Journal of Development Economics 146 (2020) 102488

Contents lists available at ScienceDirect

Journal of Development Economics

journal homepage: www.elsevier.com/locate/devec

Regular Article

The effects of fuel standards on air pollution: Evidence from China[☆]

Pei Li[a], Yi Lu[b], Jin Wang[c,*]

[a] Department of Public Finance, Zhejiang University, Hangzhou, 310027, China
[b] School of Economics and Management, Tsinghua University, Beijing, 100084, China
[c] Division of Social Science, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong

ARTICLE INFO

JEL classification:
L51
Q53
Q58

Keywords:
Fuel standards
Sulfur content
Vehicle emissions
Pollution
China

ABSTRACT

This paper examines the causal relationship between China's fuel standards and air pollution using a difference-in-differences design and regression discontinuity analyses. Combining data on staggered prefectural implementation of the regulations with hourly station-level pollution data, we show that the enforcement of high-quality gasoline standards significantly improved air quality, especially in terms of fine particles and ozone. The average pollution across all pollutants was reduced by 12.9%. The new gasoline standard's net benefit is estimated to be about US$26 billion annually. These findings demonstrate the effectiveness of precise standards in reducing air pollution in a developing country setting.

1. Introduction

Ninety-two percent of the world's population lives in places where the average air quality is beyond the World Health Organization's suggested limits for pollutants. Those in Asia, Africa and the Middle East are disproportionately exposed to high concentrations of air pollution (WHO, 2015). With rapid urbanization, road transport, in part due to low quality fuel and numerous old and poorly maintained vehicles, has come to contribute from 12% to 70% of the air pollution in major developing cities, imposing alarming health costs on the public (WHO, 2011). The sheer size of the population and particularly the growing urban mass in less-developed countries make the welfare consequences of targeted environmental regulations to, for example, improve fuel quality of high importance. Yet, systematic empirical evidence on the extent to which such policies can effectively address automobile pollution remains scant.

In an important contribution, Auffhammer and Kellogg (2011) examines the effects of gasoline content regulations on ambient pollutant concentrations in the United States. But evidence from better-developed economies may not readily be transferred to the developing world given the different contexts and institutions. For example, developing countries face constraints on their regulatory capacity to enforce fuel standards. There is also industry opposition, and governments have limited financial resources and staff (International Council on Clean Transportation, 2010).

This study attempts to bridge this gap by documenting novel empirical evidence and measuring the benefits of fuel standards in the largest developing country, China. More precisely, it investigates the relationship between China's fuel standards, which specify lower sulfur content, and levels of air pollution. There are three reasons for this focus. First, China has become the world's largest automobile market. There has been an unprecedented increase in the country's vehicle ownership, shown in Fig. 1. Intensive oil consumption is associated with a large number of externalities (Parry et al., 2007). China's cities rank among the most polluted in the world (World Bank, 2007; Greenstone and Hanna, 2014; Chen et al., 2013; Cropper, 2010). That makes regulating

---

[☆] We would like to thank the editor, Prof. Andrew Foster and two anonymous referees for their comments and suggestions that helped to substantially improve this paper. We also thank participants at the 2016 Green Cities in Asia workshop and the 2018 China meeting of the Econometric Society for helpful discussions. Li acknowledges the financial support of China's National Social Science Fund (grant no. 18ZDA096). Wang acknowledges the financial support of a research grant from the Hong Kong University of Science and Technology (no. SBI18HS06). All errors remain our own. Lu acknowledges the financial support of China's National Natural Science Fund (grant no. 71973078).

[*] Corresponding author.
E-mail addresses: bengcome@gmail.com (P. Li), luyi@sem.tsinghua.edu.cn (Y. Lu), sojinwang@ust.hk (J. Wang).

https://doi.org/10.1016/j.jdeveco.2020.102488
Received 29 April 2019; Received in revised form 25 March 2020; Accepted 31 March 2020

fuel content of great policy relevance. Second, uniform fuel standards have gradually been introduced in Chinese cities, which provides a new opportunity to understand their environmental implications. The different standards together with accurate data allow overcoming methodological obstacles that have impeded progress in this field.

Third, Chinese fuel standards generally follow the practices of Europe, Japan and the United States. Very similar regulations facilitate meaningful international comparisons of policy effectiveness. Apart from the detailed insights into the role of policy in the current setting, some general lessons could emerge which would enrich our view of fuel content regulations in a broader context.

The study's analyses exploit a compelling quasi-experiment: changes in standards for fuel sold in Chinese cities. According to official statistics from China's Ministry of Environmental Protection (MEP), exhaust from motor vehicles contributes a quarter to a third of particulate matter (*PM*) air pollution throughout the country (MEP, 2013). Against this background, since 2013 China has gradually tightened its gasoline standards, followed by diesel standards upgrading in certain cities. The introduction of low-sulfur fuel has aimed at reducing emissions from the motor vehicle fleet substantially, and the regulations need to be strictly enforced by the retailers.

The study's analyses involve merging data on prefecture-level regulations with hourly pollution data from 1492 air quality monitoring stations in 337 prefectures for the three years 2013–2015. In addition to the air quality index (AQI), a composite measure of pollution, data on suspended particulates less than 10 μm in diameter (PM_{10}) and less than 2.5 μm ($PM_{2.5}$) and on ozone (O_3) concentrations are analyzed. Those pollutants are particularly related to fuel composition, and are among the most harmful to human health. This then marks the first time that high-quality data on fine particulates and ozone have been used to research Chinese environmental issues (Pope and Dockery, 2013).

We focus on precisely estimating the effect on air pollution of the first change in gasoline standards implemented during the period studied. Two estimation strategies are applied. In the first, both temporal and geographic variations in the implementation of the new gasoline standard are exploited to identify its effects. The empirical analysis compares daily changes in the local concentrations of air pollutants between prefectures implementing the new standard earlier (the treatment group) and later (the control group). The validity of the difference-in-differences (DD) methods applied and the causal interpretation of the results rely on the assumption that prefectures that adopted the new standard later are proper counterfactuals for what would have happened to earlier adopters in the absence of the change. A large number of other variables are included in the analyses, including monitoring station fixed effects, day fixed effects, station-specific seasonality, weather conditions, fuel prices and subsequent changes in fuel quality standards. It is also necessary to remove the confounding influence of other on-going policies aimed at curbing air pollution by directly controlling for them. To further address the selected nature of the enforcement dates, we control for the differences in the trends in outcomes between early adopters and later ones depending on the key determinants in the timing of the fuel standard changes (i.e., the historical pollution level and growth), a technique used by Gentzkow (2006). Beyond that, a series of robustness tests verify that the treatment and control prefectures are comparable in terms of their pre-regulation trends in the outcomes.

The study's second set of analyses follows the lead of work by Davis (2008) and by Auffhammer and Kellogg (2011) and the review by Hausman and Rapson (2018). A regression discontinuity in time (RDiT) framework is applied to identify the effect of gasoline standards on air pollution. After the new regulations came into force, all the retailers in the treatment prefectures were assumed to have immediately switched to supplying only gasoline meeting the new standards. That is assumed to have created a discontinuity in tailpipe emissions. In effect, the abrupt change makes other factors smooth on the day of implementation of the new regulations, so the day just before the new regulations can serve as a good counterfactual to the day the new regulations came into force. To validate concerns about causal identifications in the RDiT context, serial correlation in the daily pollutant data is evaluated to recover both the short-run and long-run policy effects. In addition, a number of auxiliary sensitivity checks provide reassuring results.

The analyses yield several main results. First, the enforcement of more stringent fuel standards significantly improved air quality. The average pollution across all pollutants was reduced by 12.9%. The progression from a lower to a higher standard led to an average fall of 9.4 $\mu g/m^3$ in the concentration of $PM_{2.5}$. Reassuringly, the long-run policy effect implied by the RD estimate aligns well with that from the DD estimate. Such improvement points to the importance of fuel standards in mitigating vehicles' environmental adversities.

Second, there are some intra-day fluctuations, but these are well explained by the atmospheric chemistry of the pollutants, ensuring the credibility of the main results. Overall, local governments' environmental protection efforts appear to be an important institutional factor in enhancing the environmental benefits of a higher fuel standard, suggesting their importance in China's decentralized system of environmental governance. The new standards are more effective in cities with more active government pollution reduction efforts.

Third, a back-of-the-envelope analysis further suggests that the net benefit of adopting the new gasoline standard is significant, at US$26 billion annually. Using health-based studies of particulate air pollution from China, the improved air quality generates US$21.44 billion annually in health benefits from reduced mortality and US$8.65 billion annually in health benefits from reduced morbidity. The upgrading involves a cost increase (measured at consumer prices) of about US$3.99 billion.

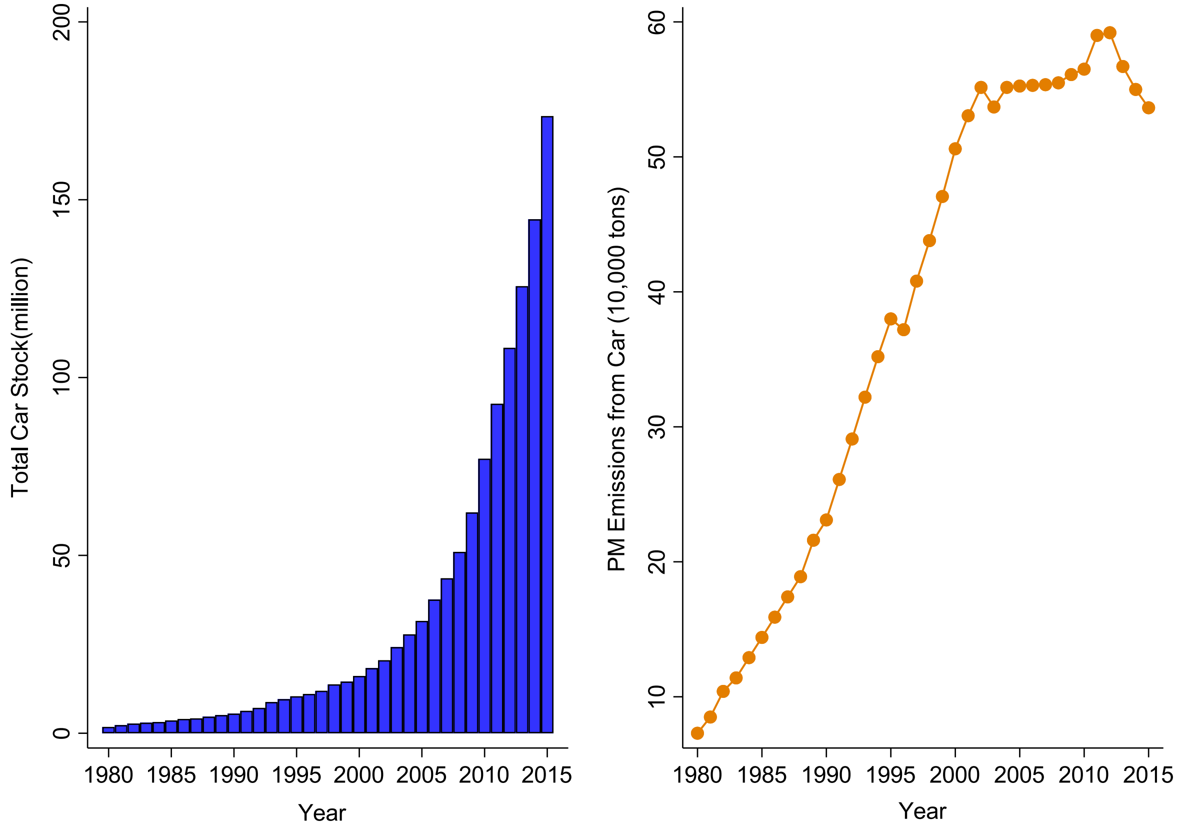
This study contributes to several strands of literature assessing the impact of environmental regulations. It relates to a growing body of work which emphasizes the technological aspects of environmental policies (Copeland and Taylor, 2004). Scholars have previously demonstrated negative relationships between air pollution outcomes and regulatory measures that tightened vehicle emission standards (Kahn, 1996; Kahn and Schwartz, 2008; Greenstone and Hanna, 2014).[1] While the impact of vehicle emission standards on air pollution takes effect only gradually through turnover in the vehicle fleet, the results of this study show that new fuel standards can immediately affect all vehicles on the road and influence air quality dynamics in the near as well as the long term.

Closest in spirit to this work is that published by Auffhammer and Kellogg (2011). They showed that the effectiveness of American gasoline content standards depends on flexibility in choosing a compliance mechanism. Flexible federal gasoline standards did not improve air quality, but accurately targeted, inflexible regulations in California did so significantly. These results from China echo theirs and highlight the role of precise standards in reducing air pollution. Auffhammer and Kellogg took note of then-recent changes in federal and Californian gasoline regulations aiming to restrict the sulfur content, but they chose to focus on studying the earlier regulation of volatile organic compounds (*VOCs*) and nitrous oxides (NO_x) from gasoline. This study has extended their work to China by examining the effect on air pollution of gasoline sulfur content requirements, a key fuel quality parameter, and by including for the first time $PM_{2.5}$ levels as a main outcome.

Another strand of research has focused on regulatory policies designed to reduce the scale of pollution. In particular, several studies

---

[1] Kahn (1996) as well as Kahn and Schwartz (2008) have shown that both government regulation and innovation by automakers have significantly reduced regional air pollution in the United States caused by driving. The environmental impact of regulation should increase with time as the share of pre-regulation vehicles on the roads declines. Greenstone and Hanna (2014) have shown that India's air pollution regulations requiring catalytic converters for new vehicles improved air quality and thus reduced death rates.



Notes: Using the data extracted from China's annual Vehicle Emission Control reports for 2010–2016, this figure plots the total car stock (millions) and related particulate matter emissions (10,000 tons). The PM emissions from cars are as estimated in the Ministry of Environmental Protection's nationwide air pollution source analysis, which is aimed at identifying sources of ambient air pollution and quantifying their contributions.

Fig. 1. Motor vehicle ownership and particulate matter emissions in China.

have examined regulations targeting gasoline consumption and emissions (Parry et al., 2007; Jacobsen, 2013; Anas et al., 2009).[2] Others have looked at the relationship between driving restrictions and air quality, including Davis (2008), Wolff (2014), as well as Viard and Fu (2015).[3] While administrative restrictions on the use of the vehicle fleet are known to be very costly and compliance is known to be a critical issue, perhaps more emphasis should be placed on cleaner technologies, of which cleaner fuel is a prominent example. The findings of this study suggest that improving fuel standards could be an efficient policy tool, since compliance can be more strictly enforced.

The rest of the paper proceeds as follows. Section 2 lays out the fuel standard reform background, followed by a description of the empirical strategy and data in section 3. Section 4 presents the empirical results. The last section concludes.

2. Fuel standards in China

2.1. Fuel composition, vehicle emissions and pollutant formation

It has been well documented that gasoline's sulfur content is among the most relevant determinants of vehicle emissions.[4] Higher-sulfur gasoline generates more sulfur compounds in the exhaust and significantly impairs the effectiveness of emission control systems (EPA, 2000, 2014). Gasoline vehicles depend crucially on catalytic converters to reduce tailpipe emissions of harmful pollutants such as nitrous oxides (NO_x), carbon monoxide (*CO*), and volatile organic compounds (*VOCs*) that include precursors for ozone and secondary *PM* in the atmosphere. There is ample evidence that the presence of even a tiny amount of sulfur in fuel has a measurable impact on catalyst efficiency, an effect known as sulfur inhibition.[5] The consequent degradation of its catalytic converter causes emission levels to significantly exceed a vehicle's full useful life emission standards. Moreover, this negative effect is known

---

[2] Parry et al. (2007) have shown that among a set of policy instruments in the U.S., gasoline taxes reduced a greater number of important externalities than did fuel economy standards. Further advances have been made by Jacobsen (2013), who has specifically studied the mechanisms and welfare implications of fuel economy standards. Anas et al. (2009) compared the effectiveness of a congestion toll and a fuel tax in reducing traffic congestion as well as gasoline consumption and emissions in Beijing.

[3] They have shown that the policy outcome varies depending on the context, from no effects in Mexico, to significant effects in Germany and Beijing. Such differences crucially depend on the behavioral responses of the drivers influenced (compliance versus compensating responses).

[4] The sulfur content of fuel is also one of the most important characteristics affecting diesel vehicles' NO_x and *PM* emissions (International Council on Clean Transportation, 2010). During combustion, sulfur in diesel fuel converts into direct particulate matter and sulfur dioxide emissions, which can lead to secondary particle formation.

[5] Sulfur and sulfur compounds attach or "adsorb" to the precious metal catalysts that are required to convert these emissions. Sulfur also blocks sites on the catalyst designed to store oxygen that is necessary to optimize the conversion of emitted NO_x (EPA, 1998, 1999, 2000, 2014).

Table 1
Fuel standards roadmap.

Stage		Standard	Maximum sulfur (ppm)	Standard Issued on	Implementation
China	Gasoline III	GB 17930-2006	150	Dec 6, 2006	Phased-in by Dec 31, 2009
	Gasoline IV	GB 17930-2011	50	May 12, 2011	Phased-in by Dec 31, 2013
	Gasoline V	GB 17930-2013	10	Dec 18, 2013	Phased-in by Dec 31, 2017
China	Diesel III	GB 19147-2009	350	June 12, 2009	Phased-in Jan 1, 2010–Jul 1, 2011
	Diesel IV	GB 19147-2013	50	Feb 7, 2013	Phased-in by Dec 31, 2014
	Diesel V	GB 19147-2009	10	June 8, 2013	Phased-in by Dec 31, 2017

Notes: This table presents the key dates and road map for improving China''s nationwide gasoline and diesel standards from III to IV and then to V. The maximum sulfur content is specified for each new standard.
Source: Intl. Council on Clean Transportation

to be irreversible for one or more pollutants, implying a profound long-term emission impact of high sulfur fuels (EPA, 2000; 2014).[6]

A large number of epidemiological studies have linked air pollutants with adverse health effects. *PM* in particular pose a great risk to humans because the particles can penetrate deep into the lungs and remain there for long periods. They diffuse readily into indoor environments, and are transported over long distances (Pope and Dockery, 2006). They are associated with increased incidence of lung cancer and with respiratory and heart disease mortality, and are known to aggravate asthma seriously (Pope and Dockery, 2013; Greenstone, 2004; Parry et al., 2007). Ozone is another important air pollutant which damages both human health and agricultural crops.

It is noteworthy that the air pollutant concentrations observed at different locations depend on more than the quantity of various emissions. Extensive studies have shown that wind speed, temperature and rainfall all play important roles in determining local pollutant levels. Indeed, most air pollutants, including ozone and fine particulates, exhibit pronounced seasonal patterns because of weather (Bharadwaj et al., 2017). It is therefore important to control for weather and seasonality in any attempt to identify the impact of fuel standards on local pollutant concentrations.

2.2. Fuel content regulations

China has primarily followed the European Union's fuel standards since the late 1990s. After a decade of practice in addressing worsening air pollution, especially in urban areas, China decided to strengthen the hazardous materials control standards for vehicle gasoline and diesel. In May 2011 the China IV gasoline standard was issued specifying a maximum of 50 ppm (parts per million) of sulfur in gasoline. That standard had been phased in by the end of 2013. In early 2013 the State Council issued a further directive calling for the nationwide introduction of ultra-low-sulfur fuels (10 ppm) by the end of 2017 (State Council, 2013). The eastern coastal cities and mega-cities served as the starting point before the requirement was expanded to other areas. That directive was translated into formal regulations over the course of 2013 (ICCT, 2013; 2014). Ultimately, three new standards were issued: China IV diesel (50 ppm) in February 2013 to be phased in by December 31, 2014; China V diesel (10 ppm) in June 2013 to be phased in by December 31, 2017; and China V gasoline (10 ppm) in December 2013 to be phased in by December 31, 2017.[7] Together, those standards constituted a road map for improving China's nationwide fuel standards, as shown in Table 1.

Following the central government's directive, China's provinces have been revising their regulations to implement the new fuel standards. The provincial authorities (including the Department of Environmental Protection and Transport, the Development and Reform Commission, and the Economic and Information Technology Commission) were deeply involved in this process. They consulted various stakeholders such as provincial branches of major state-owned oil companies (Sinopec and PetroChina, which dominate the Chinese market and run both refineries and retail stations), automobile manufacturers, research institutes and quality inspection institutions. In deciding on the local implementation, one important consideration was to limit exhaust emissions from road vehicles. But local refining capacity was also a critical factor, as they were tasked with producing and supplying higher quality fuel without shortages.[8] Some provinces moved faster than others. For twenty-nine provinces, each decided its own effective date for the standards and then applied them simultaneously to all of its prefectures. Jiangsu and Guangdong provinces chose to extend the new standards to their prefectures gradually. A list of major prefectures were selected to adopt the new standards first, followed by the rest in the second stage. Although prefectures and regions in China may implement fuel quality standards according to their own timelines, fuel price changes are tightly regulated by the National Development and Reform Commission (NDRC), the nation's top economic planner. To compensate for the required refinery upgrades and increased production costs of cleaner fuels, the NDRC announced a new pricing policy. The wholesale prices of China IV gasoline and diesel were increased by ¥290 and ¥370/ton, respectively. The prices of China V gasoline and diesel were raised by a further ¥170 and ¥160/ton. Upon the implementation of the new standards, retail stations raised their gasoline and diesel prices accordingly.

Fig. 2 shows the gradual upgrading of the China III gasoline and diesel fuel quality standards to IV and then to V in Chinese prefectures. By the end of 2013, 25.5% of prefectures had implemented the gasoline IV standard, and 1.5% of prefectures had adopted diesel IV.[9] Those ratios had increased to 100% and 22%, respectively, by the end of 2014. By 2015 all of China's prefectures were supplying only gasoline IV and diesel IV. As for the ultra-low-sulfur fuels, the gasoline V and diesel V standards are still in the process of implementation. By 2013, 3% of prefectures had started to supply gasoline V, while only 0.3% of prefectures had adopted diesel V. Those ratios had increased to 12.5% and 3.9% in the following year. By the end of the period studied in 2015, 14.5% of prefectures had adopted gasoline V and 14.2% of prefectures had adopted diesel V.

2.3. Chinese standards in a global context

Table 2 displays China's gasoline standards in an international con-

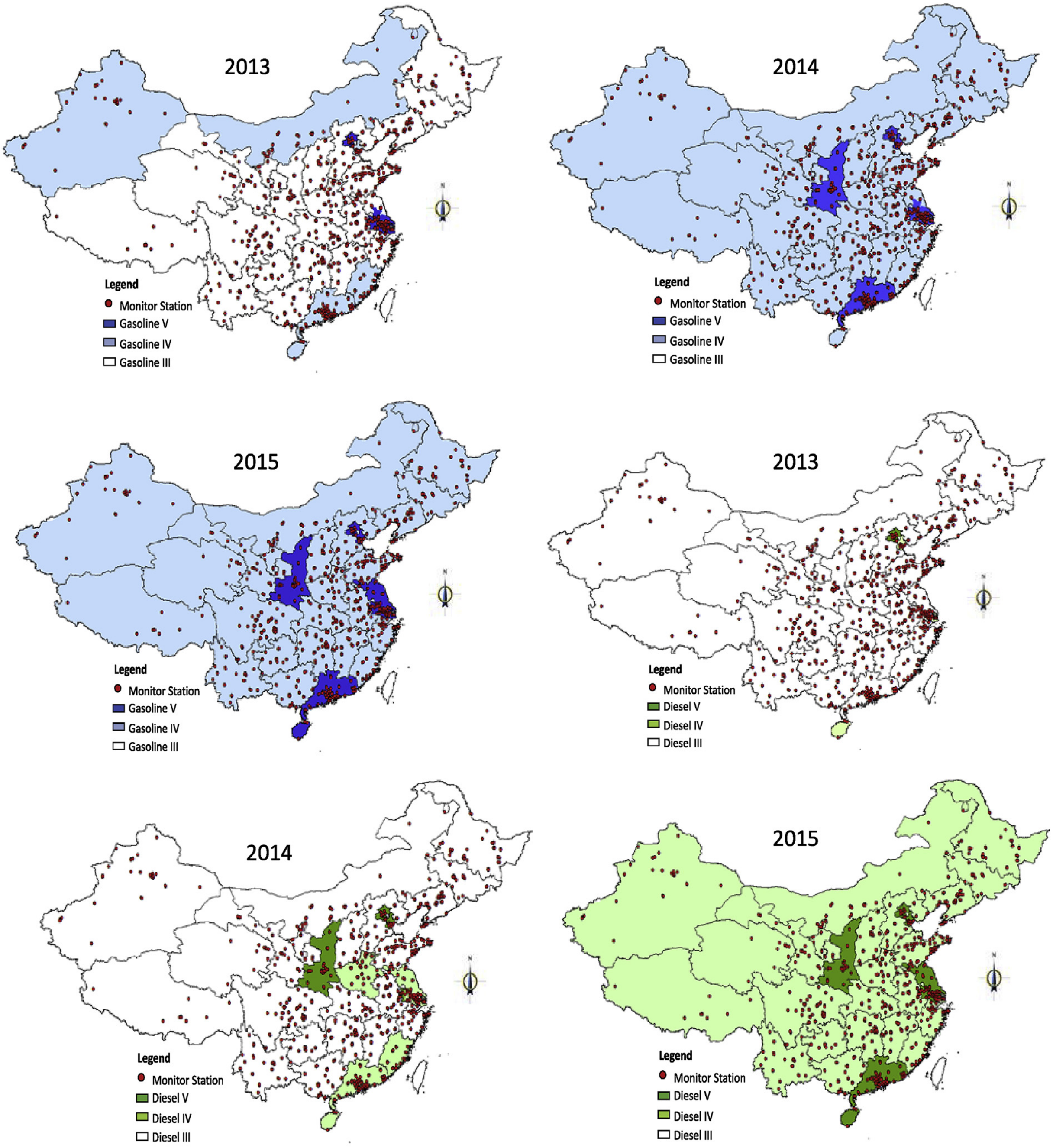
---

[6] The sulfur build-up on the catalysts due to past exposure to high sulfur fuel may be irreversible even when the current fuel carries less sulfur.

[7] In addition to sulfur reductions, the progression from China III to China V gasoline standards involves a reduction in maximum permitted manganese levels and reductions in minimum octane requirements. The progression from China III to V diesel standards involves changes in the required cetane content.

[8] For example, the Xinjiang Autonomous Region started to implement the gasoline IV standard from December 21st, 2013. By the end of November, five local Sinopec and PetroChina refineries had finished upgrading their facilities and were tasked with producing and supplying gasoline IV to local retail stations in Xinjiang after the regulation became effective.

[9] The last wave of cities had implemented the gasoline IV standard on January 1st, 2014.



Notes: This figure displays the gradual implementation of higher fuel standards in China. From 2013 to 2015, the gasoline and diesel fuel standards have been upgraded from III to IV and then to V. The fuel standards information was assembled from circulars issued by the provincial Development and Reform Commissions on implementing gasolines and diesel standards in the various provinces.

Fig. 2. Implementation of gasoline and diesel content regulations in China.

text, comparing them with those in Europe and the United States.

The fuel properties specified in European standards include maximum lead and sulfur content for gasoline, the cetane number for diesel, and sulfur content and fatty acid methyl esters for biodiesel. The standards are updated periodically with mandatory reductions in sulfur content. A timeline for achieving sulfur level reductions in Europe was established in 1999. In 2000, the Euro 3 standard mandated a maximum of 150 ppm of sulfur in gasoline and 350 ppm in diesel. From 2005, Euro 4 mandated a maximum of 50 ppm of sulfur in both gasoline and diesel. In 2009, Euro 5 mandated a maximum of 10 ppm of sulfur in gasoline.

In the US, the EPA created a federal reformulated gasoline (RFG) specifically aimed at reducing ozone pollution, and California's Air Resources Board (CARB) has had its own gasoline programs since the 1990s primarily regulating *VOCs* and NO_x from gasoline combustion. The RFG standard caps the benzene content of gasoline at 1 percent by volume. Limits on total evaporation of *VOCs*, toxic air pollutants and NO_x emissions were set based on a complex model, but the RFG standard grants refiners flexibility in deciding which specific *VOCs* to remove from their gasoline (Auffhammer and Kellogg, 2011). By contrast, the CARB has mandated more stringent reductions in concentrations of highly reactive *VOCs* than the RFG standard. It established

Table 2

International gasoline standard comparison.

Fuel Standards	China III	China IV	China V	Euro III	Euro IV	Euro V	EPA RFG	EPA Tier 2		EPA Tier 3	CARB Phase III
Fuel Property/Year implemented	by 2009	by 2013	by 2017	2000	2005	2009	1995–2000s	2004	2006	2017	2003
Research Octane, min.	97–90	97–90	95–89	95–91	95–91	95–91	NS				NS
Motor Octane, min.	88–85	88–85	90–84	85–81	85–81	85–81	NS				NS
Aromatics, vol%, max.	40	40	40	42	35	35	25				25
Olefins, vol%, max.	30	28	24	18	18	18					6
Benzene, vol%, max.	1	1	1	1	1	1	1				0.8
Sulfur, ppm, max.	150	50	10	150	50	10		120/300 (average/max)	30/80 (average/max)	10	20
RVP, kPa	88 Winter 72 Summer max.	42-85 Winter 40–68 Summer max.	45-85 Winter 40–657 Summer max.	60/70 max.	60/70 max.	60/70 max.	48 kPa (7.0 psi) max				48.2/47.6 max (7 psi)
Lead, mg/L, max.	5	5	5	5	5	5	ND				
Manganese, mg/L, max.	16	8	2	NS	NS	MMT<6 (by 2011) MMT<2(by 2014)	ND				ND
Oxygen, % m/m	2.7 (max.)	2.7 (max.)	2.7 (max.)	2.7 (max.)	2.7 (max.)	2.7 (max.)	2				1.8–2.2

NS = Not specified; ND = Nondetectable.

Notes: This table displays the gasoline standard of China in an international context. In terms of the key parameters, the Chinese gasoline standards III, IV and V are compared with those in Europe and the United States (including California). EPA RFG is the US EPA's reformulated gasoline. EPA Tier 2 and Tier 3 are the US EPA's Gasoline Sulfur programs. CARB is the California Air Resources Board. RVP is Reid vapor pressure.

Sources: Intl. Council on Clean Transportation (2010); Intl. Council on Clean Transportation (2014); U.S. Environmental Protection Agency (EPA) website.

limits for sulfur, benzene, olefins, aromatic hydrocarbons, oxygen, and Reid vapor pressure. From the early 2000s, the EPA regulations started to shift their focus to sulfur requirements, slightly later than in Europe. Under the Tier 2 program, refiners and importers of gasoline were given an overall sulfur cap of 300 ppm, with an annual average sulfur level of 120 ppm in 2004. The standard was reduced to 30 ppm with an 80 ppm cap in 2006. Since 2017, the Tier 3 program lowers the sulfur content of gasoline to a maximum of 10 ppm. Evidence has yet to be developed on the effectiveness of these sulfur programs.

Overall, Table 2 shows that China's fuel quality standards mostly, but not entirely, followed European precedents. Both the Chinese and European fuel specifications pay particular attention to sulfur content. A central question is how well low-sulfur fuel addresses air pollution problems. A prospective examination of their environmental impact in China is the main goal of this study.

3. Estimation strategy

3.1. Data

Analyzing changes in pollution levels in response to changes in fuel standards involved assembling a data set containing matched indicators of fuel standards implementation, fuel prices, air pollution, weather conditions, socioeconomic conditions in the prefectures and vehicle registration records.

The fuel standards information was assembled from circulars issued by the provincial Development and Reform Commissions on the implementation of standards IV and V in the various provinces. The fuel price data were compiled from NDRC circulars.[10] The fuel market is not fully integrated in China due to geography, distance, and different levels of regional development. There are also regional price policy variations. Twenty-four provinces apply a province-wide price in the prefectures they administer. Another six provinces apply prefecture-specific prices. Nationally, however, fuel price changes are tightly regulated by the NDRC. During the period studied, the top planner adjusted fuel prices 52 times in response to changes in international crude prices. In the circulars they issued they specified incremental price changes and effective dates for all provinces.

Air pollution data are published hourly and daily by the MEP.[11] The data for 2013 to 2015 came from 1492 monitoring stations in 337 prefectures. Following the implementation of new ambient air quality standards (MEP, 2012), data on fine particulates and O_3 became publicly available for the first time in 2013. An air quality index (AQI) was developed based on the hourly and daily observations of SO_2, NO_2, CO, PM_{10}, $PM_{2.5}$, and ozone. This was a notable shift from the previous index which considered only SO_2, NO_2, and PM_{10}. All 337 prefectures were required to disclose their once-classified air quality data beginning in 2015. The AQI scale ranges from 0 to 500. It is further divided into six ranges: 0–50, 51–100, 101–150, 151–200, 201–300 and 301–500. In public reports these are termed good, moderate, unhealthy for sensitive groups, unhealthy, very unhealthy, and hazardous, respectively.

Those Chinese data were supplemented by weather data from the China Meteorological Data Sharing Service System (http://cdc.cma.gov.cn/). Those data include daily readings from 820 weather stations in China during the 2013 to 2015 period. The meteorological variables were aggregated to the prefecture level by averaging the daily readings of all the weather stations within a prefecture. The indicators used were temperature, precipitation, humidity, duration of sunshine and wind speed. Weather stations for which there were fewer than 100 records per year were winsorized.

A comprehensive data set of monthly car registrations by model is compiled by the State Administration of Industry and Commerce. It documents the month and prefecture of sale, the brand and model of the vehicle registered, as well as major attributes such as fuel type, nominal oil consumption, rated power, and engine size. 65.42 million new cars were sold from 2013 to 2015. Those data were aggregated to the month-prefecture level and used to check the robustness of the findings of the regression discontinuity (RD) analyses.

An indicator of local governments' overall environmental efforts was constructed from 3432 government annual work reports from the 286 prefecture-level cities from 2001 to 2012. Other data on exhaust emission regulations were obtained from the official documents published by prefectural transportation and public security bureaus. Those documents clearly state the exact dates when the prefectural governments began enforcing exhaust emission standard IV for gasoline vehicles and required the scrapping of yellow label vehicles—gasoline vehicles not meeting Chinese gasoline standard I and diesel vehicles not meeting Chinese diesel standard III.

Detailed variable definitions and descriptive statistics are presented in Table 3.

3.2. Estimation framework

Two estimation techniques are applied to quantify the effects of the fuel content regulations on air pollution: a DD specification, and an RD framework.

3.2.1. DD estimation

Temporal and regional variations in the implementation of the regulations are exploited to conduct a DD estimation. Specifically, air pollution outcomes in the prefectures where the new fuel content regulations were being enforced are compared with those in prefectures where they were not (the first difference) before and after the implementation of the regulations (the second difference).

The analyses focus on the shift from gasoline standard III to standard IV for several reasons. First, because standard III had been implemented in China by 2009, so there had been two years without any fuel content regulation changes before the upgrading to gasoline standard IV. That provides a relatively long window to check the comparability between the treatment and control groups in the pre-treatment period. Second, the sequence of later fuel standard reforms (including the upgrading from IV to V and two sequential diesel standard upgrades from III to IV and then to V) was highly correlated with that of the gasoline standard IV reform. As a result, even if the treatment and control groups were comparable before the gasoline standard IV reform, that change would make the two groups different given the potential policy impact. That would violate the parallel trend assumption for the later reforms, making it more difficult to quantify any air pollution effects.

To contain any possible contamination of the gasoline standard IV reform's effect from the later fuel content reforms, those later reforms are included as explicit controls in the DD estimation. Specifically, the DD specification is:

$$y_{iscd} = \beta \cdot Gasoline4_{cd} + \boldsymbol{\rho}\mathbf{LPolicy}_{cd} + \boldsymbol{\lambda}_s + \boldsymbol{\lambda}_d + \boldsymbol{\lambda}_i + \boldsymbol{\gamma}\mathbf{X}_{scd} + \varepsilon_{iscd}, \quad (1)$$

where i, s, c, and d denote pollutants (AQI, $PM_{2.5}$, PM_{10}, or O_3), monitoring stations, cities, and days, respectively; y_{iscd} is the logarithm of the daily average concentration; $Gasoline4_{cd}$ is a dummy variable indicating whether city c has upgraded to gasoline standard IV by day d; λ_s is the set of station fixed effects, controlling for all time-invariant differences among the monitoring stations within a city, including topographic features; λ_d is the set of day fixed effects, controlling for the daily shocks common to all cities (e.g., monetary policy and exchange rate changes);

---

[10] Source: http://www.sdpc.gov.cn/zcfb/zcfbtz (in Chinese).

[11] The MEP directly operates 1492 state-level monitoring stations nationwide, which ensures that the data collected are independent of local government oversight, thus preventing interference and helping to provide better accuracy and authenticity. Fig. A1 in the appendix verifies that there is no clear bunching at important thresholds in the data used, further alleviating concerns about data manipulation.

Table 3
Variable definitions and descriptive statistics.

Variable	Definition	Mean	SD
Fuel Standard Regulations and Prices			
Gasoline IV	= 1 if a prefecture has upgraded its gasoline content regulation from standard III to standard IV; = 0 otherwise.	0.714	0.452
Gasoline V	= 1 if a prefecture has upgraded its gasoline content regulation from standard IV to standard V; = 0 otherwise.	0.148	0.355
Diesel IV	= 1 if a prefecture has upgraded its diesel content regulation from standard III to standard IV; = 0 otherwise.	0.522	0.500
Diesel V	= 1 if a prefecture has upgraded its diesel content regulation from standard IV to standard V; = 0 otherwise.	0.069	0.253
Fuel_Price	Average retail price of gasoline and diesel (thousand yuan/ton)	8.410	0.972
Vehicle Emission Regulations			
Gasoline_Car IV	= 1 if a prefecture has upgraded its gasoline vehicle emission regulation from standard III to standard IV; = 0 otherwise.	0.898	0.303
Yellow_Car	= 1 if a prefecture has started to phase out yellow label vehicles; = 0 otherwise.	0.613	0.487
Air Pollutant Variables			
AQI	Air Quality Index	87.732	57.002
PM2.5	$PM_{2.5}$ concentration (μg/m3)	58.764	49.085
PM10	PM_{10} concentration (μg/m3)	99.344	74.952
O3	O_3 concentration (μg/m3)	55.347	35.825
Weather Variables			
Sunshine_time	Prefecture daily average sunshine hours (0.1 h)	54.454	40.538
Temperature	Prefecture daily average temperature (0.1°)	149.545	105.889
Wind_Speed	Prefecture daily average wind speed (0.1 m/s)	22.132	27.349
Rainfall	Prefecture daily average rainfall (0.1 mm)	56.214	186.933
Humidity	Prefecture daily average humidity (1%)	69.117	17.449
Gov environmental_effort	Average count of "green" keywords in the government work reports for 2001 to 2012 divided by total word count ($\times 10^3$)	0.271	0.087
Car Model Sales Variables			
Total_car sales	Prefecture monthly total sale of car models (thousands)	4.692	6.621
Horsepower	Prefecture average horsepower of cars sold (kw)	81.318	5.462
Fuel consumption	Prefecture average of the nominal fuel consumption of the cars sold (litres/100 km)	7.330	0.343
Displacement	Prefecture average displacement of cars sold (litres)	1.513	0.115
Curb weight	Prefecture average of the curb weight of the cars sold (1000 kg)	1.298	0.129
Other Variables			
Winter heating	= 1 if a prefecture had been experiencing the winter heating period; = 0 otherwise.	0.172	0.377

Notes: This table's data sources are described in full in Section 3.1. Fuel standard regulations, vehicle emission regulations, fuel prices and weather variables are all at the prefectural level. The "green" keywords include "environmental protection" (huan-jing-bao-hu or huan-bao). The air pollutant variables are by monitoring station.

and ε_{iscd} is the error term. The standard errors are clustered by prefecture.

$\mathbf{LPolicy}_{cd} = \{Gasoline5_{cd}, Diesel4_{cd}, Diesel5_{cd}\}$, where $Gasoline5_{cd}$ is a dummy variable indicating whether city c had upgraded to gasoline standard V by day d. Similarly, $Diesel4_{cd}$ is a dummy variable indicating whether city c's diesel standard had progressed from III to IV; and $Diesel5_{cd}$ is a dummy variable indicating whether city c had upgraded its diesel standard from IV to V.

The parameter of interest is β, which is expected to be negative, reflecting the effect of better quality gasoline on air pollution. To capture the overall effect of the gasoline standard IV reform, the study follows the approach of Duflo et al. (2013) and uses stacked data. Specifically, the four pollutants are pooled together along with the pollutant dummies λ_i, which control for any time-invariant pollutant heterogeneity.

The identification exploits daily variations among the cities, so any potential bias could only have arisen from omitted variables on the day level. One primary threat is seasonality. Specifically, if the implementation of new gasoline and diesel content regulations corresponded with specific weather conditions, the estimates could be mistakenly attributed to policy effects. While the day fixed effect has effectively controlled for all national average seasonality, station-specific seasonality is addressed by including three sets of controls in $\mathbf{X}_{scd}$: $\lambda_s \times \mathbf{Day\ of\ Week}_d$, where $\mathbf{Day\ of\ Week}_d = \{Monday, Tuesday, \ldots, Sunday\}$ and $\lambda_s \times \mathbf{Week\ of\ Month}_d$, where $\mathbf{Week\ of\ Month}_d = \{Week1, \ldots, Week5\}$; and $\lambda_s \times \mathbf{Month\ of\ Year}_d$, where $\mathbf{Month\ of\ Year}_d = \{January, \ldots, December\}$. As weather conditions are well known to significantly influence pollution levels, concerns about prefecture-specific weather conditions are further dealt with by adding a series of weather variables in $\mathbf{X}_{scd}$: $Temperature_{cd}$, $Rainfall_{cd}$, $Wind\ Speed_{cd}$, $Sunshine\ Time_{cd}$ and $Humidity_{cd}$. Finally, as fuel prices changed a number of times during the period studied, a $FuelPrice_{cd}$ term is included in X_{scd} to isolate the effects of the regulatory changes.

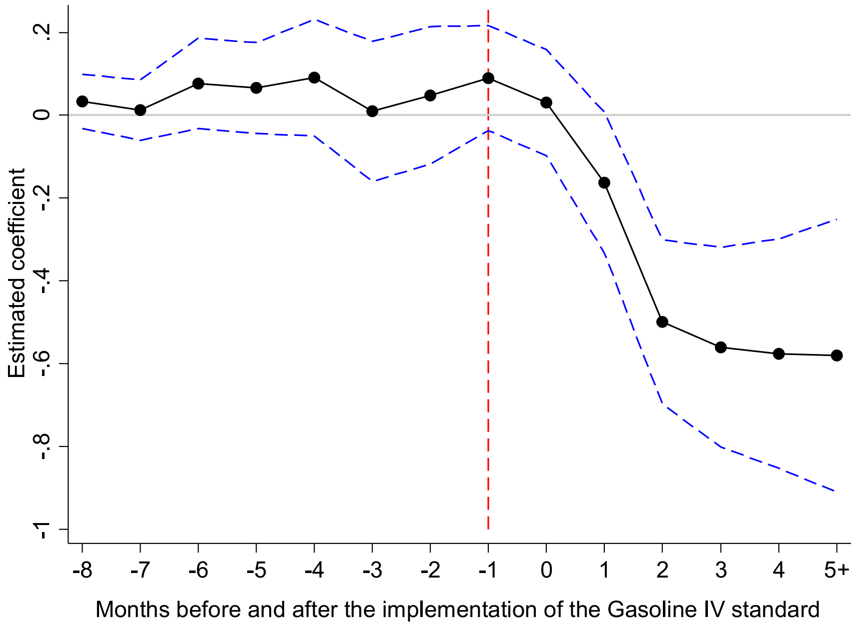
A second threat to the identification is that if there were other ongoing reforms around the time of the gasoline standard IV reform the estimates might also reflect those confounding factors. To address this concern, government documents setting out environmental protection policies around the time of the gasoline standard IV reform are examined carefully. A prominent policy change was a change in vehicle emission regulations, with two measures targeting new and old vehicles separately. Exhaust emission standard IV was imposed for new gasoline vehicles. For old vehicles, China adopted one of the world's most ambitious voluntary scrapping programs aimed at gasoline vehicles not meeting Chinese gasoline standard I and diesel vehicles not meeting Chinese diesel standard III (termed "yellow label vehicles"). Two additional controls were included to remove any confounding effect of those changes. $GasolineCar4_{cd}$ indicates whether city c had adopted exhaust emission standard IV for gasoline vehicles by day d, and $YellowCar_{cd}$ indicates whether city c had started to phase out yellow label vehicles.

Third, provinces differed how promptly they implemented gasoline standard IV. As described in Section 2.2, one important consideration was to limit exhaust emissions, so the reform's timing could have been related to historical pollution levels and their growth. This concern is addressed by applying the approach proposed by Gentzkow (2006) and explicitly controlling for the trends in air pollution caused by these pre-existing covariates to isolate the treatment effect. Specifically, the average air pollution value during the two months prior to the policy's implementation and changes in the pollution level during those two

Table 4
Effects of fuel standard regulations on air quality (DD estimates).

	Log(daily stacked average pollutant concentration)	
	(1)	(2)
Gasoline IV	−0.040**	−0.129**
	(0.018)	(0.053)
Station dummies	Yes	Yes
Day dummies	Yes	Yes
Station dummies × Day of week dummies	Yes	Yes
Station dummies × Week of month dummies	Yes	Yes
Station dummies × Month of year dummies	Yes	Yes
Weather condition controls	Yes	Yes
Gasoline V	Yes	Yes
Diesel IV	Yes	Yes
Diesel V	Yes	Yes
Gasoline_Car IV	Yes	Yes
Yellow_Car	Yes	Yes
Fuel_Price	Yes	Yes
Pollutant dummies	Yes	Yes
Time polynomial interactions		Yes
Adjusted R^2	0.389	0.407
No. of observations	36,37,066	19,93,591

Notes: This table presents estimates of DD regressions of the logarithm of daily average stacked pollutant concentrations on the implementation of the gasoline IV standard and other control variables. The weather condition controls are the daily average temperature, rainfall, humidity, sunshine time and wind speed for each prefecture. All of the specifications control for any station-specific seasonality and the indicators of the gasoline V standard, the diesel IV standard, the diesel V standard, exhaust emission standard IV for gasoline vehicles, the voluntary scrapping programs, as well as fuel prices. Column 2 additionally includes a fourth-order polynomial in time interacted with pretreatment pollution levels and their growth. Reported in parentheses are robust standard errors clustered by prefecture. ***, ** and * represent significance at the 1%, 5% and 10% levels of confidence, respectively.



Notes: The figure plots the estimated coefficients of a series of β_t and their 95% level confidence intervals for each month within the event window with the same set of controls as in equation (1). The omitted time category is more than 8 months before gasoline standard IV's adoption.

Fig. 3. Tests for parallel trends.

months are interacted with a fourth-order polynomial in time respectively and then included in the main specification.

To further verify the validity of the DD's identification assumptions, a series of robustness checks are performed. They include examining parallel pre-treatment trends, including an additional control variable representing the local government's environmental-improvement efforts, a flexible event study analysis (see Deshpande and Li, 2019), and an instrumental variable (IV) analysis following the lead of Freyaldenhoven et al. (2019).

Table 5
Effects of fuel standard regulations on air quality: Robustness tests of DD.

	Log(daily stacked average pollutant concentration)		
	With additional controls	Flexible DD	IV
	(1)	(2)	(3)
Gasoline IV	−0.131**	−0.083**	−0.354***
	(0.053)	(0.034)	(0.024)
Station dummies	Yes	Yes	Yes
Day dummies	Yes	Yes	Yes
Station dummies × Day of week dummies	Yes	Yes	Yes
Station dummies × Week of month dummies	Yes	Yes	Yes
Station dummies × Month of year dummies	Yes	Yes	Yes
Weather condition controls	Yes	Yes	Yes
Gasoline V	Yes	Yes	Yes
Diesel IV	Yes	Yes	Yes
Diesel V	Yes	Yes	Yes
Gasoline_Car IV	Yes	Yes	Yes
Yellow_Car	Yes	Yes	Yes
Fuel_Price	Yes	Yes	Yes
Pollutant dummies	Yes	Yes	Yes
Time polynomial interactions	Yes	Yes	Yes
Government_effort × Day dummies	Yes		
Winter heating			Yes
Adjusted R^2	0.412	0.463	
No. of observations	19,93,591	9,93,464	19,93,591

Notes: This table presents estimates of DD robustness tests of the logarithm of daily average stacked pollutant concentrations on implementation of the gasoline IV standard and other control variables. The dependent variables are the logarithms of the stacked pollutants. The weather condition controls are the daily average temperature, rainfall, humidity, sunshine time and wind speed for each prefecture. All of the specifications control for any station-specific seasonality and the indicators of the gasoline V standard, the diesel IV standard, the diesel V standard, exhaust emission standard IV for gasoline vehicles, the voluntary scrapping programs, fuel prices as well as time polynomial interactions. Column 1 additionally includes the interaction of the local government's overall environmental efforts with the day dummies. Column 2 reports a flexible event study. Column 3 exploits covariates related to the policy only through the potential confounding factors. The 2SLS estimator is reported using the one and two months leads of gasoline standard IV as the instruments for the winter heating period in each city. Reported in parentheses are robust standard errors clustered by prefecture. ***, ** and * represent significance at the 1%, 5% and 10% levels of confidence, respectively.

3.2.2. RD framework

The implementation of the gasoline standard IV can be taken as defining a discontinuity in time. The daily data and that discontinuity allow conducting an RD estimation as well (e.g., Davis, 2008; Auffhammer and Kellogg, 2011). The RD analysis focuses on a narrow window around the policy change in which unobservables are allowed to act non-linearly so long as they change smoothly at the time of the reform.

One problem with using discontinuity in time is that when the treatment effects vary over time, the global polynomial estimation of the RD may suffer from overfitting, generating a significant estimation bias in the Monte Carlo simulations. In light of potential time-varying treatment effects in this research setting, augmented local linear RD estimation is applied as suggested by Hausman and Rapson (2018). This approach allows controlling for un-smooth factors such as seasonality and wealth conditions in a narrow window around the policy change. Specifically, the estimation involves two steps. First, the outcome y_{iscd} is regressed against a set of controls including station fixed effects, station-specific seasonality (i.e., $\lambda_s \times DayofWeek_d$, $\lambda_s \times WeekofMonth_d$, $\lambda_s \times MonthofYear_d$), prefecture-specific weather conditions and fuel prices to obtain a residual $\widetilde{y}_{scd}$. In the second step, the residual $\widetilde{y}_{scd}$ is used in the nonparametric estimation (2) to obtain the parameter of interest $\widehat{\beta}_{RD}$.

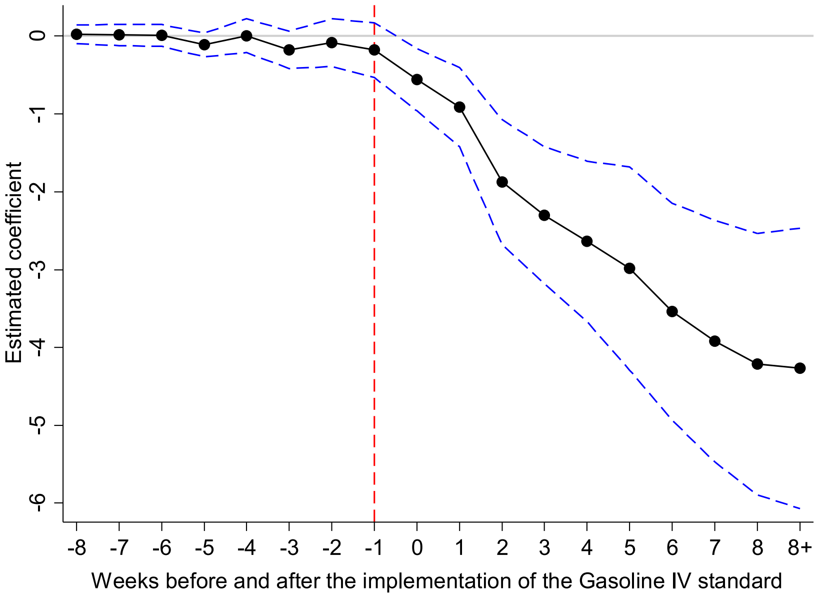
$$\min_{\alpha,\beta,\delta,\tau} \sum_{s=1}^{N} K\left(\frac{d_c - d_{c0}}{h}\right)\left[\widetilde{y}_{scd} - \delta - \tau\left(d_c - d_{c0}\right) - \beta E_c - \alpha E_c\left(d_c - d_{c0}\right)\right]^2, \tag{2}$$

where E_c indicates the implementation of the gasoline standard IV in city c (i.e., taking a value of 1 if $d_c \geq d_{c0}$ and 0 otherwise); h is the bandwidth; and $K(.)$ is a rectangular kernel function.

We calculate the optimal bandwidth h using the method developed by Imbens and Kalyanaraman (2012). To check whether the results are sensitive to the optimal bandwidth selected, alternative bandwidths are tested (see, e.g., Carneiro et al., 2015 for details). To account for any serial correlation within a prefecture over time, we estimate standard errors clustered by prefecture.

Hausman and Rapson (2018) also point out that any serial correlation in daily pollutant concentrations could also obscure the RD estimates. Specifically, if the pollutants take time to dissipate, the RD estimates might combine both short-run and long-run policy influences. To understand how strongly the study's outcome variable is serially correlated, an autoregressive AR(1) regression is conducted. The results are reported in column 1 of Table A1 in the appendix. The estimate of AR(1) is 0.681 and statistically significant, confirming strong autocorrelation in the outcome variable. That requires including a lagged outcome variable in the second step of the augmented local linear RD estimations. Specifically, the coefficient of the lagged outcome is denoted as α. $\widehat{\beta}_{RD}$ indicates the short-term effect of the policy, and its long-term effect of the policy is then $\widehat{\beta}_{RD}/\left(1-\widehat{\alpha}\right)$.

To further verify the validity of the RD estimates, several more robustness checks are performed following the suggestions of Hausman and Rapson (2018). They include testing alternative local linear bandwidths, evaluating parallel RDs using placebo dates, RDs on the control variables, and testing a donut RD specification (see also Barreca et al., 2011).



Notes: This figure plots the weekly coefficients and their 95% level confidence intervals for each week within the eight weeks before and the eight weeks after the implementation of the gasoline IV regulations from a flexible event study. A separate dataset is created for each wave of the gasoline standard IV adoption. In each dataset, cities which had adopted the new standard are regarded as the treatment group, and those which did not within the subsequent two months form the control group. All the datasets are then appended into one dataset and the analyses of the parallel time trends are re-conducted. The omitted time category is more than 8 weeks before gasoline standard IV's adoption.

Fig. 4. Tests for parallel trends (Flexible DD).

Table 6
Effects of fuel standard regulations on air quality (RD estimates).

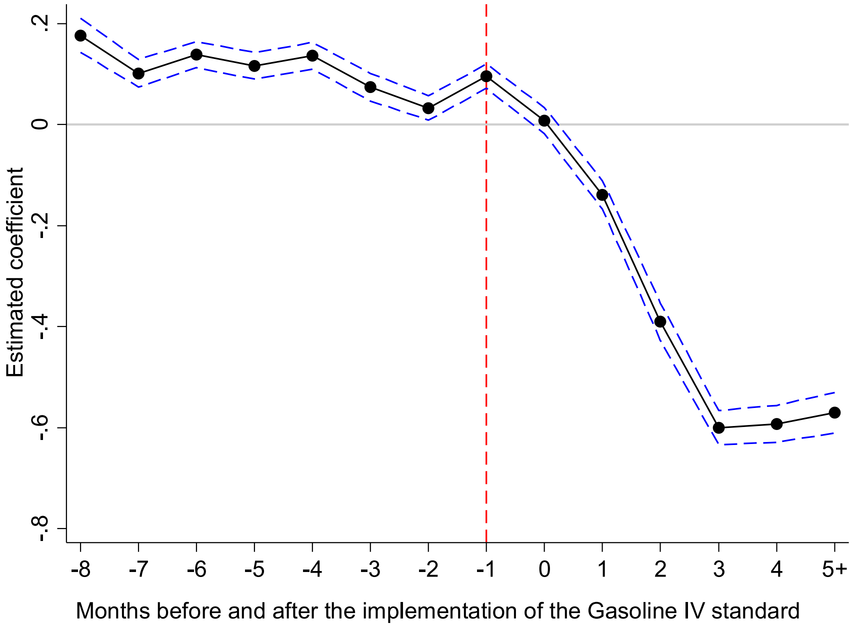
	Residualized log(daily stacked average pollutant concentration)					
	Baseline Estimates	Placebo		Donut RD dropping		
		Half a year earlier	One year later	1 day	1 day to 7 days	1 day to 21 days
	(1)	(2)	(3)	(4)	(5)	(6)
Gasoline IV	−0.066***	−0.001	0.006	−0.067***	−0.067***	−0.057***
	(0.015)	(0.018)	(0.012)	(0.015)	(0.015)	(0.022)
Day	−0.001**	0.000	−0.001***	0.001**	−0.001***	−0.001*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Gasoline IV × Day	0.001**	−0.0004	0.001***	0.001**	0.001**	0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log (lagged pollution)	0.533***	0.395***	0.466***	0.532***	0.524***	0.492***
	(0.013)	(0.017)	(0.010)	(0.013)	(0.013)	(0.013)
Optimal bandwidth	117	44	136	117	117	117
Control group mean	0.036	−0.013	−0.024	0.036	0.034	0.035
R^2	0.438	0.247	0.338	0.438	0.427	0.386
No. of observations	5,47,060	1,25,570	10,28,942	5,45,086	5,19,017	4,56,354

Notes: The dependent variables are residualized logarithms of the stacked pollutant concentrations. Column 1 reports the augmented local linear RD estimation. Columns 2 and 3 report the placebo estimates assuming the date of the policy's implementation to be a half-year earlier and one year later than the actual date. Columns 4–6 report the coefficients of a donut specification excluding the observations around the date of the policy's implementation. The method developed by Imbens and Kalyanaraman (2012) is used to calculate the optimal bandwidth. Reported in parentheses are robust standard errors clustered by prefecture. ***, ** and * represent significance at the 1%, 5% and 10% levels of confidence, respectively.

4. Findings

4.1. DD estimates

Table 4 reports the results from the DD estimation of equation (1) showing the effect of fuel standard regulations on air pollution from the stacked data. Specifically, column 1 includes controls for station fixed effects, day fixed effects, station-specific monthly seasonality, weather conditions, and other ongoing fuel standards and vehicle emission regulations. In column 2 further controls for the historical pollution level and its growth are added. The coefficient of the $Gasoline4_{cd}$ term is negative and precisely estimated, consistent with the findings of previous studies (e.g., Auffhammer and Kellogg, 2011). In terms of economic magnitude, the baseline estimate in column 2 indicates that across the four measures of air pollution, the gasoline IV standard on average reduced the pollutant concentration by 12.9%, leading to an aver-



Notes: This figure plots the 2SLS estimates of a series of β_ts as well as their 95% level confidence intervals for every one month within the event window. This approach identifies the causal effect of the policy by exploiting covariates related to the policy only through confounders. Specifically, the one and two months leads of the gasoline standard IV are used as the instruments for the winter heating period in each city. The omitted time category is more than 8 months before gasoline standard IV's adoption.

Fig. 5. Tests for parallel trends (IV).

age fall of 13.4 unit in *AQI*, 9.4 μg/m^3 in the concentration of $PM_{2.5}$, 16 μg/m^3 in the concentration of PM_{10} and 7 μg/m^3 in the concentration of O_3.

Parallel Time Trends. Despite the exhaustive set of controls included in the analyses, the comparability between the treatment and control groups, which is central to DD estimation might still be a concern. One validity check commonly used is to examine whether the treatment and control groups exhibit parallel pre-treatment trends. To this end, pre-treatment data from about 400 monitoring stations for 8 months leading up to the implementation of the gasoline IV standard are used. The post-treatment data up to 5 months after the implementation of the gasoline IV standard (the time around which most cities started to further upgrade their standards to gasoline V) are also explored. To provide a smooth graphic presentation, a series of β_t are estimated for each month within the event window with the same set of controls as in equation (1). The omitted time category is more than 8 months before the policy's adoption.

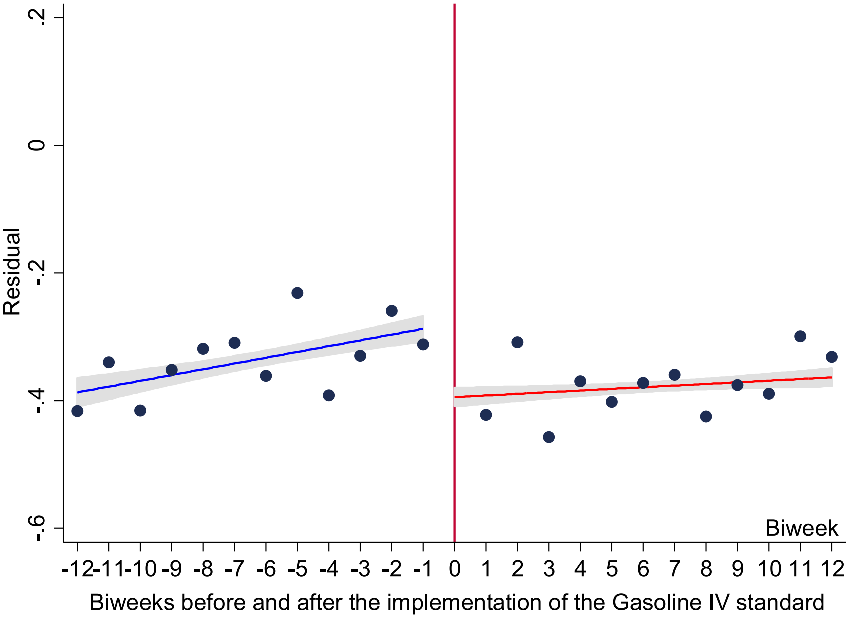
The estimated coefficients along with the 95% confidence intervals are presented in Fig. 3. There is a fairly flat trend between the treatment and control groups before the policy adoption. That inspires confidence that the control group cities provide a good counterfactual for the treatment group in the period studied. Meanwhile, there is a gradual and significant reduction in air pollution within the two months after the policy's implementation. This appears sensible, because gasoline might not turn over immediately when fuel standards change. And even if all retailers instantly switched to supplying new gasoline, some drivers would not refill for some weeks or months.

Controlling for the environmental protection efforts of the local governments. With such a vast territory, China exhibits significant regional differences in local governance. Local authorities were deeply involved in implementing the fuel standard. That raises the possibility that the governments' different speeds in implementing gasoline standard IV could be related to their willingness to protect the environment. If so, that would confound the estimates. To alleviate this concern, each local government's overall environmental efforts and regulatory attitudes are investigated by conducting a thorough content analysis of each city government's annual work report from 2001 to 2012. The report is usually presented by the city's mayor at a Communist Party meeting early in the year. It sets goals and objectives for the city government's work for the upcoming year, and the goals' accomplishment is meant to be supervised by local residents. Manually searching China Statistical Yearbooks, city government websites and local newspapers yields 3432 reports from 286 prefecture-level cities. The key words used are "environmental protection" (huan bao or huanjing baohu in Chinese). The number of times these terms were mentioned is normalized as a fraction of the total number of words in the report. A larger fraction is assumed to indicate a local government more active in enforcing environmental regulations.

That measure's interactions with the day dummies λ_d are included as additional controls in the analyses. The estimation results are reported in column 1 of Table 5. The estimated coefficient barely changes, suggesting that the results are not biased due to the environmental protection behavior of local governments.

Flexible event study. The DD analyses exploit variations in treatment timings, as all regions had adopted gasoline standard IV by the end of the period studied. As Goodman-Bacon (2019) has shown, this involves comparing early adopters with later ones. Time-varying treatment effects may then cause estimation bias depending on the weights of each group. To alleviate this concern, a flexible event study is also conducted following the lead of Deshpande and Li (2019). A separate dataset is created for each wave of the gasoline standard IV adoption. In each dataset, cities which had adopted the new standard are regarded as the treatment group, and those which did not within the subsequent two months form the control group. All the datasets are then appended into one dataset and the analyses of the parallel time trends and the DD regressions are re-conducted.

For the parallel time trend analysis, coefficients are evaluated for each week within the eight weeks before and the eight weeks after



Notes: This figure plots the relationship between a normalized time variable based on the date of the gasoline IV standard's implementation and the residual of the pollutant measures using a bin width of two weeks. The estimation method is described in Section 3.2.2. The circles in the figures represent mean values for each two-week bin; the lines have been fitted using local linear regression with the optimal bandwidth calculated using the method of Imbens and Kalyanaraman (2012). The grey areas are the 95% level confidence intervals, and the vertical line is the cutoff point for the assignment variable.

Fig. 6. Effects of fuel standards on air quality (RD Estimates).

the implementation of the gasoline IV regulations. Fig. 4 presents the estimated coefficients and their 95% confidence intervals. In the pre-treatment period there was little difference between the treatment and control groups, with the estimated coefficients close to zero. That result lends further support to using this estimation strategy, and mitigates the concerns about time-varying treatment effects in the DD setting with multiple treatment timings. In the eight weeks after the adoption of the new gasoline content regulations there is a continuous decline, which is consistent with the findings reported in Fig. 3.

Regression results using that appended dataset are presented in column 2 of Table 5. The estimated coefficients remain negative and statistically significant. Their magnitude is smaller than in the baseline estimate in column 2 of Table 2, primarily due to the short term of the effect (8 weeks).

IV estimation. In a recent study, Freyaldenhoven et al. (2019) argue that even with some pre-trends in the DD analysis, researchers can still use additional information to correct for any potential bias from the pre-existing differences between the treatment and control groups. Specifically, they propose to use covariates that are not affected by the focal policy but by confounders. Those covariates can then help condition out the pre-trends caused by confounders between the treatment and control groups. This is implemented in an IV estimation. That methodology is applied in this study. During the winter, northern China switches on a centralised heating system to provide heat to local homes, which is mostly coal-based. Winter heating is well known to significantly increase air pollution (e.g., Almond et al., 2009; Chen et al., 2013; Ebenstein et al., 2017). However, the centralised heating policy is certainly not affected by the upgrading of fuel standards. Hence, following the lead of Freyaldenhoven et al. (2019), an indicator for the winter heating period in each city $WinterHeating_{cd}$ is added into the DD specification (1), using one and two months leads of the enforcement of gasoline standard IV as its instruments.

The parallel time trend is presented in Fig. 5. The pattern is similar to that of Fig. 3. There were some fluctuations, but no clear signs of a deterministic trend in the air pollution before the new gasoline IV standard took effect. The trend starts to decline within three months after the policy's enforcement and then stabilizes. The IV regression results are reported in column 3 of Table 5. The estimated coefficient of interest is still negative and statistically significant. The magnitude is larger than in the baseline estimate, indicating that any potential bias from the pre-existing confounders tended to underestimate the effect.

4.2. RD estimates

Consider first the visual relationship between a normalized time variable (the assignment variable; $\widetilde{d}_c = d_c - d_{c0}$) and the residualized outcome. Imbens and Lemieux (2008) and Lee and Lemieux (2010) suggest that the selection of bin size trades off precision in calculating average outcome values against proximity to the cutoff point. Fig. 6 reports the RD figure using a bin width of two weeks. Specifically, the sample is collapsed into bi-weekly bins and the residuals of the nonparametric estimation as described in Section 3.2.2 are used. The circles in the figure represent mean values for each two-week bin; the lines have been fitted through a local linear regression with the optimal bandwidth calculated using the method of Imbens and Kalyanaraman (2012). The grey areas are the 95% confidence intervals, and the vertical line is the cutoff point for the assignment variable. With this bin width, the plots are smooth on either side of the cutoff value, but at the same time show a clear drop in pollutants at the cutoff point. That further confirms the findings from the DD estimation. Better quality gasoline does indeed reduce air pollution.

The regression results using the stacked daily data are reported in column 1 of Table 6. The estimate of β, the parameter of interest, is negative and statistically significant, further confirming the pattern of Fig. 6 and the DD estimates in Table 4. In terms of magnitude, the short-term effect is 6.6%, and the long-term effect is 14.1% $(= 0.066/(1 - 0.533))$. Note that the DD baseline estimate is 12.9%. Given the gradual turnover and refilling of gasoline, it is reasonable that the short term effect implied by the RD estimate should be smaller than that from the DD estimate which essentially captures the average effect in the post-treatment period. Meanwhile, the long-term effect implied by the RD estimate aligns well with the DD estimate. The DD and RD estimations use different control groups and different identifying assumptions, but the consistent results from the two estimation techniques lend support to the conclusion that the new gasoline content regulations achieved their primary goal of reducing air pollution.

Alternative bandwidth. The RD technique's local linear estimation requires the calculation of an optimal bandwidth. The method developed by Imbens and Kalyanaraman (2012) is used. To check whether the findings are sensitive to the optimal bandwidth selected, alternative bandwidths from $h^* - 10$ to $h^* + 10$ in intervals of 2 are tested. Estimates using those alternative bandwidths are plotted in Fig. 7. The estimates remain stable, suggesting that the results are not driven by a particular bandwidth.

Placebo test using alternative dates. To verify the RD specification, Hausman and Rapson (2018) suggest using "fake" treatment dates. Two such placebo tests are conducted. In the first the date of the policy's implementation is assumed to be a half-year earlier than the actual date; in the second placebo it is assumed to be one year later than the actual timing. Given that there were no real policy changes at those two placebo dates, the main outcome should be smooth across the cutoff. Estimation results are reported in columns 2 and 3 of Table 6.[12] Neither estimated coefficient of the *Gasoline IV* term is statistically significant and the magnitudes are close to 0, again lending support to the RD specification.

Donut specification. An RD involving time may be confounded by the effects of sorting and anticipation. Knowing the policy is to be implemented, individuals may react in advance to take advantage of it. To address such concerns, a donut specification is evaluated following the suggestion of Hausman and Rapson (2018) and applying Barreca's methods (Barreca et al., 2011). Specifically, there are three exercises which exclude the observations at the date of the policy's implementation, the observations within 7 days around the implementation date, and the observations within 21 days around the date. The estimation results are reported in columns 4–6 of Table 6. The key coefficients remain negative and statistically significant. The magnitudes barely change, suggesting that the RD specification is not contaminated by any anticipation effect.

RDs on the controls. Another robustness check is to examine the smoothness of control variables such as weather conditions. As the policy only affects air pollution, weather variables should be smooth at the policy implementation date. The coefficients of AR(1) regressions using weather as the outcome are reported in Table A2 in the appendix. There is significant autocorrelation in the weather data, so an augmented local linear RD estimation is evaluated with lagged outcome as an additional control variable. The results are reported in Table 7. Rain, humidity, suntime, wind and temperature are examined separately. None of them exhibit any statistically significant discontinuity at the cutoff point, and the estimated magnitudes are all close to zero. The confirmation of smooth weather conditions across the cutoff further reassures about the validity of the RD specification.

Price increase. To reflect the costs of fuel upgrading, wholesale fuels prices were increased by the NDRC when the new fuel was introduced. This raises the possibility that the discontinuity in fuel prices might itself constitute an omitted variable in the RD estimation. For example, vehicle owners might have driven less in response to the price increase, generating less air pollution. However, the price increased by only about 3% from gasoline standard III to IV. Such a small change may not have had a significant impact on driving behavior. Nonetheless, to check for any potential bias, fuel prices are included in the RD estimations. Also, monthly new vehicle registration records by prefecture are collected spanning 2013 to 2015 (because there is no comprehensive data on daily driving and fuel consumption). Whether the fuel upgrade had any effect on the attributes of the cars purchased is then assessed to shed light on the regulation's effects on drivers' preferences about driving and fuel consumption.

As Table 8 shows, no significant relationship is found between the new standard and total car purchases or and of the attributes studied. These results are reassuring because they confirm that reduced driving as a result of the price changes is not important in the context studied.

4.3. Heterogeneous effects

The DD and RD estimations will now be used to discuss some temporal and spatial differences in the results.

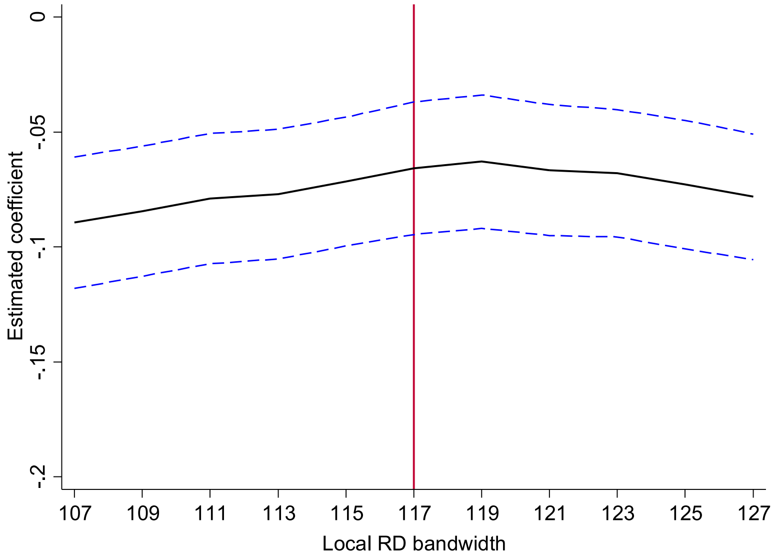
Temporal heterogeneity. Hourly data on air pollution allow examining intra-day fluctuations. To do so, each day is divided into eight time periods: 12–2 am, 3–5 am, …, 9–11 p.m. Fig. 8 reports estimated coefficients for each time period separately. The pollution reduction effects of the fuel content regulations persist throughout the day. There are, however, some interesting intra-day fluctuations. The effect of fuel standards on ambient concentrations seems to follow an inverse U-shaped curve. Pollution mitigation starts to decline after daybreak and is at its weakest in the evening hours. The effect picks up again later on. The regulations are most effective in reducing particulates during the early morning and late evening hours. This is probably associated with daily variations in the depth of the boundary layer and with other anthropogenic emissions (Zhang and Cao, 2015). Taken together, the hourly estimates are consistent with the atmospheric chemistry of the pollutants. This further validates the research design and gives more credibility to the baseline results.

Spatial heterogeneity. Under a regionally decentralized authoritarianism regime like China's one might anticipate the role of government in tackling environmental issues to be especially prominent (Xu, 2011; Coxhead, 2019). With a combination of political centralization and economic (administrative) regional decentralization, China exhibits significant regional differences in local governance. While implementing the new fuel standards, cities with decentralized authority oversaw monitoring and enforcement of violations and other activities. So the local institutional environment could closely interact with various stakeholders in shaping environmental outcomes.

The theoretical arguments on the role of local government can work in both directions. One line emphasizes an essential role for the visible hand of government in providing decentralized solutions. Proactive governments could devote more resources to improve regulatory capacity. They work closely with various parties including local refineries, retail stations, transportation bureaus and drivers to enhance compliance. In such circumstances a larger policy impact should be observed in areas with a more environmentally-active government. The other line instead implies that good local governance might in part substitute for environmental standards. For example, to the extent that there is already stringent management of anti-pollution measures, the policy effects of fuel standards might be smaller in more environmentally-active prefectures. Equivalently, the marginal environmental return to using high quality fuel might be less in areas with better governance.

---

[12] The AR(1) regressions using these two placebo tests are reported in columns 2 and 3 of Table A1 in the appendix. Both show strong autocorrelation in the outcome variable.



Notes: This figure plots β_{RD} coefficients estimated using bandwidths from h*-10 to h*+10 in intervals of 2 for a set of pollutant measures. The vertical line is the optimal bandwidth h* calculated using the method of Imbens and Kalyanaraman (2012). The dashed lines show the 95% level confidence intervals on each coefficient. The estimation method is described in full in Section 3.2.2.

Fig. 7. Sensitivity test on the choice of bandwidth.

Table 7
Effects of Fuel Standard Regulation on Weather Conditions (RD estimates).

	Rain	Humidity	Suntime	Wind	Temperature
	(1)	(2)	(3)	(4)	(5)
Gasoline IV	0.110	−0.005	−0.042	−0.005	−0.005
	(0.078)	(0.005)	(0.038)	(0.010)	(0.008)
Day	−0.001	0.000	0.000	0.000	0.000
	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)
Gasoline IV × Day	−0.007***	0.000	0.000	0.000	0.000
	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)
Log (lagged pollution)	0.140***	0.003	0.001	0.001	−0.009
	(0.010)	(0.006)	(0.008)	(0.008)	(0.006)
Optimal bandwidth	61	217	402	143	212
Control group mean	0.099	0.001	0.004	−0.011	0.001
R^2	0.035	0.001	0.000*	0.001	0.001
No. of observations	2,94,904	10,15,436	17,96,100	6,80,376	8,79,860

Notes: This table presents estimates of RD regressions of the residualized logarithms of weather conditions against implementation of the gasoline IV standard. All specifications use the augmented local linear RD estimation with the lagged outcome as an additional control variable. The method developed by Imbens and Kalyanaraman (2012) is used to calculate the optimal bandwidth. Reported in parentheses are robust standard errors clustered by prefecture. ***, ** and * represent significance at the 1%, 5% and 10% levels of confidence, respectively.

Motivated by these reasonings, we investigate to what extent local government's overall environmental efforts and regulatory attitudes influenced the environmental outcomes from the policy. Separate policy effects are predicted by estimating a separate RD equation for each station (cf. Auffhammer and Kellogg, 2011). To provide a smooth illustration, the government effort index is divided into 30 bins and the average value of the policy effect is calculated for each bin. The policy effect coefficients (short-run and long-run effects) are plotted against the measure of government efforts in Fig. 9a and b.

Cities whose work reports more frequently mentioned environmental protection exhibited larger effects of the higher-quality fuels. Reassuringly, the pattern holds for both the short run and long-run effects. These results point to the importance of local governance in realizing the benefits of environmental initiatives. The within-country exercise also helps enrich scholarly understanding of the different environmental policy outcomes often observed in developed and developing countries. It seems that they can at least partially be attributed to institutional gaps.

Following a similar approach, Fig. A2 displays how other spatial features influenced the effect. Larger effects are found in more mountainous areas, and in prefectures with a greater density of population, and car ownership. These patterns are consistent with the science of pollutant formation, and corroborate findings reported from the US (Auffhammer and Kellogg, 2011).

4.4. *Cost-benefit analyses*

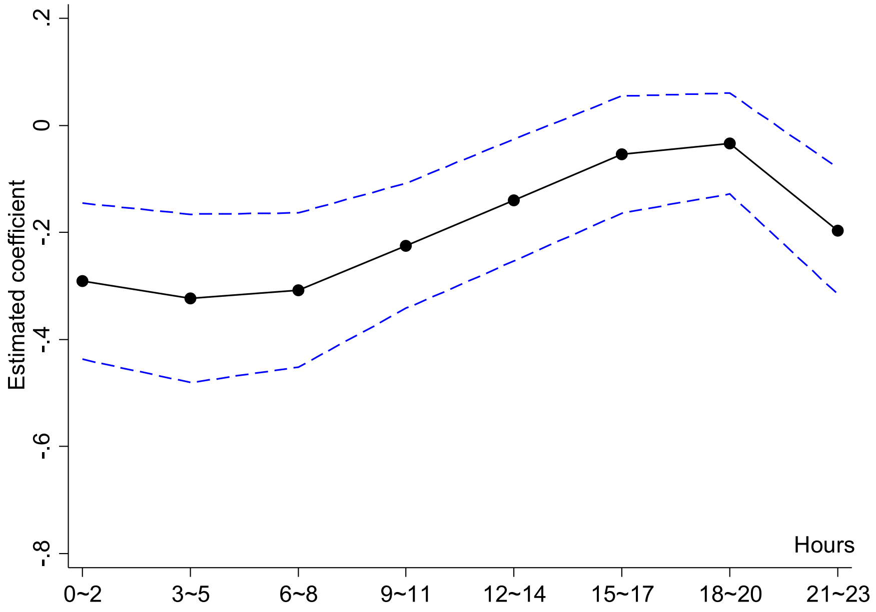
We use the estimates to conduct “back of the envelope” benefit-cost analyses for the fuel quality upgrading.

Benefits. The standards' primary benefits are assumed to be health improvements associated with air pollutant reductions, especially reductions in particulate matter (Wolff, 2014). There is evidence from the European Union that *PM* is the most lethal air pollutant, with an impact much greater than that of the second most deadly air pollutant, ozone (Watkiss et al., 2005). The analysis therefore focuses on inferring

Table 8
Effects of fuel standard regulations on car model sales.

	Log(Total_CarSales)	Log(Horsepower)	Log(OilConsumption)	Log(Displacement)	Log(CurbWeight)
	(1)	(2)	(3)	(4)	(5)
Gasoline IV	0.013 (0.024)	0.002 (0.002)	0.001 (0.002)	0.005 (0.003)	0.001 (0.005)
Prefecture dummies	Yes	Yes	Yes	Yes	Yes
Month dummies	Yes	Yes	Yes	Yes	Yes
Prefecture dummies × Year dummies	Yes	Yes	Yes	Yes	Yes
Weather condition controls	Yes	Yes	Yes	Yes	Yes
Gasoline V	Yes	Yes	Yes	Yes	Yes
Diesel IV	Yes	Yes	Yes	Yes	Yes
Diesel V	Yes	Yes	Yes	Yes	Yes
Gasoline_Car IV	Yes	Yes	Yes	Yes	Yes
Yellow_Car	Yes	Yes	Yes	Yes	Yes
Fuel_Price	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.976	0.906	0.835	0.648	0.335
No. of observations	9241	9241	9241	9241	9241

Notes: This table presents estimates of DD regressions of major attributes of car models sold by month and prefecture on implementation of the gasoline IV standard. The data are from the State Administration of Industry and Commerce. The weather condition controls are the daily average temperature, rainfall, humidity, sunshine time and wind speed by prefecture. All of the specifications control for the indicators of the gasoline V standard, the diesel IV standard, the diesel V standard, exhaust emission standard IV for gasoline vehicles, the voluntary scrapping programs, as well as fuel prices. Reported in parentheses are robust standard errors clustered by prefecture. ***, * and * represent significance at the 1%, 5% and 10% levels of confidence, respectively.



Notes: This figure plots the impact of the gasoline IV standard on the logarithm of stacked pollutants through the day: 12-2am, 3-5am, …, 9-11pm. The coefficients βs in equation (1) are estimated for each time period separately. The dashed lines show the 95% level confidence intervals on each coefficient.

Fig. 8. The hourly effects of fuel standards on air quality.

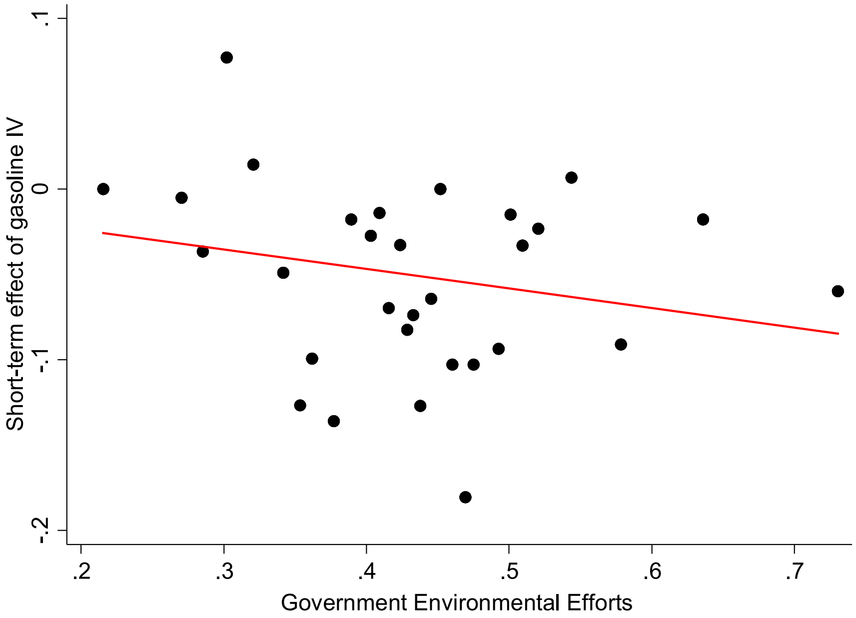
any health benefits related to reducing *PM* pollution.[13]

A number of recent studies have applied new data and approaches to estimate the mortality and morbidity impacts of mitigating air pollution in China. Ebenstein et al. (2017) examine the mortality impact of PM_{10} in China for different age groups and find that a 10 unit increase in PM_{10} raises cardiorespiratory mortalities by 8% on average. The monetized mortality cost based on the Value of a Statistical Life (VSL) is US$13.4 billion from a 10 μg/m^3 increase in PM_{10}. Barwick et al. (2018) provide the first comprehensive analysis of morbidity costs in China based on the universe of credit and debit card spending. They report that in monetary terms a permanent reduction of 10 μg/m^3 in daily $PM_{2.5}$ would lead to total annual savings of US$9.2 billion (in 2015 terms) in healthcare spending. Applying the estimates from column 2 of Table 4 (i.e., an average 9.4 μg/m^3 reduction in $PM_{2.5}$, and 16 μg/m^3 reduction in PM_{10} concentration), the gasoline IV standard's implementation implies US$21.44 billion in health benefits for China from reduced mortality and US$8.65 billion from reduced morbidity

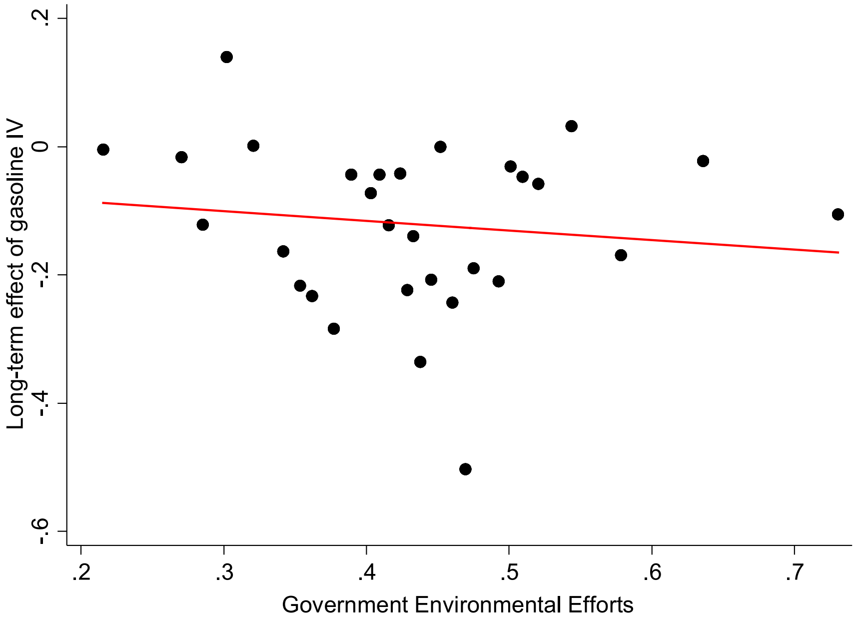
---

[13] Many studies have examined the health effects of PM10 exposure in China, but much less attention has been paid to $PM_{2.5}$, largely due to data availability (see He et al., 2016; Viard and Fu, 2015). In fact, $PM_{2.5}$ is known to be a better predictor of PM-driven acute and chronic health effects than the levels of coarser particles (Schwartz et al., 1996; Cifuentes et al., 2000; Pope and Dockery, 2006; Matus et al., 2012).

a Local Governments' Efforts and the Standard's Short-term Effect (RD Estimates)



b Local Governments' Efforts and the Standard's Long-term Effect (RD Estimates)



Notes: This figure plots the policy's RD coefficients (short-run and long-run) against the measure of government environmental protection efforts. The government efforts index is divided into 30 bins, and the average value of the policy effect is calculated for each bin.

Fig. 9. (a) Local governments' efforts and the standard's short-term effect (RD Estimates). (b) Local governments' efforts and the standard's long-term effect (RD Estimates).

(Table 9). Regressions reported in Table A3 in the online appendix show that when epidemiological data are used to infer the impacts of improved air quality on both mortality and morbidity the total predicted benefit is US$27.15 billion. So health benefits estimates using different approaches are fairly similar.

Costs. The new standards of course also have their costs. The NDRC announced that the wholesale price of China IV gasoline would be increased by ¥290/ton (¥0.21/litre).[14] The costs related to fuel upgrading were to be shared by the oil companies and consumers

---

[14] The NDRC has explained that the increases were based on comprehensive investigations and audits of Sinopec and PetroChina refineries which had already been upgraded to produce China IV and V fuel, as well as experience from Beijing and Shanghai.

Table 9
Valuation of the health benefits associated with air quality improvement.

Health impact	The monetized mortality and morbidity costs in China (US$ billions)	Reference	Health benefits (US$ billions)
Mortality	from a 10-μg/m3 increase in PM10 13.40	Ebenstein et al. (2017)	with PM10 reduced by 16 μg/m3 21.44
Morbidity	from a 10 μg/m3 increase in daily PM2.5 9.20	Barwick et al. (2018)	with PM2.5 reduced by 9.4 μg/m3 8.65
Total			30.09

Notes: This table presents estimates of the health benefits associated with air quality improvement. Based on existing health-based assessment of particulate air pollution in China, the health impacts of improved air quality as a result of fuel standards are inferred. The mortality costs denote cardiorespiratory mortalities from Ebenstein et al. (2017). The morbidity costs are estimated healthcare spending from Barwick et al. (2018).

(roughly, 30% vs. 70%), but the cost increase was not fully passed through to consumer prices, at least according to the NDRC. With the simplifying assumption that the price changes applied to all gasoline consumed without considering the gradual expansion of the use of higher quality fuels across China, a progression from gasoline standard III to IV would then have induced cost increases of approximately ¥25.18 billion (US$3.99 billion) for the consumers.

The envelope theorem can be applied to Chinese fuel consumers, but no simple envelope theorem applies to producers, in part due to Chinese oil refineries being neither perfectly competitive nor pure monopolists (Weyl and Fabinger, 2013). Formally quantifying the loss to producers from fuel upgrading would be very difficult and we choose not to pursue it (Ganapati et al., 2020).

Were the price adjustments more than adequate to cover the costs, or did they fall short of paying for the required refinery upgrades and increased production costs for China IV gasoline? The answer critically depends on the economic realities of the oil refineries, which in China are not transparent. Nonetheless, these rough estimates suggest that the net benefits of adopting the new gasoline standard are on the order of US$26 annually.

Policy Comparisons. We compare the impact of more stringent fuel standards with that of the driving restrictions applied in some Chinese cities (Viard and Fu, 2015), and with that of California's gasoline regulations (Auffhammer and Kellogg, 2011). In Beijing, restricting the number of vehicles driven each day has led to a 21% reduction in PM_{10} (a 30.8 $\mu g/m^3$ drop from an average level of 147 $\mu g/m^3$). That translates into 1114 fewer deaths and 15.3 million fewer restricted-activity days annually. The estimated benefits could be as much as ¥3.05 billion as opposed to costs of ¥519 million annually. Gasoline standard IV decreased PM_{10} by 12.9% (16 $\mu g/m^3$). That implies that while restricting the number of vehicles driven each day is known to be very costly and compliance is difficult to enforce, perhaps more emphasis should be placed on developing and implementing cleaner fuel and emissions technologies.

The CARB's gasoline regulations, by reducing ozone, have been estimated to save 660 lives in California each year. Valued in 2008 US dollars, the CARB imposed a cost of $1.2–1.6 billion per year, i.e., a cost of $1.8–2.4 million per life saved. That is much less than the EPA's VSL (US$6.45 million). Overall, the sheer size of China's population yields a much larger effect from improved fuel standards than the CARB can claim, confirming the importance of precise regulation in a large, developing country context.

5. Conclusions

Air pollution is currently China's most severe environmental problem, with the population increasingly experiencing prolonged and dangerous smog events. Extreme concentrations of fine particulates and ground-level ozone pose deadly threats to human health. With the dramatic growth of the private motor vehicle fleet, vehicle exhaust has become a major source of ambient air pollution in Chinese cities. In this paper, we assess quantitatively the importance of fuel quality and shed light on how low-sulfur fuels may help address air pollution.

Taking advantage of the roll-out and strict enforcement of new gasoline standards across China, this study has shown that cleaner fuels do indeed translate into better air quality. The adoption of higher gasoline standards significantly reduces local air pollutant concentrations, including those of *PM* and ozone. Further, using published estimates, the improved air quality translates into significant health impacts. Local governments' environmental efforts and regulatory attitudes influence the effectiveness of fuel standards in reducing pollution. These findings constitute the first compelling evidence about the benefits of Chinese fuel standards.

These findings about Chinese fuel standards greatly extend scholarly understanding of how this type of initiative affects the environment in a large developing country. Given the relevance of China's institutional setting, these Chinese findings may provide important policy insights for other less-developed countries such as India, which recently implemented new Bharat Stage fuel norms restricting sulfur content. This study focuses on the technical aspects of regulations aimed at reducing vehicle pollution. Government officials' political incentives in advancing this environmental agenda are also an important topic, but that difficult enquiry is left for future research.

Declaration of competing interest

The author declares that he has no relevant or material financial interests that relate to the research described in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jdeveco.2020.102488.

References

Almond, Douglas, Chen, Yuyu, Greenstone, Michael, Li, Hongbin, 2009. Winter heating or clean air? Unintended impacts of China's Huai River policy. Am. Econ. Rev. 99 (2), 184–190.

Anas, Alex, Timilsina, Govinda, Zheng, Siqi, 2009. An Analysis of Various Policy Instruments to Reduce Congestion, Fuel Consumption and CO2 Emissions in Beijing. World Bank Policy Research Working Paper 5068. World Bank, Washington, DC.

Auffhammer, M., Kellogg, R., 2011. Clearing the air? The effects of gasoline content regulation on air quality. Am. Econ. Rev. 101 (6), 2687–2722.

Barreca, Alan I., Guldi, Melanie, Lindo, Jason M., Waddell, Glen R., 2011. Saving babies? Revisiting the effect of very low birth weight classification. Q. J. Econ. 126 (4), 2117–2123.

Barwick, P.J., Li, S., Rao, D., Bin Zahur, N., 2018. The morbidity cost of air pollution: Evidence from consumer spending in China. NBER Working Paper.

Bharadwaj, P., Gibson, Matthew, Zivin, Joshua Graff, Neilson, Christopher A., 2017. Gray matters: fetal pollution exposure and human capital formation. J. Assoc. Environ. Resour. Econ. 4, 505–542.

Carneiro, Pedro, Løken, Katrine V., Salvanes, Kjell G., 2015. A flying start? Maternity leave benefits and long-run outcomes of children. J. Polit. Econ. 123 (2), 365–412.

Chen, Y., Ebenstein, A., Greenstone, M., Li, H., 2013. Evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River policy. Proc. Natl. Acad. Sci. Unit. States Am. 110, 12936–12941.

Cifuentes, L.A., Vega, J., Kopfer, K., Lave, L.B., 2000. Effect of the fine fraction of particulate matter versus the coarse mass and other pollutants on daily mortality. J. Air Waste Manag. Assoc. 50, 1287–1298.

Copeland, Brian R., Taylor, M. Scott, 2004. Trade, growth, and the environment. J. Econ. Lit. 42 (1), 7–71.

Coxhead, Ian, 2019. Environmentalism with Chinese characteristics—a review of Matthew E. Kahn and Siqi Zheng's blue skies over Beijing: economic growth and the environment in China. J. Econ. Lit. 57 (1), 161–179.

Cropper, Maureen L., 2010. What are the health effects of air pollution in China? In: Heal, Geoffrey M. (Ed.), Is Economic Growth Sustainable?. Palgrave Macmillan, Basingstoke, Hampshire. 10–46. Houndmills.

Davis, L.W., 2008. The effect of driving restrictions on air quality in Mexico city. J. Polit. Econ. 116 (1), 38–81.

Deshpande, Manasi, Li, Yue, 2019. Who is screened out? Application costs and the targeting of disability programs. Am. Econ. J. Econ. Pol. 11 (4), 213–248.

Duflo, Esther, Greenstone, Michael, Pande, Rohini, Ryan, Nicholas, 2013. Truth-telling by third-party Auditors and the response of polluting firms: experimental evidence from India. Q. J. Econ. 128 (4), 1499–1545.

Ebenstein, Avraham, Fan, Maoyong, Greenstone, Michael, He, Guojun, Zhou, Maigeng, 2017. New evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River policy. Proc. Natl. Acad. Sci. Unit. States Am. 114 (39), 10384–10389.

Environmental Protection Agency, 1998. EPA Staff Paper on Gasoline Sulfur Issues. EPA 420-R-98-005.

Environmental Protection Agency, 1999. Tier 2/Sulfur Regulatory Impact Analysis. Appendix B: Irreversibility of Sulfur's Emission Impact.

Environmental Protection Agency, 2000. Final Rule for Control of Air Pollution from New Motor Vehicles: Tier 2 Motor Vehicle Emissions Standards and Gasoline Sulfur Control Requirements.

Environmental Protection Agency, 2014. Tier 3 Motor Vehicle Emission and Fuel Standards.

Freyaldenhoven, Simon, Hansen, Christian, Shapiro, Jesse M., 2019. Pre-event trends in the panel event-study design. Am. Econ. Rev. 109 (9), 3307–3338.

Ganapati, Sharat, Shapiro, Joseph S., Walker, Reed, 2020. Energy cost pass-through in US manufacturing: estimates and implications for carbon taxes. Am. Econ. J. Applied Economics, forthcoming.

Gentzkow, M., 2006. Television and voter turnout. Q. J. Econ. 121 (3), 931–972.

Goodman-Bacon, A., 2019. Difference-in-Differences with Variation in Treatment Timing. Vanderbilt University. Working Paper.

Greenstone, Michael, 2004. Did the clean air act cause the remarkable decline in sulfur dioxide concentrations? J. Environ. Econ. Manag. 47 (3), 585–611.

Greenstone, Michael, Hanna, Rema, 2014. Environmental regulations, air and water pollution, and infant mortality in India. Am. Econ. Rev. 104 (10), 3038–3072.

Hausman, Catherine, Rapson, David S., 2018. Regression discontinuity in time: considerations for empirical applications. Ann. Rev. Resour. Econ. 10 (1), 533–552.

He, G., Fan, M., Zhou, M., 2016. The effect of air pollution on mortality in China: evidence from the 2008 Beijing Olympic Games. J. Environ. Econ. Manag. 79, 18–39.

Imbens, G., Lemieux, Thomas, 2008. Regression discontinuity designs: a Guide to practice. J. Econom. 142 (2), 615–635.

Imbens, Guido, Kalyanaraman, Karthik, 2012. Optimal bandwidth Choice for the regression discontinuity estimator. Rev. Econ. Stud. 79 (3), 933–959.

International Council on Clean Transportation, 2010. Overview of China's Vehicle Emission Control Program: Past Successes and Future Prospects. http://www.theicct.org/sites/default/files/publications/Retrosp_final_bilingual.pdf.

International Council on Clean Transportation, 2013. China Announces Breakthrough Timeline for Implementation of Ultra-low Sulfur Fuel Standards. Policy Update. March. Retrieved from http://www.theicct.org/sites/default/files/publications/ICCTupdate_CH_fuelsulfur_mar2013_rev.pdf.

International Council on Clean Transportation, 2014. China V Gasoline and Diesel Fuel Quality Standards. Policy Update. Jan. Retrieved from http://www.theicct.org/sites/default/files/publications/ICCTupdate_ChinaVfuelquality_jan2014.pdf.

Jacobsen, Mark R., 2013. Evaluating US fuel economy standards in a model with producer and household heterogeneity. Am. Econ. J. Econ. Pol. 5 (2), 148–187.

Kahn, Matthew E., 1996. New evidence on trends in vehicle emissions. Rand J. Econ. 27 (1), 183–196.

Kahn, Matthew E., Schwartz, Joel, 2008. Urban air pollution progress despite sprawl: the 'greening' of the vehicle fleet. J. Urban Econ. 63 (3), 775–787.

Lee, David S., Lemieux, Thomas, 2010. Regression discontinuity designs in economics. J. Econ. Lit. 48 (2), 281–355.

Matus, K., Nam, K.M., Selin, N.E., Lamsal, L.N., Reilly, J.M., Paltsev, S., 2012. Health damages from air pollution in China. Global Environ. Change 22 (1), 55–66.

Ministry of Environmental Protection, 2013. National Environmental Statistic Report (In Chinese). March 16. Retrieved from http://zls.mep.gov.cn/hjtj/qghjtjgb/201503/t20150316_297266.htm.

Ministry of Environmental Protection, 2012. T.M.o.E.P., Technical Regulation on Ambient Air Quality Index Daily Report GB3095-2012.

Parry, Ian W.H., Walls, Margaret, Harrington, Winston, 2007. Automobile externalities and policies. J. Econ. Lit. 45 (2), 373–399.

Pope, C.A., Dockery, D.W., 2006. Health effects of fine particulate air pollution: lines that connect. J. Air Waste Manag. Assoc. 56, 709–742.

Pope, C.A., Dockery, D.W., 2013. Air pollution and life expectancy in China and beyond. Proc. Natl. Acad. Sci. Unit. States Am. 110 (32), 12861–12862.

Schwartz, J., Dockery, D.W., Neas, L.M., 1996. Is daily mortality associated specifically with fine particles? J. Air Waste Manag. Assoc. 46, 927–939.

State Council, 2013. Ten Measurements of the Atmospheric Pollution Prevention Action Plan (In Chinese). June 14. Retrieved from http://www.gov.cn/ldhd/2013-06/14/content_2426237.htm.

Viard, V. Brian, Fu, Shihe, 2015. The effect of Beijing's driving restrictions on pollution and economic activity. J. Publ. Econ. 125, 98–115.

Watkiss, P.S., Pye, S., Holland, M., 2005. CAFE CBA: Baseline Analysis 2000—2020. Report to the European Commission's Director General for the Environment, Brussels.

Weyl, E. Glen, Fabinger, Michal, 2013. Pass-through as an economic tool: principles of incidence under imperfect competition. J. Polit. Econ. 121 (3), 528–583.

Wolff, H., 2014. Keep your clunker in the suburb: low-emission zones and adoption of green vehicles. Econ. J. 124 (578), 481–512.

World Bank, 2007. Cost of Pollution in China: Economic Estimates of Physical Damages. World Bank, Washington, D.C.

World Health Organization, 2011. Urban Transport and Health. World Health Organization, Geneva.

World Health Organization, 2015. Ambient Air Pollution: A Global Assessment of Exposure and Burden of Disease. World Health Organization, Geneva.

Xu, Chenggang, 2011. The fundamental institutions of China's reforms and development. J. Econ. Lit. 49 (4), 1076–1151.

Zhang, Y.L., Cao, F., 2015. Fine particulate matter ($PM_{2.5}$) in China at a city level. Sci. Rep. 5, 14884.