to game for the award.

Our second approach is comparing the reported API with visibility and AOD data. Given the imperfect comparison of the three air pollution proxies, we will investigate their statistical correlations after controlling for other factors that may correlate with them in different ways (e.g. weather). To the extent that the underreporting of the API generates a random number below but close to 100, we expect the correlation between the API and visibility/AOD to be lower when the API is close to 100. However, if the underreporting is monotonic to the actual API – for example, report 105 as 100, 104 as 99, and 101 as 96 – gaming does not necessarily predict a lower correlation between the API and visibility/AOD.

4 Tests of Discontinuity

This section examines discontinuity of API, visibility and AOD data. To the extent that one cannot misreport the API far away from the truth, we expect misreporting to result in less density on the right of 100 than on the left.

Figure 1A plots the histogram of the API in our whole sample, where bins are defined over the complete range of the API (0 to 500) with the finest bin width of 1 (the API is reported in integers). The plot shows likely discontinuity at 50, 100 and 500: the number of observations jumps from 1,408 for 50 to 2,034 for 51; from 1,367 for 99 and 1,005 for 100 to 509 for 101, and from 2 for 499 to 126 for 500.

Our first test of discontinuity is the Burgstahler and Dichev test (BDT) for each bin (Burgstahler and Dichev 1997). In particular, for any bin (*j)* that is not at the boundaries, we can compute a BDT statistics by comparing the bin's observed relative frequency ($\hat{p}_j$) with the average of frequencies of adjacent bins ($\hat{p}_{j-1}$and $\hat{p}_{j+1}$):

$$BDT_j = \frac{\frac{\hat{p}_{j-1} + \hat{p}_{j+1}}{2} - \hat{p}_j}{\sqrt{var(\frac{\hat{p}_{j-1} + \hat{p}_{j+1}}{2} - \hat{p}_j)}}$$

where n is the total number of observations and

$$var\left(\frac{\hat{p}_{j-1}+\hat{p}_{j+1}}{2} - \hat{p}_j\right) = \frac{1}{n}\hat{p}_j\left(1-\hat{p}_j\right) + \frac{1}{4n}\left(\hat{p}_{j-1} + \hat{p}_{j+1}\right)\left(1 - \hat{p}_{j-1} - \hat{p}_{j+1}\right) + \frac{1}{n}\hat{p}_j\left(\hat{p}_{j-1} + \hat{p}_{j+1}\right).$$

According to Burgstahler and Dichev (1997) and Takeuchi (2004), BDT_j conforms to a standard normal distribution if the true distribution underlying the data is continuous. Obviously, the power of the BDT depends on sample size and bin width. Using Monte Carlo simulation, Takeuchi (2004) shows that the test is powerful over moderate sample size (n>500) and is able to detect discontinuity for a small jump in a continuous distribution if the sample size is large (n>5000). Our sample size is 126,688 for the API and visibility, and 63,948 for the city-day average of the AOD.

One potential caveat of the BDT is that its value is proportional to the square root of sample size. When sample size is huge (e.g. 2.6 million observations for our point-specific AOD data) and the value of each raw data point is limited to a small number of decimal points, each computed BDT can exceed the critical value even if the underlying distribution is continuous. This is partly why we choose to focus on city-day average of the AOD rather than point-specific AOD. While this choice may create some smoothness on the AOD, it is arguably a better comparison with the API, not only because they are both at the city-day level but also because the API by construction is an average of station- and time-specific data within a city-day.

Figure 1B draws the BDT against each API value. The two dashed lines correspond to 2.58 and -2.58, the critical values for the 99% confidence in a standard normal distribution. As in Figure 1A, the BDT confirms significant discontinuity in the neighborhood of 50, 100 and 500. The large BDT at 499 is not surprising given the truncation of the API at 500. The BDTs at 49 (4.39), 50 (5.10)), and 51 (-5.81) are consistent with the frequency jump up from 50 to 51 in the raw data. We do not have good explanations for this: one possibility is that the observations below and around 50 are relatively small thus introduce measurement errors. Another possibility is that 50 is the threshold for the MEP's categories of "excellent" (grade I) and "good" (grade II) air quality but cities do not have incentives to improve from "good" to "excellent" given the policy focus is on 100 instead of 50. In comparison, the BDTs at 99 (-3.94), 100 (-1.75) and 101 (9.3), are consistent with the sharp drop of frequency from 99 and 100 to 101.

Figures 1C and 1D repeat the BDT graph for before and after 2003 separately. In 2000-2003, we observe a marginal discontinuity around 50 and an obvious discontinuity towards the maximum (500) due to censoring. However, after 2003, the BDTs around 100 become highly significant. One potential explanation is that in 2003 the model city evaluation started to require at least 80% of days with blue sky and therefore motivated cities to report the API at or below 100.

The BDT statistics use only the data frequency of the bins adjacent to the study bin. A more general test introduced by McCarry (2008) employs all the data to the left and right of a potential break point and smooth the histogram by running local linear regressions on the two sides separately. If there is no discontinuity at the break point, the predicted density at the break point should be close to each other when the prediction is conducted from the left and right separately. This yields a discontinuity estimate (log difference in the two predicted densities at the break point) and the corresponding standard error and t statistics. Figure 2 presents the API histogram with smoothed densities to the left and right of 100, where left stops at 100 and right starts at 101. Following McCarry (2008), we set the bin size as 1 and bandwidth as 15. Below Figure 2, we report the discontinuity estimate, standard error and t-test for the full sample of the API and the subsamples of 2000-2002, 2003-2006, 2007-2009, model cities, and non-model cities. By model cities, we mean all the $\geq y-1$ observations of a city if that city has won the model city award in year y. Within model cities, we further distinguish model cities in the years preparing for and getting the award ($y-1$ and y) versus after winning the award ($> y$).

Not only does the McCarry test confirm the API discontinuity at 100, it shows that the discontinuity estimate has more than doubled from 2000-2002 to 2003-2006 and only declines slightly after 2007. Similarly, the discontinuity estimate for model cities is almost tripled that of non-model cities, although within model cities the estimate is similar in the years before and after winning the award. These patterns support the argument that local governments have more incentives to underreport the API once the central government emphasizes the number of blue sky days in model city evaluation.

Is it possible that the API discontinuity around 100 is driven by local governments targeting 100 in real air quality? If the answer is yes, to the extent that visibility and AOD are correlated with the API, we shall observe some discontinuity for these two variables as well, especially in the days where the API is not far from 100. Conversely, if the API discontinuity at 100 is driven by gaming, visibility and API should not demonstrate any discontinuity because they are used in neither mass media nor model city evaluation.

Figures 3-6 repeat the Burgstahler and Dichev test of discontinuity for visibility and AOD. Contrary to our expectation, visibility demonstrates significant discontinuity at integers, especially the multiples of five. This is easy to explain because visibility is based on human eyes and manual report. Figure 4 separates the BDT of visibility by the range of API (40-80, 80-120, 120-160). If the API discontinuity at 100 reflects real air quality and visibility is a valid proxy of air quality, we may observe more discontinuity of visibility when the API is between 80 and 120, as compared to other ranges of the API. This conjecture is not supported by Figure 4. As for the city-day average of AOD, Figure 5 shows no obvious discontinuity at any particular value and Figure 6 confirms this finding by the ranges of the API.[22] These patterns are consistent with the lack of gaming incentives in the AOD data.

[22] In unreported graphs, we plot the BDT against a fine grid of point-specific AOD. Probably due to the huge sample size (2.6 million) and the limited decimal points in the AOD data, we have the BDT exceeding the critical value everywhere and across all ranges of the API. Even if we take the BDT

Above all, the observed API distribution reveals significant discontinuity at 100 and such discontinuity is consistent with potential gaming in response to the definition of blue sky days and the model city policies.

5. Regression results on model city incentives

This section uses regressions to detect how cities respond to the targeted number of blue sky days in model city evaluation. Because the pressure for city c to manipulate the API of day d in year y depends on the target and the previously achieved number of blue sky days, we define

$$Pressure_{cyd} = \frac{target\ \#\ of\ blue\ sky\ days - realized\ blue\ sky\ days}{\#\ of\ days\ in\ the\ calendar\ year - day\ of\ year}.$$

For example, if a city has realized 16 blue-sky days in January 2003, then its pressure on February 1 is (365*80%-16)/(365-32)=0.829. If February 1 is not a blue-sky day, then its pressure on February 2 is (365*80%-16)/(365-33)=0.831.

Several adjustments are necessary. First, *Pressure* is coded 0 before 2003, because the blue-sky day target was introduced in November 2002 and became effective in 2003. Second, if the targeted number of blue sky days is achieved before the end of year, then *Pressure* is coded 0 after the achievement. Third, *Pressure* is coded 0 if the above formula yields a value larger than 1 as it is impossible to achieve the target then. We control for such cases by a separate dummy variable. Fourth, the above formula is not well defined for the last day of a calendar year. We code *Pressure* 0 for the last day of year, because it should be zero if the target has been realized before the last day. If the

---

values literally, it implies discontinuity everywhere, which is inconsistent with the special discontinuity of API around 100.

target has not been met, there is no way to meet it unless the city is only one day short of the target.[23]

To capture the potential change of incentives over time, we define:

$Model_{cy} = 1$ if city c is announced as a model city in year y;

$Modellag_{cy} = 1$ if city c is announced as a model city in year y-1; and

$Modelahead_{cy} = 1$ if city c is announced as a model city in year $y + 1$.

Given the discontinuity of the API around 100, we would like to understand whether cities systematically underreport an above-100 API number to be below 100. To the extent that underreporting must be discreet to avoid public attention, it predicts a lower probability to report right above 100 (defined as between 101 and 105, inclusive) and a higher probability to report right below 100 (defined as between 96 and 100, inclusive).

At the first glance, it seems straightforward to regress the dummy of the reported API falls into [96,100] or [101,105] on $Pressure, Model, Modellag, Modelahead$ and their interactions for the full sample. However, this regression is likely to generate bias because $Pressure$ is defined by the API in previous days of the same year and there could be serial correlations within a city-year. To address this, we focus on the last quarter of each year and take $Pressure$ as of September 30 as a predetermined variable. By the end of the third quarter, if the city is on track to reach the targeted number of blue sky days for the model city award, it should have achieved 75% of the target on September 30. So we further define:

$Close_{cy} = 1$ if city c's # of blue sky days $\geq 75\%$ of target by 9/30 of year y.

[23] This happens only once in our data for Yinchuan in 2003 and its last day API read is 101 in 2003. This city did not win the model city award until 2011.

Then we run the regressions:

$$1_{96\leq API\leq 100} = \alpha_{1c} + \alpha_{1d} + \theta_{1c} \cdot d + \gamma_1 \cdot X_{cyd} + f_1(Pressure_{cyd}, Close_{cy}, Model_{cy}, Modellag_{cy}, Modelahead_{cy}) + \varepsilon_{1cyd}$$

$$1_{101\leq API\leq 105} = \alpha_{2c} + \alpha_{2d} + \theta_{2c} \cdot d + \gamma_2 \cdot X_{cyd} + f_2(Pressure_{cyd}, Close_{cy}, Model_{cy}, Modellag_{cy}, Modelahead_{cy}) + \varepsilon_{2cyd}$$

where α_{1c} and α_{2c} are city fixed effects, α_{1d} and α_{2d} are date fixed effects, θ_{1c} and θ_{2c} are city-specific time trends, and X are control variables including city-day weather and socioeconomic indicators such as GDP growth rate, GDP per capita, industrial production, population, energy consumption, number of private vehicles and a dummy for regular heating season if heating is provided by the city. Functions $f_1(.)$ and $f_2(.)$ include the linear terms of $Pressure_{cyd}$, $Close_{cy}, Model_{cy}, Modellag_{cy}$, $Modelahead_{cy}$, and their interactions . Our main interest is the three-way interaction of $Pressure, Close$ and the status of model city. If a city wins the model city award by gaming the API, gaming should be more apparent in the year immediately before the award, when it is close to the target, and when it is subject to a greater pressure to reach the target. Errors are clustered by city.

The linear probability[24] results reported in Table 2 confirm this expectation. In particular, we present results for three dependent variables: a dummy of the API at or below 100 for a city-day, a dummy of the API in [96,100], and a dummy of the API in [101,105]. For each dependent variable, we report the regression with and without city-

[24] We use linear probability model because every regression includes a large number of date fixed effects.

specific trend and socioeconomic indicators, while always controlling for weather, date fixed effects and city fixed effects.

As shown in Table 2, the three way interactions of $Pressure, Close$ and the status of model city suggest that, condition on being close to the yearly target as of September 30, the higher the pressure to reach the target, the more likely for the API to be at or below 100 (especially in [96,100]) and the less likely for the API to be in [101,105] in the years of and right before receiving the award (as compared to at least 2 years before receiving the award). It is natural to argue that this is because the awarded cities reached the target by actual air quality improvement from right above 100 to right below 100. However, if this were the case, we should see similar rather than opposite patterns after the city has won the award. In Section 7, we will further show that this is not because model cities have significantly improved or reduced air quality after winning the award.

For robustness, we repeat the exercise of Table 2 using October 31 instead of September 30 as the cutoff date.[25] This makes the regression sample smaller because we focus on the last two months of the year. The three way interactions of $Pressure \cdot Close \cdot Modelahead$ are similar to that of Table 2: significantly positive for the API in [96,100] and significantly negative for the API in [101,105]. In comparison, $Pressure \cdot Close \cdot Model$ is only significant (and negative) for the API in [101,105] while $Pressure \cdot Close \cdot Modellag$ is not significant in any specification. These findings confirm the argument that model cities in the year right before the award are more likely to have the API in [96,100] and less likely in [101,105].

[25] Close is redefined as >=83% because days before October 31 account for 83% of a year.

As a second robustness check, we use the 2000-2002 data to conduct a placebo test, assuming that the 80% blue sky days cutoff were effective since 1/1/2001. As shown in Table 3, this hypothetical policy does not create the gaming patterns as that of Table 2.

6. Comparison of API, visibility and AOD

This section examines whether visibility and AOD are less correlated with API when the API are reported to be close to 100. In particular, we define

$Left = 1$ if API is in [96,100],

$Right = 1$ if API is in [101,105],

and run the regressions:

$$\ln(Visibility_{cyd}) = \alpha_{1c} + \alpha_{1d} + \theta_{1c} \cdot d + \gamma_1 \cdot X_{cyd} + \beta_{11} \cdot \ln(API_{cyd}) + \beta_{12} \cdot Left_{cyd}$$
$$+\beta_{13} \cdot Right_{cyd} + \beta_{14} \cdot \ln(API_{cyd}) \cdot Left_{cyd} + \beta_{15} \cdot \ln(API_{cyd}) \cdot Right_{cyd} + \epsilon_{1cyd};$$
$$\ln(\overline{AOD}_{cyd}) = \alpha_{2c} + \alpha_{2d} + \theta_{2c} \cdot d + \gamma_2 \cdot X_{cyd} + \beta_{21} \cdot \ln(API_{cyd}) + \beta_{22} \cdot Left_{cyd}$$
$$+\beta_{23} \cdot Right_{cyd} + \beta_{24} \cdot \ln(API_{cyd}) \cdot Left_{cyd} + \beta_{25} \cdot \ln(API_{cyd}) \cdot Right_{cyd} + \epsilon_{2cyd}.$$

As before, control variables X_{cyd} include weather and socioeconomic indicators. Errors are clustered by city. We present the regression results of visibility in Table 4 and AOD in Table 5. Both Tables include two samples: the full sample and the subsample in which TSP is reported as the dominant pollutant. As we expect, the API is negatively correlated with visibility and positively correlated with AOD. In both tables, the coefficients of $\ln(API) \cdot Left$ and $\ln(API) \cdot Right$ are never significantly different from zero. Results are qualitatively similar if we use absolute API, visibility and AOD instead of log.

These results are inconsistent with the prediction that the API is less correlated with visibility and AOD data when it is reported close to 100. However, they do not rule out all types of gaming. It is possible that the extent of underreporting the API below 100 is not random (e.g. shifting down the API by a constant). Another possibility is that the underreporting of the API could be random but go much further below 95.

7. Evaluating the mode city policies

Given the permanency of model city award and the evidence of API gaming around 100, a remaining question is whether model city policies are effective in motivating local governments in real air quality improvement. Specifically, we have two predictions: first, if the model city award is granted to acknowledge air quality improvement of a city, model cities should have more air quality improvement right before winning the award than at least two years before the award. Second, the lax reexamination policy implies that model cities could reduce efforts in air quality improvement after winning the award.

Before testing these two predictions, Table 6 summarizes the API by city type and city status in terms of model city award. In particular, we distinguish three city types: cities that did not win any model city award by 2010 (total 22 cities), cities that won the model city award in our sample period of 2000-2009 (total 9 cities), and cities that won the award at or before 1999 (total 6 cities). For cities that won the award during our sample, we further distinguish their observations in at least 2 years before the award, 0-1 years before the award, and all years after the award. As we expect, cities that won the award earlier tend to have lower average API and more counts of blue sky days. Similarly, within the cities that won the award during our sample, API and blue sky days improve over time. However, except for 2001 and 2002, both API and count of blue sky days improve constantly nationwide. This is why we need the below regressions to better examine whether a model city has better API or blue sky days during and after winning the award, in addition to city fixed effects and national changes over time.

Table 6 also shows that the probability of API falling in [96,100] increases during the award winning years but decreases afterwards. Given the fact that the overall density of API declines in the range of 96 to 105, the non-monotonic change is inconsistent with actual air quality improvement over time but consistent with more incentives to report the API in [96,100] right before winning the award. In comparison, the probability of API falling in [101,105] declines both during and after the award winning years, which could be consistent with gaming or actual improvement.

To test whether the model city award is granted to acknowledge air quality improvement of a city or motivates model cities to reduce air quality protection after the award, we regress each air quality measure on the timing of model city award while controlling for city fixed effects, date fixed effects, weather, and socioeconomic indicators as stated above. Effectively, the default control cities are those that do not win any model city award before the end of our sample period. Comparing with these control cities, a city that won the award in our sample period should have better air quality right before winning the award and worse air quality after winning the award if the model city award is effective. More specifically, we define

$Prepare_{cyd} = 1$ if city c won the model city award in year y or y+1,

$After_{cyd} = 1$ if city c won the model city award in or before year y-1, and run regressions at the city-day level:

$$\ln(API_{cyd}) = \alpha_{1c} + \alpha_{1d} + \theta_{1c} \cdot d + \gamma_{11} \cdot X_{cyd} + \beta_{11} \cdot prepare_{cyd} + \beta_{12} \cdot After_{cyd} + \epsilon_{1cyd}$$

$$1(API_{cyd} \leq 100) = \alpha_{2c} + \alpha_{2d} + \theta_{2c} \cdot d + \gamma_{21} \cdot X_{cyd} + \beta_{21} \cdot prepare_{cyd} + \beta_{22} \cdot After_{cyd} + \epsilon_{1cyd}$$

$$\ln(Visibility_{cyd}) = \alpha_{3c} + \alpha_{3d} + \theta_{3c} \cdot d + \gamma_{31} \cdot X_{cyd} + \beta_{31} \cdot prepare_{cyd} + \beta_{32} \cdot After_{cyd} + \epsilon_{2cyd}$$

$$\ln\left(\overline{AOD}_{cyd}\right) = \alpha_{4c} + \alpha_{4d} + \theta_{4c} \cdot d + \gamma_{41} \cdot X_{cyd} + \beta_{41} \cdot prepare_{cyd} + \beta_{42} \cdot After_{cyd} + \epsilon_{3cyd}.$$

Results presented in Table 7 show that model cities do not significantly improve air quality in the 0-1 years right before winning the model city award, nor do they reduce air quality after winning the award. This finding is consistent across all three measures of air quality as well as the dummy of whether a day is blue sky or not. In light of the significant (and temporary) air quality improvement found in Beijing around the Olympic Games (Chen et al. 2011), we have rerun the regressions without Beijing, without other Olympic related cities, and without data of 2008 and 2009. Results are similar.

What could explain these results with the evidence of API gaming around 100? We can think of several explanations: first, the model city award does not generate any significant improvement of API, visibility and AOD throughout the year, although it generates incentives to game the API around 100 towards the end of the year when the city faces higher pressure to reach the targeted number of blue sky days; second, the gaming of the API around 100 is dominated by the overall air quality changes that a city incurs before, during, and after winning the model city award.

Evidence so far casts doubt on the effectiveness of the model city award in air quality improvement. If the central government anticipates this ex ante, why would they stipulate and implement the model city policy? We speculate a few reasons: first, air quality is only one of many criteria in the model city evaluation; second, the award is not designed to encourage environmental protection efforts, rather it is given according to the cross-sectional difference across cities. In our data, if we compute the average API of year 2001 for every city and correlate its rank with the order of receiving the model city award, the spearman rank correlation is 0.42 and significant with 99% confidence. There is also a third possibility that the model city policy is so effective that it encourages every city to improve air quality all the time. We believe this explanation is unlikely given the lax reexamination policy after a city wins the model city award.

8. Conclusion

Overall, this paper focuses on two regulatory measures that China has adopted to incentivize air quality improvement: one is publishing daily API for major cities since 2000; the other is linking the extent of blue sky days (defined by $API \leq 100$) to the evaluation of model city award. Using daily API, visibility and AOD data from 37 large cities during 2000-2009, we show that officially reported API has a significant discontinuity at the threshold of 100. This discontinuity is more pronounced after 2003, and model cities are less likely to report right above 100 when they face more pressure to reach the cutoff by the end of the year before they won the award. These patterns suggest data gaming around the threshold of blue sky days.

That being said, we also find significant correlation of the API with two alternative measures of air pollution – namely visibility reported by the China Meteorological Administration (CMA) and aerosol optical depth (AOD) derived from NASA satellites. These findings suggest that some improvement on the count of "blue sky days" could be subject to data manipulation, but the API does contain useful information about air pollution.

Interestingly, after we control for weather, national trend and city-specific factors, there is no statistically significant improvement of API, visibility and AOD immediately before or after a city won the model city award. This implies that model city is not awarded to acknowledge significant air quality improvement within a city, which in turn casts doubt about the role that the model city policy has played in providing incentives for a city to improve air quality.

9. References

Andrews, Steven (2008) "Beijing Plays Air Quality Games" *Far Eastern Economic Review* (uploaded on July 4, 2008). Accessed at http://www.feer.com/essays/2008/august/playing-air-quality-games on January 10, 2010.

Becker, Randy and Vernon Henderson (2000) "Effects of Air Quality Regulations on Polluting Industries" *The Journal of Political Economy,* 108 (2): 379-421.

Brook, Dann and Burnett (1997) "The relationship among TSP, PM_{10}, $PM_{2.5}$, and inorganic constituents of atmospheric particulate matter at multiple Canadian locations" Journal of the Air & Waste Management Association ISSN 1096-2247 1997, vol. 47, n^{o}1, pp. 2-19 (44 ref.)

Burgstahler, D. and I. Dichev (1997) "Earnings management to avoid earnings decreases and losses" Journal of Accounting and Economics, 24, 99-126.

Chay, Kenneth and Michael Greenstone (2003) "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession" *Quarterly Journal of Economics*, 118(3).

Chay, Kenneth; Carlos Dobkin and Michael Greenstone (2003) "The Clean Air Act of 1970 and Adult Mortality" Journal of Risk and Uncertainty, 27(3).

Che, Huizheng; Xiaoye Zhang; Yang Li; Zijiang Zhou; and Zhenline Chen (2006) "Relationship Between Horizontal Extinction Coefficient and PM10 Concentration in Xi'An, China During 1980-2002" *China Particuology* 4(6): 327-329.

Chen, Ye; Hongbin Li and Li-An Zhou (2005) "Relative Performance Evaluation and the turnover of Provincial Leaders in China" *Economics Letters*, 88(3): 421-25.

Chen, Yuyu; Ginger Zhe Jin; Naresh Kumar and Guang Shi (2011) "The Promise of Beijing: Evaluating the Impact of the 2008 Olympic Games on Air Quality", NBER working paper #16907.

Congleton, R. (1992) "Political Institutions and Pollution Control" *Review of Economics and Statistics* 74: 412–21.

Chu, D. A.; Y. J. Kaufman; G. Zibordi; J. D. Chern; J. Mao; C. Li; and B. N. Holben (2003) "Global Monitoring of Air Pollution over Land from EOS-Terra MODIS." Journal of Geophysical Research, 2003. 108(D21): p. 4661.

Currie, Janet, and Matthew Neidell (2005) "Air Pollution and Infant Health: What Can We Learn from California's Recent Experience?" *Quarterly Journal of Economics* 120(3): 1003–30.

Currie, Janet; Matthew Neidell and Johannes F. Schmieder (2009) "Air pollution and infant health: Lessons from New Jersey" *Journal of Health Economics*, 28(3): 688-703.

Davis, Lucas W. (2008) "The Effect of Driving Restrictions on Air Quality in Mexico City" *Journal of Political Economy* 116(1): 37-80.

Dean, J.M.; Lovely, M.E. and Wang, H. (2009) "Are foreign investors attracted to weak environmental regulations? Evaluating the evidence from China" *Journal of Development Economics*, 90(1): 1-13.

Ederington, J.; Levinson, A. and Minier, J. (2005) "Footloose and Pollution-Free*" The Review of Economics and Statistics* 87(1), 92-99.

Edin, Maria (2003) "State Capacity and Local Agent Control in China: CCP Cadre Management from a Township Perspective" *China Quarterly*, 173: 35-52.

Fan, Yin-Qi and Chun Li (2008) "Visibility Trends in Beijing, Tianjin and Hebei Province During 1980-2003" *Plateau Meteorology* 27(6): 1392-1400.

Fisman, Raymond, and Shang-Jin Wei. 2004. "Tax Rates and Tax Evasion." *Journal of Political Economy*, 112: 471–496.

Fisman, Raymond and Shang-Jin Wei. 2007. "The Smuggling of Art and the Art of Smuggling: Uncovering the Illicit Trade in Cultural Property and Antiques," NBER Working Paper No. W13446.

Greenstone, Michael (2002) "The Impacts of Environmental Regulations on Industrial Activity: Evidence from the 1970 and 1977 Clean Air Act Amendments and the Census of Manufactures" *Journal of Political Economy*, 110(6): 1175-1219.

Greenstone, Michael (2004) "Did the Clean Air Act Cause the Remarkable Decline in Sulfur Dioxide Concentrations?" *Journal of Environmental Economics and Management*, May 2004.

Guinot, Benjamin (2008) "Olympic Games: Beijing Air Quality" accessed at http://www.ambafrance-cn.org/奥林匹克运动会：北京大气质量,4570.html on January 10, 2010.

Gupta, P., S. A. Christopher, et al. (2006). "Satellite remote sensing of particulate matter and air quality assessment over global cities." *Atmospheric Environment* **40**(30): 5880-5892.

Henderson, J. Vernon (1996) "Effects of Air Quality Regulation" *The American Economic Review*, 86 (4): 789-813.

Huang, Wei; Jianguo Tan; Haidong Kan; Ni Zhao; Weimin Song; Guixiang Song; Guohai Chen; Lili Jiang; Cheng Jiang; Renjie Chen; and Bingheng Chen (2009) "Visibility, air quality and daily mortality in Shanghai, China" *Science of the Total Environment*, 407 (2009): 3295–3300.

Jacob, Brian and Steven Levitt. 2003a. "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating," *Quarterly Journal of Economics*, 118(3): 843-877.

Kaufman, Y. J., N. Gobron, et al. (2002). "Relationship between surface reflectance in the visible and mid-IR used in MODIS aerosol algorithm - theory." *Geophysical Research Letters (0094-8276)* **29**(23): 31-31 to 31-34.

Kaufman, Y. J., D. Tanre, et al. (2002). "Aerosol radiative impact on spectral solar flux at the surface, derived from principal-plane sky measurements." *Journal of the Atmospheric Sciences* **59**(3): 635-646.

Keller, W. and Levinson, A. (2002) "Pollution Abatement Costs and Foreign Direct Investment Inflows to U.S. States" The Review of Economics and Statistics, 84(4): 691-703.

Kumar, Naresh and Andrew Foster (2009), "Air quality interventions and spatial dynamics of air pollution in Delhi and its surroundings," *International Journal of Environment and Waste Management*, 4(1/2): 85-111.

Kumar, N., A. Chu, and A. Foster, "Remote sensing of ambient particles in Delhi and its environs: estimation and validation." *International Journal of Remote Sensing*, 2008. 29(12): p. 3383-3405.

Kumar, N., A. Chu, and A. Foster, "An Empirical Relationship between PM2.5 and Aerosol Optical Depth in Delhi Metropolitan*." Atmospheric Environment*, 2007, 41(21): p. 4492-4503.

Kumar, N. (2010). "A Hybrid Approach for Predicting $PM_{2.5}$ Exposure." *Environmental Health Perspective* **118**(10).

Kumar, N., A. Foster, S.N. Tripathi, A. Dasguptai, and A. D. Chu "Predicting Air Quality using MODIS data" (2009), web.ccs.miami.edu/~nkumar/EST_Final_Version.pdf.

Kumar, N., A. D. Chu, et al. (2011). "Satellite Remote Sensing for Developing Time and Space Resolved Estimates of Ambient Particulate in Cleveland, OH." *Aerosol Science and Technology* 45(9): 1090-1108.Levitt, Steven D. and Chad Syverson. 2008. "Market Distortions When Agents Are Better Informed: The Value of Information in Real Estate Transactions," *Review of Economics and Statistics* 90(4): 599-611.

Levy, R. C., L. A. Remer, et al. (2007a). "Global aerosol optical properties and application to Moderate Resolution Imaging Spectroradiometer aerosol retrieval over land." *Journal of Geophysical Research-Atmospheres* **112**(D13): -.

Levy, R. C., L. A. Remer, et al. (2007b). "Second-generation operational algorithm: Retrieval of aerosol properties over land from inversion of Moderate Resolution Imaging Spectroradiometer spectral reflectance." *Journal of Geophysical Research-Atmospheres* **112**(D13).

List, John A.; Daniel L. Millimet; Per G. Fredriksson and W. Warren McHone (2003): "Effects of Environmental Regulations on Manufacturing Plant Births: Evidence from a Propensity Score Matching Estimator" *The Review of Economics and Statistics*, 85(4): 944-952.

Maskin, Eric and Chenggang Xu (2001) "Soft Budget Constraint Theories: From Centralization to Market" *Economics of Transition*, 9(1): 1-27.

McCarry, Justin (2008) "Manipulation of the Running variable in the Regression Discontinuity Design: A Density Test" *Journal of Econometrics*, 142(2), February.

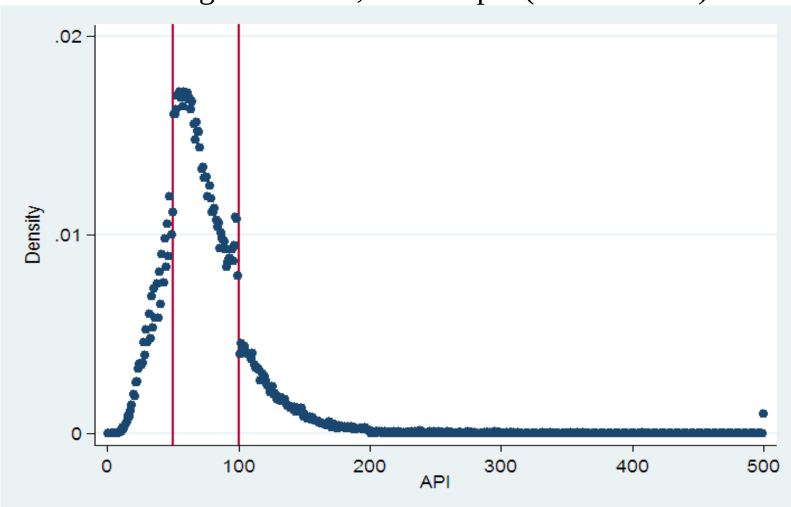
Michaely, Roni, and Kent Womack. 1999. "Conflict of Interest and the Credibility of Underwriter Analyst Recommendations," *Review of Financial Studies*, 12 (4): 653–686.

Mishra, Prachi, Petia Topalova, and Arvind Subramanian. 2007. "Policies, Enforcement, and Customs Evasion: Evidence from India," IMF Working Paper No 07/60.
Millimet, Daniel L.; John A. List and Thanasis Stengos (2003): "The Environmental Kuznets Curve: Real Progress or Misspecified Models?" The Review of Economics and Statistics, 85 (4): 1038-1047.
Murdoch, J. and T. Sandler (1997) "The Voluntary Provision of a Pure Public Good: The Case of Reduced CFC Emissions and the Montreal Protocol" *Journal of Public Economics* 63: 331-50.
NASA. (2010). "the Level 1 and Atmosphere Archive and Distribution System." Retrieved July 3, 2010, from http://ladsweb.nascom.nasa.gov/.
NCDC. (2007). "National Climatic Data Center." from http://www.ncdc.noaa.gov/oa/ncdc.html.
Oats, Wallace and Paul Portney (2003) "The Political Economy of Environmental Policy" in K.G. Maler and J. Vincent edited Handbook of Environmental Economics (Amsterdam: Elsevier) pp. 325-354.
Ozkaynak H, et al. (1985) "Relationship between aerosol extinction coefficients derived from airport visual range observations and alternative measures of airborne particle mass" JAPCA. 35 (1985):1176–85.
Saez, Emmanuel. 2010. "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy* 2: 180-212.
Simonich, Stacil L. Massey (2009) "Response to Comments on 'Atmospheric Particulate Matter Pollution during the 2008 Beijing Olympics'" *Environmental Science & Technology* (43): 7590-7591.
Slemrod, Joel. 1985. "An Empirical Test for Tax Evasion," *Review of Economics and Statistics* 67(2): 232-238.
Snyder, Christopher and Owen Zidar. 2009. "Resume Padding Among Economists?" Unpublished.
StataCorp (2010). STATA/SE Version 10.1. College Station, Texas 77845, StataCorp LP.
Takeuchi, Y (2004) "On a statistical method to detect discontinuity in the distribution function of reported earnings" *Mathematics and Computers in Simulation*, 64(1): 103-111.
Tang, Xiaoyan; Min Shao; Min Hu; Zifa Wang; and Junfeng (Jim) Zhang (2009) "Comment on 'Atmospheric Particulate Matter Pollution during the 2008 Beijing Olympics'" *Environmental Science & Technology* (43): 7588.
Tsui, Kai-yuen and Youqiang Wang (2004) "Between Separate Stoves and a Single Menu: Fiscal Decentralization in China" *China Quarterly*, 177: 771-90.
United Nations Environment Programme (2009) "Independent Environmental Assessment Beijing 2008 Olympic Games." Accessed at http://www.unep.org/publications/UNEP-eBooks/BeijingReport_ebook.pdf.
Wang, Wentao; Toby Primbs; Shu Tao; and Stacil L. Massey Simonich (2009) "Atmospheric Particulate Matter Pollution during the 2008 Beijing Olympics" *Environmental Science & Technology* (43): 5314-5320.
Wheeler, D. (2004) "Racing to the Bottom? Foreign Investment and Air Pollution in Developing Countries" *The Journal of Environment and Development* 10(3): 225-245.

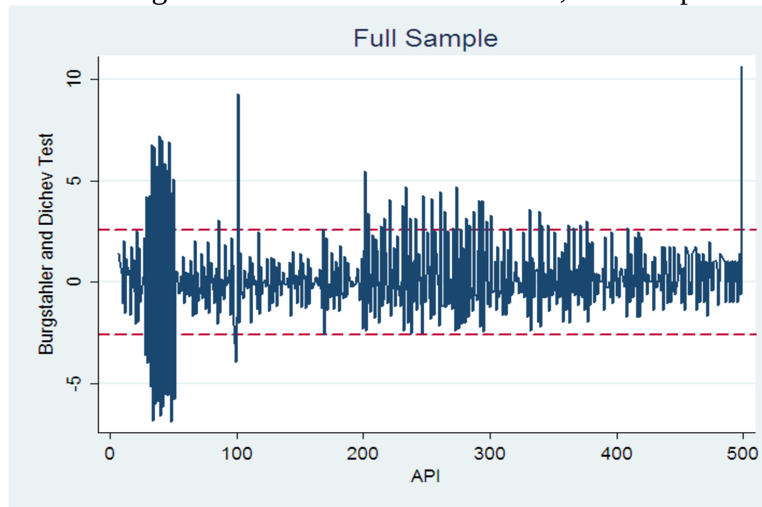
Whiting, Susan H. (2000) Power and Wealth in Rural China: The Political Economy of Institutional Change. Cambridge University Press.
Xu, Chenggang (2011) “The Fundamental Institutions of China’s Reforms and Development” *Journal of Economic Literature*, 49(4): 1076-1151.
Yao, Xiaohong; Xiaohong Xu; Kelly Sabaliauskas and Ming Fang (2009) “Comment on ‘Atmospheric Particulate Matter Pollution during the 2008 Beijing Olympics’” *Environmental Science & Technology* (43): 7589.
Zeng, D. Z. and Zhao, L. (2009) “Pollution havens and industrial agglomeration” *Journal of Environmental Economics and Management* 58(2): 141-153.
Zinman, Jonathan and Eric Zitzewitz. 2009. "Wintertime for Deceptive Advertising?" http://www.dartmouth.edu/~ericz/wintertime.pdf.
Zitzewitz, Eric (2011) “Forensic Economics”*Journal of Economic Literature forthcoming.*

Figure 1: Histogram and Burgstahler and Dichev Test (BDT) of discontinuity for the API

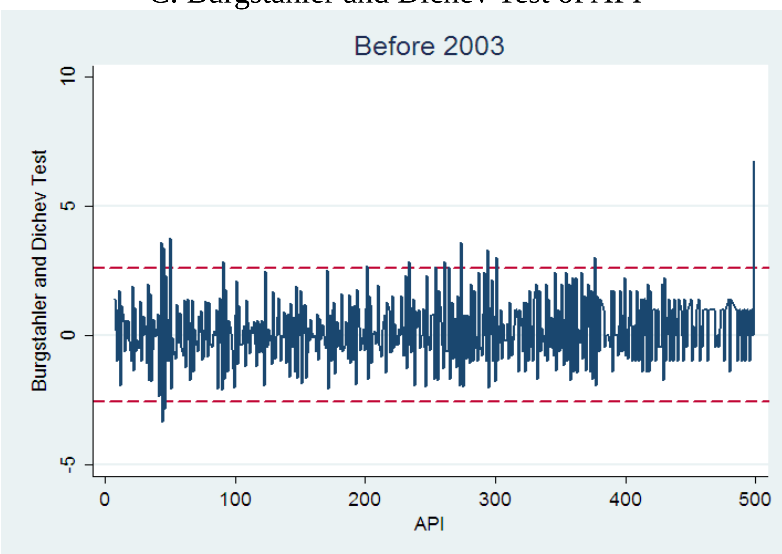
A: Histogram of API, full sample (bin width =1)



B: Burgstahler and Dichev Test of API, full sample



C: Burgstahler and Dichev Test of API



D: Burgstahler and Dichev Test of API

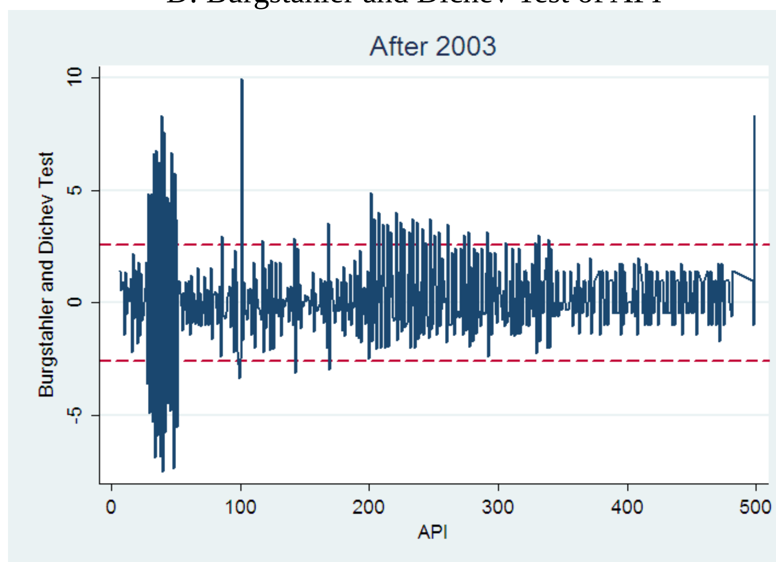
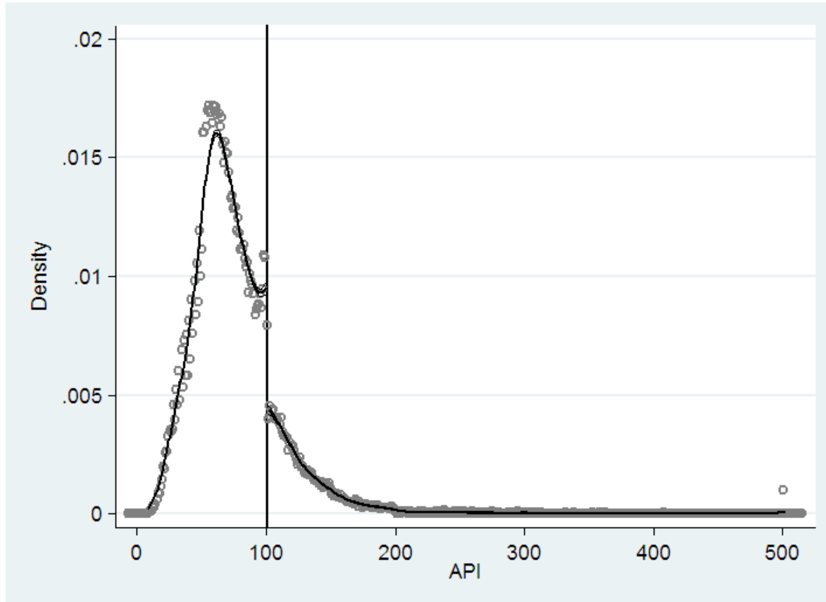


Figure 2: McCarry test of the API discontinuity around 100 (bin size =1, bandwidth=15)



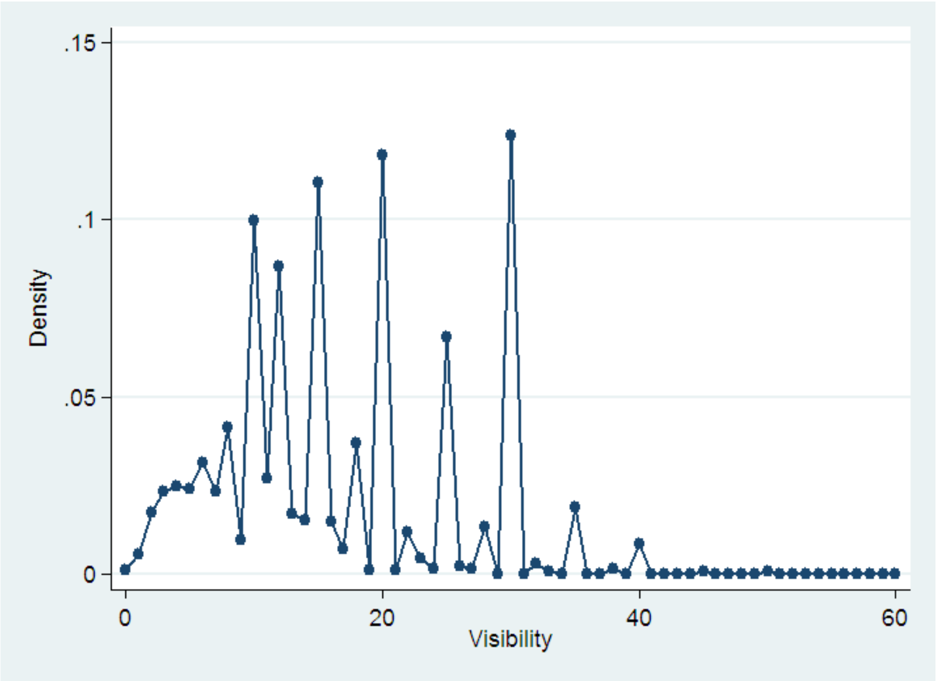
	Full sample	2000-2002	2003-2006	2007-2009	Model cities	Model cities 0-1 years before award	Model cities after award	Non-model cities
Discontinuity estimate	-0.776	-0.409	-0.958	-0.889	-1.568	-1.497	-1.589	-0.536
Standard error	0.029	0.052	0.044	0.060	0.068	0.139	0.078	0.033
T-statistics	-26.744***	-7.901***	-21.787***	-14.941**	-23.005***	-10.752***	-20.335***	-16.358***

*** p<0.01. If a city won the model city award at year y, the sample of "model cities" includes its daily observations in years at or later than y-1, the sample of "model cities 0-1 years before award" includes its daily observations in years y and y-1, and the sample of "model cities after award" includes its daily observations in years from y+1 and on. The sample of "non model cities" includes every observation that is not in the sample of "model cities."

Figure 3: Histogram and Burgstahler and Dichev Test of Discontinuity for Visibility

Visibility ranges from 0 to 60 kilometers. Bin width = 1 kilometer.

Histogram of Visibility



Burgstahler and Dichev Test of Discontinuity

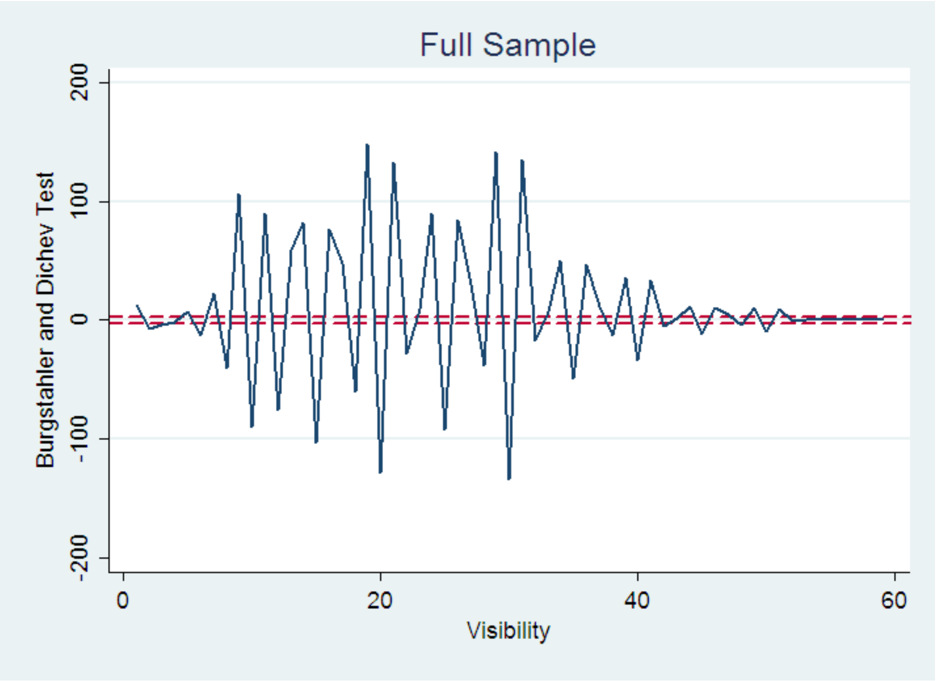
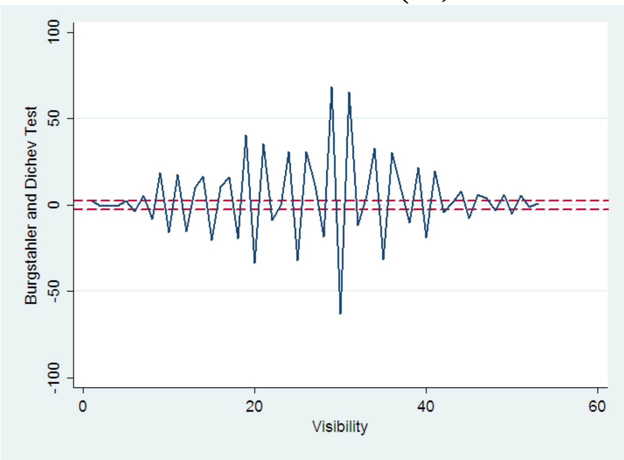
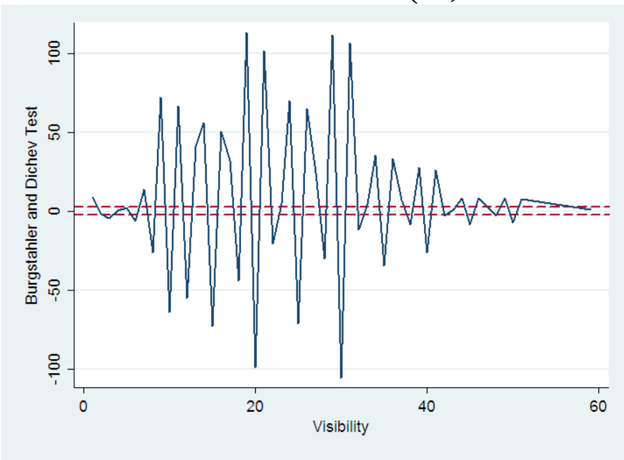


Figure 4: Burgstahler and Dichev Test of Discontinuity for Visibility by Ranges of API

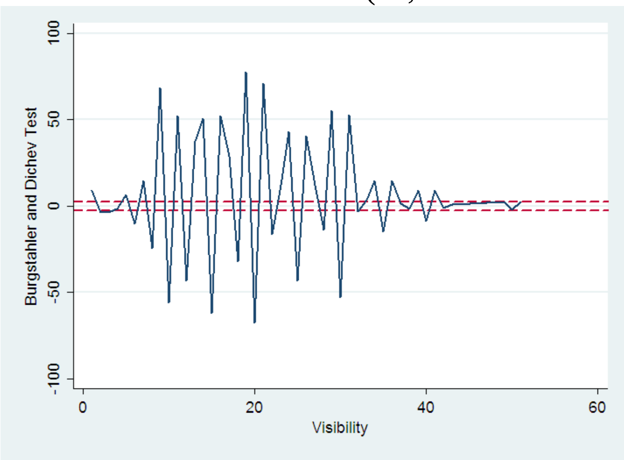
When the API is between 0 and 40 (12,287 observations)



When the API is between 40 and 80 (68,997 observations)



When the API is between 80 and 120 (35,223 observations)



When the API is between 120 and 500 (10,181 observations)

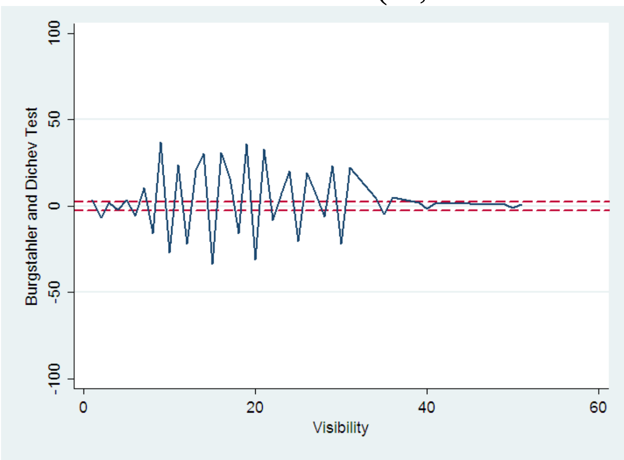
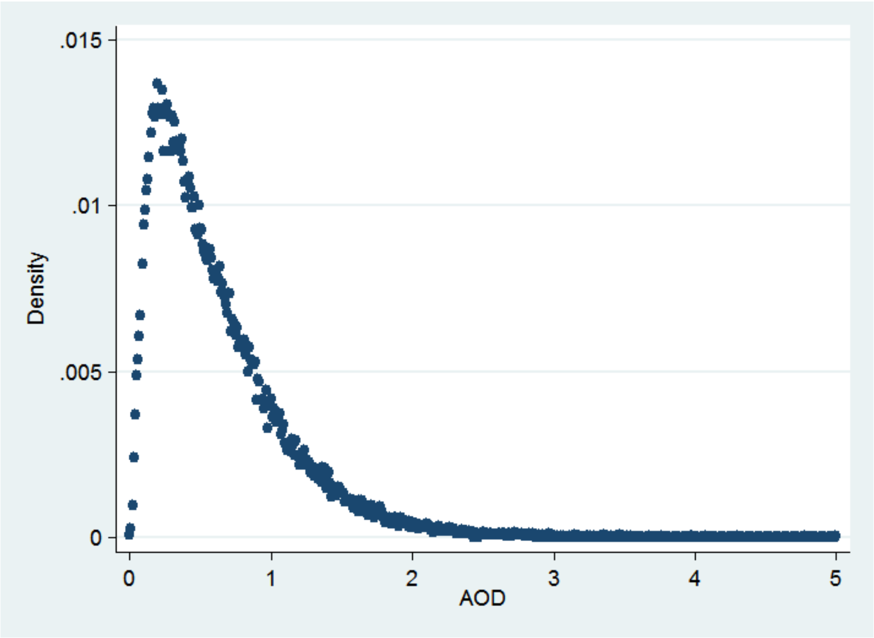


Figure 5: Histogram and Burgstahler and Dichev Test of Discontinuity for city-day average of AOD

Bin width = 0.01.

Histogram of city-day average of AOD



Burgstahler and Dichev test of discontinuity

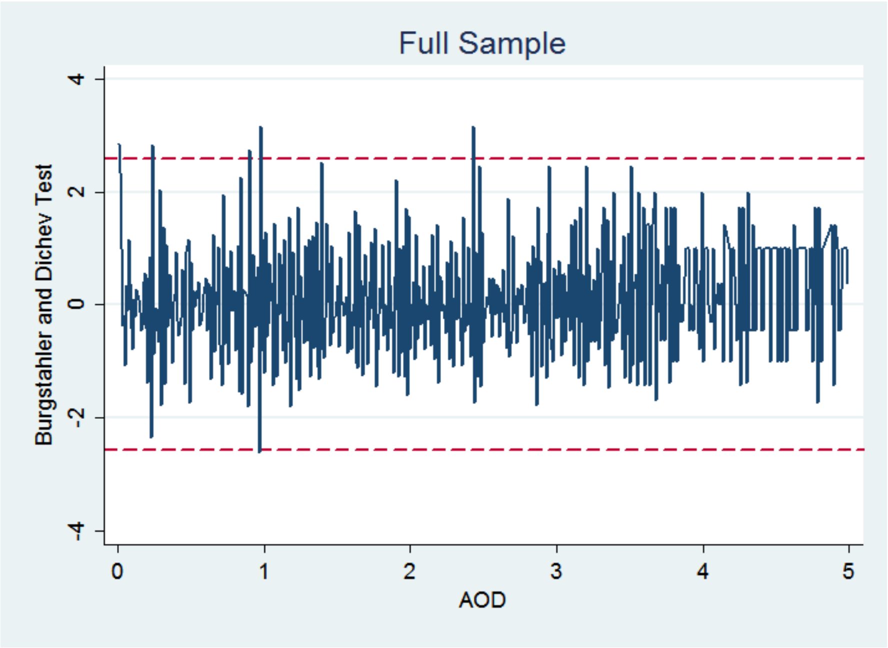
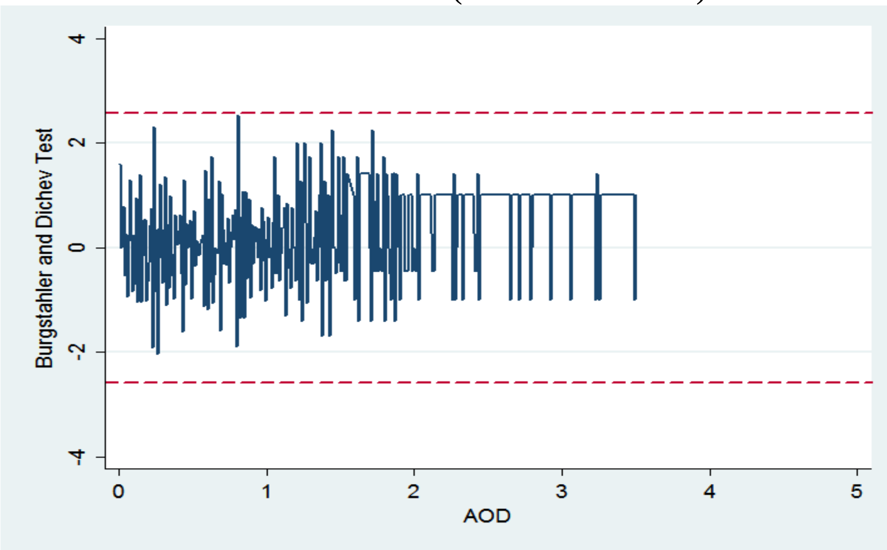
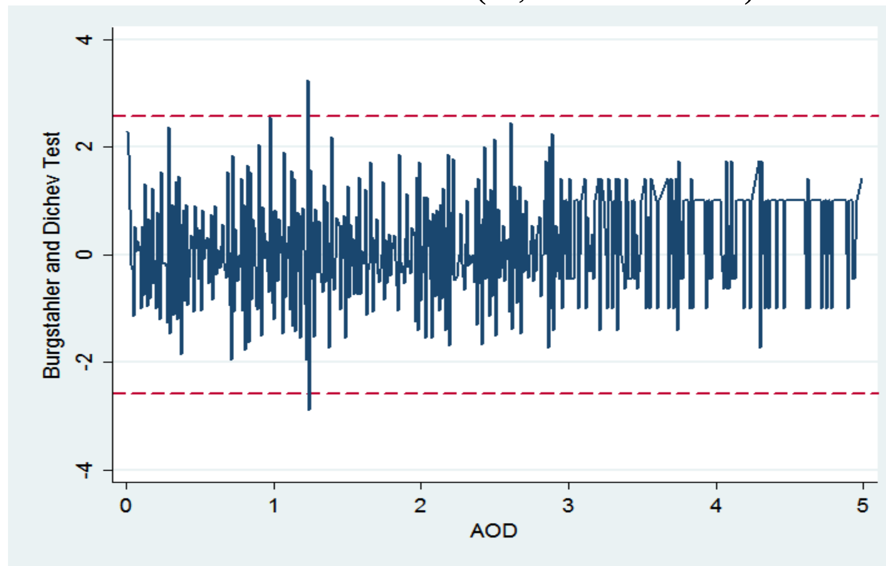


Figure 6: Burgstahler and Dichev Test of Discontinuity for City-Day Average of AOD by Ranges of the API

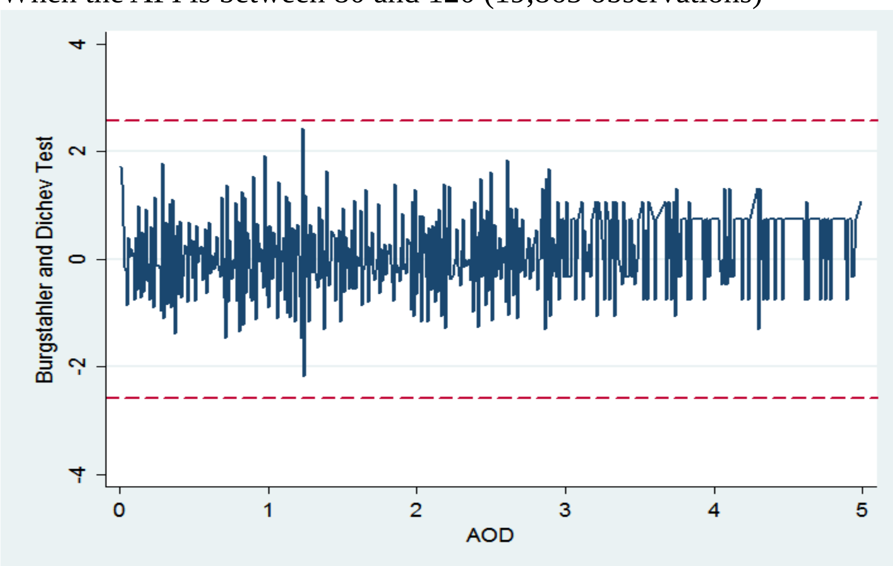
When the API is between 0 and 40 (3924 observations)



When the API is between 40 and 80 (35,070 observations)



When the API is between 80 and 120 (19,865 observations)



When the API is between 120 and 500 (5,090 observations)

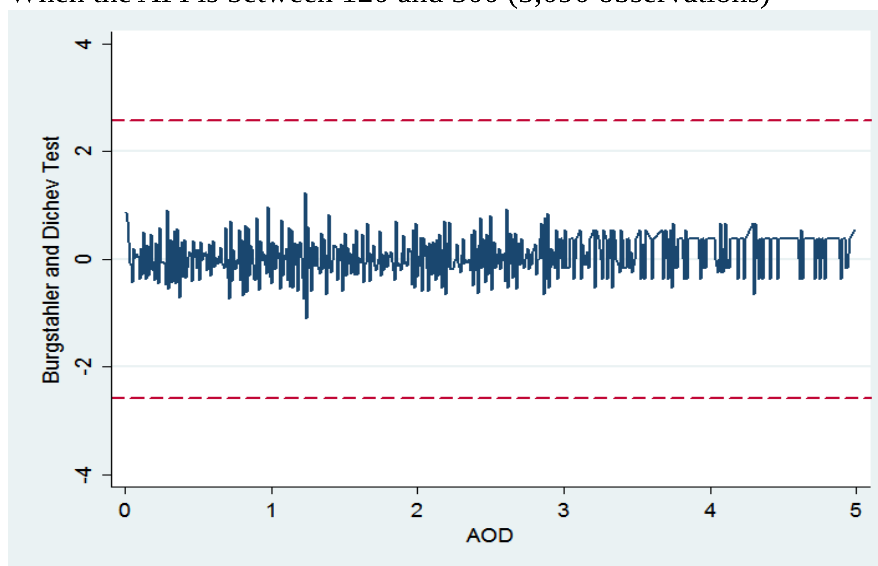


Table 1: MEP cutoff points for different levels of API

API	Pollutant intensity (μg/m³)			Air quality level	Air Quality condition	Notes of health effects
	TSP	SO_2	NO_2			
500	1000	2620	940	V	Heavy pollution	Exercise endurance of the healthy people drops down, some will have strong symptoms. Some diseases will appear.
400	875	2100	750			
300	625	1600	565	IV	Moderate pollution	The symptoms of the patients with cardiac and lung diseases will be aggravated remarkably. Healthy people will experience a drop in endurance and increased symptoms.
200	500	250	150	III	Slightly polluted	The symptom of the susceptible is slightly aggravated, while the healthy people will have stimulated symptoms.
100	300	150	100	II	Good	Daily activity will not be affected.
50	120	50	50	I	Excellent	Daily activity will not be affected.

Source: The first four columns are taken from the MEP website. The last three columns are copied from Table 2.2 of UNEP (2009).

Table 2: Regression results on pressure to meet the target of model city (observation = city-day)

Linear probability model, observations from the 4th quarters only, 2000-2009.

Pressure0930 = $\frac{target\ \#\ of\ blue\ sky\ days - realized\ blue\ sky\ days}{\#\ of\ days\ in\ the\ calendar\ year - day\ of\ year}$ as of September 30. Close=1 if Pressure0930>75%.

	(1) **API<=100**	(2) **API<=100**	(3) **API in [96,100]**	(4) **API in [96,100]**	(5) **API in [101,105]**	(6) **API in [101,105]**
pressure0930_close_model	0.129	0.265**	0.156**	0.157**	-0.042**	-0.066**
	(0.115)	(0.098)	(0.076)	(0.063)	(0.019)	(0.029)
pressure0930_close_modellag	-0.076	0.043	-0.269*	-0.279*	-0.068**	-0.064
	(0.201)	(0.098)	(0.143)	(0.140)	(0.033)	(0.041)
pressure0930_close_modelahead	0.194*	0.463***	0.276***	0.273***	-0.097***	-0.112***
	(0.103)	(0.124)	(0.071)	(0.060)	(0.025)	(0.029)
close_model	-0.213**	-0.124*	-0.064	-0.062	0.008	0.006
	(0.080)	(0.073)	(0.053)	(0.051)	(0.013)	(0.012)
close_modellag	-0.014	-0.011	0.146	0.155*	0.027*	0.018
	(0.137)	(0.087)	(0.094)	(0.090)	(0.016)	(0.019)
close_modelahead	-0.242*	-0.289**	-0.113**	-0.095***	0.048**	0.053***
	(0.132)	(0.111)	(0.043)	(0.031)	(0.021)	(0.016)
pressure0930	0.109***	0.076**	-0.018	-0.017	0.012	0.013
	(0.038)	(0.030)	(0.016)	(0.018)	(0.011)	(0.012)
Close	-0.075***	-0.016	-0.026***	-0.015	0.014**	0.013*
	(0.026)	(0.026)	(0.007)	(0.009)	(0.006)	(0.007)
Model	0.139***	0.001	-0.005	-0.001	-0.009	0.004
	(0.037)	(0.035)	(0.010)	(0.019)	(0.007)	(0.007)
Modellag	0.099	0.039	0.020*	0.023***	-0.006	0.001
	(0.063)	(0.047)	(0.012)	(0.005)	(0.007)	(0.007)
Modelahead	0.120	0.057	-0.016	-0.026**	0.003	0.004
	(0.124)	(0.088)	(0.019)	(0.013)	(0.015)	(0.009)
Weather	Y	Y	Y	Y	Y	Y
Date FE	Y	Y	Y	Y	Y	Y
City FE	Y	Y	Y	Y	Y	Y
City-specific trend		Y		Y		Y
Socioeconomic		Y		Y		Y
Observations	31688	31688	31688	31688	31688	31688
R-squared	0.330	0.358	0.061	0.066	0.062	0.067

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Socioeconomic factors include GDP growth rate, GDP per capita, industrial production, population, energy consumption and vehicle number by city and year, as well as the dummy for regular heating season.

Table 3: Placebo test on response to a hypothetical target of 80% blue sky days effective 1/1/2001 (observation = city-day)

Linear probability model, observations from the 4th quarter only, 2000-2002.

Pressure0930 = $\frac{target\ \#\ of\ blue\ sky\ days - realized\ blue\ sky\ days}{\#\ of\ days\ in\ the\ calendar\ year - day\ of\ year}$ as of September 30. Close=1 if Pressure0930>75%.

	(1) **API<=100**	(2) **API<=100**	(3) **API in [96,100]**	(4) **API in [96,100]**	(5) **API in [101,105]**	(6) **API in [101,105]**
pressure0930_model	-0.579***	-0.447	-0.032*	0.042	0.038*	-0.055
	(0.154)	(0.341)	(0.018)	(0.373)	(0.020)	(0.107)
pressure0930_modellag	-0.361***	-0.034	0.007	0.037	-0.002	0.055
	(0.118)	(0.479)	(0.022)	(0.542)	(0.017)	(0.156)
pressure0930_modelhead	0.113	0.281	0.063***	0.063	-0.018	0.010
	(0.134)	(0.302)	(0.019)	(0.349)	(0.022)	(0.090)
pressure0930	0.038	0.039	-0.033***	-0.051*	0.011	-0.003
	(0.066)	(0.089)	(0.011)	(0.026)	(0.014)	(0.026)
model	0.197**	0.222	0.020***	-0.067	-0.013*	0.101
	(0.074)	(0.459)	(0.005)	(0.587)	(0.007)	(0.142)
modellag	0.127**	0.031	0.024**	-0.025	-0.003	0.018
	(0.057)	(0.071)	(0.010)	(0.025)	(0.008)	(0.021)
modelhead	-0.075	-0.080*	-0.020***	-0.036	0.005	0.025*
	(0.074)	(0.041)	(0.007)	(0.037)	(0.008)	(0.013)
Weather	Y	Y	Y	Y	Y	Y
Date FE	Y	Y	Y	Y	Y	Y
City FE	Y	Y	Y	Y	Y	Y
City-specific trend		Y		Y		Y
Socioeconomic		Y		Y		Y
Observations	10117	10117	10117	10117	10117	10117
R-squared	0.427	0.458	0.054	0.061	0.050	0.059

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Socioeconomic factors include GDP growth rate, GDP per capita, industrial production, population, energy consumption and vehicle number by city and year, as well as the dummy for regular heating season.

Table 4: Correlation between API and Visibility (observation = city-day)

	Full sample			Sample with TSP as the dominant pollutant		
	Ln(visibility)	Ln(visibility)	Ln(visibility)	Ln(visibility)	Ln(visibility)	Ln(visibility)
Ln(API)	-0.381***	-0.381***	-0.404***	-0.420***	-0.418***	-0.442***
	(0.034)	(0.034)	(0.035)	(0.031)	(0.030)	(0.031)
Ln(API)*Left		0.127	-0.009		0.093	0.029
		(0.359)	(0.373)		(0.368)	(0.383)
Ln(API)*Right		-1.323*	-1.166*		-0.920	-0.798
		(0.720)	(0.669)		(0.713)	(0.666)
Left		-0.589	0.031		-0.441	-0.148
		(1.641)	(1.705)		(1.682)	(1.753)
Right		6.151*	5.415*		4.259	3.695
		(3.341)	(3.103)		(3.305)	(3.087)
City FE	Y	Y	Y	Y	Y	Y
Date FE	Y	Y	Y	Y	Y	Y
City-specific trend			Y			Y
Socioeconomic			Y			Y
Observations	126684	126684	126684	92379	92379	92379
R-squared	0.561	0.561	0.570	0.572	0.572	0.581

Standard errors are clustered by city. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Socioeconomic factors include GDP growth rate, GDP per capita, industrial production, population, energy consumption and vehicle number by city and year, as well as the dummy for regular heating season.

Table 5: Correlation between API and city-day average of AOD (observation = city-day)

	Full sample			Sample with TSP as the dominant pollutant		
	Ln(AOD)	Ln(AOD)	Ln(AOD)	Ln(AOD)	Ln(AOD)	Ln(AOD)
Ln(API)	0.415***	0.410***	0.427***	0.368***	0.358***	0.237***
	(0.042)	(0.042)	(0.045)	(0.038)	(0.037)	(0.029)
Ln(API)*Left		0.417	0.555		0.598	0.424
		(0.508)	(0.496)		(0.493)	(0.361)
Ln(API)*Right		0.771	0.675		0.137	0.477
		(0.907)	(0.923)		(1.074)	(0.660)
Left		-1.877	-2.503		-2.688	-1.912
		(2.324)	(2.269)		(2.256)	(1.655)
Right		-3.556	-3.110		-0.598	-2.189
		(4.203)	(4.277)		(4.978)	(3.058)
City FE	Y	Y	Y	Y	Y	Y
Date FE	Y	Y	Y	Y	Y	Y
City-specific trend			Y			Y
Energy			Y			Y
Socioeconomic			Y			Y
Observations	63948	63948	63948	50672	50672	50672
R-squared	0.638	0.639	0.641	0.659	0.659	0.585

Standard errors are clustered by city. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Socioeconomic factors include GDP growth rate, GDP per capita, industrial production, population, energy consumption and vehicle number by city and year, as well as the dummy for regular heating season.

Table 6: Summary of API before, during and after a city won the Model City Award (observation = city-day)

	City type		
	Did not win model city award before 2010	Won model city award between 2000 and 2009	Won model city award at or before 1999
Number of cities per type	22	9	6
API			
>= 2 years before winning the award	80.21	82.09	
	(42.34)	(39.37)	
0-1 years before winning the award		77.52	
		(32.28)	
After winning the award		74.34	56.67
		(28.68)	(26.37)
Blue sky day? (=1 if API<=100)			
>= 2 years before winning the award	80.89%	79.47%	
	(0.39)	(0.40)	
0-1 years before winning the award		85.91%	
		(0.35)	
After winning the award		89.18%	96.34%
		(0.31)	(0.19)
API in [96,100]			
>= 2 years before winning the award	4.81%	5.91%	
	(0.21)	(0.24)	
0-1 years before winning the award		7.21%	
		(0.26)	
After winning the award		6.60%	2.00%
		(0.25)	(0.14)
API in [101,105]			
>= 2 years before winning the award	2.77%	2.50%	
	(0.16)	(0.16)	
0-1 years before winning the award		1.80%	
		(0.13)	
After winning the award		1.34%	0.48%
		(0.12)	(0.07)

Table 7: Model City Award and Air Quality Improvement (observation=city-day)

	(1) Ln(API)	(2) Ln(API)	(3) Ln(visibility)	(4) Ln(visibility)	(5) Ln(AOD)	(6) Ln(AOD)	(7) API<=100	(8) API<=100
0-1 years before winning model city award	-0.040	-0.029	-0.037	-0.050	0.046	0.134	0.023	0.028
	(0.027)	(0.043)	(0.029)	(0.034)	(0.109)	(0.098)	(0.037)	(0.041)
After winning model city award	-0.037	-0.041	0.048	0.033	0.008	0.018	0.003	0.016
	(0.035)	(0.058)	(0.058)	(0.067)	(0.075)	(0.110)	(0.046)	(0.049)
Weather	Y	Y	Y	Y	Y	Y	Y	Y
Date FE	Y	Y	Y	Y	Y	Y	Y	Y
City FE	Y	Y	Y	Y	Y	Y	Y	Y
City-specific trend		Y		Y		Y		Y
Socioeconomic		Y		Y		Y		Y
Observations	126688	126688	126684	126684	126706	126706	126688	126688
R-square	0.524	0.547	0.533	0.540	0.479	0.495	0.292	0.318

Standard errors are clustered by city and in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Socioeconomic factors include GDP growth rate, GDP per capita, industrial production, population, energy consumption and vehicle number by city and year, as well as the dummy for regular heating season.