How Do Consumers Respond to Fuel Standard Regulation: Evidence from Gas Stations at City Boundaries

Yucheng Wang Lei Zhang Liyang Zhou*

This draft: Jan 2022
First draft: May 2021

Abstract

This paper estimates consumers' response to fuel standard regulation and willingness to pay for lower-emission gasoline. We exploit the unique market structural of China and policy-induced emission standard changes at city borders where consumers can freely choose from higher or lower-emission gasoline. Using high-frequency gas station-level data, our identification strategy compares sales volume of gas stations contiguous to each side of the border before and after one side experiences exogenous fuel standard reforms. We find evidence that consumers respond positively to fuel standard improvement and substitute higher-emission gasoline with lower-emission one. After controlling for price effects and gas station characteristics, our preferred specification shows that enforcing higher fuel standards increases relative sales at gas stations on the treated side of the boundary by 14%. This estimation corresponds to consumers' WTP for higher emission standards as about 0.345 CNY per liter (0.204 US$ per gallon), which is amounted to 4.7% of the total gasoline price. Mechanism analysis indicates that consumers care about environmental value of gasoline as the result of green preferences. A back-of-the-envelope calculation suggests that private welfare gains from higher gasoline standards are about 49.44 billion CNY per year, even without accounting for benefits from environmental and health improvements. Our findings highlight the importance of considering consumers' private value from emission standards when designing environmental regulation.

Keywords: Fuel Emission Standard, Willingness to Pay, Gasoline Market, Price Elasticity

*Wang and Zhou are from the Department of Economics, University of Pittsburgh. Zhang is from the Planning and Engineering Institute of China National Petroleum Corporation (CNPC). We are extremely grateful to Randall Walsh for his guidance and support on this project. We also thank Nicholas Flores, Roger von Haefen, Massimo Filippini, Davide Cerruti, Santosh Anagol, Andy Ferrara, Arie Beresteanu, Hui Li, Ron Yang, Zhiwen Wang and seminar/conference participants at the University of Pittsburgh, 2021 NEUDC, 15th NAUEA, 22nd CU Environmental and Resource Economics Workshop, Camp Resources XXVII, EMEE summer workshop, CES 2021 annual conference for their helpful comments and suggestions. Wang and Zhou declare no conflict of interest associated with CNPC. The results in this paper do not reflect the views of the CNPC. This draft is still preliminary and incomplete. All errors are our own.

1 Introduction

On-road transportation is one of the major sources of emission and air pollution. Automobiles contribute 28% of greenhouse gas emissions in the U.S. (EPA, 2014), up to 30% of air pollution in European cities, and up to 69% of air pollution in major developing cities (WHO, 2011). As a result, clean fuel standards or gasoline content regulations have been widely adopted around the world to reduce the carbon, sulfur, and ozone intensity of transportation fuels.[1]

A related question that often grabs the headlines and lies at the center of policy discourse is: would consumers be willing to pay a higher price for lower-emission gasoline? Economics theory regards emission as the *pure* negative externality and suggests that self-interested consumers ignore their impact on the environment when purchasing gasoline. However, recent studies and surveys have documented the existence of 'green consumerism' or 'green preferences' that motivate consumers to pay more in purchasing more environmental-friendly products (Ambec and De Donder, 2022; Wichman, 2016; World Value Survey, 2014). Therefore, the answer to the above question remains empirically ambiguous. Meanwhile, the consumers' willingness to pay (WTP) for cleaner gasoline is the key parameter to evaluate the cost and benefit of fuel standard policies. It is also important for the government because it would shape public attitudes toward fuel regulations. This paper provides the first evidence on changing purchasing behaviors in response to higher gasoline standards and estimates consumers' WTP for lower-emission gasoline.

Given its theoretical and policy importance, surprisingly few rigorous studies estimate consumers' preference for gasoline emission levels. Two main empirical challenges are related to the lack of credible estimations. First, a specific geographic region (e.g., state or province) usually implements a unified environmental standard, which makes it challenging to find a context where consumers could freely choose from fuels with different emission standards. Second, when emission standards change, the retail price and other non-emission-related characteristics are likely to change simultaneously. Therefore, it is challenging to separate consumers' preferences over the environmental value of gasoline from those over other product attributes and retailers' characteristics.

We overcome these challenges by using new data and employing a unique empirical context and market structure from the retail gasoline market of China. First, we collect longitudinal data of gas stations at a daily frequency from a major gasoline retailer in China. The granularity of the data allows us to explore the rich spatial dimension of variations. Specifically, we focus on the consumers'

[1]For example, the Washington Clean Fuel Standard, Oregon Clean Fuels Programs, and California Low Carbon Fuel Standard

purchasing behavior at city boundaries between Beijing and Hebei province, where the switching cost across borders is sufficiently low so that consumers can freely choose between high or low standard gasoline. The boundary research design here is in the similar spirit of Ito (2014). Because adjacent gas stations on different sides of the boundary are generally located beside the same main road, the potential customers of these stations are likely to be identical groups of commuters who cross the border by that road. Therefore, the research design help to partial out confounding factors related to the spatial dimension.

Second, during our sample period, there are several policy-induced reforms to improve the fuel standard for gasoline sold on either side of boundaries and thus reduce its emission level. The reforms increase only gasoline's environmental value while not increasing those non-environmental values, e.g., engine power and combustion efficiency. These spatial differences and quasi-experimental variations in fuel standards are crucial to identifying consumers' preferences on lower-emission gasoline. The intuition of our identification strategy is to compare relative sales gap across stations that are contiguous to either side of the boundary before and after exogenous fuel standard reforms. In addition, the high-frequency feature of our data allows us to detect consumers' purchasing choices around the exact dates of new gasoline being sold. We also include extensive fixed effects to control for confounding shocks.

Third, we take advantage of the simple-supply side structure of China's retail gasoline market, including government-controlled supply, state-owned retailers, and highly regulated prices. The high-intervened nature of the market simplifies our analysis from the general equilibrium effects of reforms and provides us with policy-induced price variations to estimate the price elasticity. These price coefficients are further used in the main estimation of standard reforms to control for the effect of concurrent price changes.

To connect our empirical findings to consumers' preferences parameters, we develop a simple model of gasoline demand. Consumers are assumed to get utility directly from the environmental value of gasoline and make discrete purchasing choices over stations on either side of boundaries. The tractability of the model allows us to derive a linear estimable equation from it, which translates policy-induced changes in relative market share across boundaries into revealed preference for different emission standards of gasoline. Therefore, our regression coefficients have direct interpretation as structural parameters instead of average treatment effects. This revealed preference approach is in the same spirit to the recent works such as Ito and Zhang (2020) and Houde and Myers (2021). To the best of our knowledge, our work is the first paper to estimate consumers' preference for environmental value in the gasoline market. Given estimated prices and emission standard parameters, the WTP is calculated as a monetary-equivalence utility from lower-emission gasoline.

We first estimate the price elasticity of gasoline purchasing using exogenous price shocks as instrument variables (IV). we find a sizable purchasing elasticity at the boundaries, which equals -3.744. This magnitude is consistent with previous papers, which find that consumers are very price-sensitive about their purchasing behavior when gasoline products are not highly differentiated from each other and costs from switching between gas stations are sufficiently low (Houde, 2012). We then use estimated coefficients in price regression to calculate the price-adjusted sales. To estimate the net-of-the-price response to standard reform, we use price-adjusted sales as the dependent variable of the main estimation. In our benchmark results, we find evidence that consumers respond positively to fuel standard improvement and substitute higher-emission gasoline with lower-emission one. After controlling for extensive fixed effects, the enforcement of higher fuel standards increases relative sales at gas stations on the treated side of the boundary by about 14.0%. This result corresponds to consumers' WTP for higher emission standards as about 0.345 CNY per liter, which accounts for 4.7% of the total gasoline price. Pre-treatment coefficients of event-study estimation suggest that the parallel trend assumption holds well. We track the effects up to at least 30 weeks after the reform and find that consumers' choice remains stable and persist by then. Counter to the traditional wisdom, our findings suggest that people are willing to pay for improving environmental quality. We argue this pro-environmental behavior could be driven by or warm-glow utility from using more environmental-friendly products or green preference. Although surprising at first glance, this conclusion is further supported by several consumer surveys and the World Value Survey, which indicates that China ranks highly in the world for the percentage of green consumers.

Our benchmark results are robust to using different event-time windows, distance-to-boundary bandwidths, the definition of *de facto* reform dates, and even estimation methods. In addition, we conduct a host of placebo tests. We find no effects on gas stations that are located far from boundaries. To shut down free substitution, we consider boundaries with differentiated products, where the Beijing side sells gasoline while the Hebei sides sell ethanol gasoline (E10). The fuel standard reforms generate no effects on those placebo boundaries. Results of fake reform dates and placebo fuel types further rule out the potential confounders from unobserved aggregated shocks and calendar-year effects. We examine all three gasoline standard improvements that happened during our sample period: one in Beijing and two in Hebei. The result of all three reforms consistently confirm our finding: the relative share increases in the treated side no matter it is Beijing or Hebei. In addition, heterogeneity analysis suggests that estimated effects are more prominent for stations adjacent to main roads, premium gasoline, and fuel standard reforms in Hebei (the initial lower standard region). A back-of-the-envelope calculation suggests that private welfare gains from higher gasoline standards are about 49.44 billion CNY per year, even without accounting for public benefits from environmental and health improvements.

Previous works have paid considerable attention to examining the effects of gasoline emission regulations as they impose substantial costs on retailers and consumers. One strand of existing studies estimates the impact on air pollution (Auffhammer and Kellogg, 2011; Li et al., 2020) and heath outcomes (Marcus, 2017) from a program evaluation perspective. Another strand of works studies the effects on prices and supply-side market structure (Anderson and Elzinga, 2014; Brown et al., 2008; Chakravorty et al., 2008). However, the impact on consumers has not been studied.[2] Despite the media and regulators' interest in consumers' WTP for lower-emission gasoline, and several surveys and news polls have been conducted on this issue,[3] no estimation using rigorous econometric tools has been provided due to data limitations and empirical challenges. This paper breaks new ground to the literature by studying consumers' behavioral responses to gasoline emission standards and providing the first estimates of WTP for cleaner gasoline. We take advantage of the gas station-level micro data, which allow us to explore rich spatial variation of gasoline standard. This type of data has been increasingly available to researchers, and have already been proved to be useful in studying spatial production differentiation (Houde, 2012) and demand elasticity of gasoline (Levin et al., 2017). In terms of methodology, we contribute to the literature about consumers' preference and awareness over goods with different energy efficiency or environmental value. The existing papers exploit experimental intervention in the lab or field (Allcott and Taubinsky, 2015). And non-experimental papers mainly count on household survey data (Jacobsen et al., 2012; Kotchen and Moore, 2007a; Sundt and Rehdanz, 2015). We develop an empirical framework to uncover consumers' value for cleaner gasoline using non-experimental settings and observational market sales data. Even though our empirical context is China, the research design that use gas station-level data and spatial and temporal variations of gasoline standard regulation around territory boundaries can apply to more general contexts.

Due to the importance of gasoline consumption to air pollution, our paper also indirectly speaks to

[2]In Li et al. (2020), they mention that consumers' behavior response is worth considering when evaluating the effects of gasoline standard regulations, but they assume it away because of lack of daily fuel consumption data.

[3]For example, based on a news poll, voters in Washington state, on average, are willing to pay 16.5 cents per gallon for gas if it meant a "significant reduction" in air pollution.

the emerging literature about the WTP for clean air. The existing papers estimate WTP for clean air from defensive investment (Ito and Zhang, 2020; Liu et al., 2018; Zhang and Mu, 2018) or pollution-induced migration behaviors (Bayer et al., 2009; Freeman et al., 2019). The underlying assumption is that people's utility from clean air derives from their values on health or amenities. Our finding complements those studies by showing that people have strong WTP for clean air even without assuming health and amenities in the utility function: consumers gain direct utility from reducing the production of air pollution other than the avoidance behaviors. This perspective has importance policy implication. In a paper that studies the same gasoline standard reform in China, Li et al. (2020) finds the reform reduced the "average pollution across all pollutants" by 12.9%. They estimate the public benefit from reducing lifetime mortality and morbidity to be about 30.09 billions US$. In contrast, our back-of-the-envelope calculation indicates that the private WTP for reducing air pollution from gasoline consumption is quantitatively important. Our findings highlight the importance of considering consumers' private value from emission standards when designing environmental regulation.

Another contribution of our paper is that we provide an estimation of the gasoline price elasticity using high-frequency gas-station-level data. A large body of early works on the price elasticity of gasoline use cross-section or highly aggregated data and find an inelastic demand response to the price change (Brons et al., 2008; Dahl and Sterner, 1991; Espey, 1998; Goodwin, 1992; Hughes et al., 2008; Park and Zhao, 2010; Small and Van Dender, 2007). In a recent paper, Levin et al. (2017) works with a daily city-level dataset to include an extensive set of fixed effects in regression. They find the price elasticity much larger than studies using lower-frequency data. In the industrial organization (IO) literature, Houde (2012) develops a Hotelling-style model with spatial Differentiation and estimates the price elasticity an order of magnitude large than the previous papers. He highlights that when considering the geographical proximity of gas stations, consumers are super sensitive to the price difference. Our paper bridges the gap between the "reduced form" method and the structural model with locations. We use high-frequency data and focus on the choice among stations near the territory boundary. Exploiting policy-induced price variation, our paper identifies a sizeable price elasticity for gasoline that is close to the structural estimation. We confirm the importance of the spatial dimension of this question and show that the estimation highly hinges on the source of price variation for identification.

The remainder of the paper proceeds as follows: Section 2 introduces China's fuel regulations and emission reforms background. Section 3 develops a simple gasoline demand model to guide the empirical analysis. Section 4 describes the data and sample construction and presents some descriptive evidence. Section 5 conducts the empirical analysis and present the results. Section 6 discusses the policy implications. The last section concludes.

2 Gasoline Market and Emission Standard Reform in China

China retail gasoline market has several features that make it a unique context to exploit the questions we are interested in. First, the retail market is a duopoly market predominated by two state-owned companies. The variations between their products are very small. Second, the retail price is limited by a price ceiling ("guided price") set by the government, which leads to exogenous price shocks on the supply side. Third, there are substantial regional differences in gasoline standards between the regions. Meanwhile, several reforms that exogenously change the standards in some of the regions (Li et al., 2020), providing spatial and temporal variations in fuel standard. In this section, we introduce in details about these features.

2.1 Duopoly Market and Regulated Price

China retail gasoline market is dominated by two major state-owned companies: *China National Petroleum Corporation* (*PetroChina*) and *China Petroleum & Chemical Corporation* (*Sinopec*).[4] These two brands together take up over 90% of the market share in the region we focus on. As vertical-integrated corporations, both *PetroChina* and *SINOPEC* hold effective control over their supply chain and product quality. The refining, supply, and transportation of gasoline in each province are managed by the provincial branch so that the price and the characteristics of gasoline vary little between different gas stations owned by the same brand in the same province. As part of the evidence, we show in section 4 that the price variation in our sample mainly comes from the time series rather than from different gas stations in cross-section data.

Meanwhile, the retail gasoline price is regulated by *National Development and Reform Committee* (NDRC). Since 2008, NDRC has been setting the price ceiling (the "guided price") for refined oil products and announced it to the public regularly. In specific, between 2008 to 2013, NDRC adjusts the gasoline price ceiling every twenty-two working days, based on crude oil prices in Singapore, New York, and Rotterdam.[5] After early 2013, NDRC adjusts the gasoline price ceiling every ten working

[4] Besides *PetroChina* and *Sinopec*, consumers have very few options and have to choose from either well-known foreign brands that only have limited numbers of gas stations in the region, such as Shell, or unknown local brands that sell gasoline with an unstable quality because of the also unstable supply chain.

[5] The price will only be adjusted if the weighted average price of refined oil in these three places fluctuates by more than 4% for 22 consecutive working days, otherwise NDRC will announce the price to stay the same.

days, to anchor to the Brent, Dubai, and Minas crude oil prices. The markup is included in the price ceiling based on the transaction cost, taxation, and profit for oil companies. [6] For gasoline companies, it's a common practice to follow the announced price ceilings, which makes the price ceilings likely to be the focal points of market prices (Zhang et al., 2020).

2.2 Gasoline Emission Regulation in Beijing and Hebei

To improve gasoline quality and reduce air pollution, China has implemented a series of fuel standard reforms since 1999. The reforms established a system of fuel emission standards of different generations, based on European regulations.[7] In our sample period, Beijing and Hebei held standards from III to V. Following the European standards, a higher generation index represents a more stringent standard on gasoline emission. The reforms allow regional differences in the implementation of the standards, as we have seen in the case of Beijing and Hebei. Except for the historic reasons mentioned before, a critical reason for the heterogenous timings is the local refining capacity (Li et al., 2020). To be specific, the provincial government will not start upgrading until the provincial branches of *PetroChina* and *Sinopec* can ensure a sufficient capacity of supply for higher standard gasoline.

The region we study has huge demand on gasoline consumption. Beijing, as the capital of China, has 21.89 million residents and more than 6 million vehicles. Meanwhile, Hebei, the province surrounding Beijing, has 74.61 million residents and 20.57 million vehicles.[8] As the most developed region of China, Beijing offers more job opportunities, and have higher housing and renting prices compared to Hebei. As the results, a considerable amount of people chooses to reside in Hebei and commute to Beijing for daily work. This strategy is enabled by the road network connecting the two provinces, which includes roads of different classes from highways to motorways with varying quality.

The gasoline emission standards substantially differ between Beijing and Hebei. Beijing implement more stringent environmental regulations compare to the rest of China. For example, to prepare for the 2008 Olympic Games in Beijing, the government devoted large efforts to alleviate the air pollution and gasoline emission. For this reason, Beijing became the first region to employ the emission standard IV since the beginning of 2008. During our sample period, Beijing's standard is higher than Hebei until 2015, when standard V is implemented in Hebei. We present more details when discussing the timeline

[6] This reform is implemented on March 26th, 2013, when the state council issued *"Notification on the Implementation of Retail Oil Price and Taxation Reform"* and modified the price ceiling enacting rules.

[7] See Appendix **??** for comparison between standards of China and other countries and regions including Europe.

[8] Source: 7th population census and Ministry of Public Security's report.

of reforms in section 2.3.

The more stringent environmental regulation in Beijing is also associated with travel restriction policies. To be specific, Beijing has forbidden vehicles that failed to meet the vehicle emission standard of the city from entering the region within the sixth ring road of it. And vehicles satisfying the standard still need an entry permit before entering the city if they have non-Beijing plates. The permit costs no fee, though it can take several minutes for drivers to fill out a form and provide identity information of their vehicles and their own. The restriction could be strengthened when important activities are held in Beijing, such as the "two meetings" in the spring of each year. We address these confounding policies in section 3.3.

2.3 Three Emission Standard Reforms

Within our sample period (2012-2016), there are three standard reforms implemented in the two sample regions: the first one is in Beijing (Beijing IV $\longrightarrow$ V) and the last two are in Hebei (Hebei III $\longrightarrow$ IV, IV $\longrightarrow$ V). Each reform improved the fuel standard in the treated region by one generation. We use the Hebei III-IV reform for our main estimation, and test the impacts of the rest two for robustness check. The first reform, implemented in Beijing in May 2012, improved the fuel emission standard in Beijing from IV to V. The detailed information of the new standard was announced to the public on May 7th, 2012 and the gasoline that meets the new standard should become available from May 31st, 2012. The second and the third reforms increased the standard in Hebei (and most other provinces in China) from III to IV and from IV to V. The new standard was established on the same day for all the cities within the province Hebei. In particular, standard IV gasoline should become available to consumers on January 1st, 2014, and standard V gasoline should become available on January 1st, 2016. However, for two Hebei reforms, the gas stations have started to provide higher standard gasoline before the official announced dates. The *de facto* starting date varies across gas stations and generates transition periods during which both lower and higher standard gasoline are sold (III and IV in Hebei III $\longrightarrow$ IV, IV and V in Hebei IV $\longrightarrow$ V). We reconstruct the treatment time for the Hebei reforms, and the details are presented in section 4.1.3. And though we don't know why the product release dates are moved forward in practice, in section 5.3 we show that this change of date doesn't affect the robustness of our estimation.

The main goal to emission reforms is to reduce air pollution. The reforms strengthened the restrictions on hazardous materials content, such as sulfur, aromatics, and benzene, for all types of refined gasoline. Figure 2 plots the timeline of the three reforms in our sample, with the changes of the maximum parts per million (ppm) of sulfur the refined gasoline is allowed to contain as the impact of these reforms. The horizontal axis gives the time, with the dash lines indicating the beginning and the end of the transition periods. The vertical axis gives the different standards, and the sulfur content restrictions for each standard are on the horizontal lines. The red line demonstrates the standard in Beijing and the blue line the standard in Beijing. As shown, the implementation of standard III, IV, and V lowers the maximum ppm of sulfur in gasoline to 150, 50, and 10 respectively.

The reforms also changed gasoline characteristics other than hazardous materials, such as the research octane number (RON). RON is used to measure the fuel efficiency of gasoline and identify the grades of gasoline in China.[9] As the RON changed after the IV $\longrightarrow$ V reform, the grade numbers also changed the treated regions. In particular, before the IV $\longrightarrow$ V reform, there are three major grades of refined gasoline in China, #90, #93, and #97, each indicating its RON. [10] After the IV-V reform in Beijing or Hebei, the #90, #93, and #97 gasoline in the region were renamed to #89, #92, and #95, to reflect the reduction in RON due to the stricter limitation on sulfur and manganese. Our sample contains data of both #93-#92 and #97-#95 gasoline, therefore, to avoid confusion, in this paper we use "regular" and "premium" gasoline when referring to #93-#92 and #97-#95 gasoline.

The standards of diesel also changed along with the gasoline reforms in Beijing and Hebei, though the implementation dates are not all the same. Figure A.2 displays the timeline for the diesel reforms, following the same manner of Figure 2. For the Beijing IV $\longrightarrow$ V and Hebei IV $\longrightarrow$ V reforms, the diesel upgrading dates are the same as the gasoline reforms, but for the Hebei III $\longrightarrow$ IV reform, the upgrading for diesel is about one year later than the gasoline, at the end of 2014, possibly because there is not sufficient refining capacity for diesel at the time of the gasoline reform. We use this delay as an opportunity to conduct a placebo test and test how diesel sales change at the time of Hebei III-IV gasoline reform. The results are presented in section 5.3.

[9] Each vehicle model has its own recommended gasoline grades given by the manufacturer. In general, consumers tend to follow the recommendation.

[10] The #90 gasoline is the counterpart to the regular gasoline in the US, the #93 the plus/special, and the #97 the premium. One might find the numbers shown on the US pumps for regular, plus/special, and premium gasoline are lower than these grades, this is because US pumps usually show the Motor Octane Number(MON), which tends to be than RON due to different measuring.

3 Empirical Model of Gasoline Demand

The empirical challenge in measuring consumers' private willingness to pay for higher gasoline standards is to distinguish consumers' preferences over environmental standards from those over price, other product attributes, and gas station characteristics. To overcome this problem, we develop a simple gasoline demand and gas station choice model to guide our estimable equation. Two features of our framework are worth noting. First, we model consumers' problem as a discrete choice over gas stations and abstract away the amount of gasoline consumption per purchase. This modeling choice is consistent with the nature of our data that the sales volume is aggregated at gas station level.[11] Second, we capture consumers' preference over gasoline in a "reduced-form" way, which takes the form close to the difference-in-difference specification.

3.1 A Discrete Choice Model over Gas Stations at City Boundaries

In our setting, the product refers to gasoline type j sold in gas station k at city boundaries.[12] The indirect utility of consumers i from consuming gasoline type j in gas station k, $U(p_{jkt}, \boldsymbol{\Lambda_{jt}}, G_k, \xi_{jkt}; \theta)$, is specified as,

$$\begin{gathered} u_{ijkt} = \alpha\, p_{jkt} + \boldsymbol{\Lambda_{jt}} + G_k \cdot \gamma + \xi_{jkt} + \varepsilon_{ijkt}, \\ i = 1, ..., I, \quad j = H, L \quad k = 0, 1, ..., K \end{gathered} \tag{1}$$

and,

$$\boldsymbol{\Lambda_{jt}} = \beta \cdot D_j \cdot \mathbb{1}(t \geq T^*) + \lambda \cdot \mathbb{1}(t \geq T^*) + \phi \cdot D_j \tag{2}$$

where $\alpha\, p_{jkt}$ is the disutility of price; $\boldsymbol{\Lambda_{jt}}$ is consumers' utility from gasoline type j, which captures both environmental value as well as a utility from other environmental-unrelated characteristics.[13] G_k is a vector of time-invariant gas-station level characteristics for gasoline station k, such as location, the number of pumps, convenience store area.[14] ξ_{jk} is the unobserved product-level characteristics. ε_{ijk} is

[11]The underlying assumption here is that the distribution of quantity per purchase is i.i.d across all gas stations. And the number of purchases (within a day) is big enough so that the law of large number could apply. We regard this as a reasonable assumption given our context.

[12]The dimension of products space is k instead of $2k$, since the choice of gas station nest the gasoline type.

[13]In our context, environmental value is higher when the fuel type is enforced with a more stringent emission standard, in other words, cleaner fuel. In contrast, environmental-unrelated characteristics include all value differences between gasoline types irrelevant to the emission standard, e.g., damage to the engine, explosion resistance, combustion efficiency, or even local protectionism and loyalty.

[14]The utility from gas station characteristics is not necessarily assumed to be time-invariant. In section 5, we allow for gas station-by-day-of-week fixed effects and interactions between gas station characteristics and time trends.

consumers' idiosyncratic utility shock. For the parameters of interest, α is the marginal disutility for price or marginal utility of money; γ is the marginal utility for gas station characteristics.

The specification of $\boldsymbol{\Lambda_{jt}}$ is central to our empirical design. As there are only two gasoline types in the market (H, Beijing and L, Hebei), we model $\boldsymbol{\Lambda_{jt}}$ in a "reduced-form" manner, which takes the form similar to the "difference-in-difference" specification.[15] Specifically, $\boldsymbol{\Lambda_{jt}}$ is a function of: (1) fuel type D_j, which indicates the time-invariant fuel characteristics. D_j equals to 1 if $j = L$ (Hebei) and equals to 0, otherwise; (2) time dummy $\mathbb{1}(t \geq T^*)$, which indicates post-reform period where T^* is the time spot of policy enforcement; and (3) the interaction term. λ captures the time-related utility shifter, and ϕ captures the value gap of all characteristics between two products that do not change with the reform. The parameter of interest β captures the changes of relative value on new L (Hebei) gasoline relative to the old one. If $\beta > 0$, the reform increases the WTP for L and thus increases sales volume of product L relative to H. As we claim in section 2, the only policy-induced change is the emission standard of gasoline; therefore, β captures consumers' preference on higher-standard (lower-emission) gasoline after reforms. The (marginal) willingness to pay (WTP) for higher emission standard is the monetary-equivalence (marginal) utility, which can be calculated by $-\beta/\alpha$.[16] Note that, in this framework, we cannot distinguish whether $\boldsymbol{\Lambda_{jt}}$ derives from preference on *pure* or *impure* public goods.[17]

Since our data cover only the gas stations of *PetroChina*, we define the potential market for *PetroChina* in our empirical exercise.[18] Therefore, the market size M is the total potential consumers for *PetroChina* gas stations during a specific period of time (one day in our data). Accordingly, the outside option ($k = 0$) represents those whose do not to purchase gasoline from any *PetroChina* station (e.g., do not drive). The reservation utility is defined as,

$$u_{i0t} = \delta_{0t} + \xi_{0t} + \varepsilon_{i0t}. \tag{3}$$

[15]This product-space approach provides a more direct interpretation of estimated policy effects to structural parameters compared with the widely used characteristic-space approach, such as in Berry et al. (1995).

[16]Modeling $\boldsymbol{\Lambda_{jt}}$ in a "reduced-form" way allows us to connect the WTP directly to the change in expected consumer surplus due to the standard reform. Specifically, $\triangle E[CS] = marketshare_L \cdot (WTP - \triangle p_L)$. See Appendix E for details.

[17]See Kotchen (2005); Kotchen and Moore (2007a) for details about *pure* and *impure* public goods model of environmentally friendly consumption. The *pure* public goods preference in the utility function is pure warm-glow while *impure* public goods preference is impurely altruistic (Andreoni, 1990).

[18]Due to the data limitation, we split gas stations of *PetroChina*, and *Sinopec* together with other non-state-owned brands into two distinct potential markets. Note that the validity of this distinction does not require consumers to be immobile across two potential markets but requires the relative size of two markets to be stable during the reform period. Given the highly regulated nature of the retail gasoline market in China, we regard this to be a reasonable assumption.

3.2 Choice Probabilities and Sales Volume

Consumers make discrete purchasing choices over gas stations located within a bandwidth h around city boundaries. It is reasonable to expect that the substitution patterns cross city boundaries and within city are substantially different. We restrict consumers' choice set to stations near the boundary since the cross-boundary choice is our main object of interest in this paper.[19] We discuss the choice of h in more detail as they become relevant for our empirical analysis.

We assume that for all consumer i the error term ε_{ijk} is distributed according to a $i.i.d.$ Type I extreme-value distribution. Since we consider only homogeneous consumers, the standard conditional logit model apply here.[20] Combining equation 1 and 3, and the distributional assumption on ε_{ijk}, the market share of gasoline station k, $s_k = Q_k/M$, is given by,

$$\begin{aligned} s_{jkt} &= \frac{exp\Big(\delta_{jkt} + \xi_{jkt}\Big)}{\sum_{k'=0}^{K} exp\Big(\delta_{j'k't} + \xi_{j'k't}\Big)}, \quad k = 1, ..., K \\ s_{0t} &= \frac{1}{\sum_{k'=0}^{K} exp\Big(\delta_{j'k't} + \xi_{j'k't}\Big)}. \end{aligned} \tag{4}$$

where $\delta_{jkt} = \alpha\, p_{jkt} + \beta\, \boldsymbol{\Lambda_{jt}} + G_k \cdot \gamma$ is the mean utility of consuming gasoline j in station k, and $\delta_0 + \xi_0$ is normalized to 0. Based on Berry (1994), the inversion of marker-share equation is,

$$ln\,(s_{jkt}) - ln\,(s_{0t}) = \delta_{jkt} + \xi_{jkt}, \quad j = H, L \quad k = 1, ..., K \tag{5}$$

Since market size M and outside option share s_0 do not vary within the market, these two terms are absorbed in the constant term (Ito and Zhang, 2020).[21] Combining equation 1, 2 and 5, we could

[19] An alternative way to allow for different substitution patterns across locations is to use the nested logit model and divide gas stations into three groups: near city boundaries, non-boundaries Beijing region, and non-boundaries Hebei region. We keep only the near city boundaries group to ease our analysis. And we also show our results to be robust to changing bandwidth h.

[20] This type of model capture the probability of choice conditional on prices, product characteristics, and consumers characteristics. See McFadden et al. (1973) for details. In the appendix, we discuss more about our modeling choice and potential impacts of allowing random coefficient logit model on our results.

[21] Note that, here, for expositional purposes, we regard all gas stations in our sample as in a single market. However, our estimation does not rely on the single-market assumption. It is easy to relax the assumption by further dividing potential demand in multiple regional sub-market based on geographical location. For the muli-market cases, these two terms do not vary within each sub-market and would be absorbed by the sub-market fixed effect (Ito and Zhang, 2020). For example, later in our empirical exercise, we include the road-segment fixed effect, which means each road segments could be regarded as a separated market. This feature takes advantage of the strong longitudinal feature of our data.

write down the sales equation for gas station k as,[22]

$$log(Q_{jkt}) = \beta_0 + \alpha\, p_{jkt} + \beta \cdot D_j\, \mathbb{1}(t \geq T^*) + \lambda \cdot \mathbb{1}(t \geq T^*) + \phi \cdot D_j + G_k \cdot \gamma + \xi_{jkt} \tag{6}$$

where, $\beta_0 = log(M) + log(s_0)$. Equation 6 is our main estimating equation for the empirical analysis.

3.3 Identification and Estimation

We briefly discuss the estimation strategy and source of identification. The challenging to identification is to seperate consumers' preference emission-related characteristics from non-emission-related value and gasoline station characteristics. The recent literature to study consumers' WTP for energy efficiency or environmental value confront similar issue (Allcott and Taubinsky, 2015; Ito and Zhang, 2020; Sallee et al., 2016). Existing papers deal with the issue either by exploiting experimental or quasi-experimental variation for identification.[23] In our paper, we exploit policy-induced exogenous variations of increasing the environmental value. The central assumption is that the enforcement of higher-quality gasoline standards changes the environmental value of gasoline while keeping the non-environmental value invariant. In other words, there are no time-variant unobservables. The validation of this assumption is supported by facts illustrated in section 2 that upgrading gasoline standard in China does not substantially change non-emission-related characteristics. In addition, we restrict the sample to a short event time window around the reform dates to get rid of the shifting of systemic factors, for instance, commuters' route choice and new vehicle purchasing that could potentially bias the estimates. Another important assumption is that the reforms are orthogonal to any pre-determined trend. This assumption could be violated if the government strategically chooses the timing of the reform. We validate this assumption by testing the parallel trends as in the standard difference-in-difference method and discuss the potential concerns in section 5. With these assumptions, the estimable equation 6 maps the observed changes of sales volume in Hebei relative to Beijing, controlling for the price effect, to the structural parameter of interest.

Another empirical challenge in our context is to separately identify β from α. This problem arises

[22] Using $ln\,(s_{jkt}) - ln\,(s_{0t})$, $ln\,(s_{jkt})$, or $ln\,(Q_{jkt})$ as the left hand side of the regression equation would give same estimation of α and β but different estimation of the coefficient on the intercept.

[23] For example, most similar to our context, Ito and Zhang (2020) identifies the environmental value of high-efficiency air purifiers using a plausibly exogenous variation of air pollution as the result of geographical discontinuity of the heating policy. Allcott and Taubinsky (2015) identifies the information value of utility between two light bulbs with different energy efficiency by conducting an information nudges experiment. Sallee et al. (2016) exploits the variation of future gasoline price to distinguish the value of fuel economy from other factors that affect the level of used vehicle prices.

due to the lack of enough price variation at the gas station level. To see this, decompose the price p_{jkt} into and policy-induced price fundamental (terms in bracket) and within-gasoline-type price adjustment for each gas station, $\tilde{p}_{jkt}$. Specifically,

$$p_{jkt} = \left[\rho_1 \cdot D_j \cdot \mathbb{1}(t \geq T^*) + \rho_2 \cdot \mathbb{1}(t \geq T^*) + \rho_3 \cdot D_j\right] + \epsilon^p_{jkt} \tag{7}$$

where ρ_1 is the price increases that reflect the heightened refining costs after implementing more stringent environmental standards. Substituting equation 7 into 6, it is clear that the identification of α depend solely on the variation of ϵ^p_{jkt}. As we have shown in section 2, the retail price of the gasoline market in China is highly regulated by the government, and the within-gasoline-type price variation is generally small. This could cause a potential multicollinearity problem between p_{jkt} and $D_j\, \mathbb{1}(t \geq T^*)$. For the extreme case where $\epsilon^p_{jkt} = 0$, the price is perfectly collinear with $\boldsymbol{\Lambda_{jt}}$. Consequently, β is not identifiable from α.[24] At the same time, if ϵ^p_{jkt} is correlated with unobserved errors, the estimation of α will be biased, which leads to a biased β.[25]

To deal with this problem, we construct a two-step procedure. In the first step, we estimated α using an exogenous variation of price. We use other policy-induced price shocks and also international crude oil prices as the instruments for gasoline prices. Given estimated $\hat{\alpha}$, we move $\hat{\alpha} p_{jkt}$ to the left-hand side of the regression equation. The new dependent variable, $log(Q_{jkt}) - \hat{\alpha}\, p_{jkt}$ can be regarded as the "price-adjusted-gasoline-consumption", or the net value of fuel after accounting for its price. We use the two-step procedure as our preferred method, but also report the results using one-step method.[26]

4 Data and Descriptive Statistics

In this section, we first describe each data source and construction of data sets. Then, we provide summary statistics about gas station attributes and price variations. We end by presenting descriptive evidence about the consumer's purchasing behavior at the border.

[24]Even though no perfectly multicollinearity problem and regularity condition $E[\epsilon^p_{jkt} \cdot \xi_{jkt}] = 0$ holds, still, this will lead to a problem of large asymptotic variance. We illustrate this problem using a Monte Carlo simulation in figure A.4.

[25]To see this, substitute equation 7 into 6 and denote the coefficient of the new interaction term as B. it is easy to show that $\beta = B - \hat{\alpha} \cdot \rho$. As long as the assumptions mentioned in the first paragraph of this section are satisfied, B and ρ are unbiasly estimated. However, Identification of α depends on ϵ^p_{jkt} to be orthogonal to the error term. If α is biased, β will also be biased.

[26]As we show in section 5, the direction of estimated α and β are the same in both methods, while the magnitude using the one-step method is much larger. Therefore, we regard the two-step method as a more conservative estimation.

4.1 Data Source

4.1.1 Retail Gasoline Data

We compile a dataset from two main data sources: the administrative data of *PetroChina* gas stations, and Geographic Information System (GIS) data of transportation from the Ministry of Transport of China. The administrative gas-station data includes daily prices and sales volumes of gasoline and diesel, for universe gas stations owned by *PetroChina* in Beijing and seven cities in Hebei from January 1st, 2012 to December 31st, 2017. The cities in Hebei include all five cities adjacent to Beijing and two cities not adjacent to Beijing. Among the five adjacent ones, four sell pure gasoline, and one sells ethanol gasoline (E10, a blend of 10% ethanol and 90% gasoline). The two cities not adjacent to Beijing both sell ethanol gasoline. [27] The data also includes the geographical features and specific characteristics of gas stations, such as addresses, longitudes and latitudes, whether there is a nearby convenience shop, and whether provides engine oil replacing service.

Figure 3 plots the map of our sample region and the locations of gas stations. The red region in the center is Beijing, and the surrounding blue ones are cities of Hebei, among which the lighter ones sell ethanol gasoline and are therefore excluded from our main estimation. The *PetroChina* gas stations in our sample are represented by the black dots. Within the sample period, *PetroChina* owns 880 gas stations in the sample region, 180 in Beijing, and 700 in Hebei.

Figure 1 shows the price fluctuation within the sample. Our sample consists of 3,733,444 unique prices in 1826 days from the 880 gas stations, and we can observe an immediate price change after each reform was implemented, except after the Beijing IV $\longrightarrow$ V reform. This is because the international crude oil of the time is at a high price, leading to a relatively high gasoline retail price in China. As a result, the NDRC decided it's better to delay the price change to "stabilize the market and people's living standards". With the international crude oil price falling, the gasoline price in Beijing finally jumped up in August 2012, three months after the new standard was established. Therefore, the 2012 reform provides us two unique shocks in the sample: the first one only improved the emission standard and therefore the environmental value of the gasoline, without changing the price; the second one only increased the gasoline price, with no change in emission standard or environmental value. This is a foundation for part of our estimation strategies, as we are going to use the second price shock to estimate the price elasticity of demand. More details are provided in Chapter 5 when talking about the

[27] Among the five adjacent ones, the four selling pure gasoline are Zhangjiakou, Chengde, Tangshan, and Langfang, and the one selling ethanol gasoline is Baoding. The two cities not adjacent to Beijing selling ethanol gasoline are Shijiazhuang and Handan.

identification.

4.1.2 GIS Data about Road Networks

Our GIS data comes from the Ministry of Transport of the PRC and includes all levels of road network information within the two provinces. The roads, based on quality and other characteristics, are classified as highways, national roads, provincial roads, and so on. As we are mostly interested in the behavior of the commuters, we focus on the highways and national highways as they connect Beijing and Hebei and are frequently used by the commuters. Like the highways in the US, the gas stations near the highways and national highways are usually located in the service area and can be easily found every several miles. In section 5.3 we test the heterogeneity between the gas stations near and away from the highways and use it to further explore the mechanism of consumer behavior.

4.1.3 Transition Periods and Treatment Dates of Reforms

The granular nature of our data to identify the exact date of implement new emission standard for each stations. As mentioned in section 2.3, during two Hebei reforms, stations start to supply new gasoline a bit earlier than announced policy implementation dates. Therefore, we define the treatment date as de facto date when Hebei gas stations near the boundary start selling new gasoline. The transition period is defined as the days during which both lower and higher standard gasoline (or diesel in the diesel reform) are available at the gas station. The beginning and the end and therefore the length of the transition period vary between the gas stations. The transition period for a given reform is defined as the union of all transition periods of all gas stations. In specific, the transition period starts when there is at least one gas station starts selling the higher standard gasoline and ends when all gas stations have stopped selling lower standard gasoline and only provide higher standard gasoline. After the transition period of each reform is determined, we use the beginning date of it as the treatment date. In the benchmark, we set same treatment date for all stations and check the robustness when using different start dates for each station.

4.1.4 Price Cap Adjustments and Holidays

We also manually collect the announcement of the retail gasoline price cap adjustment from NDRC circulars.[28] The information includes the specific price changes and effective dates for all provinces. There are 176 times price adjustments in total during our sample period, with 80 times being upward

[28]Source: https://www.ndrc.gov.cn/xwdt (In Chinese).

adjustment and 96 times being downward adjustment. Holiday and festival arrangements are also collected from NDRC.

4.1.5 POI data of non-*PetroChina* gas stations

To control for the competition from gas stations of other brands, we collected the geographic location data for all non-*PetroChina* gas stations in our sample region from the Baidu Map API.[29] We then calculate the geometric distance from these gas stations to the *PetroChina* gas stations in our sample, and use the numbers of non-*PetroChina* gas stations within a certain radius from a given *PetroChina* gas station as the proxy of the competition it would face. On average, each *PetroChina* gas station has 5.27 non-*PetroChina* gas stations within a radium of 5km.

4.2 Summary Statistics and Price Variations

Table 1 provides the summary statistics of gasoline retail data and gas station characteristics of the gas stations located within 30 km to the boundary, grouped by Beijing and Hebei. We report summary statistics for all 254 gas stations within the 30 km bandwidth in column 1, the 141 on the Beijing side in column 2, and the 113 on the Hebei side in column 3.

Rows 1 to 2 present the prices of regular and premium gasoline, with the ethanol gasoline (E10) included, as its price doesn't differ from the pure one of the same grade. Gas stations in Beijing sell no premium E10 and a very small fraction of regular E10, while in Hebei only gas stations in certain cities sell E10, regular and premium. As shown, regular gasoline in Beijing is averagely 0.235 CNY/liter more expensive than that in Hebei. For premium, the price gap is 0.435 CNY/liter, due to the higher refining cost for a higher grade. Rows 3 to 7 present the gas-station-level attributes, including distance to the boundary, distance to the nearest gas station in the other province, distance to the nearest gas station in the same province, whether near the highways, and whether provides engine oil replacement service. Compared to the gas stations in Hebei, the ones in Beijing are significantly further away from the border, and much fewer of them are near the highways or providing engine oil replacements. We observe trivial and statistically insignificant differences between gas stations in Beijing and Hebei in the terms of distances to the nearest gas station, no matter in the same or the other province, suggesting similar geographical densities of gas stations in the two provinces.

As mentioned above, the retail price in China's gasoline market is regulated by the price cap, and each gas station could be regarded as a price taker. Therefore, the price variation for the same type

[29]Baidu Map API is one of the two most widely used navigation software in China.

of gasoline with a specific time window is relatively small. Figure 4 confirms this pattern. Panel (a) plots the distribution of residuals of the price regression with only date fixed effect, and panel (b) plots the one with date and region fixed effects. Together, two figures imply cross-section price variation comes mainly from the cost difference between region-related standards instead of the pricing strategy difference between each gas station. This fact is crucial for our econometric method that the observed price could be potentially colinear with fuel type indicators. We further discuss this problem in section 3.3 and adopt a two-step estimation approach to address it.

4.3 Descriptive Evidence about Consumer's Choice at the Boundaries

Before formal econometric analysis of reforms, we present descriptive evidence about consumers' choice between two regions and explain how estimations below relate to the observed sales gap at the boundaries.

Beijing and Hebei are closely connected by road networks. There are a large number of highways or motorways that one could easily commute between the two provinces by vehicle. With the high living cost in Beijing, many people choose to live in Hebei and work in Beijing. The cross-boundary cost of switching from one gas station to another is almost zero when driving along the highway. Therefore, commuters make purchasing choices across gas stations in two regions in our model. Figure 7 shows policy-induced discontinuity of price at the border for adjacent regions both selling pure gasoline. Panel (a) stands for before the enforcement of emission standards III $\longrightarrow$ IV in Hebei, and Panel (b) is after the reform. The gas stations in Beijing are presented on the left-hand side of the vertical line at zero. The retail gasoline price in Beijing is higher than in Hebei due to the extra cost for higher standards from the refining process. The price variation between gas stations within the same region is relatively small due to price regulation. Therefore, the cross-boundary price difference provides the main variation for the sales price. The price gap is narrowed down after the reform as the result of increasing Hebei gasoline prices.

Figure 8 plot the daily log sales of regular gasoline for each gas station at the geographic boundary. Panels (a) and (b) show borders, both sides of which sell pure gasoline, while Panels (c) and (d) show borders where stations in Beijing sell pure gasoline and stations in Hebei sell ethanol gasoline. The upper two Panels show a striking discontinuity at the boundary that the sales volume jump upward in Hebei stations just crossing the boundary. We call it the sales gap in the rest of this paper. This evidence suggests that the difference in gasoline makes commuters "bunch" at the Hebei side. The findings is robust to controlling a set of observable covariates (as shown in Appendix figure A.3). The statistical significance of this discontinuity is established using parametric and non-parametric, respectively, in Appendix table A.4 and A.5. Such a sales gap is not detected for ethanol boundaries in the lower two Panels. Recall in Section 2 we discussed that ethanol gasoline is considered a much less substitutable good to pure gasoline, as mixing the two types of fuels or switching from one to another frequently would cause damage to the engine and gas tank. Therefore, the lower two Panels serve as placebo tests to validate that the sales gap is driven by production differentiation instead of other social and economic difference between regions.

Figure 8 provides graphical evidence for consumers' gasoline choice behavior at the boundary and also directly relate to our main empirical analysis. Panels (a) and (b) indicate that the sales gap exists before and after the III $\longrightarrow$ IV reform. Therefore, the intuitive interpretation of our empirical analysis below is to examine how emission standard reforms change the sales gap, which is directly related to the parameter β. The theoretical hypothesis is that if consumers do not value environmental standards, the sales gap will narrow as the price of Hebei gasoline increases. However, if consumers have high WTP for environmental value and its magnitude outweighs the dis-utility from price increases, the sales gap will be enlarged.

5 Empirical Analysis and Results

The target of our empirical analysis is to separately estimate consumers' (dis)utility from price (α) and environmental value (β), which allows us to calculate WTP for higher emission standards. We use the estimating equations derived from the gasoline demand model and implement the two-step estimator as described in section 3. We begin by presenting the price parameter estimation and using these first-step results to construct the new dependent variable of the second step. We then present the main results and the robustness checks for our β estimation.

5.1 Estimation of Price Parameter

5.1.1 Empirical Strategy

We first estimate structural parameter α and the price elasticity of demand, which is interesting in itself, but more importantly, will be used to construct the dependent variable in the second step.

The price elasticity of gasoline has been extensively studied by the literature, while no consensus has been reached. The potential reasons include different frequencies of data aggregation and a lack of exogenous variations to serve as credible instruments of the retail gasoline price.[30] Previous studies attempt to improve the accuracy of estimation on three margins: first, utilizing supply-side instrumental variables, e.g., crude oil production shock (Hughes et al., 2008) and state-level gasoline tax (Davis and Kilian, 2011; Li et al., 2014); second, exploiting higher frequency and more microdata with extensive fixed effects in estimating equations (Levin et al., 2017); third, using fully structured demand model and structural estimation (Houde, 2012). By summarizing the existing studies, we conclude that the distinction between gasoline consumption and gasoline purchasing is central to understanding the demand responses. The impact on total usage is relatively negligible in the short run as consumers are less likely to change their commuting behaviors fundamentally (e.g., switching to public transit or buying more efficient vehicles.). Therefore, the purchasing elasticity dominates the short-run demand response since consumers may delay or postpone purchasing decisions after a price change. This results in the estimation using daily-level data being an order of magnitude large than monthly- or quarterly-level data (Levin et al., 2017). In addition, when gas station-level data instead of city-level data is being used, the estimation tends to be enlarged for another order of magnitude since consumers are extremely sensitive to price differences between gas stations in close proximity (Houde, 2012). Since we are interested in the short-run response to changing gasoline standards, The central parameters studied in this paper are the purchasing elasticity (on price and emission standard). Therefore, we assume that the total gasoline consumption is inelastic, and the reforms do not generate general equilibrium effects on market size.[31]

Taking advantage of the unique feature of the retail gasoline market in China, we use several policy-induced price shocks that are orthogonal to the demand side unobservables as instrument variables (IV). We use two different sets of IV as well as the OLS estimation to show the robustness of our results.

Price shock on August 28th, 2012 — The first set of instruments comes from a policy-induced gasoline price surge in Beijing. As shown in figure 1, the sudden enlarge of the price gap happened on August 28th, 2012, when the gasoline price in Beijing side jumped upward. This price surge is to compensate for the refine cost increases caused by the enforcement of a higher emission standard in Beijing (Beijing IV $\longrightarrow$ V) two months ago. NRDC postponed the price adjustment for the sake of political stability. According to the anecdotal evidence, the standard reform in Beijing came at a time when the retail gasoline price, as well as the international crude oil price, were at their all-time peak (as shown in figure A.1). As a result, NRDC was worried that further increasing the gasoline price might

[30]See Levin et al. (2017) for discussion about the severity of these problems and a decomposition of potential sources of bias.

[31]A similar setting in studying the gasoline market is also used by Houde (2012).

incur consumer dissatisfaction. They therefore postponed the price adjustment to facilitate the smooth implementation of the standard reform.[32] The price surge provides an attractive set of instruments as it only changes the price gap at the border while leaving the other gasoline attributes constant. In addition, the timing of this price surge is less likely to be driven by any other demand-side shifts, and its occurrence is unpredictable for consumers. Therefore, we regard it as our preferred instrumental variables.[33].

We employ a two-stage-least-square (2SLS) estimator. In the first-stage, we capture the price change using a difference-in-difference specification. The second stage equation is based on the demand equation derived in section 3. The estimating equations are,

$$\textit{Second-stage:} \quad log(Q_{jkt}) = \alpha_0 + \alpha \hat{P_{jkt}} + \boldsymbol{X_{kt}} \cdot \delta + \eta_t + \rho_k + u_{jkt} \tag{8}$$

$$\textit{First-stage:} \quad P_{jkt} = \psi_0 + \psi_1 \cdot D_j \, \mathbb{1}(t \geq T^{**}) + \psi_2 \cdot \mathbb{1}(t \geq T^{**}) + \psi_3 \cdot D_j + v_{jkt} \tag{9}$$

where in the first stage regression, $\mathbb{1}(t \geq T^{**})$ is an indicator for post 08/28/2012; D_j is a dummy for the gas station j locating in Hebei. The interaction term $D_j \, \mathbb{1}(t \geq T^{**})$ captures the effects of this price shock. We calculate the fitted value of price from the First-stage and use it as the explanatory variable in the Second-stage. Gas-station level controls $\boldsymbol{X_{kt}}$, time fixed effects η_t, and gas station fixed effect ρ_k are included in the second-stage regression. The value of gasoline, $\boldsymbol{\Lambda_{jt}}$ is absorbed by ρ_k as it does not change during the event window. P_{jkt} is the observed price, and α is the estimated marginal disutility of price in equation 1. When replacing P_{jkt} with $log(P_{jkt})$, the *log-log* estimation gives the interpretation as price elasticity.

5.1.2 Results

Table 2 presents the results of the first-stage regression estimating equation. In columns 1 to 3, we use log price as the dependent variable and add in sequence the date fixed effect, the gas station fixed effect, and the interaction between the distance polynomial and time trend. In columns 4 to 6, we use price as the dependent variable and add the controls in the same manner. This price shock increases

[32] Anecdotal evidence to support it

[33] One caveat of this instrumental variable is that we observe a price surge only for regular gasoline. For premium gasoline and diesel, the reason for not observing such a price surge is that the price difference was already large (see Appendix for detail). Therefore, we assume the price elasticity is the same for all fuels in the benchmark. We show the results using other instrumental variables to validate this assumption.

the *log* gasoline price in Beijing by 1.86% to 2.15% and increases the absolute level of gasoline price in Beijing by 16% to 17.8%, relative to Hebei. All coefficients are statistically significant at the 1% level. To explore the dynamic effects and test for the parallel trend assumption of difference-in-difference, we use a flexible event-study specification that replace the interaction term $D_j\,\mathbb{1}(t \geq T^{**})$ with a series of interactions between D_j and date dummies within the event window $\sum_{t\in W} D_j\,\mathbb{1}(\text{date} = t)$, where W is the span of time window. Figure 5 plots the estimates and 95% confidence interval of the flexible first-stage specification. The omitted day category is Aug. 28th of 2012. It is clear that parallel trend assumption on price gap between two regions perfectly holds, which confirm that the timing of government price adjustment is not driven by other confounding factors. Price in Beijing increases suddenly by about 2% right after Aug. 28th, 2012. Figure 5 also implies a strong first-stage correlation.

The benchmark IV estimations are shown in table 3. The time window used here is 30 days before and after the price shock.[34] Using various specifications, the estimated $\hat{\alpha}$ ranges from -0.725 to -0.550, and the price elasticity is about -2.76. All of the estimates are statistically significant at the 1% level. All columns control for gas station and date fixed effects. And the estimates are robust no matter whether we include the interaction between a polynomial function of distance to boundary and time trend, which implies the results shown here are not driven by some gas stations that are in proximity to boundaries. Interestingly, the estimated elasticity of our preferred specification is close to those results using fully structural model with spatial differentiation (Houde, 2012), and is much larger than the "reduced form" estimations (Levin et al., 2017). Since our source of identification is the changing price gap across boundaries, consumers are sensitive to price changes due to the low switching costs among alternative choices. Therefore, this results highlight the importance of location to gasoline elasticity.

We also compare IV with OLS estimations to better understand the role played by spatial and temporal factors. The OLS estimations with different controls and fixed effects are shown in table A.7. OLS results without date fixed effects are an order of magnitude smaller than the IV estimates, but are consistent with the previous gasoline demand elasticity estimations using high-frequency consumption data (Levin et al., 2017). However, when both gas station and date fixed effects are included (as shown in columns 1 and 4 of table 3), IV and OLS results become similar in magnitude. These findings support that, when the time-fixed effects are controlled, the price variation across gas stations is rather exogenous to demand-side unobservables as the result of the highly regulated nature of China's retail gasoline.

[34]When calculating the price-adjusted sales for the second-step regression, we use $\hat{\alpha}$ by the corresponding event window. For example, in our main second-step results shown in table 4, we use an even window of 70 days. IV estimations using 70 days window are shown in Appendix Table A.9.

The flexible event-study estimates on gasoline sales are shown in figure 6. Even though the sales data are a bit noisier, the parallel trend assumption is not violated, and no clear evidence of a deterministic trend in sales volume before the price surge. This implies that neither the timing for price adjustment is chosen on the basis of the demand-side factor nor consumers anticipate the price shock. The estimates for interaction terms become significantly positive right after Aug. 28th of 2012 and last for at least one month after the price shock. Together, these figures confirm the validity of our IV-DD strategy of estimating the price elasticity.

NRDC's price adjustment announcement — An alternative set of instruments are the regulated price adjustment announcement by NRDC of China. As mentioned above, the NRDC adjusts its "guide" prices twice or three times a month to anchor the international crude oil price. We show sequences for NRDC "guide" prices in Beijing and Brent crude oil price in figure A.1, wave patterns of the two sequences are synchronized, while the "guide" prices are less fluctuated. Since the international market equilibrium is not likely to correlate with the demand of the Beijing and Hebei gasoline market, the price adjustment announcement can be regarded as an exogenous supply-side shifter. We include dummy variables that indicate observes are within one week of the adjustment and its interaction with the magnitude of price adjustment in the left-hand side of first-stage regression. We do not focus on the first-stage results as they are mechanical. The second-stage results are shown in table A.8. The estimated elasticity is -1.9 for regular gasoline and -2.3 for premium gasoline. The IV estimations using NRDC announcement are slightly smaller than those using Aug. 28th, 2021 price surge with 30 days event window. Similar results obtained by different IV strategies make us confident about the accuracy of our elasticity estimation. When calculating the price-adjusted sales in the second-step estimation, we use $\hat{\alpha}$ estimated from the price surge IV but check the robustness of our final results for a wide range of α that cover the results using different methods in this section.

5.2 Consumers' Response to the Emission Standard Reform and WTP

5.2.1 Empirical Strategy

Having estimated the price parameter $\hat{\alpha}$, we calculate the price-adjusted sales volume as: $\widetilde{log(Q_{jkt})} = log(Q_{jkt}) - \hat{\alpha}\, P_{jkt}$, where P_{jkt} is observed price of fuel j at gas station k.[35] We use it as the dependent variable in the second-step estimation. We study the reform in Hebei that improves from gasoline

[35] This operation is equivalent to moving the $\hat{\alpha}\, P_{jkt}$ term to the left-hand side of the empirical demand equation. Bayer et al. (2009) use the similar two-step method when they estimate the coefficient of housing services price.

standard III to standard IV in our baseline results.[36] It is because previous study has estimated the impact of this reform on air pollution as about 12.9%, which gives us a sense about the magnitude of the reform (Li et al., 2020).

Specification 1: The estimating equations 10 is the empirical analogy of the log-linear gasoline demand model 6, which takes the form similar to the difference-in-difference specification,

$$\widetilde{log(Q_{jkt})} = \beta_0 + \beta \cdot D_j \cdot \mathbb{1}(t \geq T^*) + \lambda \cdot \mathbb{1}(t \geq T^*) + \phi \cdot D_j + \boldsymbol{X_{kt}} \cdot \delta + \eta_t + \rho_k + \xi_{jkt}, \quad \text{for } |distance_k| < h;\ |t - T^*| < \tau \tag{10}$$

where D_j equals 1 for Hebei, and equals 0 otherwise; $\mathbb{1}(t \geq T^*)$ is a indicator function for the post-reform period; $\boldsymbol{X_k}$ is the vector of gas station characteristics; ρ_k is gas station fixed effect that captures time-invariant gas station unobservables; ξ_{jkt} is the error term. The identification exploits daily-level variation among gas stations, therefore potential bias could only arise from daily level omitted variables. The primary threatens are seasonality and holiday-related events. For example, if the implementation of new emission standard overlaps with holidays or extreme weather conditions, we could mistakenly attribute these seasonality effects to the policy effects. Therefore, we include date fixed effects η_t to controls for the seasonality. Meanwhile, to allow for the seasonality effects vary across gas stations, we include two sets of controls: $\rho_k \times$ *day-of-week*, where *day-of-week* $= \{Monday, ..., Sunday\}$; and the interactions of gas station characteristics and time trend. To capture the changing of unobservables over time, we include in $\boldsymbol{X_k}$ the interaction term between the distance-to-boundary polynomial function and the post reform indicator, $\boldsymbol{f}(distance_k) \cdot \mathbb{1}(t \geq T^*)$.[37] This idea is borrowed from the regression discontinuity (RD) approach, which use the polynomial function of running variable to control of the unobservables near the cut-off. Including $\boldsymbol{f}(distance_k) \cdot \mathbb{1}(t \geq T^*)$ plays a similar role as to control for interaction of pre-existing covariates with time trends proposed by Gentzkow (2006). We check validity of RD assumption by showing that after controlling for polynomial function $f(\cdot)$, no covariates "jump" at the boundary in Appendix Figure A.5 and Table A.3.

Since cross-boundary choice behavior is our main interest, we restrict our sample to gas stations near the city boundaries. We choose a distance bandwidth h as 30 km around the boundary in the benchmark. We show the estimates are robust to a wide range of different h. Following the existing

[36] We show the results of other two reforms happened during our sample period in the robustness check.

[37] $\boldsymbol{f}(\cdot)$ is a polynomial function of order 2: $\boldsymbol{f}(distance) = \mu_1\, distance + \mu_2\, distance^2 + \mu_3\, distance \times D_j + \mu_4\, distance^2 \times D_j$. Including the interaction between distance polynomial and indicator variable D_j allows for a more flexible fitting curve that the slope differs at the two sides of the provincial border. Following Gelman and Imbens (2019), we do not use polynomial of order higher than 3 to avoid putting too much weight on observations that are far from the border.

studies that use daily-level data, we further restrict the sample within an event time window τ. Avoiding using a sample far from the reform date would reduce the probability of unobserved shocks to confound our estimates (Hausman and Rapson, 2018).[38] The coefficient of interest is β, which is a structural parameter capturing how much consumers value higher emission standards.[39] We interpret it as the environmental value since the reforms only change environment-related characteristics of gasoline.

Specification 2: In our context, both gasoline standard reforms in Hebei (III $\longrightarrow$ IV and IV $\longrightarrow$ V) were implemented at the end of the calendar year. A severe threat to our identification arises if other shocks and events change relative gasoline consumption behavior in two regions around the same period as of gasoline reforms. For example, the new year festival could increase the outflow traffic volume from Beijing, and "two meetings" that happen at the beginning of March might increase the traffic inflow to Beijing.[40] To control for these potential calendar-year confounding factors, we pool together the data of the treated year and the previous year, and use the previous year as the control group.[41] The assumption here is that all confounding shocks at the end of the calendar year generate the same impact every year, which could be ruled out using the control year. Our Specification 2 takes the form close the triple difference (DDD) estimator,

$$\begin{aligned} \widetilde{log(Q_{jkt})} &= \beta_0 + \beta_1 \cdot D_j \cdot \mathbb{1}(t \geq T^*) \cdot treated\_year_t + \beta_2 \cdot D_j \cdot \mathbb{1}(t \geq T^*) \\ +\lambda_1 \cdot \mathbb{1}(t \geq T^*) \cdot treated\_year_t &+ \lambda_2 \cdot \mathbb{1}(t \geq T^*) + \phi_1 \cdot D_j \cdot treated\_year_t + \phi_2 \cdot D_j \\ &+ \boldsymbol{X_{kt}} \cdot \delta + \eta_t + \rho_k + \xi_{jkt}, \quad \text{for } |distance_k| < h;\ |t - T^*| < \tau \end{aligned} \tag{11}$$

where, *treated_year* equals one if the observation is within the event window of the reform date for "treated" year. We include gas stations-by-event-year fixed effects as well as all other controls in Specification 1. The coefficient of interest β_1 can be interpreted as the relative gap between β estimated

[38] Since we only use a sample within a small window, gas station-by-month and by-year fixed effects are not included here.

[39] Even though our estimating equation has a similar formula as the canonical difference-in-difference estimator, the interpretation is fundamentally different. Since the reform changes the relative utility between fuel L (the "treated unit") and H (the "control unit") and impact sales volume of both gasoline types, the assumption of stable unit treatment value assignment (SUTVA) is apparently violated. Therefore, our estimand should be interpreted as structural parameter instead of the Local Average Treatment Effect (LATE) We elaborate this issue in Appendix D.3.

[40] The "two meetings" are known as the National People's Congress (NPC) and the Chinese People's Political Consultative Conference (CPCC). During the "two meetings", the traffic inflow to Beijing will surge, and the government will start to implement driving restrictions and additional security checks several weeks ahead. These events might alter commuters' travel behaviors.

[41] For example, for the III $\longrightarrow$ IV reform and event window $\tau = 70$, we keep sample range from Sept. 24th, 2013 to Feb. 11th, 2014 and from Sept. 24th, 2012 to Feb. 11st, 2013, and pool them together.

from Specification 1 using treated year and control year sample, respectively.

Before moving to the results, we briefly discuss other events and policies that could potentially confound our estimates. First, several festivals are within our event window (e.g., the National Day and Lunar New Year). If the effects of festivals on gasoline consumption are identical across the year, they should be ruled out in Specification 2. However, if these effects differ between treated and control years, it could bias our estimates. Therefore, we further control the effects of festivals by including $D_j \cdot festival_t$ in Specification 1 and $D_j \cdot festival_t \cdot treated\_year$ in Specification 2. In addition, there are two confounding policies that are directly related to vehicle emission and could potential bias the effects of III $\longrightarrow$ IV reform: implementing vehicle exhaust emission standard IV and scrapping the "yellow label vehicles".[42] Based on the circulars of Beijing transportation, both policies took effect in Beijing on April 11th, 2014. Therefore, we restrict our event window in the benchmark results to ensure these two policies do not impact our analyses. In addition, we examine consumers' response to both III $\longrightarrow$ IV and IV $\longrightarrow$ V reform, and the results are consistent across two reforms. Therefore, our finding is not affected by those confounding policies.

5.2.2 Benchmark Results

Table 4 reports estimation of β, which shows consumers' value on improving gasoline emission standard III $\longrightarrow$ IV.[43] The sample includes both types of gasoline, regular and premium. Across all columns, we cluster the standard error to the gas station level. We also report the cluster bootstrap standard error in the bracket to deal with the potential estimation error in the first step. We use -0.376 as the value of $\hat{\alpha}$, which is the IV estimation with a full set of controls and fixed effects and the same event window (70 days).[44] As we illustrate in Appendix D.3, the estimated β increases with absolute value of α. Therefore, we use a relatively conservative estimate of price coefficient. We also show the results using different $\hat{\alpha}$ in the robustness check. The corresponding WTP is calculated as $-\beta/\alpha$.

Columns 1 through 3 reports the results of Specification 1. Column 1 uses the parsimonious specification of estimating equation 10, where fixed effects and controls are not included. Column 2 controls

[42]The exhaust emission standard IV policy of Beijing restrict the vehicle with exhaust standard lower than IV from entering Beijing (Source: http://fgcx.bjcourt.gov.cn:4601/law?fn=lar787s428.txttruetag=859titles=, in Chinese). The "yellow label vehicle" refers to gasoline vehicles that do not meet the national exhaust emission standard I and diesel vehicles that do not meet the standard III. The policy forcing to scrap 1 million "yellow label vehicle" every year (Source: http://fgcx.bjcourt.gov.cn:4601/law?fn=lar810s109.txttruetag=999titles=, in Chinese).

[43]Appendix figure A.13 plots a graphical presentation of the benchmark estimation and explains how does β relate to the cross-section sales gap at the boundary

[44]The IV results using 70 days event window are shown in Table A.9.

for date fixed effects and gas station fixed effects. The coefficient of interest is about 0.11, and is positive and statistically significant. In column 3 we further add gas station-by-day-of-week fixed effects and $\mathbb{f}() \times Post$ to account for time-variant unobservables. After including these additional controls, the magnitude of the result drops, and the estimation becomes close to significant. It implies that the time-variant gas station unobservables confound our raw estimates. However, the direction is still consistent with previous columns.

Columns 4 through 6 take care of the end-of-calendar-year effects using Specification 2. If the other confounding events happened both years, our specification would cancel out its impact on outcome variables. Therefore, the variable of interest is now the triple interaction terms $D \times Post \times treated\_year$. Fixed effects and controls follow a pattern similar to columns 1 to 3. We also include the treated year dummy and its interactions with the Hebei and post-reform dummies. The results are in general robust to various fixed effects and controls. Column 6 includes a full set of controls and is our preferred specification. The estimation shows that after controlling for the concurrent price change and other confounding factors, the enforcement of higher fuel standards increases the sales gap (Hebei$-$Beijing) the border by 13.8 log point.[45] The results are statistically significant at 1% level and robust to use clustered bootstrap standard errors. Correspondingly, consumers' WTP for higher emission standards is 0.205 US$ per gallon (0.345 CNY per liter)[46], which accounts for 4.66% of the total retail gasoline price.

5.2.3 Parallel Pre-Trends

A potential threat to our identification is that government chooses the timing of the standard reform based on, or coincides with, other demand- or supply-side factors that are not captured by our controls. It means the sales gaps between two boundaries exhibit secular trends before the reform. Even though our method is not a pure diff-in-diff estimator, violation of the parallel pre-trends assumption would still prevent us from interpreting our structural parameters to consumers' preference on higher emission standard. Therefore, we exploit a flexible event-study to check the validity of the pre-trends assumption. We first aggregate our data at weekly level to provide a smooth graphic presentation. Then, we replace the triple interaction term $D_j \cdot \mathbb{1}(t \geq T^*) \cdot treated\_year_t$ in equation 11 with a set of interactions between $D_j \cdot treated\_year_t$ and week dummies within the event window $\mathbb{1}(t = Week_i)$, where $i \in \{-20, ..., 30\}$. Figure 9 presents estimated coefficients along with 95% level confidence intervals for each week within

[45] This approximates to $\exp(13.8) - 1 \approx 14.8$ percentage point.

[46] Within our sample period, the exchange rate of USD to RMB was quite stable and fluctuated from 0.15 to 0.16 Dollar per Yuan. In this paper, we always use 0.16 Dollars per Yuan for calculation.

the event-study window.[47] The omitted time category is the week before the transition period. The grey region represents the transition period when the gas stations in Hebei sell both III and IV gasoline. Figure shows that estimates before the reform are small and statistically indistinguishable from zero, implying no significant pre-reform trend of sales gap. This finding strongly supports that the parallel trend assumption is not violated. Two pieces of anecdotal evidence further support that reform timing is quite exogenous. First, both Hebei's III to IV and IV to V reforms are officially announced to take effect on Jan. 1st. Since it is a typical date for Chinese government to implement its new policies or reforms, this timing choice is not likely to be driven by any confounding factors. Second, we interviewed the staff of *PetroChina* about the refining capacity. As far as our information, unified refinery plants of *PetroChina* supplies retail gasoline for both Beijing and Hebei, meaning there is no regional difference in refining capacity. And the timing of the enforcement of higher emission standards are not subject to supply-side technology constraints, e.g., the desulfurization and catalytic hydrotreating capacity.

5.2.4 Dynamic Effects

Figure 9 is also informative to the dynamic pattern of consumers' responses. First, the relative sales gap remain unchanged during the transition period even though some Hebei stations have started to supply gasoline with the new standard. Therefore, it is the announced reform dates, instead of the *de facto* implementation dates, that matters. It confirms that consumers make purchasing decisions based on their ex-ante knowledge about gasoline quality. Second, the estimated coefficients become significantly positive after the official implementation of the new standard and fully cover the Hebei region. The results are stable and persist for at least 30 weeks after the reform, implying that our benchmark estimation is unlikely to be confounded by other shocks within a short event window. We also present the event-study estimates based on Specification 1 in Appendix figure A.9. The estimated effects surge from four weeks after officially implementing the new standard and peak at six weeks. The period of the mountain-shape pattern overlaps with the driving restrictions before the "two meetings", which could be an potential confounding event. To confirm this, we conduct a bunch of placebo tests in Appendix figure A.10, including two placebo years (end of 2012 and 2014); gas stations that are *not* located near the border; gas stations located near the border in *ethanol region*. We detect such "peak" for all placebo groups, which suggests that the driving restriction of "two meetings" indeed confounds our main results under Specification 1. These results support to use Specification 2 as our preferred

[47] We plot the dynamic effect of reform on gasoline price in the Appendix Figure A.6. The price increment is symmetric across all stations. As shown in Appendix Figure A.7, most stations maintain their rank in price after the reform.

specification in this study.

5.2.5 Transition Period

As mentioned before, the official implementation date of the new standard is Jan.1st, 201. However, most Hebei gas stations in our sample started to supply new gasoline in December 2013.[48] In the benchmark, we use the same T^* across all gas stations. Since both demand and supply-side factors could potentially change during the transition period, we perform two robustness checks to show our results are not sensitive to the definition of the transition period. First, we drop the sample in whole transition period and use the officially announced date as T^*. The results are shown in Appendix table A.10. The estimates across all columns are in the same direction and become larger in magnitude than benchmark results. It suggests that including the transition period provides a rather conservative estimation. Second, since our data allow us to identify the *de facto* switching date of gasoline standard for each station, we use the *de facto* dates as T^*.[49] As shown in Appendix table A.11, the results are similar to the benchmark.

5.2.6 One-step Estimation

In the main estimation, we adopt the two-step approach to address the multicollinearity problem mentioned in section 3.3. In this section, we further use the one-step approach to probe the robustness to methodology choices. To this end, use $log(Q)$ as the dependant variable, we estimate α and β simultaneously in one regression. Fixed effects and covariates controlled in columns 1 to 6 in Table 5 are the same as in the corresponding columns in Table 4. Columns 1 to 3 display the results based on Specification 1. The one-step results for Specification 1 is sensitive to include fixed effects. In columns 2 and 3, the price coefficient α is incorrectly estimated since the result becomes positive and statistically insignificant. Columns 4 to 6 present the one-step counterparts for Specification 2. The estimated β across all columns are positive statistically significant. The magnitude are consistently larger than those from the two-step approach. In column 5, the magnitude of α is close to what we get in the first step IV estimation, even though not statistically significant. Correspondingly, the estimation of β is close to the two-step results. Therefore, consistent with our theoretical analysis, β is biased in the opposites direct as α. When α is correctly estimated, β is not biased. We present the event-study estimation

[48]We define the transition periods starting from new gasoline being available untill no old gasoline being sold in Hebei.

[49]For each station, the *de facto* date is the date when new gasoline accounts for over 90 % of total gasoline sales within a single day.

using one-step method in Appendix Figure A.8. The results are consistent with Figure 9, while the standard errors are larger. These results illustrate that our benchmark findings are not manipulated by methodology and support our choice of using a two-step approach.

5.3 Robustness Checks and Heterogeneous Effects

5.3.1 Sensitivity

In this section, we explore the sensitivity of the benchmark estimation on values of $\hat{\alpha}$ and examine the robustness of our preferred results to different distance bandwidths, the event time windows.Table 8 presents the second-stage results using different $\hat{\alpha}$ range from -0.1 to -0.6. As $\hat{\alpha}$ the price elasticity becomes larger, the β estimated increases from 6.65% to 17.1%. The relationship between β and α is consistent with our theoretical prediction. Meanwhile, the corresponding WTPs are relatively stable for a wide range of $\hat{\alpha}$. Together with one-step estimations, these results suggest our main findings are not sensitive to the choice of method.

Figure 11 shows how β changes with varying bandwidths of distance to the boundary. We start from 10 km to guarantee enough gas stations in our sample. For each bandwidth, we conduct two estimations: the first one uses the same distance bandwidth in both steps of estimation; and the second one alters bandwidth for the second stage while keeping $\hat{\alpha}$ same as the benchmark. The first result is demonstrated by the black line, and the second one the grey line. In the second estimation, β is robust to changing distance bandwidths. In the first estimation, β is slightly sensitive only when the bandwidth is lower than 25 km. Specifically, β becomes higher as the bandwidth gets narrower. The comparison between the two lines suggests that the estimated price elasticity decreases with distance bandwidths. It implies that consumers closer to the boundary are more sensitive to the price gap, which is consistent with the existing studies about spatial differentiation in the retail gasoline market (Houde, 2012).

Figure 10 displays the estimation of β using different time windows, ranging from 30 days to 100 days round the reform date. The β increases with the length of the event window and peaks at the 55 days. As we mention above, the first four weeks after the reform are considered as transition period and no effects are detected during the transition period. As the results, extending post-reform time window would increase the estimated average effects. The β then decreases till the 70 days time window and becomes stable from their, which implies persistent long-run effect. Therefore, we use 70 days as our preferred event window.

5.3.2 Placebo Tests

As falsification tests, we present four set of placebo results. First, we consider the boundaries where the Beijing side sells gasoline while the Hebei sides sell ethanol gasoline (E10). As discussed in section 4, consumers are less likely to mix the two types of fuels or switch from one to another frequently since it could cause gas tank damage. Therefore, fuels on the two sides of these boundaries are less substitutable. We expect no effects of the standard reform on gas stations that sell ethanol gasoline. The results is shown in column 1 of Table 7. Second, we look at those non-boundary gas stations that are located more than 30 km away from the boundary. Since consumers will not substitute between gas stations far from the border, we expect no effects on non-boundary stations, either. The result is shown in column 2. Third, we use diesel as a placebo fuel type to rule out the other aggregate shocks. Recall that the emission standard of diesel are no changed during the gasoline reform. Therefore, we should detect no effects for diesel sales during the gasoline reform. The result is shown in column 3. Finally, we use the end of 2012 and 2014 as placebo event date when no reform happened on either side of the boundary. The estimation use Specification 1 and is shown in column 4. Across all columns, the estimated coefficient of placebo test is small in magnitude and statistical insignificant. In Appendix Figure A.11, we present the event-study results using ethanol boundary and non-boundary sample.

5.3.3 Results of Different Reforms.

As mentioned in section 2, there are three different standard reforms implemented within our sample period. In the main estimation, we focused on the III to IV reform in Hebei at the end of 2013. In this section, we examine the effects of other two reforms using the same specification. The results are presented in Table 9. In columns 1 and 2, we test the consumers' response to the first reform in our sample, implemented on June 21st, 2012, that increases the standard in Beijing from IV to V. This reform differs from the two in Hebei in several aspects that readers should keep in mind when comprehending the results. First, there is no cost-induced price change associated with the reform, with the reason mentioned in Section 2. Second, our sample period starts from Jan. 1st, 2012, which means we cannot use the previous no-reform year as the control group in Specification 2. As a compromise, we use the same dates in the year after to control the calendar-year effect. The assumption here is that no shocks happened within the time window around the middle of 2013. But as there might be unobserved changes after the Beijing reform in 2012, which affect the 2013 outcomes, we should interpret the results with caution. The coefficients of interest are both negative in the first two columns, though no statistical significance is achieved in column 1. Recall that we define the sales gap as Hebei minus Beijing. The negative estimations suggest that gas stations in Beijing become more attractive to consumers relatively to Hebei after the reform. The results are consistent with our benchmark findings that consumers care about the environmental value and respond positively to standard upgrades. In columns 3 and 4, we test the effects of standard reform IV to V in Hebei, which happened at the end of 2015. We use the end of 2014 as the control year in Specification 2. The coefficients using two specifications are positive and statistically significant at a 1% level, indicating that consumers at boundaries further move to Hebei in respond to its second gasoline standard upgrade. Overall, the estimates across three different reforms indicate that consumers behave consistently. In comparison, the effects of the Beijing reform are less significant than Hebei reforms. One possible explanation is the asymmetric effects of the standard upgrade, as Beijing's standards are always higher than Hebei's. Another potential explanation is the widespread media coverage on air pollution since 2013 increased public awareness on pollution issues and WTP for environmental standard. (Ito and Zhang, 2020).[50]

5.3.4 Heterogeneity by Grades

In this section, we investigate the heterogeneous effects on different gasoline grades. The first two columns of Table 6 shows the results using regular and premium gasoline sample separately. The result on both grades are significant and similar in magnitude, which implies that our benchmark results are not driven by a specific grade of gasoline. Specifically, the coefficient of regular gasoline is 0.124, corresponding to the WTP as 4.471% of the total regular gasoline price. The estimation of premium gasoline is a bit large, corresponding to the WTP of 5.197% of price. In practice, premium gasoline consumers are more likely to drive more luxurious vehicles and thus wealthier, these heterogeneous effects provide suggestive evidence that WTP for environmental value increases with wealth level. However, as not all stations sell regular and premium gasoline simultaneously, we avoid pushing too much on this comparison.[51]

[50]Since 2013, the Chinese government has started its PM2.5 information disclosure program, and Chinese media remarkably increase the coverage of air pollution issues. The impacts of environmental information exposure on public awareness and WTP for clean air have been documented by recent literature (Gao et al., 2021; Ito and Zhang, 2020; Wang and Zhang, 2021).

[51]In our sample, nearly one-third of gas stations only sell regular gasoline, most of these stations are located in Hebei region.

5.3.5 Heterogeneity by Distance to Main Roads

To further understand which gas stations are driving main results, we investigate the heterogeneity effects of gas station locations. The results are presented in Table 10, with columns 1 and 3 using Specification 1 and columns 2 and 4 using Specification 2. Columns 1 and 2 include gas stations close to the expressways and national highways (main roads to connect two regions). Most of these gas stations are located in the rest area beside main roads. The potential customs of these stations are commuters traveling between two regions. Columns 3 and 4 use gas stations that are far from main roads[definition?]. The potential customers of these stations are more likely to be local residents. The effects on gas stations near the main roads are much larger and more statistically significant than those not near the main roads. The reason could be that commuters, on average, have smaller across-boundary switching costs than local residents. These results suggest that consumers' response to standard gasoline reform is more prominent for commuters, and the stations beside the main roads mainly drive our empirical findings.

5.4 Mechanisms Analysis

By far, our empirical findings suggest that consumers are willing to pay higher price for gasoline with a higher environmental standard. In this section, we shed light on the potential interpretations on this result. The main mechanism we try to argue is that people gain private utility from consuming more environmental-friendly gasoline — green preferences. The existence of such pro-environmental behaviour in consumers' decision making have been documented in retail electricity market (Ma and Burton, 2016); green electricity programs (Kotchen and Moore, 2007a); fuel-efficient vehicle sales (Kahn, 2007); hybrid vehicle purchases (Sexton and Sexton, 2014); electric vehicles purchases (Hidrue et al., 2011), and etc. In supporting this mechanism, we first provide direct evidence to show that consumers' responses are not driven by factors other than environmental value. Next, we conduct an international comparison using consumers survey data to show that Chinese consumers have high valuation of environment protection and strong preference for environmental-friendly goods. Finally, we rule out the impact of other potential mechanisms.

5.4.1 Environmental Value or Other Attributes

In this section, we argue that consumers' behavior are not likely to be driven by other attributes and misperceptions. First, suppose that the emission standard reforms also changed other attributes of gasoline, in that case, consumers' preference towards higher standard gasoline might reflect their preference other than environmental value, e.g., higher miles per gallon and lower damage to the engine. However, engineering literature and technical documents suggest that gasoline with more stringent emission standards has lower RON and lower combustion efficiency — and thus lower miles per gallon, instead of higher. It implies that rational people tend not to purchase high-standard gasoline if they only consider non-environmental related attributes. Second, if consumers have systematic misperceptions or inattention about the content of emission standard reform, the observed response might be different from their optimal response (Allcott and Taubinsky, 2015). In our study, this is unlikely to be the case. The government announce information on improving gasoline standards long before its implementation (detailed information and cite) and reform has received a great deal of media attention. In addition, our data show that each gas station in Hebei set transitional period, range from several days to several weeks, to make sure consumers are fully aware of the changing emission standard. Moreover, anecdotal evidence confirms consumers' belief about reforms to lower the non-environmental utility of gasoline in terms of fuel efficiency or engine power. For example, an article published by Sina Finance reports that the sulfur decreased by 70% after the III-IV reform at the beginning of 2014, the MPG (Mile per Gallon) has reduced by 2%.[52]. Another one published by Sina Vehicle said that the lower MPG and higher cost from the reforms "are frustrated". In Appendix ?, we present the collection of newspaper cover stories discussing that the decreasing MPG and engine power after the reform led to negative feedback of consumers complaining about the new standard gasoline. The articles are all collected from the most influential websites in China.

Third, we directly test consumers' valuation on non-environmental attributes by looking at the standard reform III $\longrightarrow$ IV of diesel. In our context, over 90% of diesel vehicles are cargo and trucks for business usage (Ministry of Ecology and Environment of PRC, 2016). Therefore, most diesel consumers are profit-chasing companies and are less likely to act pro-environmental behaviors than gasoline consumers.[53] So our hypothesis is, if it is the green preference that mainly determines consumers' response, the impact of diesel reforms should not be as large as gasoline reforms. However, if reforms also improve other attributes that consumers value, such as higher fuel efficiency or less damage to the engine, diesel truck owners would respond similarly or even more strongly than gasoline vehicle owners. In column 3

[52] *Sina* is among the largest portal websites in China, and *people.cn* is a large-scale information interaction platform constructed by People's Daily, the most authoritative newspaper in China.

[53] The question of what shapes pro-environmental behavior is thought to be a complex one. Altruism, empathy, and pro-social behavior are among the major factors that drive the pro-environmental behavior (Kollmuss and Agyeman, 2002). In general, our assumption here is consistent with the pro-social behavior models of pro-environmental behavior.

of Table 6, we investigate the impact of diesel reform using the same specification as in benchmark. The results show that, after the reform, the price-adjusted sales gap of diesel between Hebei and Beijing decreased 11.8%, and it is statistically insignificant. Compared to gasoline consumers, the response of diesel consumers support that the environmental value, rather than other characteristics like fuel efficiency, is the major driving factor of consumers' shifting purchasing behavior after the reform.

5.4.2 Do Chinese Consumers Care about the Environmental Value?

We use various of consumer survey data to show that consumers in China have a strong intention to pay higher prices for environmental-friendly goods. According to Global Sustainability Study 2021 shown in A.12 , 43% of respondents in China had a total or significant lifestyle change in purchase behavior, which turned their way around to be sustainable in the past five years. And 51% always or only buy sustainable alternatives if available, ranking the second highest and the highest, among the 17 countries in the survey. Global Sustainability Study 2021 is conducted several years later than our sample period, and its sample might contain a high share of younger generations with high education and income levels. We also look at the World Value Survey, which contains multiple waves conducted every five years in 120 countries. We specifically use the data from Wave 6 conducted from 2010 to 2014.[54] From Figure 12 we can see 56.6% of respondents in China believe protecting the environment should be given priority, even if it causes slower economic growth and some loss of jobs, implying they value environment over the economics outcomes and utility. This ratio is higher than those in most developed countries listed in the table. It is only lower than that in Sweden, a Northern European country known for its high valuation of environment protection. India has a green consumer ratio as high as China, suggeatin

5.4.3 Cross-grades Choices

In our demand model, we assume no utility term for cross-grades difference, and therefore consumers make cross-boundary purchasing choices only within the same gasoline grades (regular or premium). However, the change of cross-grades choice is a mechanism preventing us from interpreting consumers' responses solely as their preference for environmental value. For example, if consumers who would have bought regular gasoline from Beijing choose to buy premium gasoline from Hebei, we would get the same empirical findings. In practice, changing gasoline grade is not common since drivers tend

[54]Wave 6 is conducted in varying years for different countries: in China, it was in 2013, the year of III to IV reform in Hebei; for rest countries listed here, Spain, Sweden, United States, 2011; India, Netherlands, Poland, 2012; Germany, 2013.

to choose the gasoline grade that is recommended for their vehicle. To further address this concern, we check whether the sales gap between regular and premium gasoline changed significantly after the reform . The results are shown in Table 11. In columns 1 and 2, we redefine D as an indicator of regular gasoline and use Specification 2 on Beijing and Hebei samples, respectively. We do not find evidence of changing the regular to premium ratio that is close to significant for both regions. In column 3, we keep only the treated year sample and check whether the relatively regular vs. premium gap between two sides of boundary changes. The result shows that the relative gap remains statistically constant after the reform. In our sample, *de facto* starting date of supplying new regular is earlier than premium gasoline for a given station. Exploiting this fact, in the last column, we examine the cross-grades purchasing behavior of Hebei consumers during the transition period. The result indicates no evidence of consumers switching to Standard IV regular from Standard III premium gasoline. These results suggest that cross-grades purchasing is not a big concern for interpreting our main results as WTP for environmental value.

6 Policy Implications and Comparisons

6.1 Cost-Benefit Analysis of Standard Upgrades

To evaluate the social benefit and cost of emission standard reforms, we provide a brief back-of-envelope calculation and discuss the policy implication of the WTP estimated. Under our preferred specification, the estimated WTP of the gasoline III to IV reform for the regular gasoline is 0.345 CNY/liter (0.204 US$/gallon), which accounts for about 4.65% of its total pre-reform price. The reported total gasoline consumption of China is 107.89 million metric tons in 2013 and kept gradually increasing since then.[55] Therefore, the aggregated consumers' WTP for reform is 49.44 billion CNY (7.91 billion US$) per year.[56] Note that it is consumer's private value from using cleaner gasoline, the green preference, without accounting for benefits from environmental and health improvements. In a paper that studies the same gasoline standard reform in China, Li et al. (2020) estimate the public benefit from reducing lifetime mortality and morbidity to be about 30.09 billions US$. Our finding highlights that consumers' private value is quantitatively important for evaluating the social benefit of environmental regulations.

[55] Source: *U.S. Energy Information Administration (EIA)*

[56] Calculation: $49.44 = 0.10789 \times 1328 \times 0.345$. The volumetric mass density of gasoline (and diesel) fluctuates varies by grades and ambient temperature, here for simplicity, we use 1 tonne = 1382 liter from BP p.l.c.'s Statistical Review of World Energy.

Based on our estimation, the relative price increase associated with the reform is 0.183 CNY per liter, accounting for about 2.48% of the gasoline price. The gap between WTP and price increment implies that the consumer surplus increases after the reform. However, it doesn't necessarily suggest socal welfare gains, since the increasing refining cost do not one-to-one pass through to consumers. In practice, the state-owned retailer shoulder part of the cost increment of the new gasoline, which can be regarded as the government subsidy. Therefore, to estimate the social cost of new gasoline we need to infer the refining cost of gasoline emission standard. According to the NDRC statement, the cost increase from standard upgrades should be shared between retailers and consumers, with the retailers taking 20%-30% of the burden.[57] If we take the upper bound and assume consumers shoulder 70% of the cost change, then the overall cost increase should be 0.261 CNY per liter, which is still lower than the private WTP. The aggregated social cost per year is thus 37.40 billion CNY (5.98 billion US$) per year. Across our calculation, the benefit-cost ratio is likely to be a lower bound. Our results suggest that consumers' aggregate WTP outweighs the increase in refining costs. when considering only consumers' private value, the gasoline standard reform in China tends to improve the social welfare.

In this section, we only consider the direct comparison between aggregated WTP and social cost for new gasoline, which we regard as the first-order welfare effects. In the Appendix E, we further discuss the distortion effects as the results of government subsidy. The idea is that the subsidy distort the market equilibrium by pushing more consumers to purchase the Hebei gasoline that is dirtier in the absolute sense.

6.2 Fuel Standard Policies in Other Countries

6.2.1 Gasoline Standards Reform in European and Indian

As mentioned in section 2.3, the Chinese gasoline standards are constructed based on European gasoline standards. Besides, India also builds their standards based on European ones. As shown in Table A.1, the differences between Chinese, European, and Indian standards are minor for almost all specifications. Therefore, we conduct a cross-country comparison to inform the magnitude of aggregated WTP in China. The Europe Union implemented its gasoline standard III to IV reform in 2005, and India in 2017. Consumers' altitude toward environmental protection and willingness to consume environmental-friendly products may differ across countries. Unfortunately, There is no existing litera-

[57]The NDRC statement (in Chinese) mentioned that the cost changes should be shared by firms and consumers. The ratio is confirmed by the director of department of price of NDRC (Source: http://www.nea.gov.cn/2013-12/23/c_132989019.htm, in Chinese).

ture investigating the WTP for cleaner gasoline in Europe or India. Instead, we adjust the consumers' WTP in this paper with consumers' green preference measured the World Value Survey (WVS). We use the proportion of pro-environmentalism respondents as a proxy of consumers' valuation of environmentalism in each country. The results are shown in Panel A of Table 12. The aggregated WTP depends on consumers' green preference and the total gasoline consumption.

Based on the WVS, 58.4% of Indian consumers regard protecting the environment a priority over economic growth, which might due to the severe environmental issues in India. The proportion of EU is 44.0%. The total gasoline consumption in the reform year is 17.26 and 61.77 million metrics ton for India and EU, respectively, while the number in China is 107.89. As the result, the aggregated WTP is much higher for gasoline standard reform in China than EU and India. This simple comparison illustrate that as the world's second largest consumer of gasoline, China benefits substantially from improving gasoline standard.

6.2.2 Low Carbon Fuel Programs in North America

To reduce greenhouse gas emissions and pursue carbon neutrality, U.S. and Canadian government also implemented low carbon fuel standard programs. Unlike China and E.U., the clean fuel programs only target carbon emission reduction and have no requirement on air pollutants such as sulfur dioxide or Particular Matter. The ways to curb carbon pollution in these programs include blending bio-fuel and constructing a credit transacting market for fuel providers. To inform how large is the cost-benefit ratio of China's gasoline standard reform, we compare it with Oregon Clean Fuels Program (CFP) in 2020 and Canada Clean Fuel Standard (CFS), 2021 to 2040. Note that the benefit we calculate for China is for consumers' private WTP, while the social benefit of the other two is mainly from reducing social costs of Greenhouse carbon. The cost of improving the gasoline standard is much higher in China ($0.154/gallon) compared with Oregon states of U.S. ($0.037/gallon). The reason is that gasoline reforms are more substantial in China and involve upgrading the refinery equipment and technology. The estimated benefit-cost ratio for Oregon CFP is about 2.2, and for Canada, CFS is about 1.4. Based on the estimation of the fuel program in North America, our benefit-cost ratio of about 1.3 should be regarded as in a reasonable range.

6.3 WTP for Clean Energy from Existing Studies

To better understand the magnitude of our estimation, we compare our results with the WTP of green energy and other eco-label goods in the existing literature in Table 13. Note that the interpretation of WTPs are not same as they measure for different energy products in different countries. The purpose of this comparison is to check whether our estimated WTP lies within a reasonable range among the existing studies. To make such a comparison meaningful, we calculate the ratio of WTP to the price or monthly bill for most studies. Guo et al. (2020) finds urban residents on average are willing to pay 65 CNY (10 USD) each year to improve air quality to WHO standards, implying a WTP of 260 CNY each year for a household of four people. Among green electricity literature, the existing papers find households are willing to pay 5% to 16% of monthly bill more to use more renewable electricity, depending on the context and the proportion of renewable energy source (Guo et al., 2014; Kotchen and Moore, 2007b; Mozumder et al., 2011; Roe et al., 2001). In addition, previous papers consumers' WTP for Eco-label as about 10.06% and 15.5% for vehicle and toilet paper, respectively. Compared to these estimated WTP, our estimation as 4.65% of gasoline price is a relatively small numbers. This could because that the environmental improvement is more substantial in their context or we adopt lower-bound estimation in our paper.

7 Conclusion and Discussion

Fuel standards and gasoline content regulations are widely adopted by policymakers to reduce emissions and air pollution. An important but understudied question for policy design is measuring consumers' private willingness to pay for higher gasoline standards. This paper presents the first estimation of consumers' willingness to pay for gasoline with higher emission standards. Our empirical context is the retail gasoline market in China, where the government regulates the retail price, the state-own monopoly retailers provide homogeneous gasoline services within the same region. We study the several gasoline standard reforms that improve the regional emission standard. Using high-frequency gas-station level data, our revealed-preference approach investigates consumers' purchasing behavior before and after the reform around the city boundaries where they can freely choose from higher or lower-emission gasoline.

We find strong evidence that consumers care about the environmental value of gasoline and substitute higher-emission gasoline with lower-emission one. Controlling for the concurrent price change and other confounding factors, the relative market share for the treated side of boundaries increases by about 14%, implying that consumers' WTP for higher emission standards is about 3.894% of the gasoline price. Our results highlight the importance of considering consumers' private value when designing environmental regulation. The cost-benefit analysis suggests that private gains from the higher environmental value of the new gasoline are about 49.44 billion CNY per year, which is higher than the refining cost of the reform, 37.40 billion CNY per year. It implies that even without accounting for benefits from environmental and health improvements, the social benefit of improving the gasoline standard is sizeable enough to compensate for social costs.

We want to discuss two key issues that readers should be careful with when interpreting our results. First, we cannot distinguish green consumerism observed in this paper to consumers' preferences on *pure* or *impure* public goods. The difference is whether consumers purely care about their contribution to the environment, in other words, purely warm-glow. The distinction between the two is informative to the design of regulation tools and green markets (Kotchen, 2006; Kotchen and Moore, 2007a). With individual-level transaction data, this question can be better explored.

Second, even though our paper emphasizes the importance of consumer WTP in the social benefit of cleaner gasoline emission standard program, we don't want to over-interpret our cost-benefit analysis as the rigorous welfare analysis. To inform the optimal gasoline standard regulation, the researcher needs a fully specified social welfare function to take into consideration other effects than merely the aggregated private utility, such as the externality. For example, our estimated WTP for standard upgrades of diesel consumers is indistinguishable from zeros, meaning that the green preference of diesel users is much lower than gasoline consumers. The reason is that most diesel vehicles in China are business cargoes and trucks owned by private companies that are sensitive to their cost. Although the private value from upgrading diesel standards is low, social welfare improvement could potentially be large, giving China's current low diesel standard and huge environmental damage. Therefore, developing a quantitative framework that incorporates both private WTP and public benefit for environmental standard improvement is an important topic for further research.

References

Allcott, H., Mullainathan, S., and Taubinsky, D. (2014). Energy policy with externalities and internalities. *Journal of Public Economics*, 112:72–88.

Allcott, H. and Taubinsky, D. (2015). Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market. *American Economic Review*, 105(8):2501–38.

Ambec, S. and De Donder, P. (2022). Environmental policy with green consumerism. *Journal of Environmental Economics and Management*, 111:102584.

Anderson, S. T. and Elzinga, A. (2014). A ban on one is a boon for the other: Strict gasoline content rules and implicit ethanol blending mandates. *Journal of Environmental Economics and Management*, 67(3):258–273.

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The economic journal*, 100(401):464–477.

Auffhammer, M. and Kellogg, R. (2011). Clearing the air? the effects of gasoline content regulation on air quality. *American Economic Review*, 101(6):2687–2722.

Bayer, P., Keohane, N., and Timmins, C. (2009). Migration and hedonic valuation: The case of air quality. *Journal of Environmental Economics and Management*, 58(1):1–14.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, pages 841–890.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, pages 242–262.

Bjørner, T. B., Hansen, L. G., and Russell, C. S. (2004). Environmental labeling and consumers' choice—an empirical analysis of the effect of the Nordic Swan. *Journal of Environmental Economics and Management*, 47(3):411–434.

Borchers, A. M., Duke, J. M., and Parsons, G. R. (2007). Does willingness to pay for green energy differ by source? *Energy Policy*, 35(6):3327–3334.

Brons, M., Nijkamp, P., Pels, E., and Rietveld, P. (2008). A meta-analysis of the price elasticity of gasoline demand. a sur approach. *Energy Economics*, 30(5):2105–2122.

Brown, J., Hastings, J., Mansur, E. T., and Villas-Boas, S. B. (2008). Reformulating competition? gasoline content regulation and wholesale gasoline prices. *Journal of Environmental economics and management*, 55(1):1–19.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326.

Chakravorty, U., Nauges, C., and Thomas, A. (2008). Clean air regulation and heterogeneity in us gasoline prices. *Journal of Environmental Economics and Management*, 55(1):106–122.

Dahl, C. and Sterner, T. (1991). Analysing gasoline demand elasticities: a survey. *Energy economics*, 13(3):203–210.

Davis, L. W. and Kilian, L. (2011). Estimating the effect of a gasoline tax on carbon emissions. *Journal of Applied Econometrics*, 26(7):1187–1214.

Deryugina, T., MacKay, A., and Reif, J. (2020). The long-run dynamics of electricity demand: Evidence from municipal aggregation. *American Economic Journal: Applied Economics*, 12(1):86–114.

EPA, U. (2014). Inventory of us greenhouse gas emissions and sinks: 1990–2012. *The Air Pollution Consultant; New York*, 24(3):1_17–1_22.

Espey, M. (1998). Gasoline demand revisited: an international meta-analysis of elasticities. *Energy* *economics*, 20(3):273–295.

Freeman, R., Liang, W., Song, R., and Timmins, C. (2019). Willingness to pay for clean air in China. *Journal of Environmental Economics and Management*, 94:188–216.

Gao, X., Song, R., and Timmins, C. (2021). The role of information in the rosen-roback framework. Technical report, National Bureau of Economic Research.

Gelman, A. and Imbens, G. (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, 37(3):447–456.

Gentzkow, M. (2006). Television and voter turnout. *The Quarterly Journal of Economics*, 121(3):931–972.

Goodwin, P. B. (1992). A review of new demand elasticities with special reference to short and long run effects of price changes. *Journal of transport economics and policy*, pages 155–169.

Guo, D., Wang, A., and Zhang, A. T. (2020). Pollution exposure and willingness to pay for clean air in urban china. *Journal of Environmental Management*, 261:110174.

Guo, X., Liu, H., Mao, X., Jin, J., Chen, D., and Cheng, S. (2014). Willingness to pay for renewable electricity: A contingent valuation study in Beijing, China. *Energy Policy*, 68:340–347.

Hausman, C. and Rapson, D. S. (2018). Regression discontinuity in time: Considerations for empirical applications. *Annual Review of Resource Economics*, 10:533–552.

Hidrue, M. K., Parsons, G. R., Kempton, W., and Gardner, M. P. (2011). Willingness to pay for electric vehicles and their attributes. *Resource and energy economics*, 33(3):686–705.

Houde, J.-F. (2012). Spatial differentiation and vertical mergers in retail markets for gasoline. *American Economic Review*, 102(5):2147–82.

Houde, S. and Myers, E. (2021). Are consumers attentive to local energy costs? evidence from the appliance market. *Journal of Public Economics*, 201:104480.

Hughes, J., Knittel, C. R., and Sperling, D. (2008). Evidence of a shift in the short-run price elasticity of gasoline demand. *The Energy Journal*, 29(1).

Huntington, H. G., Barrios, J. J., and Arora, V. (2019). Review of key international demand elasticities for major industrializing economies. *Energy Policy*, 133:110878.

International Council on Clean Transportation (2014). China V gasoline and diesel fuel quality standards.

Ito, K. (2014). Do consumers respond to marginal or average price? evidence from nonlinear electricity pricing. *American Economic Review*, 104(2):537–63.

Ito, K. and Zhang, S. (2020). Willingness to pay for clean air: Evidence from air purifier markets in china. *Journal of Political Economy*, 128(5):1627–1672.

Jacobsen, G. D., Kotchen, M. J., and Vandenbergh, M. P. (2012). The behavioral response to voluntary provision of an environmental public good: Evidence from residential electricity demand. *European Economic Review*, 56(5):946–960.

Kahn, M. E. (2007). Do greens drive hummers or hybrids? environmental ideology as a determinant of consumer choice. *Journal of Environmental Economics and Management*, 54(2):129–145.

Kleven, H. J. (2016). Bunching. *Annual Review of Economics*, 8:435–464.

Kollmuss, A. and Agyeman, J. (2002). Mind the gap: why do people act environmentally and what are the barriers to pro-environmental behavior? *Environmental education research*, 8(3):239–260.

Kotchen, M. J. (2005). Impure public goods and the comparative statics of environmentally friendly consumption. *Journal of environmental economics and management*, 49(2):281–300.

Kotchen, M. J. (2006). Green markets and private provision of public goods. *Journal of Political Economy*, 114(4):816–834.

Kotchen, M. J. and Moore, M. R. (2007a). Private provision of environmental public goods: Household participation in green-electricity programs. *Journal of Environmental Economics and management*, 53(1):1–16.

Kotchen, M. J. and Moore, M. R. (2007b). Private provision of environmental public goods: Household participation in green-electricity programs. *Journal of Environmental Economics and Management*, 53(1):1–16.

Levin, L., Lewis, M. S., and Wolak, F. A. (2017). High frequency evidence on the demand for gasoline. *American Economic Journal: Economic Policy*, 9(3):314–47.

Li, P., Lu, Y., and Wang, J. (2020). The effects of fuel standards on air pollution: Evidence from china. *Journal of Development Economics*, 146:102488.

Li, S. (2018). Better lucky than rich? welfare analysis of automobile licence allocations in beijing and shanghai. *The Review of Economic Studies*, 85(4):2389–2428.

Li, S., Linn, J., and Muehlegger, E. (2014). Gasoline taxes and consumer behavior. *American Economic Journal: Economic Policy*, 6(4):302–42.

Liu, T., He, G., and Lau, A. (2018). Avoidance behavior against air pollution: evidence from online search indices for anti-PM2.5 masks and air filters in Chinese cities. *Environmental Economics and Policy Studies*, 20(2):325–363.

Ma, C. and Burton, M. (2016). Warm glow from green power: Evidence from australian electricity consumers. *Journal of Environmental Economics and Management*, 78:106–120.

Marcus, M. (2017). On the road to recovery: Gasoline content regulations and child health. *Journal of health economics*, 54:98–123.

McFadden, D. et al. (1973). Conditional logit analysis of qualitative choice behavior.

Ministry of Ecology and Environment of PRC (2016). China vehicle environmental management annual report 2016. *Ministry of Ecology and Environment of the People's Republic of China.*

Mozumder, P., Vásquez, W. F., and Marathe, A. (2011). Consumers' preference for renewable energy in the southwest USA. *Energy Economics*, 33(6):1119–1126.

Park, S. Y. and Zhao, G. (2010). An estimation of us gasoline demand: A smooth time-varying cointegration approach. *Energy Economics*, 32(1):110–120.

Roe, B., Teisl, M. F., Levy, A., and Russell, M. (2001). US consumers' willingness to pay for green electricity. *Energy Policy*, page 9.

Sallee, J. M., West, S. E., and Fan, W. (2016). Do consumers recognize the value of fuel economy? evidence from used car prices and gasoline price fluctuations. *Journal of Public Economics*, 135:61–73.

Sexton, S. E. and Sexton, A. L. (2014). Conspicuous conservation: The prius halo and willingness to pay for environmental bona fides. *Journal of Environmental Economics and Management*, 67(3):303–317.

Simon-Kucher & Partners (2021). Global sustainability study 2021.

Small, K. A. and Rosen, H. S. (1981). Applied welfare economics with discrete choice models. *Econometrica: Journal of the Econometric Society*, pages 105–130.

Small, K. A. and Van Dender, K. (2007). Fuel efficiency and motor vehicle travel: the declining rebound effect. *The Energy Journal*, 28(1).

Sundt, S. and Rehdanz, K. (2015). Consumers' willingness to pay for green electricity: A meta-analysis of the literature. *Energy Economics*, 51:1–8.

Wang, Z. and Zhang, J. (2021). The value of information disclosure: Evidence from mask consumption in china. *Available at SSRN*.

WHO (2011). Urban transport and health: on behalf of federal ministry for economic cooperaton and development (bmz).

Wichman, C. J. (2016). Incentives, green preferences, and private provision of impure public goods. *Journal of Environmental Economics and Management*, 79:208–220.

World Value Survey, t. (2014). World values survey: Round six - country-pooled datafile version: https://www.worldvaluessurvey.org/wvsdocumen. Technical report.

Zhang, J. and Mu, Q. (2018). Air pollution and defensive expenditures: Evidence from particulate-filtering facemasks. *Journal of Environmental Economics and Management*, 92:517–536.

Zhang, X.-B., Fei, Y., Zheng, Y., and Zhang, L. (2020). Price ceilings as focal points to reach price uniformity: Evidence from a Chinese gasoline market. *Energy Economics*, 92:104950.

Figures and Tables

Figures

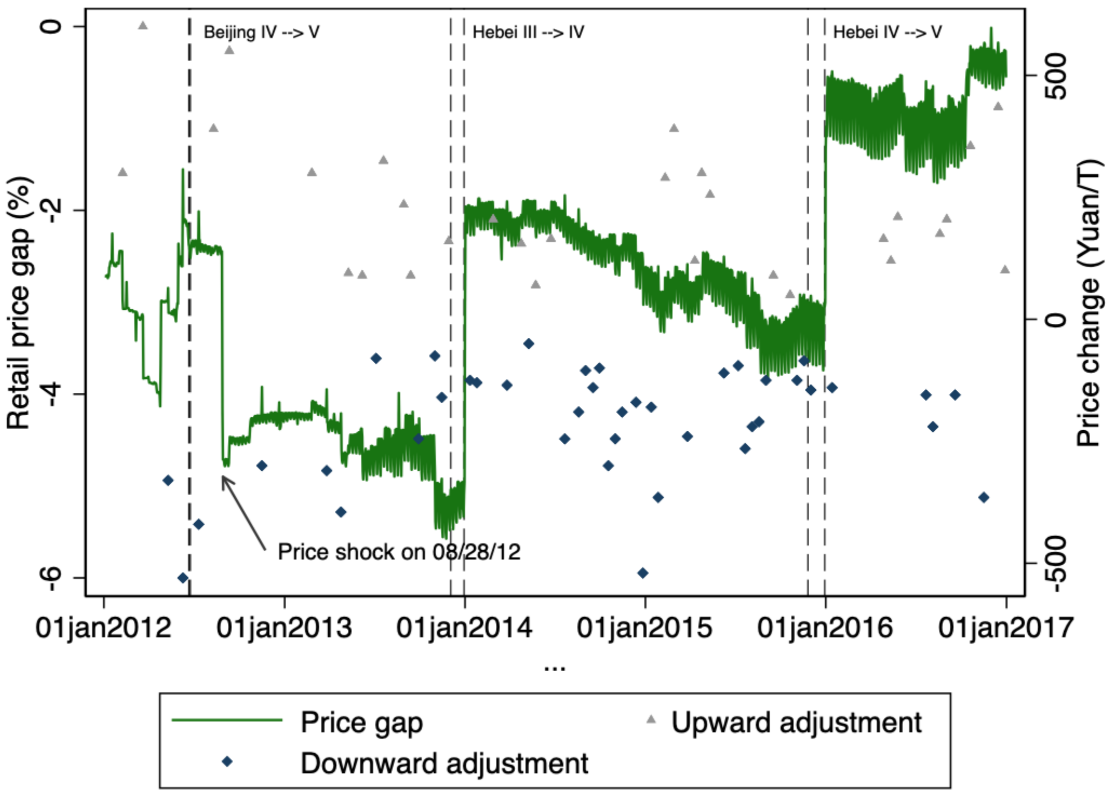


Figure 1: The Price Gap Between Hebei and Beijing and the Regulated Price Adjustments

Notes: This figure shows the fluctuation of price gap between Beijing and Hebei within our sample period. The gap is generated by subtracting the price of #93 gasoline in Beijing from the price in Hebei. The price cap adjustments data are collected manually from NDRC.

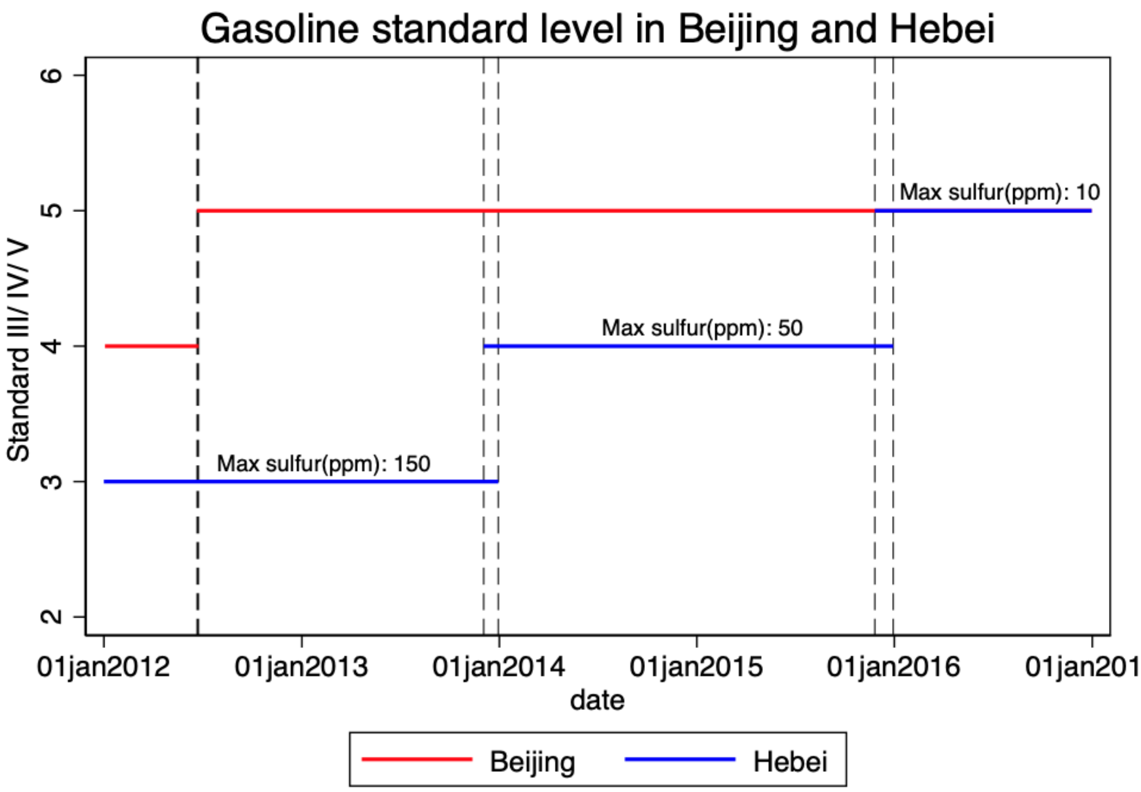


Figure 2: Gasoline Emission Standard Change Before and After the Reforms

Notes: This figure presents the timeline of gasoline standard upgrades in Hebei and Beijing, demonstrated in blue line and red line. The vertical axis shows the standards. After Hebei IV-V reform, both regions have standard V and the two lines overlaps. The maximum sulfur content allowed for each standard is shown above the line.

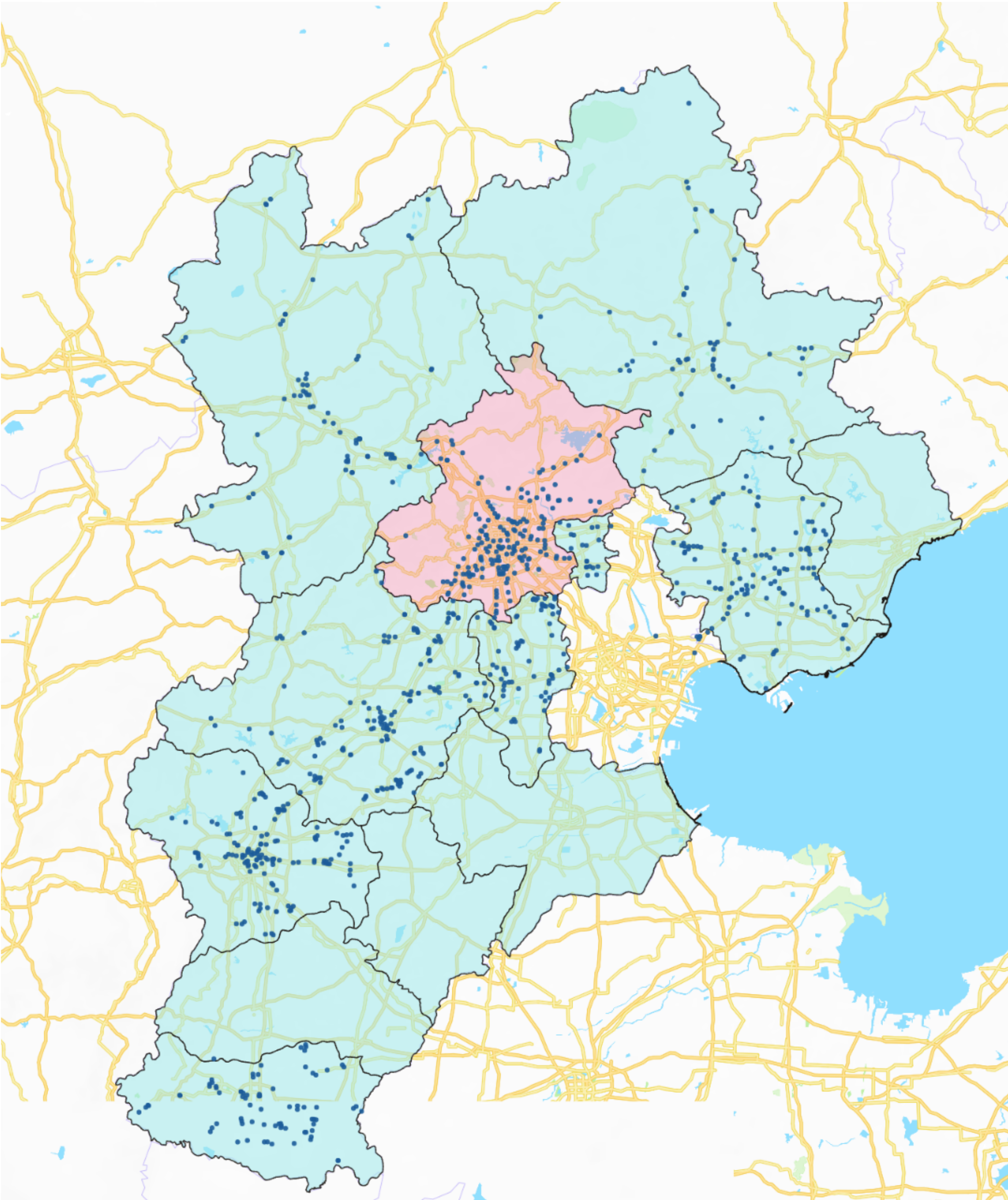


Figure 3: Geographical Distribution of Gas Stations in Beijing and Hebei

Notes: This figure displays the geographic locations of *PetroChina* gas stations in our sample. The red region is the city of Beijing; the dark blue regions are cities in Hebei selling regular gasoline (#89, #92, #95); the light blue regions are cities in Hebei selling ethanol gasoline (E10). The black dots plot the *PetroChina* gas stations in Beijing and Hebei in our sample. The yellow lines plot the road network in the regions, including highway and non-highway.

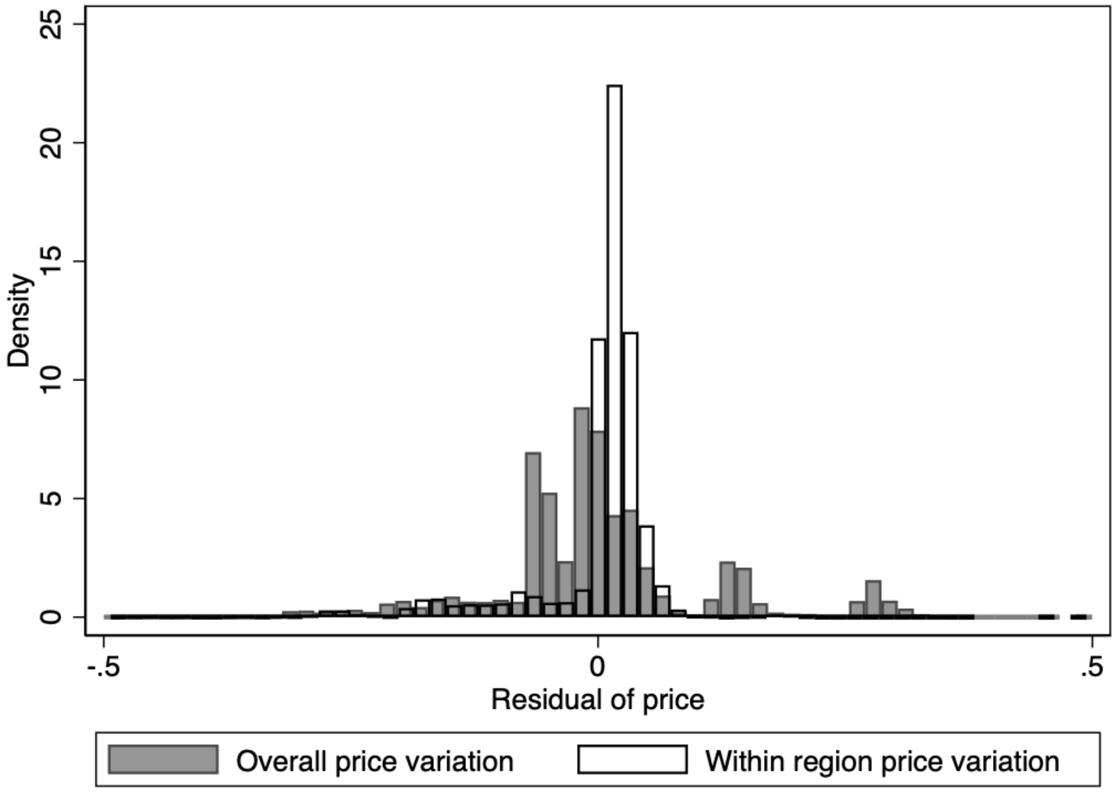


Figure 4: Overall and Within-region Price Variation

Notes: This figure plots the overall and within-region price variation. The overall price residual are calculate by including only date fixed effects; the within-region price residual includes both date fixed effects and region fixed effect. It is clear that the within-region price variation is much smaller than the overall price variation, implying that the price variation mainly comes from the different regions instead of different gas stations within the same region.

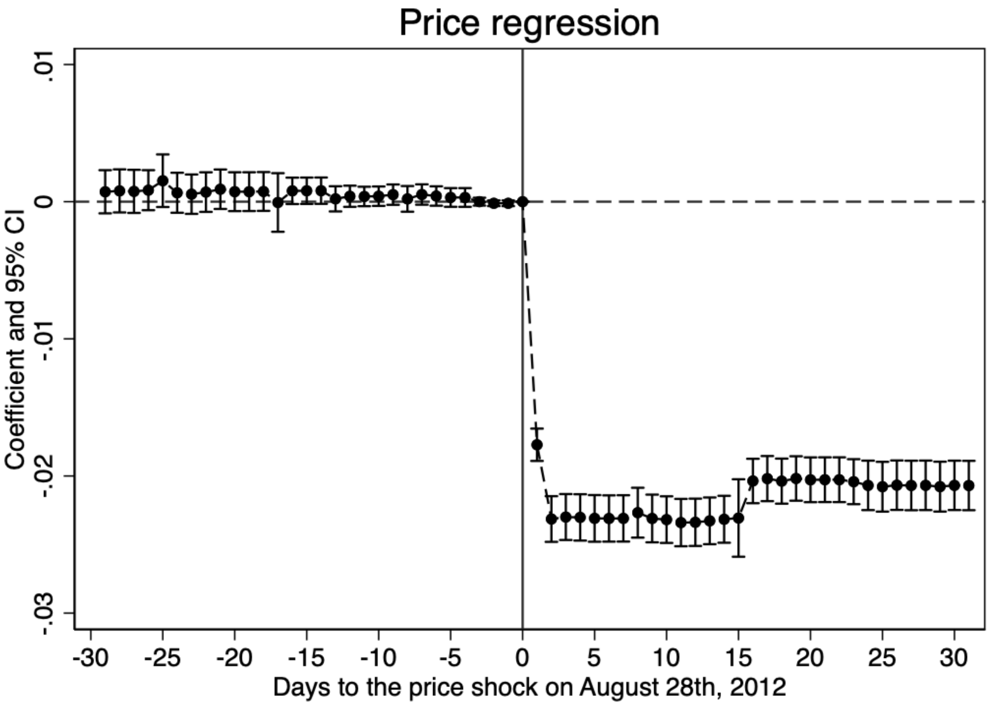


Figure 5: Event Study: Impact of the Price Shock (August 28th, 2012) on Gasoline Price

Notes: This figure presents a flexible event-study results of the First-stage regression of price elasticity estimation. The figure plots the estimated coefficients and 95% level confidence intervals of gasoline price for each day within the event-study window. The regression specification includes a series of date indicators with the Hebei dummy interactions and the omitted day category is August 12th, 2012. The sample used here is for regular gasoline.

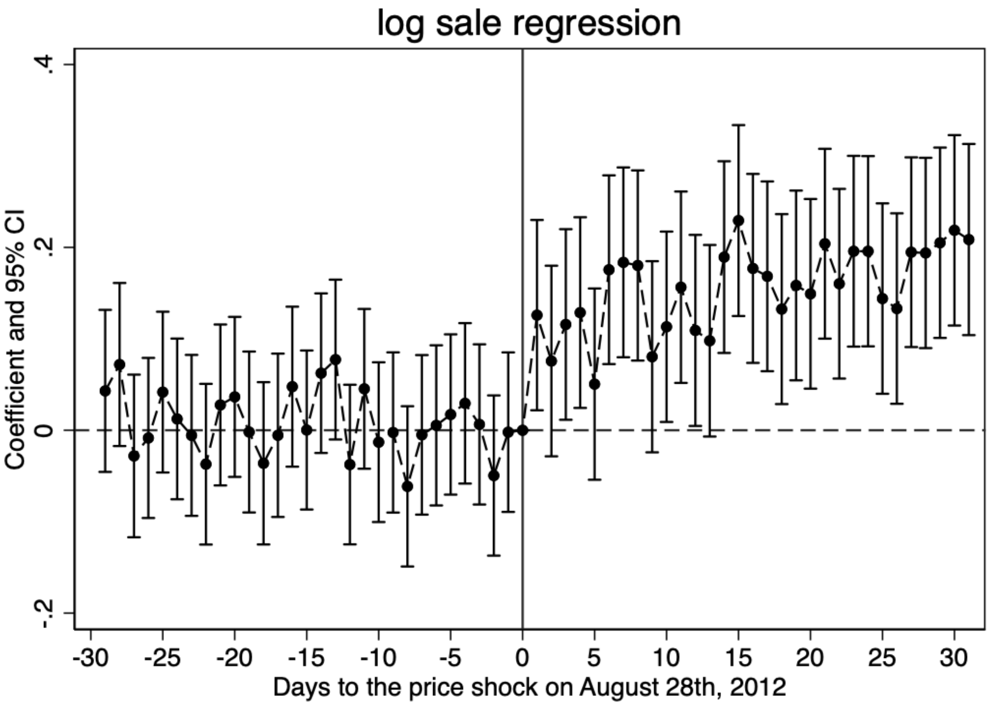
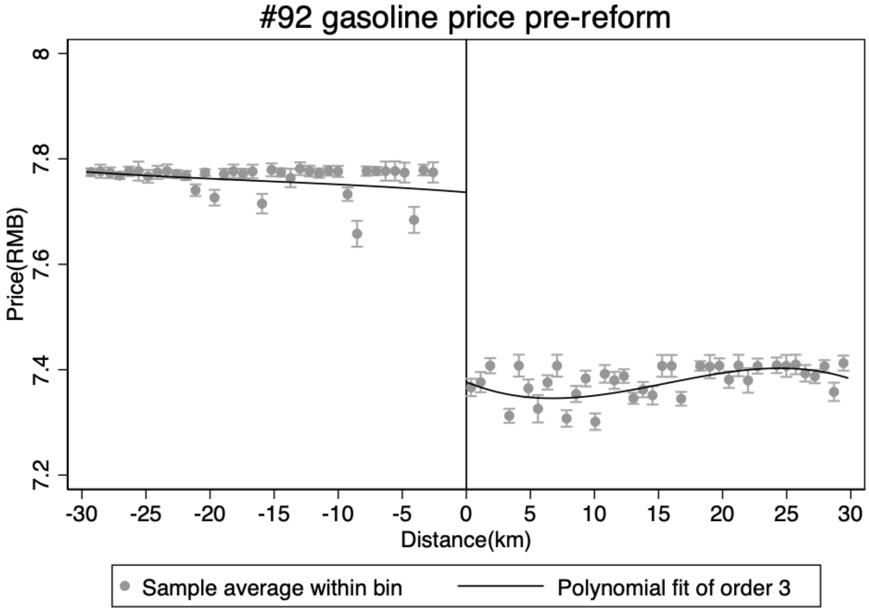
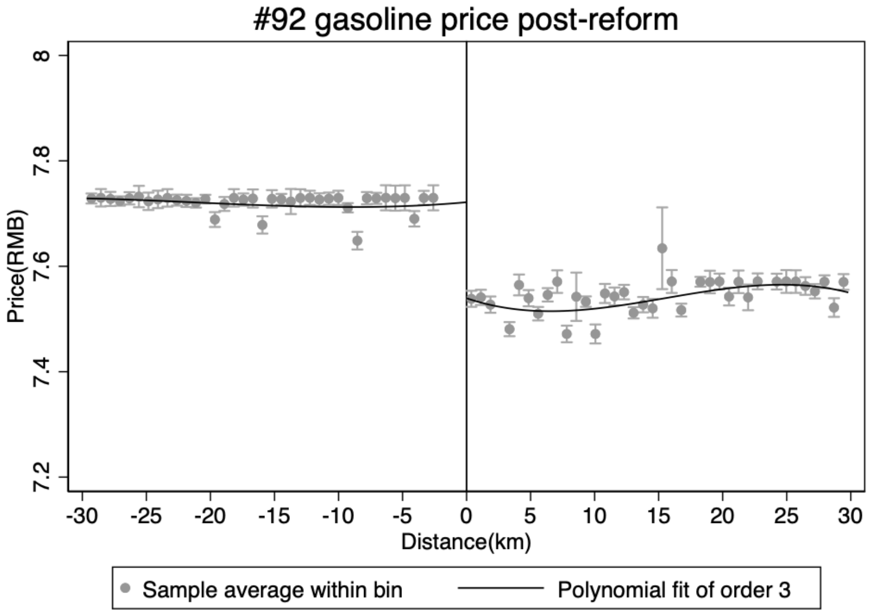


Figure 6: Event Study: Impact of the Price Shock (August 28th, 2012) on Sales

Notes: This figure presents a flexible event-study results of the Reduced-form regression of price elasticity estimation. The figure plots the estimated coefficients and 95% level confidence intervals of *log* sales for each day within the event-study window. The regression specification includes a series of date indicators with the Hebei dummy interactions and the omitted day category is August 12th, 2012. The sample used here is for regular gasoline.



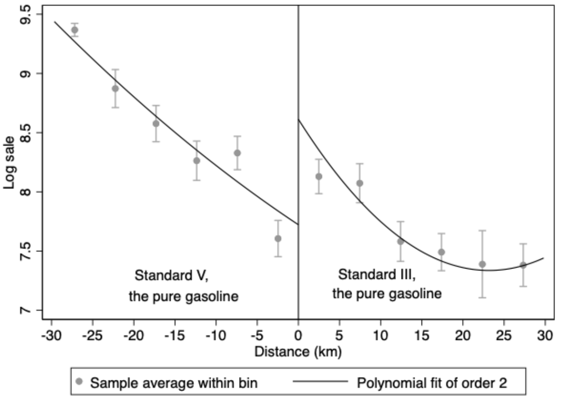
(a) Pre-reform price of #92 pure gasoline



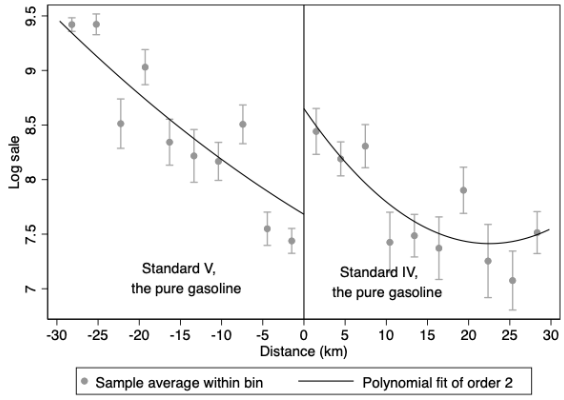
(b) Post-reform price #92 pure gasoline

Figure 7: Gasoline Price at Two Side of the Border, Before and After the Reform

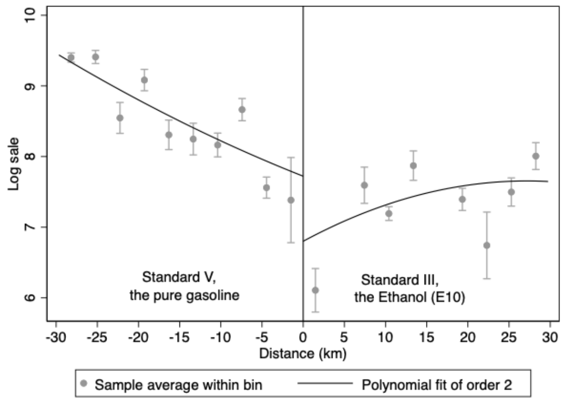
Notes: This figure plots the “jump” of price in CNY at the geographical border, before and after the enforcement of higher emission standard in Hebei. The scatters plot the local mean of the sales price for regular pure gasoline with a bin size of 1 kilometer. The X-axis is the distance from the each gas station to the city boundaries. The solid curves are the regression fitted lines for a polynomial fit of order 3. The solid vertical line represents the boundary. Negative/positive values of distance give the distance of gas station at the Beijing/Hebei side of the border respectively. The sample used in the figure is regular gasoline.



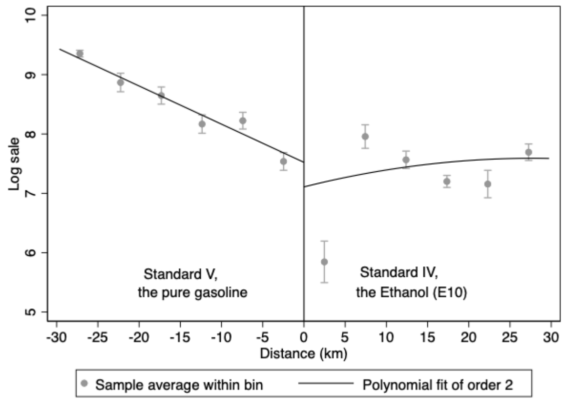
(a) Pre-reform log sales of the pure regular gasoline



(b) Post-reform log sales of the pure regular gasoline



(c) Pre-reform log sales of the ethanol regular gasoline



(d) Post-reform log sales of the ethanol regular gasoline

Figure 8: RD Plot of *log* Sales Before and After the Reform:
Both Sides of the Border Sell Regular Gasoline
V.S.
Beijing Side Sells Regular and Hebei Side Sells Ethanol Gasoline

Notes: This figure compares the RD plot of the main outcome variable of interest before and after the enforcement of higher emission standard in Hebei, between Beijing and Hebei cities selling different types of gasoline. Figure (a) and (b) present the log sales of regular gasoline in Beijing and in Hebei cities. Figure (c) and (d) present the log sales of regular gasoline in Beijing and of E10 ethanol gasoline in Hebei cities. Each of Hebei cities either sells regular gasoline or E10 ethanol gasoline, so the Hebei cities in (a), (b) and (c), (d) are different. The predicted value with each bins, the confidence interval and fitted line are shown in the figure. The solid vertical line represents the boundary. Negative/positive values of distance give the distance of gas station at the Beijing/Hebei side of the border respectively. Comparing (a), (b) and (c), (d) we find consumers "bunching" at Hebei side when both sides sell regular gasoline, while the same pattern is not found when gas stations on Hebei side sell E10. We argue that it's because consumers are less likely to switch from E10 to regular gasoline of the same grade to avoid engine and gas tank damage.

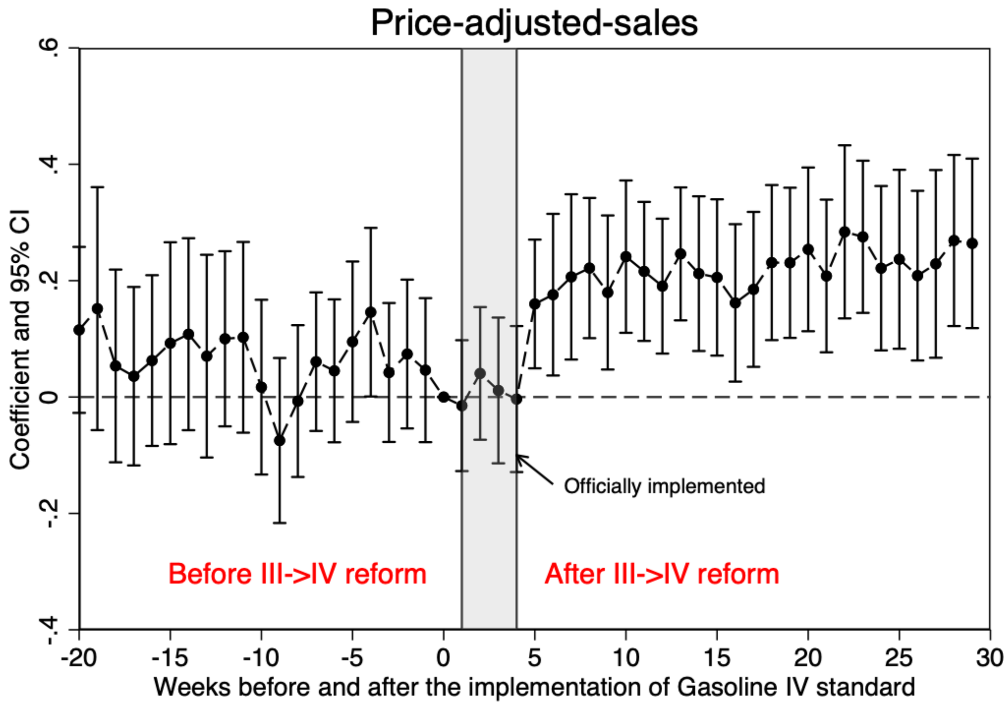


Figure 9: Tests for Parallel Trends and Dynamic Effects

Notes: This figure presents a flexible event-study of the enforcement of standard IV in Hebei. The figure plots the estimated coefficients and 95% level confidence intervals of price-adjusted sales for each week within the event-study window. The grey region represents the transition period when gas stations in Hebei sell gasoline of both standard III and IV. The sample used here are regular gasoline. The order of distance polynomial $f(\cdot)$ is 2. The distance bandwidth is 30 km. The Standard errors are clustered to gas station level.

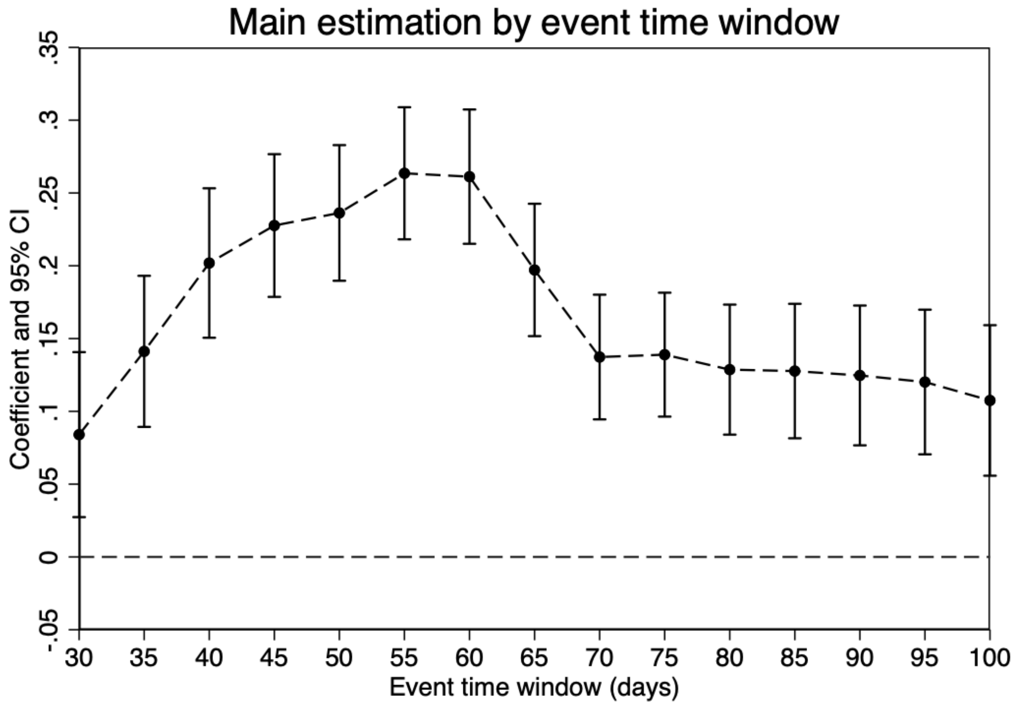


Figure 10: Two Stage Estimates Across Time Window

Notes: This figure presents how the estimate of coefficient of interest, β, changes with a time window varying from 30 to 100 days before and after the treatment day. The treatment day of each reform is defined in section 4.1.3. The time bandwidth we use in our main estimation is 70 days before and after the treatment day.

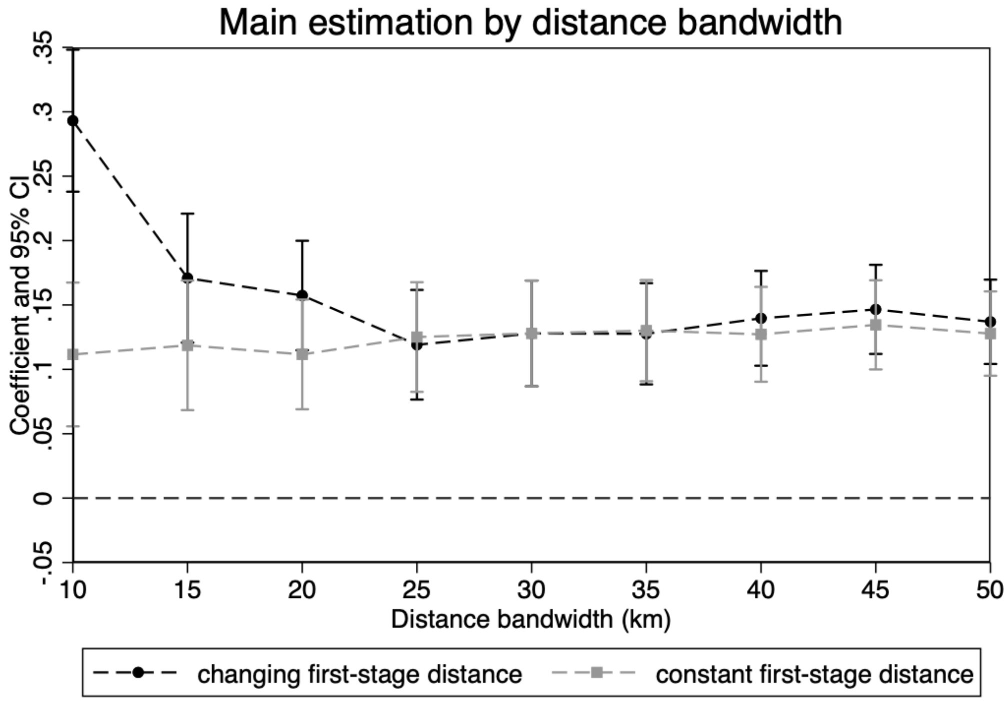


Figure 11: Two Stage Estimates Across Distance Bandwidth

Notes: This figure presents how the estimate of coefficient of interest, β, changes with a varying distance bandwidth from 10 to 50 km away from the boundary. The black dash line demonstrates the results using the same distance bandwidth in both β and α estimation, and the grey dash line demonstrates the results using changing bandwidth for only second-stage with the $\hat{\alpha}$ same as the main estimation. The bandwidth we use in our main estimation is 30 km.

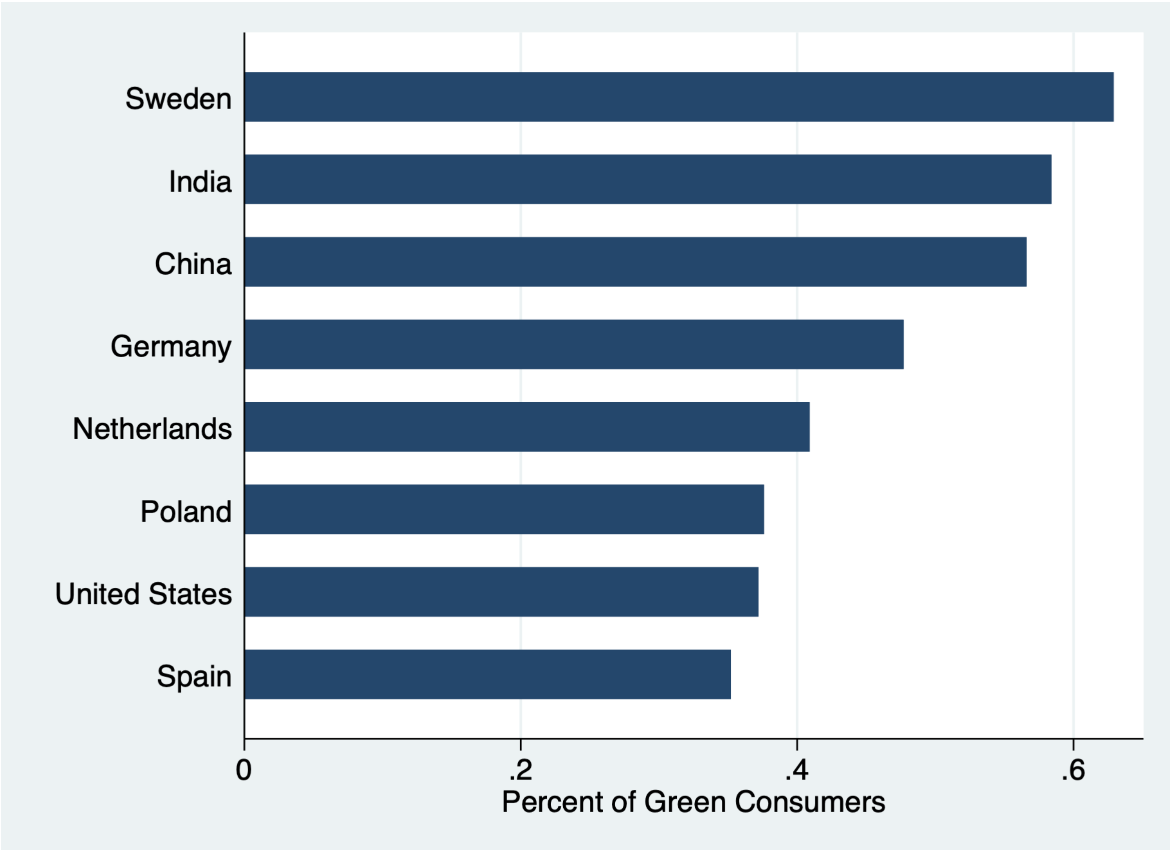


Figure 12: The Percentage of Green Consumers in World Value Survey

Notes: This figure shows the percentage of respondent who think protecting the environment should be given priority even if it causes slower economic growth and some loss of jobs when asked about the priority between environment and economic growth in World Value Survey. The results are from the wave 6 conducted in 2013 in China and Germany, in 2012 in India, Netherlands, and Poland, and in 2011 in Spain, Sweden, and United States. Source: World Value Survey.

Tables

Table 1: Summary Statistics of Gas Station Data within 30 km to the Boundary

	(1) All	(2) Beijing	(3) Hebei
Price Regular(CNY/liter)	6.904	7.006	6.769
	(.001)	(.002)	(.002)
Price Premium(CNY/liter)	7.314	7.455	7.020
	(.002)	(.002)	(.003)
Dist to border (km)	15.017	21.458	16.759
	(.523)	(.657)	(.799)
Dist to Diff (km)	21.970	22.661	21.107
	(.684)	(.816)	(1.151)
Dist to Same (km)	3.431	3.415	3.451
	(.299)	(.359)	(.499)
Near Highway(=1)	0.539	0.447	0.655
	(0.031)	(0.042)	(0.045)
Engine oil(=1)	.480	.270	.743
	(.031)	(.037)	(.041)
N	254	141	113

Notes: This table presents summary statistics of gas station level data for gas stations within the 30km bandwidth from the boundary. In our sample there are 141 gas stations in Beijing and 113 in Hebei that are located 30 km away or closer to the boundary. Price regular and price premium are prices of regular and premium gasoline sold at these gas stations; dist to border stands for the geometric distance from the gas station to the boundary; dist to diff and dist to same stand for the distance from the gas station to the nearest gas station in the other province and in the same province; near highway is a dummy that equals one if the gas station is near a highway or national highway; engine oil is a dummy that equals one if the gas station provides the engine oil replacing service. We present standard errors in parentheses.

Table 2: First Stage Regression of the Price Elasticity Estimation

	(1) *log* price	(2) *log* price	(3) *log* price	(4) price	(5) price	(6) price
$Hebei \times Post$	-0.0209***	-0.0215***	-0.0186***	-0.178***	-0.175***	-0.160***
	(0.001)	(0.001)	(0.004)	(0.007)	(0.006)	(0.028)
Date FE	Yes	Yes	Yes	Yes	Yes	Yes
Gas station FE	No	Yes	Yes	No	Yes	Yes
$f(\cdot) \times trend$	No	No	Yes	No	No	Yes
Adjusted R-squared	0.951	0.989	0.989	0.967	0.991	0.992
N	11712	11712	11712	12907	12907	12907

Notes: This table presents first-stage estimation of the price elasticity using the sample of regular gasoline (#93). The instrument variables use in this estimation is the reform-induced price shock in Beijing on August 28th, 2021. Therefore, this first stage takes the form of Diff-in-Diff regression. The dependent variables are *log* of price and level of price, respectively. The key independent variable is the interaction between fuel type dummy and the indicator for post August 28th, 2021. $f(\cdot)$ is the polynomial function of distance to boundary describe in section 5.2. The order of $f(\cdot)$ is 2. The distance bandwidth is 30 km. The time window is 30 days before and after the price shock. We present cluster standard errors in parentheses. Standard errors are clustered to gas station level. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 3: IV Regression of the Price Elasticity Estimation

	(1) OLS	(2) IV	(3) IV	(4) OLS	(5) IV	(6) IV
log price (elasticity)	-4.243***	-2.783***	-2.758***			
	(0.520)	(0.877)	(0.815)			
price (α)				-0.690***	-0.725***	-0.550***
				(0.088)	(0.228)	(0.181)
Date FE	Yes	Yes	Yes	Yes	Yes	Yes
Gas station FE	Yes	Yes	Yes	Yes	Yes	Yes
$f(\cdot) \times trend$	No	No	Yes	No	No	Yes
Adjusted R-squared	0.886	0.962	0.962	0.886	0.962	0.962
N	1257078	11269	11269	1233033	11269	11269

Notes: This table presents second-stage estimation of the price elasticity using only the sample of regular gasoline (#93). The instrument variables use in this estimation is the reform-induced price shock in Beijing on August 28th, 2021. The dependent variable is $log(Q_{jk})$. The independent variables are the predicted value of *log* price and level of price calculated in the first-stage. The estimates use $\widehat{log\text{price}}$ and $\widehat{\text{price}}$ correspond to the price elasticity and structural parameter α, respectively. $f(\cdot)$ is the polynomial function of distance to boundary describe in section 5.2. The order of $f(\cdot)$ is 2. The distance bandwidth is 30 km. The time window is 30 days before and after the price shock. We present cluster standard errors in parentheses. Standard errors are clustered to gas station level. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 4: Effects of Emission Standard Reform and WTP

	(1)	(2)	(3)	(4)	(5)	(6)
$Hebei \times Post$ (β)	0.117***	0.106***	0.0733	-0.204***	-0.0402*	-0.199***
	(0.030)	(0.022)	(0.049)	(0.038)	(0.021)	(0.038)
	[0.027]	[0.022]	[0.052]	[0.037]	[0.019]	[0.042]
$Hebei \times Post \times treatedyear$ (β)				0.246***	0.139***	0.138***
				(0.040)	(0.024)	(0.024)
				[0.037]	[0.021]	[0.022]
Date FE	No	Yes	Yes	No	Yes	Yes
Gas station FE	No	Yes	Yes	No	Yes	Yes
GS by DOW FE	No	No	Yes	No	No	Yes
$\mathbb{f}() \times Post$	No	No	Yes	No	No	Yes
X_trend	No	No	No	No	No	Yes
Adjusted R-squared	0.00408	0.866	0.880	0.00829	0.862	0.875
N	32170	32170	31767	62594	62594	61787
WTP (CNY/L)	.293	.264	.183	.614	.348	.345
WTP (%)	3.953	3.57	2.475	8.294	4.703	4.664

Notes: This table shows benchmark estimation of the effect of gasoline emission standard reform, using the two-step method described in section 3.3. The reform examined in the table is Standard III $\longrightarrow$ IV reform in Hebei. The dependent variable is $log(Q_{jk}) - \hat{\alpha}\, p_k^j$. We use -0.376 as the value of $\hat{\alpha}$, which is a conservative estimate in Table A.9. The willingness to pay (WTP) is calculated as $-\beta/\alpha$. The variable DoW stands for day-of-the-week; X_{kt} includes interactions between gas station characteristics and time trend. The order of distance polynomial $f(\cdot)$ is 2. The distance bandwidth is 30 km. The event time window is 70 days. For columns 4 to 6, we also include the treated year dummy and its interactions with the Hebei and post-reform dummies. We present cluster standard errors in parentheses and bootstrap cluster standard errors in square brackets. Standard errors are clustered to gas station level. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 5: Effects of Emission Standard Reform Using One Step Method

	(1)	(2)	(3)	(4)	(5)	(6)
$Hebei \times Post\ (\beta)$	0.367***	0.0612	0.0588	-0.257***	-0.0432*	-0.264***
	(0.038)	(0.089)	(0.100)	(0.035)	(0.025)	(0.046)
$Hebei \times Post \times treatedyear\ (\beta)$				0.706***	0.150**	0.244***
				(0.057)	(0.075)	(0.049)
Price (α)	-2.535***	0.0159	0.0155	-2.598***	-0.426	-0.859***
	(0.194)	(0.723)	(0.723)	(0.196)	(0.340)	(0.220)
Date FE	No	Yes	Yes	No	Yes	Yes
Gas station FE	No	Yes	Yes	No	Yes	Yes
GS by DOW FE	No	Yes	Yes	No	Yes	Yes
f() $\times trend$	No	No	Yes	No	No	Yes
X_trend	No	No	No	No	No	Yes
Adjusted R-squared	0.238	0.877	0.891	0.250	0.866	0.876
N	32170	32167	31767	62594	62594	61787

Notes: This table shows robustness check of the effect of gasoline emission standard reform, using the one-step method instead of two-step described in section 3.3. The reform examined in the table is Standard III $\longrightarrow$ IV reform in Hebei. The dependent variable is $log(Q_{jk})$. In one-step method, the value of $\hat{\alpha}$ is estimated with $\hat{\beta}$ simultaneously. The variable DoW stands for day-of-the-week; X_{kt} includes interactions between gas station characteristics and time trend. The order of distance polynomial $f(\cdot)$ is 2. The distance bandwidth is 30 km. The event time window is 70 days. For columns 4 to 6, we also include the treated year dummy and its interactions with the Hebei and post-reform dummies. We present cluster standard errors in parentheses. Standard errors are clustered to gas station level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Effects of Emission Standard Reform by Fuel Types

	(1) Regular	(2) Premium	(3) Diesel
$Hebei \times Post \times treatedyear$ (β)	0.124***	0.158***	-0.118
	(0.021)	(0.035)	(0.087)
Date FE	Yes	Yes	Yes
Gas station FE	Yes	Yes	Yes
GS by DOW FE	Yes	Yes	Yes
$f() \times trend$	Yes	Yes	Yes
X_trend	Yes	Yes	Yes
Adjusted R-squared	0.929	0.869	0.695
N	35482	26298	62730
WTP (Yuan/L)	0.331	0.395	-
WTP (%)	4.471	5.197	-

Notes: This table shows the robustness check estimation using gasoline of different grades and diesel, with the two-step method described in section 3.3. Column 1 and 2 report the results of gasoline with a RON equals 93 and 97. Column 3 reports the result of diesel. The reform examined in the table is Standard III $\longrightarrow$ IV reform in Hebei. The dependent variable is $log(Q_{jk}) - \hat{\alpha}\, p_k^j$. We use -0.376 as the value of $\hat{\alpha}$, which is a conservative estimate in Table A.9. The willingness to pay (WTP) is calculated as $-\beta/\alpha$. The variable DoW stands for day-of-the-week; X_{kt} includes interactions between gas station characteristics and time trend. The order of distance polynomial $f(\cdot)$ is 2. The distance bandwidth is 30 km. The event time window is 70 days. The treated year dummy and its interactions with the Hebei and post-reform dummies are included in all columns. We present cluster standard errors in parentheses. Standard errors are clustered to gas station level. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 7: Placebo Tests

	(1) Ethanol boundary	(2) Non-boundary	(3) Diesel	(4) Placebo year
$Hebei \times Post$ (β)				0.0577
				(0.038)
$Hebei \times Post \times treatedyear$ (β)	0.0476	0.00395	-0.0810	
	(0.049)	(0.020)	(0.094)	
Date FE	Yes	Yes	Yes	Yes
Gas station FE	Yes	Yes	Yes	Yes
GS by DOW FE	Yes	Yes	Yes	Yes
$f() \times trend$	Yes	Yes	Yes	Yes
X_trend	Yes	Yes	Yes	No
Adjusted R-squared	0.896	0.883	0.657	0.848
N	42863	102900	25543	55344

Notes: This table shows the placebo testes using different samples and the two-step method described in section 3.3. Column 1 and 2 report the results of gasoline sold at gas stations in Hebei cities that sell E10 only and Beijing. Column 3 reports the result of diesel. Column 4 reports the result of gasoline in 2012 and 2014 and therefore includes no treated year dummy. The reform examined in the table is Standard III $\longrightarrow$ IV reform in Hebei. The dependent variable is $log(Q_{jk}) - \hat{\alpha}\, p_k^j$. We use -0.376 as the value of $\hat{\alpha}$, which is a conservative estimate in Table A.9. The variable DoW stands for day-of-the-week; X_{kt} includes interactions between gas station characteristics and time trend. The order of distance polynomial $f(\cdot)$ is 2. The distance bandwidth is 30 km. The event time window is 70 days. For columns 1 to 3, the treated year dummy and its interactions with the Hebei and post-reform dummies are included. We present cluster standard errors in parentheses. Standard errors are clustered to gas station level. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 8: Sensitivity of Effects of Emission Standard Reform on Different Alpha

	(1)	(2)	(3)	(4)	(5)	(6)
$Hebei \times Post \times treatedyear$ (β)	0.0665***	0.0875***	0.108***	0.129***	0.150***	0.171***
	(0.020)	(0.020)	(0.021)	(0.021)	(0.021)	(0.021)
α	-0.1	-0.2	-0.3	-0.4	-0.5	-0.6
Adjusted R-squared	0.927	0.928	0.928	0.929	0.929	0.930
N	35482	35482	35482	35482	35482	35482
WTP (Yuan/L)	.665	.437	.362	.324	.301	.286
WTP (%)	8.99	5.912	4.886	4.373	4.065	3.86

Notes: This table shows how the estimation of the effect of gasoline emission standard reform changes with a varying $\hat{\alpha}$, using the preferred specification of the two-step method described in section 3.3. The reform examined in the table is Standard III $\longrightarrow$ IV reform in Hebei. The dependent variable is $log(Q_{jk}) - \hat{\alpha}\, p_k^j$. The value of $\hat{\alpha}$ we used changes from -0.1 to -0.6, in a step of -0.1. The willingness to pay (WTP) is calculated as $-\beta/\alpha$. The variable DoW stands for day-of-the-week; X_{kt} includes interactions between gas station characteristics and time trend. The order of distance polynomial $f(\cdot)$ is 2. The distance bandwidth is 30 km. The event time window is 70 days. The treated year dummy and its interactions with the Hebei and post-reform dummies are included for all columns. We present cluster standard errors in parentheses. Standard errors are clustered to gas station level. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 9: Effects of Two Other Emission Standard Reforms

	(1)	(2)	(3)	(4)
	BJ,IV $\longrightarrow$ V		HB,IV $\longrightarrow$ V	
$Hebei \times Post(\beta)$	-0.0720	0.0637	0.172***	-0.0745
	(0.051)	(0.043)	(0.059)	(0.045)
$Hebei \times Post \times treatedyear(\beta)$		-0.144***		0.147***
		(0.041)		(0.026)
Date FE	Yes	Yes	Yes	Yes
Gas station FE	Yes	Yes	Yes	Yes
GS by DOW FE	Yes	Yes	Yes	Yes
$f() \times trend$	Yes	Yes	Yes	Yes
X_trend	No	Yes	No	Yes
Adjusted R-squared	0.861	0.863	0.866	0.877
N	22211	39166	25955	54403

Notes: This table shows the estimation of the effect of different gasoline emission standard reforms, using the two-step method described in section 3.3. Column 1 and 2 examine the Standard IV $\longrightarrow$ V reform in Beijing. Column 3 and 4 examine the Standard IV $\longrightarrow$ V reform in Hebei. The dependent variable is $log(Q_{jk}) - \hat{\alpha}\, p_k^j$. We use -0.376 as the value of $\hat{\alpha}$, which is a conservative estimate in Table A.9. The variable DoW stands for day-of-the-week; X_{kt} includes interactions between gas station characteristics and time trend. The order of distance polynomial $f(\cdot)$ is 2. The distance bandwidth is 30 km. The event time window is 70 days. For columns 2 and 4, the treated year dummy and its interactions with the Hebei and post-reform dummies are included. We present cluster standard errors in parentheses. Standard errors are clustered to gas station level. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 10: Heterogeneity: Near Highway or Not

	Besides the main road		Not besides the main road	
	(1)	(2)	(3)	(4)
$Hebei \times Post$ (β)	0.153***	-0.206***	0.0439	-0.178**
	(0.051)	(0.047)	(0.066)	(0.080)
$Hebei \times Post \times treatedyear$ (β)		0.167***		0.0824***
		(0.028)		(0.027)
Date FE	Yes	Yes	Yes	Yes
Gas station FE	Yes	Yes	Yes	Yes
GS by DOW FE	Yes	Yes	Yes	Yes
$f() \times trend$	Yes	Yes	Yes	Yes
X_trend	No	Yes	No	Yes
Adjusted R-squared	0.939	0.930	0.941	0.925
N	10798	20971	7681	14511

Notes: This table shows the estimation of the effect of gasoline emission standard reforms on different gas stations, using the two-step method described in section 3.3. Column 1 and 2 examine the gas stations near a highway or a national highway. Column 3 and 4 examine the gas stations that are not close to any highway or national highway. The dependent variable is $log(Q_{jk}) - \hat{\alpha}\, p_k^j$. We use -0.376 as the value of $\hat{\alpha}$, which is a conservative estimate in Table A.9. The variable DoW stands for day-of-the-week; X_{kt} includes interactions between gas station characteristics and time trend. The order of distance polynomial $f(\cdot)$ is 2. The distance bandwidth is 30 km. The event time window is 70 days. For columns 2 and 4, the treated year dummy and its interactions with the Hebei and post-reform dummies are included. We present cluster standard errors in parentheses. Standard errors are clustered to gas station level. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 11: Effects of Emission Standard Reform on Cross-Grade Sales

	(1)	(2)	(3)	(4)
	Beijing sample	Hebei sample	Treated year	transition period
$Regular \times Post \times treatedyear$	0.0223	-0.0570		
	(0.028)	(0.043)		
$Regular \times Post \times Hebei$			0.0121	0.0785
			(0.035)	(0.134)
Date FE	Yes	Yes	Yes	Yes
Gas station FE	Yes	Yes	Yes	Yes
GS by DOW FE	Yes	Yes	Yes	Yes
$f() \times trend$	Yes	Yes	Yes	Yes
Adjusted R-squared	0.903	0.833	0.888	0.928
N	33919	27727	31767	8525

Notes: This table shows the estimation of the effect of gasoline emission standard reforms on the sales gap between regular and premium gasoline, using the two-step method described in section 3.3. We collapse gasoline with different standards by their grades and keep gas stations that sell both regular and premium gasoline. In our sample, nearly one third of gas stations only sell regular gasoline. Column 1 and 2 report the results of gas stations in Beijing and Hebei, and the treated year dummy and its interactions with the Hebei and post-reform dummies are included. Column 3 and 4 examine all gas stations kept in the treated year and in the transition period only. The transition period is defined as in 4.1.3. The dependent variable is $log(Q_{jk}) - \hat{\alpha}\, p_k^j$. We use -0.376 as the value of $\hat{\alpha}$, which is a conservative estimate in Table A.9. The variable DoW stands for day-of-the-week. The order of distance polynomial $f(\cdot)$ is 2. The distance bandwidth is 30 km. The event time window is 70 days. We present cluster standard errors in parentheses. Standard errors are clustered to gas station level. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 12: Benefit-Cost Analysis of Other Fuel Emission Policies

	(1)	(2)	(3)
Panel A: Gasoline Standards Reform in European and Indian			
Country	Green consumers[a] (%)	Gasoline consumption[b] (million MT/year)	Aggregated WTP (billion US$/year)
China	56.6%	107.89	7.91
India	58.4%	17.26	1.96
Europe	44.0%	61.77	5.36
Panel B: Low Carbon Fuel Programs in North America			
Program	Cost	Social Benefit	Private WTP
China Gasoline Emission Standard Reform, 2013	$0.154/gallon	—	$0.204/gallon
Oregon Clean Fuels Program, 2020[c]	$0.037/gallon	$0.081/gallon	—
Canada Clean Fuel Standard, 2021 to 2040[d]	$94/tCO2	$135/tCO2	—

Notes: This table shows the benefit-cost analysis of fuel policies in other countries or regions. Panel A compares the emission standards upgrading in China, India, and Europe as their standards are almost the same. Panel B lists the clean fuel standards or programs in the US and Canada, which focus on reducing Carbon intensity compared to China's emission standards upgrading.

a. The proportion of green consumers are calculated form the World Value Survey, value 6.

b. The gasoline consumption data is collected and calculated from the U.S. Energy Information Administration. MT stands for metric tons.

c. The per gallon cost and total Greenhouse Carbon reduction come from Oregon Department of Environmental Quality; The total motor gasoline consumption comes from the EIA.

d. The cost and social benefit of the Canadian Clean Fuel Standard program come from the Canada Gazette by the Department of Environment, Canada.

Table 13: Summary of Studies on WTP for Eco-friendly Goods

Paper	Region	Content of Policy	WTP
This paper	China	Gasoline standard upgrade III to IV	4.65% of gasoline price
Guo et al. (2020)	China	Improve air quality to WHO standards	65 CNY (10 USD) each year
Roe et al. (2001)	US	10% increase in the renewable energy generation capacity	6.5% of the monthly bill
Mozumder et al. (2011)	NM, US	10% increase of renewable energy in power supply	14% of the monthly bill
Borchers et al. (2007)	DE, US	A generic green energy source compared to other specific sources	8-16% of the average bill
Guo et al. (2014)	China	Renewable electricity from natural gas, hydro and biomass energy sources	4.94% of the monthly bill
Sexton and Sexton (2014)	US	The green signal provided by the distinctively designed Toyota Prius	1.86-18.26% of the car price
Bjørner et al. (2004)	Danmark	Green labels on brand choice for toilet paper	13-18% of the price.

Notes: This table lists the existing studies on WTP for renewable energy or Eco-label goods.

Appendix A Additional Results

A.1 Additional Analysis

RD Validity—

A.2 Additional Figures

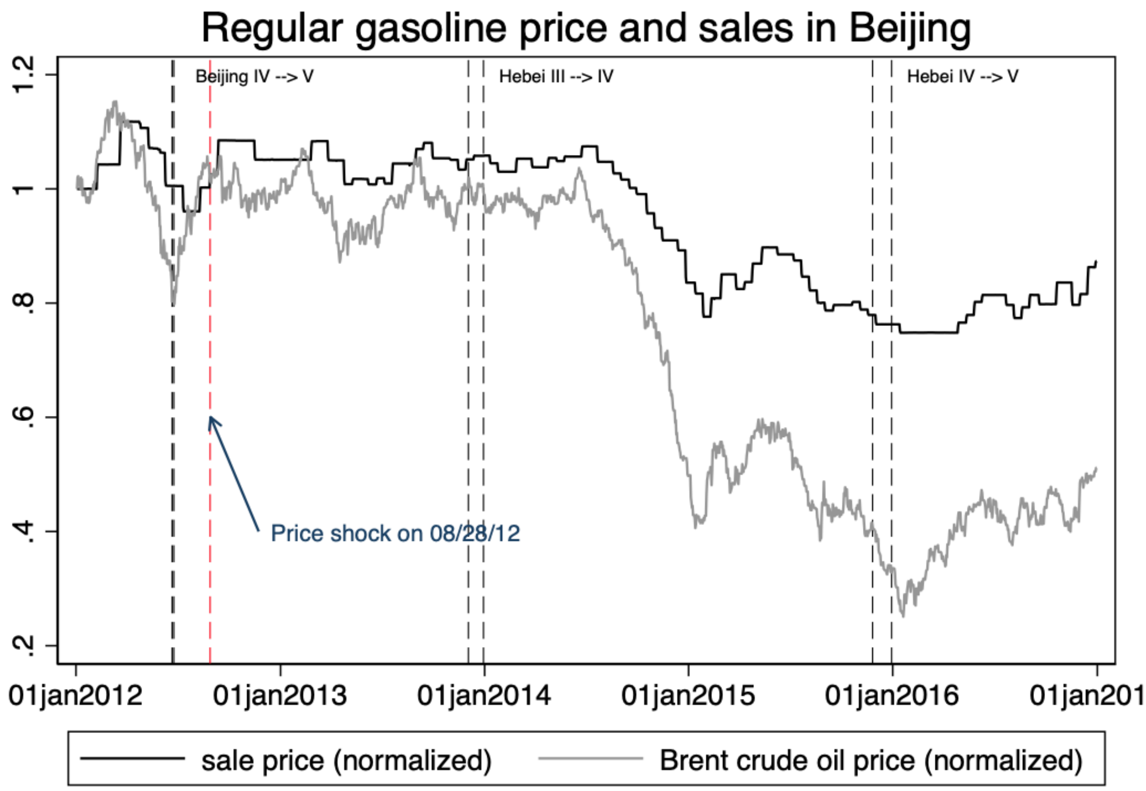


Figure A.1: The Regular Gasoline Price in Beijing and International Crude Oil Price

Notes: This figure plots the sales price of regular gasoline in Beijing within our sample period. The black line shows the normalized sales price in Beijing. The grey line shows the normalized Brent crude oil price. The black vertical dash lines demonstrate the transition period of the three different reforms. The red vertical dash line shows the price surge in Beijing on August 28th, 2012.

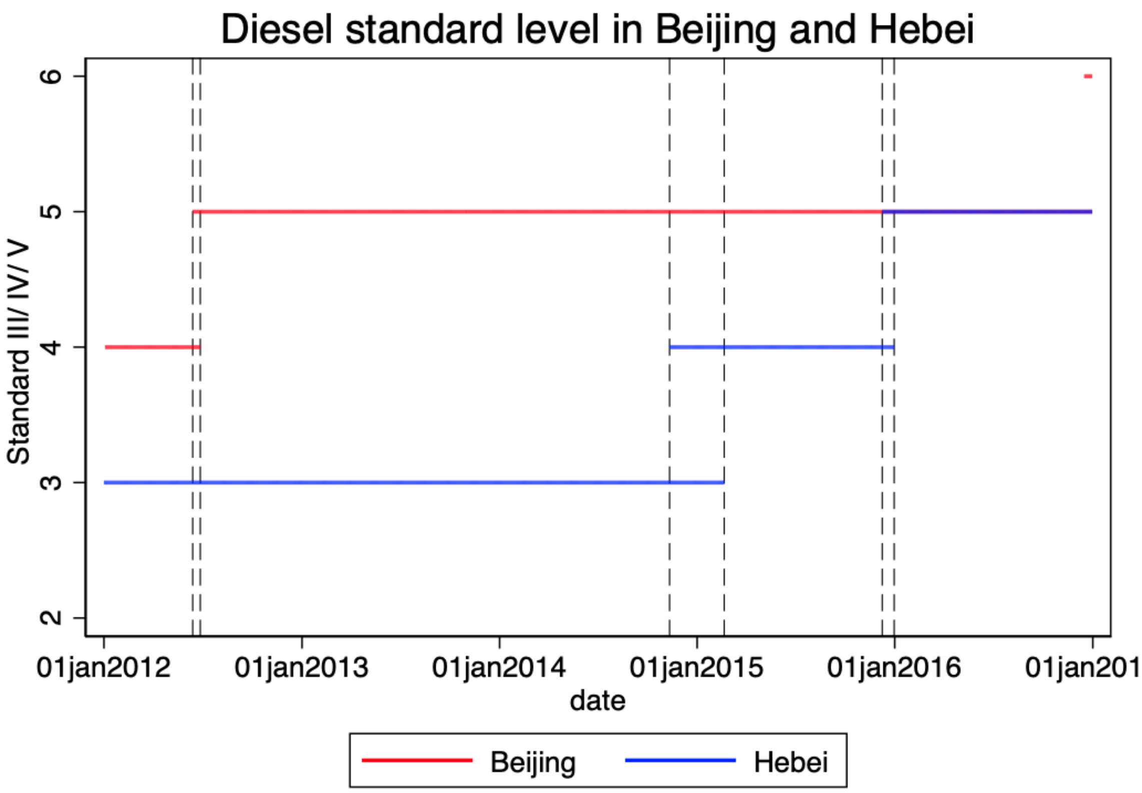
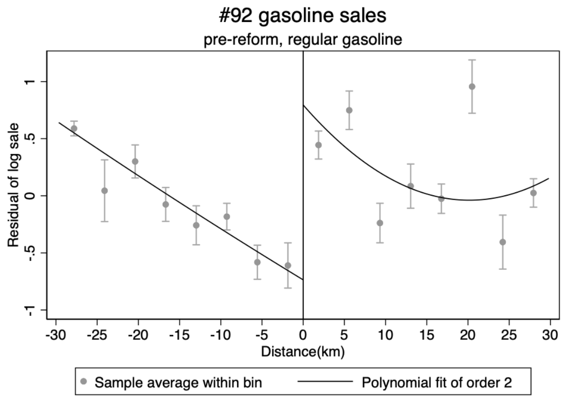
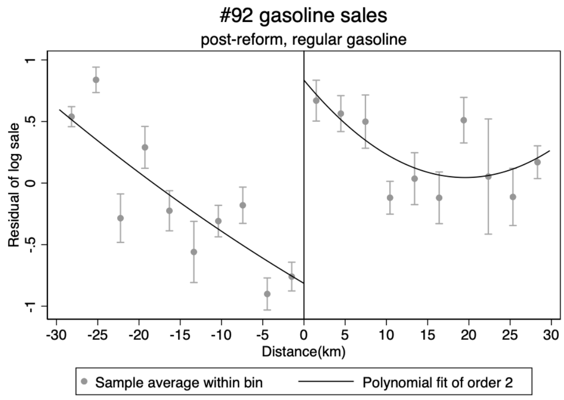


Figure A.2: Diesel Emission Standard Change Before and After the Reforms

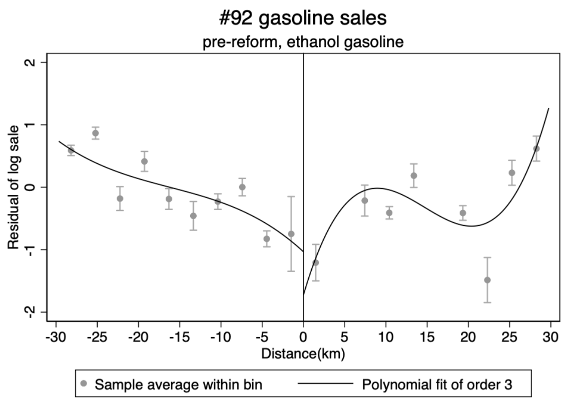
Notes: This figure presents the timeline of diesel standard upgrades in Hebei and Beijing, demonstrated in blue line and red line. The vertical axis shows the standards. After Hebei IV-V reform, both regions have standard V and the two lines overlaps.



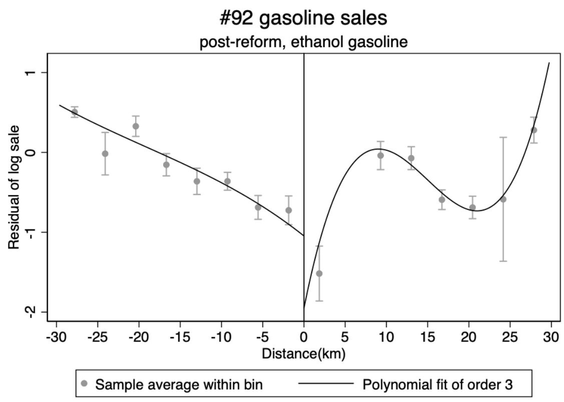
(a) Pre-reform, regular pure gasoline



(b) Post-reform, regular pure gasoline



(c) Pre-reform, regular ethanol gasoline



(d) Post-reform, regular ethanol gasoline

Figure A.3: Residual log sales of the Gasoline at Two Side of the Border Before and After the Reform, Regular and Ethanol Gasoline

Notes: This figure is the RD plot of the residual of $log(Q_{jk})$ of regular gasoline, before and after the enforcement of standard IV in Hebei. Figures (a) and (b) present the log sales of regular gasoline in Beijing and in Hebei cities. Figure (c)and (d) present the log sales of regular gasoline in Beijing and of E10 ethanol gasoline in Hebei cities. Each Hebei city either sells regular gasoline or E10 ethanol gasoline, so the Hebei cities in (a) (b) and (c) (d) are different. The predicted value with each bins, the confidence interval and fitted line are shown in the figure. The solid vertical line represents the boundary. Negative/positive values of distance give the distance of gas station at the Beijing/Hebei side of the border respectively.

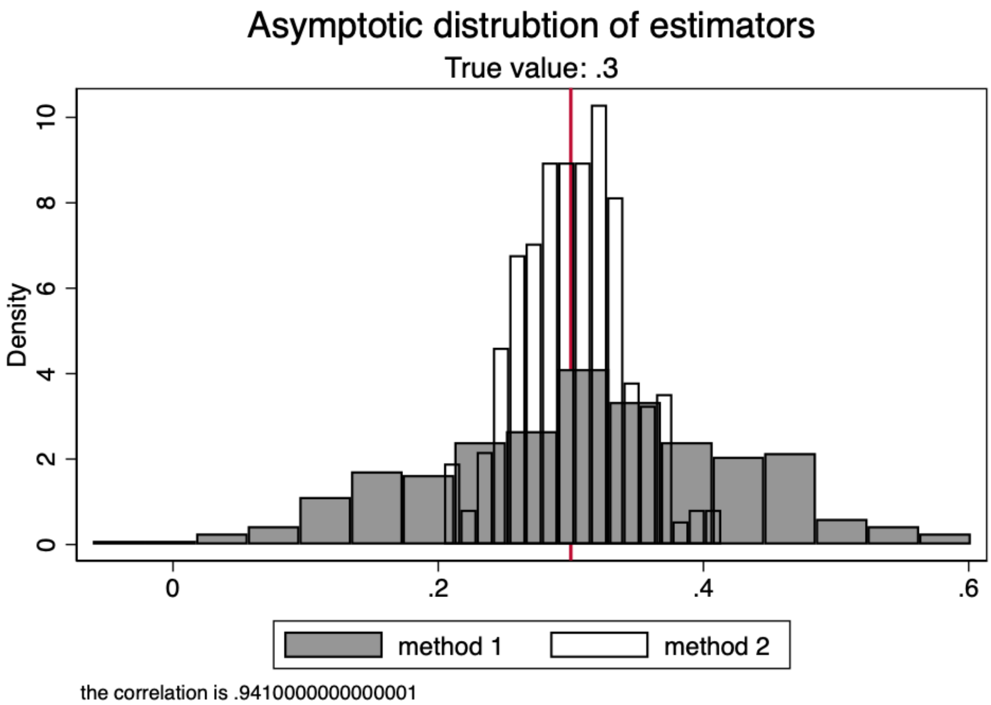
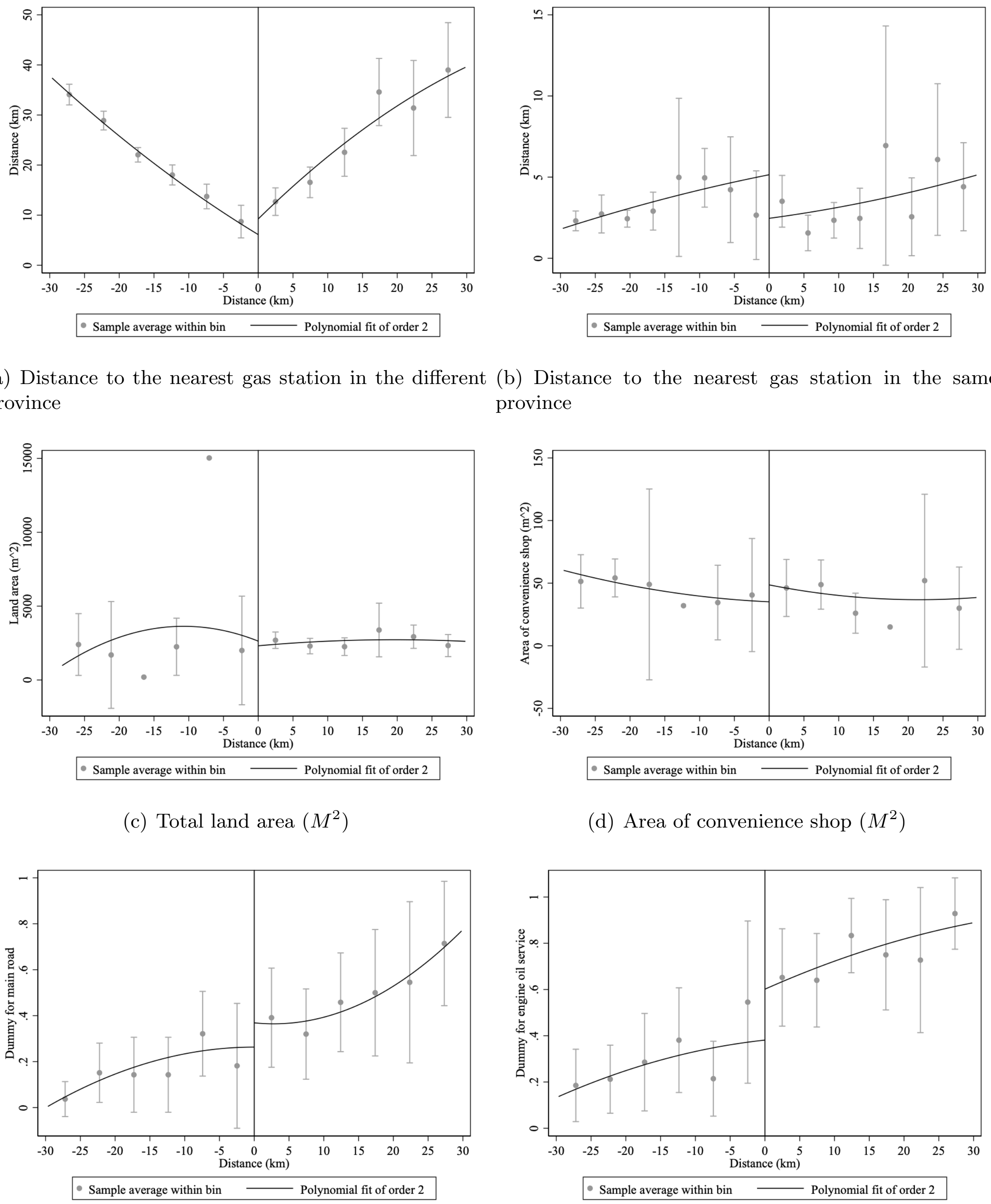


Figure A.4: Monte Carlo Simulation of One-step Estimator and Two-step Estimator

Notes: This figure plots the asymptotic distribution of the one-step and two-step estimators. The number of simulation is 300. Method 1 is one-step estimation and method 2 is two-step estimation. For the two-step method, we use the true value of α to calculate the dependent variable for the second step. When regularity condition $E[\epsilon^p_{jkt} \cdot \xi_{jkt}] = 0$ holds, both method generate consistent estimation. However the asymptotic variance is much larger for the second method.



(a) Distance to the nearest gas station in the different province

(b) Distance to the nearest gas station in the same province

(c) Total land area (M^2)

(d) Area of convenience shop (M^2)

(e) Dummy for near the main road

(f) Dummy for providing engine oil replacing service

Figure A.5: RD Plots for the Covariates (Balance Test)

Notes: This figure presents the RD plots for the gas station level characteristics. The solid vertical line represents the geographical boundary. The order of polynomial is 2. The main road is defined as interstate highways and national highways.

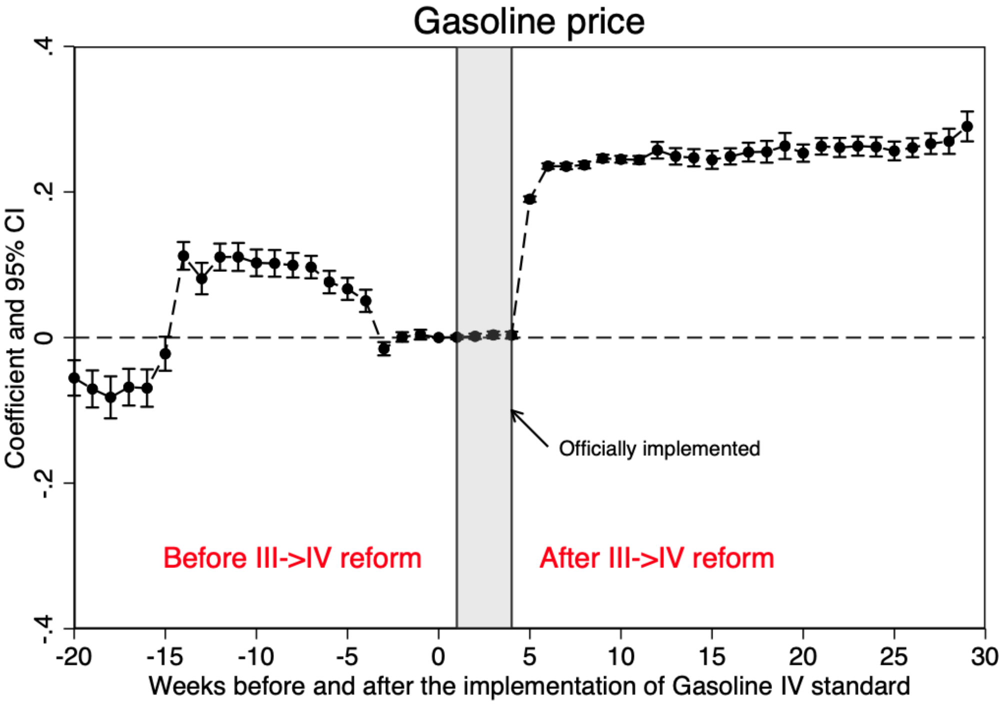


Figure A.6: Tests for Parallel Trends and Dynamic Effects on Gasoline Price

Notes: This figure presents a flexible event-study of the enforcement of higher emission standard in Hebei. The dependent variable is price. The figure plots the estimated coefficients and 95% level confidence intervals for each day within the event-study window. The grey region represents the transition period when gas station in Hebei sells both III and IV gasoline. he sample used here only includes regular gasoline sold at gas stations in Beijing and Hebei cities that sell pure gasoline. The order of distance polynomial $f(\cdot)$ is 2. The distance bandwidth is 30 km. The Standard errors are clustered to gas station level.

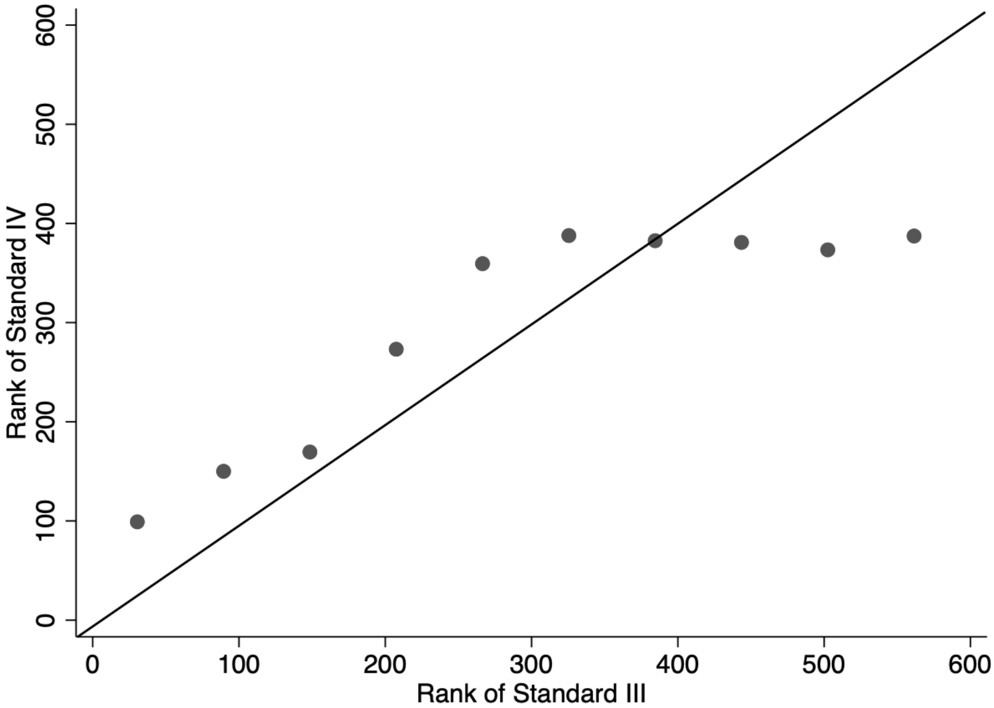


Figure A.7: Relative Price Ranking for Stations: Standard III and IV

Notes: This figure presents relative gasoline price changes across gas stations for the standard reform III and IV. The horizontal axis is the rank of price for standard III. The vertical axis is the rank of price for standard IV. For expositional purposes, We aggregate the rank into 10 bins. The straight line is the 45 degree line.

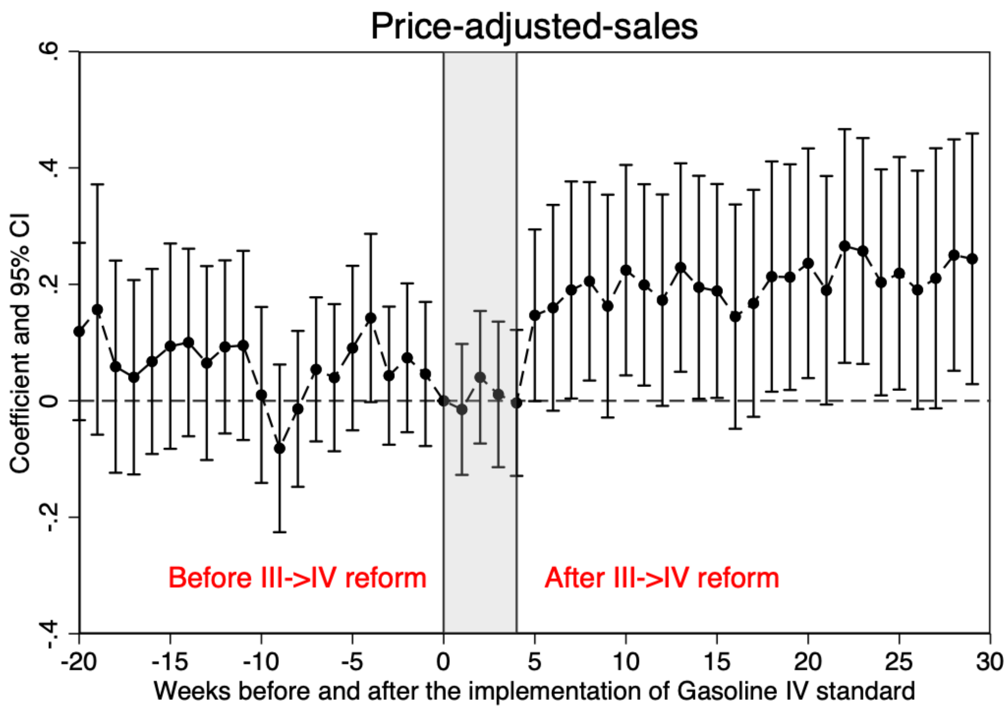


Figure A.8: Event-study Results: One-step Method

Notes: This figure presents a event-study estimation of the enforcement of higher emission standard in Hebei, using one-step method. The figure plots the estimated coefficients and 95% level confidence intervals of price-adjusted sales for each week within the event-study window. The grey region represents the transition period when gas station in Hebei sells both III and IV gasoline. The sample used here is for regular gasoline. Gas stations in both Beijing and Hebei sell pure gasoline. The order of distance polynomial $f(\cdot)$ is 2. The distance bandwidth is 30 km. The Standard errors are clustered to gas station level.

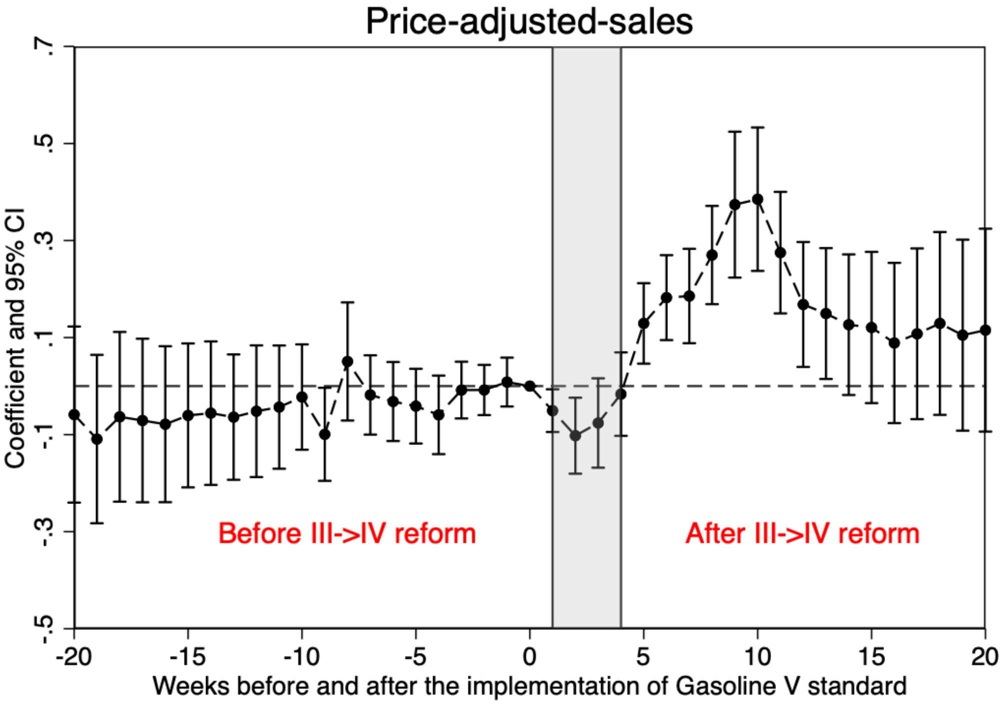
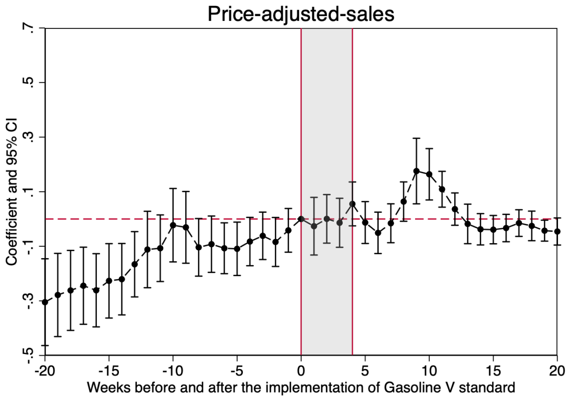
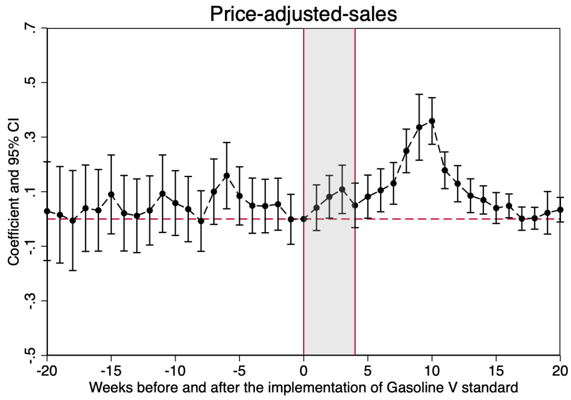


Figure A.9: Event-study Result using Specification 1: III to IV Reform

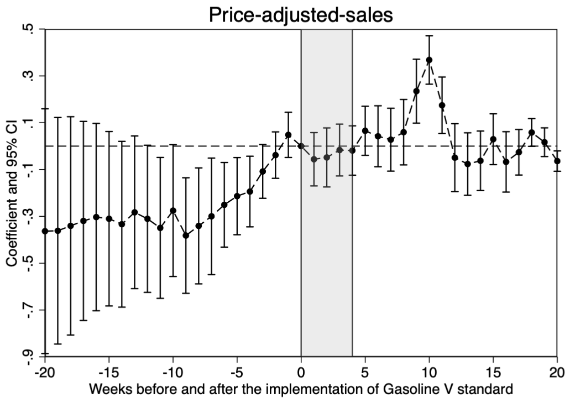
Notes: This figure presents a event-study estimation of the enforcement of higher emission standard in Hebei, based on the specification 1. The figure plots the estimated coefficients and 95% level confidence intervals of price-adjusted sales for each week within the event-study window. The grey region represents the transition period when gas station in Hebei sells both III and IV gasoline. The sample used here is for regular gasoline. Gas stations in both beijing and hebei sell pure gasoline. The order of distance polynomial $f(\cdot)$ is 2. The distance bandwidth is 30 km. The Standard errors are clustered to gas station level.



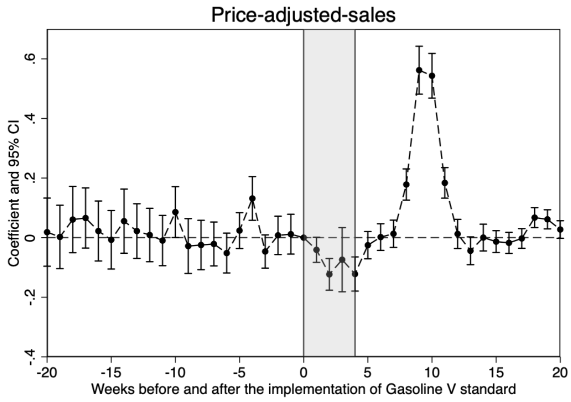
(a) Boundary and 2012 sample



(b) Boundary and 2014 sample



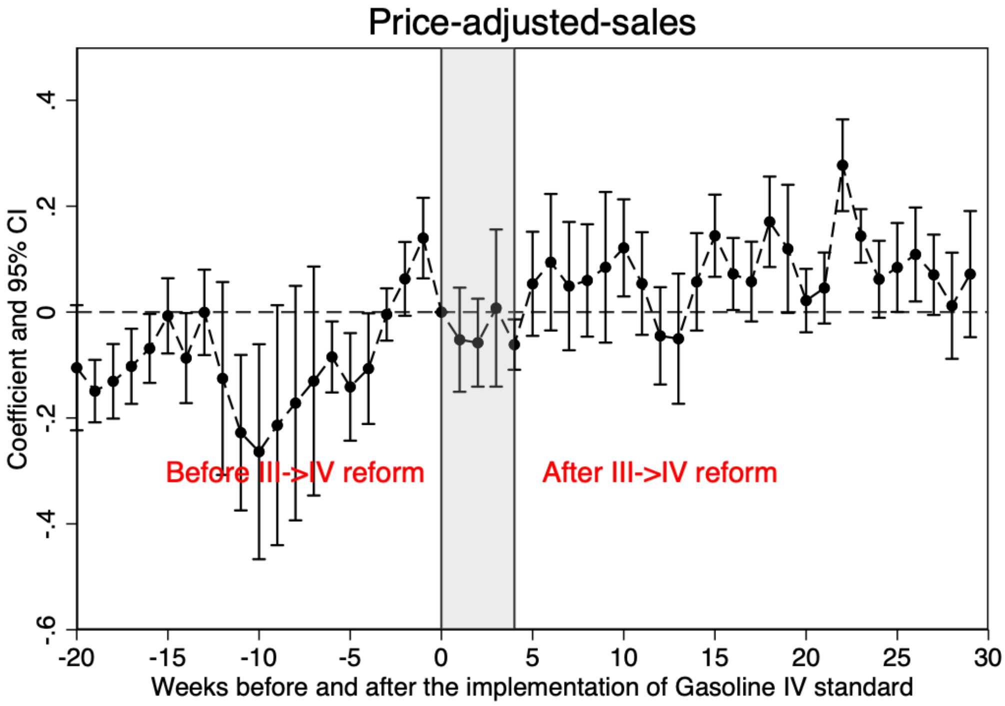
(c) Boundary and ethanol region



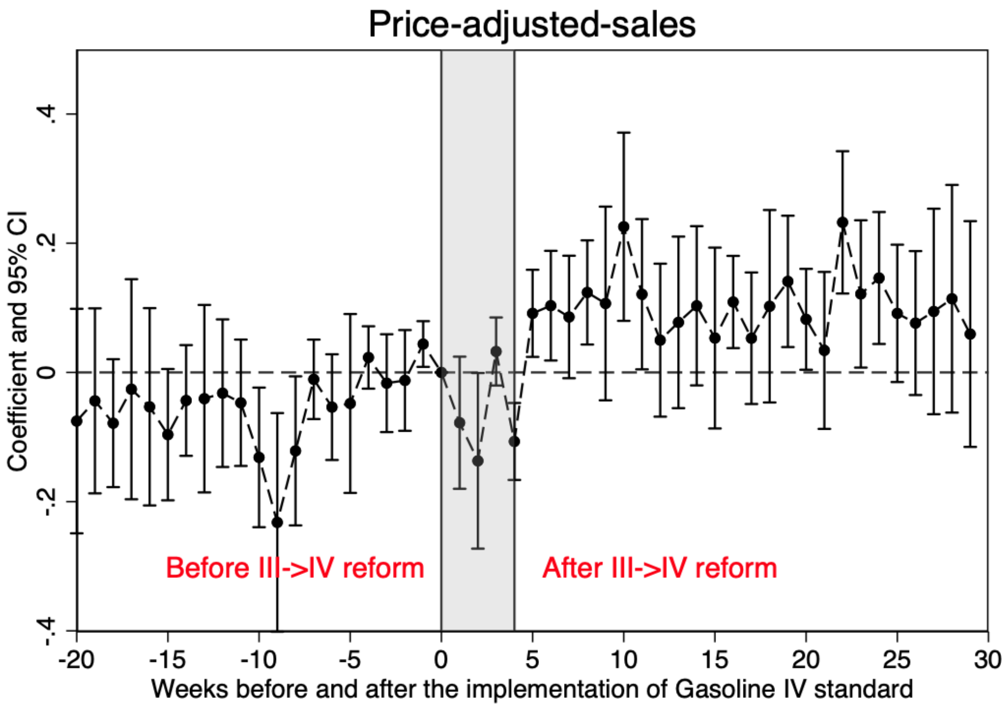
(d) Nonboundary and 2013 sample

Figure A.10: Event-study using Specification 1: Placebo Tests

Notes: This figure presents event-study placebo tests, based on the specification 1. Figure (a) and (b) preform event-study on two placebo year, 2012 and 2014, using the same sample criteria as the benchmark results. Figure (c) and (d) study the true event year (2013), however, using placebo sample. Specifically, Figure (c) studies the ethanol boundaries where the Beijing side sells gasoline and the Hebei sides sell ethanol gasoline (E10). Figure (d) studies the gas stations 30 km away from the border. The figures plot the estimated coefficients and 95% level confidence intervals of price-adjusted sales for each week within the event-study window. The grey region represents the transition period when gas station in Hebei sells both III and IV gasoline. The order of distance polynomial $f(\cdot)$ is 2. The distance bandwidth is 30 km. The Standard errors are clustered to gas station level.



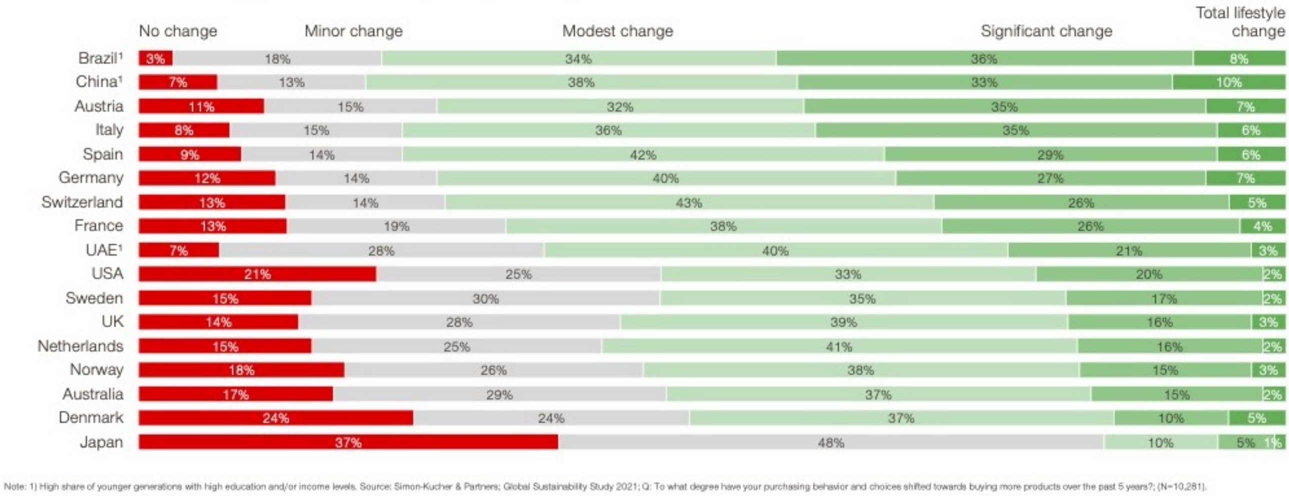
(a) Ethanol boundary sample



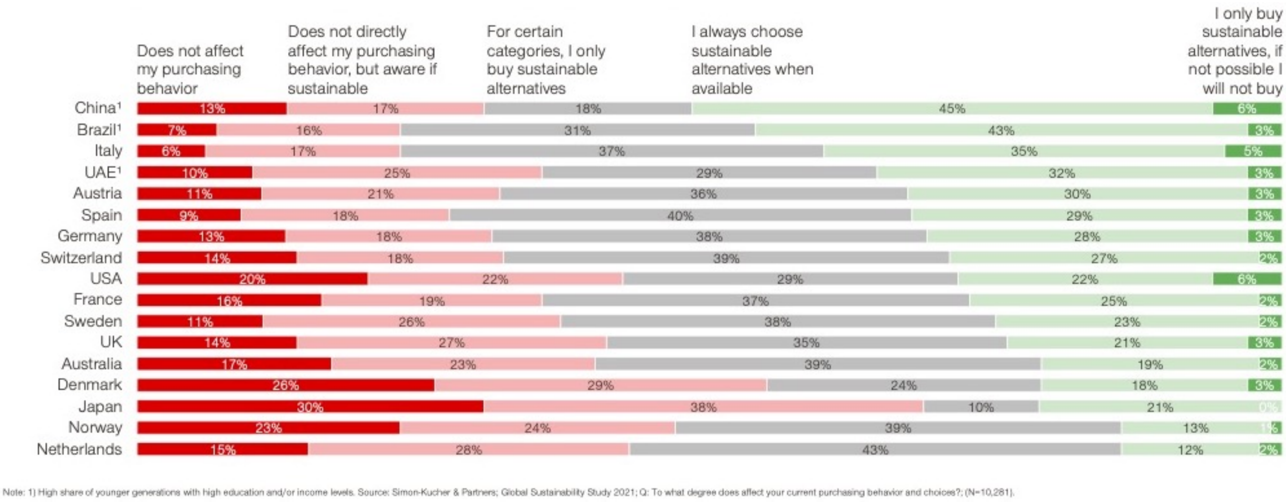
(b) Non-boundary sample

Figure A.11: Event-study: Placebo Test

Notes: This figure presents event-study placebo tests, based on the specification 2. Figure (a) studies the ethanol boundaries where the Beijing side sells gasoline and the Hebei sides sell ethanol gasoline (E10). Figure (b) studies the gas stations 30 km away from the border. The figures plot the estimated coefficients and 95% level confidence intervals of price-adjusted sales for each week within the event-study window. The grey region represents the transition period when gas station in Hebei sells both III and IV gasoline. The order of distance polynomial $f(\cdot)$ is 2. The distance bandwidth is 30 km. The Standard errors are clustered to gas station level.



(a) Shift in purchasing behavior past 5 years



(b) Current purchasing behavior

Figure A.12: Shifted and Current Purchasing Behavior Across Countries and Regions

Notes: This figure shows the percentage of respondents shifted towards or currently have an environmental-friendly purchasing behavior. Panel (a) presents the component of respondents based on to what degree they have shifted purchasing behavior and choices towards buying more products over the past 5 years. Panel (b) presents the component of respondents based on to what degree they would buy sustainable alternatives currently. Source: Global Sustainability Study 2021.

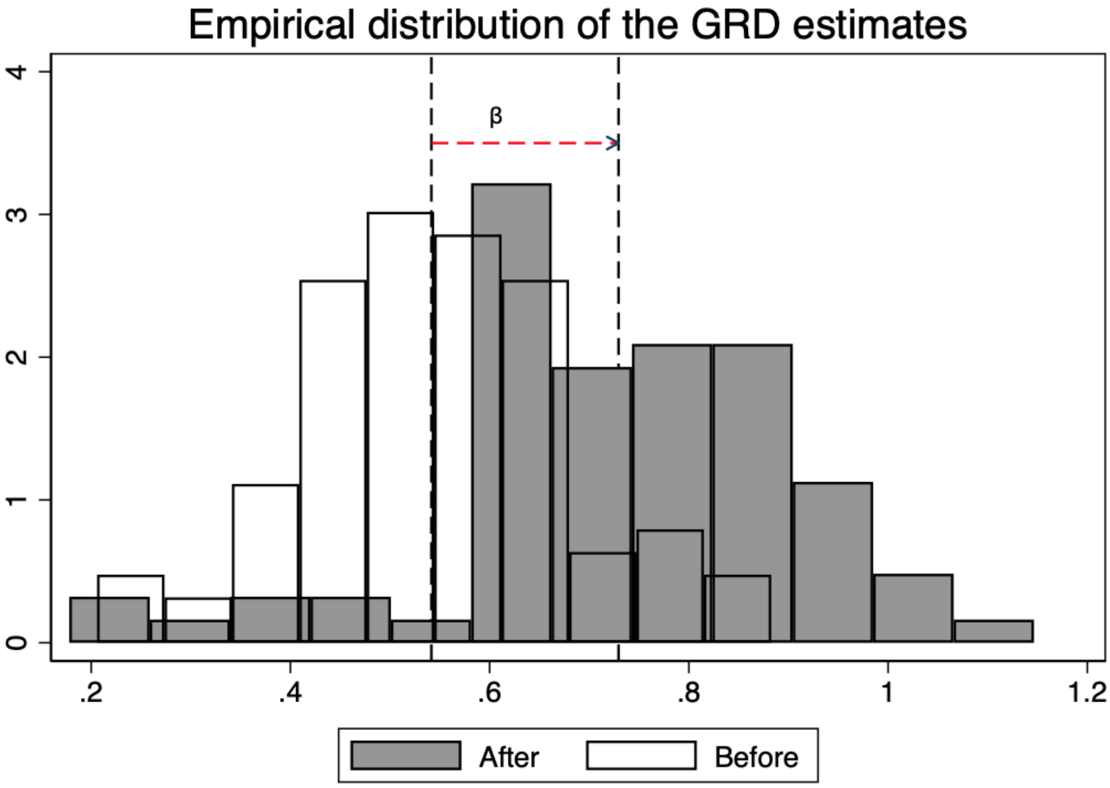


Figure A.13: Empirical Distribution of Sales Gap

Notes: This figure plots the empirical distribution of daily sales gap at boundary. Specifically, we use a geographical regression discontinuity (GRD) regress to estimate the relative sales gap. The analysis is performed on daily basis for 100 days before and after the reform. This figure plots the distribution of daily GRD estimates for pre- and post- reform. This shifting of distribution is a graphical presentation of our main estimation.

A.3 Additional Tables

Table A.1: Gasoline Emission Standard of Different Countries and Regions

Fuel Property	China III	China IV	China V	Euro III	Euro IV	Euro V	EPA RFG Summer	EPA RFG Winter	CARB III	India III	India IV
Time Implemented	2009	2013	2017	2000	2005	2009	1995-2000	1995-2000	2003	2010	2017
Research Octane (RON), min.	97-90	97-90	95-89	95-91	95-91	95-91	NS	NS	NS	91	91
Motor Octane (MON), min.	88-85	88-85	90-84	85–81	85–81	85–81	NS	NS	NS	81	81
Aromatics, vol%, max.	40	40	40	42	35	35	20.74	19.54	25	42	35
Olefin, vol%, max.	30	28	24	18	18	18	11.9	11.2	6	21	21
Benzene, vol%, max.	1	1	1	1	1	1	0.66	0.66	0.8	1	1
Sulfur, ppm, max.	150	50	10	150	50	10	71	81	20	150	50
Gum Content, mg/100ml, max.	5	5	5	NS	NS	5	5	5		4	4
Density 15C, kg/m3	N/A	N/A	720-775 (20°C)	NS	NS	720-775	NS	NS	N/A	720-775	720-775
RVP, kPa, max.	88 Winter 72 Summer	42-85 Winter 40-68 Summer	45-85 Winter 40-65 Summer	60/70	60/70	60/70	47.6 8 (6.91 psi)	82 (11.89 psi)	48.2/47.6 (7.00 psi)	60	60
Lead, mg/l, max.	5	5	5	5	5	5	13	13	ND	5	5
Manganese, mg/liter, max.	16	8	2	NS	NS	MMT¡6 (by 2011) MMT¡2 (by 2014)	NA	NA	ND	NS	NS
Oxygen, % m/m	2.7 (max.)	2.7 (max.)	2.7 (max.)	2.7 (max.)	2.7 (max.)	2.7 (max.)	2.49	2.37	1.8-2.2	2.7	2.7

Notes: This table shows the gasoline emission standards of different countries and regions. RON and MON are different ways to measure the Octane number. NS refers to Not specified; ND refers to Non-detectable. Winter is from November to April, while summer is from May to October.
Sources: International Council on Clean Transportation (2014); U.S. Environmental Protection Agency (EPA) website.

Table A.2: Emission Reduction from the Standard Reforms

	(1) III-IV	(2) III-V	(3) Difference
Olefin, vol%, max.	-6.67%	-20%	-13.33%
Sulfur, ppm, max.	-66.67%	-93.33%	-26.66%
Manganese, mg/liter, max.	-50%	-87.5%	-37.5%
Average	-41.11%	-66.94%	-25.83%

Notes: This table shows the theoretical emission reduction calculated using Table A.1. Here we only include fuel properties having improvement that can be calculated in value. Column 1 reports the reduction from China III to China IV. Column 2 reports the reduction from China III to China V. Column 3 is the difference between columns 2 and 1.

Table A.3: Balance Test on Covariates

	(1) Distance to Diff	(2) Distance to Same	(3) CVS Area	(4) Total Land Area	(5) Main Road	(6) Engine Oil Service	(7) Content Level
Hebei	3.127	-2.686	13.52	-315.1	0.106	0.220	0.0216
	(3.199)	(2.170)	(18.954)	(1978.682)	(0.193)	(0.219)	(0.284)
Bandwidth	30	30	30	30	30	30	30
Polynomial order	2nd	2nd	2nd	2nd	2nd	2nd	2nd
Adjusted R-squared	0.622	0.0308	0.0476	0.0269	0.149	0.245	0.0146
N	221	254	63	120	254	254	254
Dep mean	22.73	3.431	44.22	2624.3	0.295	0.480	1.496

Notes: This table shows the results of RD regression for gas station covariates. The independent variable is the dummy for Hebei. As for dependent variables, *Distance to Diff* represents the distance to the nearest gas station in the other province in the sample; *Distance to Same* represents the distance to the nearest gas station in the same province; *Main Road* is a dummy equals 1 if the gas station is on the side of a highway or national highway; *Engine Oil Service* is a dummy equals 1 if the gas station provides engine oil replacing service; *Content Level* is a discrete variable to indicate the total storage volume. Standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table A.4: Polynominal Regression of Sales Volume: One Step Method

	(1) Pure	(2) Pure	(3) Pure	(4) Ethanol	(5) Ethanol	(6) Ethanol
	Panel A: Before the emission standard reform					
Hebei	0.828*	1.265**	2.465*	-0.856	-0.592	1.740
	(0.433)	(0.549)	(1.309)	(0.858)	(0.792)	(1.320)
log price	2.751	2.330	-2.882	3.797	3.149	15.41
	(3.364)	(3.713)	(5.549)	(4.644)	(5.212)	(14.061)
Adjusted R-squared	0.307	0.171	0.198	0.340	0.192	0.284
N	17153	12235	6362	12936	8037	3574
	Panel B: After the emission standard reform					
Hebei	0.898**	1.309**	2.628**	-1.114**	-0.988*	0.956
	(0.401)	(0.553)	(1.274)	(0.494)	(0.533)	(1.305)
log price	4.429*	7.207**	-0.978	4.594	3.962	10.15
	(2.586)	(3.037)	(4.023)	(3.577)	(4.548)	(9.692)
Adjusted R-squared	0.285	0.127	0.180	0.352	0.215	0.336
N	17851	12461	6543	13630	8311	3836
BW	30 km	20 km	10 km	30 km	20 km	10 km
Time FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table shows polynomial RD regression for *log* sales at the border using the one step method that includes the dummy for Hebei and *log* price at the same time. The dependent variable is $log(Q_{jk})$. In the sample of columns 1 to 3, gas stations in both Beijing and Hebei sell pure gasoline. In the sample of columns 4 to 6, gas stations in Beijing sell pure gasoline and gas stations in Hebei sell E10. Panel A presents results before the emission standard reform and Panel B presents results after the reform. The Standard errors are clustered to gas station level. Standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table A.5: Local Linear Regression of Sales Volume at the Geographical Boundary: Two Step Method

	(1) LLR	(2) LLR with covariate	(3) LLR	(4) LLR with covariate
Fuel type	Pure	Pure	Ethanol	Ethanol
	Panel A: Before the emission standard reform			
Hebei	1.055**	0.862**	-0.387	-0.101
	(0.431)	(0.412)	(0.575)	(0.554)
N	100	82	185	182
BW	15.13	12.47	46.5	42.42
	Panel B: After the emission standard reform			
Hebei	1.067**	0.768*	-0.148	0.571
	(0.487)	(0.429)	(0.623)	(0.618)
N	87	79	149	144
BW	12.9	11.59	28.85	27.57

Notes: This table shows local linear regression for *log* sales at the border. To implement the LLR procedure following Calonico et al. (2014) , we collapse the panel data into cross-section data by taking average. In the sample of columns 1 and 2, gas stations in both Beijing and Hebei sell pure gasoline. In the sample of columns 3 and 4, gas stations in Beijing sell pure gasoline and gas stations in Hebei sell E10. Panel A presents results before the emission standard reform and Panel B presents results after the reform. Standard errors are clustered to gas station level. Standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table A.6: Price Elasticity Estimation: Reduced-form Regression

	(1) *log* sale	(2) *log* sale	(3) *log* sale
$Hebei \times Post$	0.0622***	0.0517**	0.135***
	(0.020)	(0.026)	(0.050)
Date FE	Yes	Yes	Yes
Gas station FE	No	Yes	Yes
$f(\cdot) \times trend$	No	No	Yes
Adjusted R-squared	0.962	0.962	0.962
N	11327	11327	11327

Notes: This table shows the price elasticity estimation using the Beijing price surge as instrument. The instrument variables used in this estimation is the large price shock in Beijing on August 28th, 2012. The dependent variable is *log* of sales. The key independent variable is the interaction between fuel type dummy and the indicator for post August 28th, 2012. $f(\cdot)$ is the polynomial function of distance to boundary describe in section 5.2. The order of distance polynomial $f(\cdot)$ is 2. The distance bandwidth is 30 km. The time window is 30 days before and after the price shock. The Standard errors are clustered to gas station level. Standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table A.7: Price Elasticity Estimation: OLS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Regular				
log price (elasticity)	-0.104	-0.348***	-9.566***	-4.705***				
	(0.076)	(0.041)	(1.221)	(0.473)				
price (α)					-0.022**	-0.048***	-1.974***	-0.690***
					(0.011)	(0.006)	(0.238)	(0.088)
				Panel B: Premium				
log price (elasticity)	-0.410**	-1.756***	-3.001	-2.811				
	(0.177)	(0.072)	(2.297)	(1.730)				
price (α)					-0.055**	-0.252***	-1.451***	-0.555***
					(0.025)	(0.010)	(0.531)	(0.127)
Date FE	No	No	Yes	Yes	No	No	Yes	Yes
Gas station FE	No	Yes	No	Yes	No	Yes	No	Yes
Fuel type FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes

Notes: This table shows the OLS estimation of price elasticity. The dependent variable is *log* of sales. The independent variables are the predicted value of *log* price and level of price calculated in the first-stage. The estimates use $\widehat{log price}$ and $\widehat{price}$ correspond to the price elasticity and structural parameter α, respectively. Panel A presents the results of regular gasoline. Panel B presents the results of premium gasoline. We present cluster standard errors in parentheses. Standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A.8: Price Elasticity Estimation: NRDC Price Adjustment as IV

	(1)	(2)	(3)	(4)
	Panel A: Regular			
log price (elasticity)	-2.070***	-1.894***		
	(0.106)	(0.092)		
price (α)			-0.292***	-0.268***
			(0.015)	(0.013)
	Panel B: Premium			
log price (elasticity)	-2.944***	-2.305***		
	(0.471)	(0.379)		
price (α)			-0.389***	-0.305***
			(0.062)	(0.050)
Date FE	Yes	Yes	Yes	Yes
Month-year FE	No	Yes	No	Yes
Fuel type FE	Yes	Yes	Yes	Yes

Notes: This table shows the price elasticity estimation using NRDC price adjustment as instrument. Since the variation of price adjustment announcement is completely at time-series dimension, we cannot control for date fixed effect. Instead, we include the Month-by year fixed effect and gas station by day-of-week fixed effect. The dependent variable is *log* of sales. The independent variables are the predicted value of *log* price and level of price calculated in the first-stage. The estimates use $\widehat{log price}$ and $\widehat{price}$ correspond to the price elasticity and structural parameter α, respectively. Panel A presents the results of regular gasoline. Panel B presents the results of premium gasoline. We present cluster standard errors in parentheses. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.9: Price Elasticity Estimation Using Beijing Price Surge as IV: 70 Days Event Window

	(1) log price	(2) log price	(3) price	(4) price
log price (elasticity)	-1.972*	-2.054**		
	(1.140)	(0.999)		
price (α)			-0.451*	-0.376*
			(0.261)	(0.192)
Date FE	Yes	Yes	Yes	Yes
Gas station FE	Yes	Yes	Yes	Yes
f() $\times$ trend	No	Yes	No	Yes
Adjusted R-squared	0.939	0.939	0.939	0.939
N	26194	26194	26194	26194

Notes: This table shows the estimation of price elasticity using the Beijing price surge as instrument and a time window of 70 days before and after the price shock. The instrument variables used in this estimation is the large price shock in Beijing on August 28th, 2012. The dependent variable is *log* of sales. The independent variables are the predicted value of *log* price and level of price calculated in the first-stage. The estimates use $\widehat{log price}$ and $\widehat{price}$ correspond to the price elasticity and structural parameter α, respectively. Panel A presents the results of regular gasoline. Panel B presents the results of premium gasoline. $f(\cdot)$ is the polynomial function of distance to boundary describe in section 5.2. The order of $f(\cdot)$ is 2. The distance bandwidth is 30 km. The time window is 30 days before and after the price shock. We present cluster standard errors in parentheses. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.10: Effects of Emission Standard Reform: Dropping Transition Period

	(1)	(2)	(3)
$Hebei \times Post \times treatedyear$ (β)	0.260***	0.244***	0.231***
	(0.040)	(0.031)	(0.034)
Date FE	No	Yes	Yes
Gas station FE	No	Yes	Yes
GS by DOW FE	No	Yes	Yes
$f() \times trend$	No	No	Yes
Adjusted R-squared	0.00759	0.468	0.470
N	56254	56254	55519

Notes: This table shows the two-step estimation of the effect of gasoline emission standard reform, with the sample within the transition period defined in **??** dropped. The reform examined in the table is Standard III $\longrightarrow$ IV reform in Hebei. The dependent variable is $log(Q_{jk}) - \hat{\alpha}\, p_k^j$. We use -0.376 as the value of $\hat{\alpha}$, which is a conservative estimate in Table A.9. The variable *DoW* stands for day-of-the-week; X_{kt} includes interactions between gas station characteristics and time trend. The order of distance polynomial $f(\cdot)$ is 2. The distance bandwidth is 30 km. The event time window is 70 days. For all columns we include the treated year dummy and its interactions with the Hebei and post-reform dummies. Standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A.11: Effects of Emission Standard Reform: *De facto* Implementation Date

	(1)	(2)	(3)
$Hebei \times Post \times treatedyear$ (β)	0.203***	0.151***	0.150***
	(0.042)	(0.027)	(0.028)
Date FE	No	Yes	Yes
Gas station FE	No	Yes	Yes
GS by DOW FE	No	Yes	Yes
$f() \times trend$	No	No	Yes
Adjusted R-squared	0.00765	0.474	0.474
N	59982	59981	59705

Notes: This table shows the two-step estimation of the effect of gasoline emission standard reform using the *de facto* implementation date from our data as T^* for the *Post* dummy. The reform examined in the table is Standard III $\longrightarrow$ IV reform in Hebei. The dependent variable is $log(Q_{jk}) - \hat{\alpha}\, p_k^j$. We use -0.376 as the value of $\hat{\alpha}$, which is a conservative estimate in Table A.9. The variable *DoW* stands for day-of-the-week; X_{kt} includes interactions between gas station characteristics and time trend. The order of distance polynomial $f(\cdot)$ is 2. The distance bandwidth is 30 km. The event time window is 70 days. For all columns we include the treated year dummy and its interactions with the Hebei and post-reform dummies. Standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A.12: One-step Estimation of Effects of Emission Standard Reform with Price Polynominal

	(1)	(2)
$Hebei \times Post \times treatedyear$ (β)	0.130***	0.107**
	(0.048)	(0.050)
Order polynominal	cubic	quadratic
Adjusted R-squared	0.881	0.880
N	61787	61787

Notes: This table shows the one-step estimation of the effect of gasoline emission standard reform with a nonlinear effect from price. The reform examined in the table is Standard III $\longrightarrow$ IV reform in Hebei. The dependent variable is $log(Q_{jk}) - \hat{\alpha}\, p_k^j$. The distance bandwidth is 30 km. The event time window is 70 days. For all columns we include the treated year dummy and its interactions with the Hebei and post-reform dummies. Standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Appendix B Data Appendix

Appendix C Additional Background Information

C.1 China Gasoline Emission Standard in Global Context

To better understand China gasoline retail market in a global context, we also provide some comparisons between China and other countries, in terms of standards and regulations. Besides the duopoly market and the price cap regulation, China's regulation and standard upgrading process are quite similar to other parts of the world. First, as mentioned, China built its standards based on European ones. In view of the fact that the European standards are also blueprints for many countries' fuel emission standards, such as India's Bharat stage emission standards, the responsiveness in China could also be comparable to those countries to some extent. Appendix Table A.1 presents the detailed requirements on each component in China and a list of countries. We can see the European and European-based standards, including China and India ones, are similar to each other, while the US standards, including the EPA federal reformulated gasoline (RFG) and California's Air Resources Board (CARB), represent a different goal, as EPA RPG specifically aimed at reducing ozone pollution and CARB primarily regulates VOC and NO_X.

Also, China is not the only country allowing regional differences in regulation implementations. As shown, in the US, California had its own gasoline standard built by CARB, which is stricter than the EPA RFG standard. In India, the more stringent Bharat emission standard would usually be piloted in National Capital Region (Delhi) and fourteen other large cities for several years before expanding nationwide. We believe this is also a common practice, especially in countries with large geographic areas where resources are unevenly distributed, which makes it important to understand the consequence and mechanism of the gasoline standards regulation.

Appendix D Additional Model Results

D.1 Derive the estimating equation

In this section, we discuss how our empirical model in section 3 can be derived from a more generalized model with consumers' utility on both gasoline and gas station attributes. The model bases on the paternalism environmental regulation framework following Allcott et al. (2014) and Allcott and Taubinsky (2015).

Consumer's Choice. We consider consumer's discrete choice over fuel types and gas stations in a single market. The consumer's choice set consists of two goods labeled as H and L. In our empirical setting, H represents gasoline with a higher emission standard (Beijing, V), while L represents gasoline with a lower emission standard (Hebei, III or IV). We use c_t^j to denote cost of gasoline product $j \in \{H, L\}$ and let $c = c^H - c^L$ denote the relative cost. We assume the relative cost is determined exogenously and allow it to vary over time. Consumer's value from gasoline comes from two parts: v^j and e^j. v^j denotes the utility from gasoline j and is assumed to be uncorrelated with emission standard or non-environmental values. We denote v as $v^H - v^L$. In other words, v captures all value difference between gasoline types that is irrelevant to the emission standard, e.g., damage on the engine, explosion resistance, combustion efficiency, or even local protectionism and loyalty. It is worth noticing here that in our homogeneous agent framework, v could also derive from differences in the transportation network, demographic, and family wealth between two regions. e^j captures consumers' value on the emission standard of gasoline type j, or environmental values; let $e = e^H - e^L$ denote the preference gap. e could be environmentalism or warm-glow value from using more environmental-friendly fuel. We assume e doesn't change over time except for the policy that we are interested in to improve emission standard in Hebei (III $\longrightarrow$ IV). The consumer also choose among gas stations $k \in K$.

Empirical Equations. In our empirical context, there are only two types of fuels, H (which is sold in Beijing) and L (which is sold in Hebei), and there is no within-region variation of gasoline characteristics. Therefore, we capture gasoline-specific characteristics e and v as a whole using discrete indicator variable of fuel type. Define D_j as a indicator function $\mathbb{1}(j = L)$, which equals 1 to if j is L, and equals to 0 otherwise. We define the gasoline-specific part of the utility function ω_j as,

$$\begin{aligned}\omega_j &= (e^H)\,(1 - D_j) + (e^L)\,D_j \quad + \quad (v^H)\,(1 - D_j) + (v^L)\,D_j \\ &= e^H - e \cdot D_j \quad + \quad v^H - v \cdot D_j\end{aligned}$$

The pre-reform utility is sepecified as,

$$u_{ijk} = \alpha\, p^{pre}_{jk} + \underbrace{(v^E - e^E)\,D_j + (v^I - e^I)\,(1 - D_j)}_{\text{gasoline attributes}} + \underbrace{G_k \cdot \gamma}_{\text{gas station attributes}} + \xi_{jk} + \varepsilon_{ijk},$$

$$i = 1, ..., I, \quad j = H, L \quad k = 0, 1, ..., K$$

By assuming T1EV distribution on individual error term, the empirical equation of *log* gasoline sales before the reform is,

$$log(Q_{jk}) = \underbrace{(v^H + e^L)}_{\text{Constant}} - \underbrace{\alpha\, p_{jk} - (v + e)\,D_j + G_k \cdot \gamma}_{\text{Utility from fuel } j \text{ in gas station } k} + \xi_k \tag{C.1}$$

The sales volumes equation is consists of three parts. The first part is a constant term that captures market size, consumer's budget, and benchmark value of fuel consumption; the second part is consumer's utility of consumption fuel type j in gas station k; the last term is the stochastic component to impact gasoline sales. Empirical sales equation C.1 makes a clear idea that consumers' relative value between H and L determines the equilibrium sales. The regulated retail gasoline price depends on the refine cost difference between two regions and gas station-specific random component a_k,

$$p^{pre}_{jk} = c^L + c\,D_j + a^{pre}_k$$

Next, we introduce the time dimension into this framework and explore the impact of the emission standard reform. The enforcement of high-quality gasoline standards in Hebei (III $\longrightarrow$ IV) improves the emission standard of fuel L, hence changes the consumer's relative value between H and L, and also the relative price. The post-reform sales equation is then,

$$log(Q_{jk}) = (\tilde{v}^H + \tilde{e}^H) - \alpha\, p^{post}_{jk} - (\tilde{v} + \tilde{e})\,D_j + \tilde{G_k} \cdot \gamma + \xi_k \tag{C.2}$$

where, $\tilde{v}^j$ and $\tilde{e}^j$ are the marginal value of fuel after the reform. Denote $\tilde{v} = \tilde{v}^H - \tilde{v}^L$, and $\tilde{e} = \tilde{e}^H - \tilde{e}^L$. The price changes come from both higher refinement cost $\tilde{c}$ and gas station specific price adjustment $\tilde{a_k}$ changes. Equation C.2 indicates that increasing of e^L changes sales of both fuel H and L as it alters the relative value. We allow gas station characteristics G_k to change overtime. In the empirical exercise in section 5, we capture these changes using the interaction between distance polynomial and time trend. The post-reform price is,

$$p^{post}_{jk} = \tilde{c}^L + \tilde{c}\,D_j + a^{post}_k$$

where $\tilde{c} = \tilde{c}^H - \tilde{c}^L$ is the post-reform relative refine cost and $\tilde{c} < c$. Incorporating equation C.1 and C.2

into an unified equation gives the estimable equation,

$$\begin{aligned} log(Q_{jkt}) = & \underbrace{(v^H + e^H + c^H + a_k^{pre})}_{\text{constant term}} + \\ & \underbrace{\left[-(\triangle v + \triangle e)\cdot D_j\, \mathbb{1}(t \geq T^*)\right] + (\triangle v^H + \triangle e^H)\cdot \mathbb{1}(t \geq T^*) + \left[-(v+e)\cdot D_j\right]}_{\boldsymbol{\Lambda_{jt}} \text{ part in equation 6}} + \\ & \alpha\, \underbrace{\left[-\triangle c \cdot D_j\, \mathbb{1}(t \geq T^*) + (a_k^{post} - a_k^{pre})\cdot \mathbb{1}(t \geq T^*) + (-c\cdot D_j)\right]}_{\text{equation 7}} + \\ & G_k \cdot \gamma + \xi_{jkt} \end{aligned} \tag{C.3}$$

where $\mathbb{1}(t \geq T^*)$ is a indicator function for the post-reform period; $\triangle v^L = \tilde{v}^L - v^L$; $\triangle e^L = \tilde{e}^L - e^L$; $\triangle v = \tilde{v} - v$; $\triangle e = \tilde{e} - e$; $\triangle c = \tilde{c} - c$.[58] T^* is the time spot when the emission standard reform happens. Rewriting equation C.3 using reduced-form notation, we get the estiamting equation 6. The identification assumption is $\triangle v = 0$, and thus our estimated β corresponds to $-\triangle e$ here.

D.2 Discussion of the Modeling Choice

Athother We briefly discuss why we are not using the "bunching" method. The key tasks of this paper is to investigate consumers' choices at the geographic boundary. In general, there are two alternative frameworks to describe the consumers' cross-border choices, the discrete choice model and the continuous first-order conditions method. The first method, borrowed from the Industrial Organization (IO) literature, studies agents' utility over a discrete production space (fuel type by the gas station in our context). The jump of exogenous parameters α, e, and v is captured by the product differentiation. A recent example of using this framework is Ito and Zhang (2020). The second method, which is widely used in the public economics literature, models agents' choice over continuous characteristic variables and derives the first-order conditions of the optimization problem to study the marginal impact of "discontinuity" or "kink" patterns of the parameters. A well-known application of this framework is the bunching and notching method to estimate an elasticity parameter (Kleven, 2016). Our paper chooses the discrete choice framework for the following three reasons: First and foremost, the problem of gas station choice is intuitively closer to the discrete choice model. Since the gas stations are distributed sparsely along the road, it is hard to distinguish the bunching mass with the non-smoothing nature of gas station distribution. Second, the bunching and notching framework requires a continuous variable in the utility function to derive the first-order condition. The natural candidate in our context is a distance measure d.[59] In this case, the bunching mass at the border is informative to the elasticity of distance from respect to the environmental value, $dlog\, d/dlog\, \Lambda$. However, It is hard to interpret the economic meaning of this elasticity and map it to the WTP that we would like to know. Third, even though the sample size of our data is large (over 1.4 million), the number of gas stations around the border is small (about 200). Therefore, our data is not fine for the inference in the bunching estimator.

D.3 Interpretation of the Parameter Estimates

Our final estimating equation 6 has a similar formula as the canonical difference-in-difference method, but the interpretation is fundamentally different. Here, the estimand is the structural parameter β that governs consumers' marginal utility from a relatively higher emission standard, instead of the reform's average treatment effect (ATE) under the potential outcome framework. Since the reform changes the relative utility between fuel L (the "treated unit") and H (the "control unit") and

[58]Note that we expect $\triangle e$ to be negative since the reform improved the emission standard for L and thus narrowed the relative gap between two types of fuel.

[59]A example of such distance measure could be the distance between the location where the driver realizes she needs to fill the tank and the location of the gas station that she finally chooses.

impact sales volume of both gasoline types, the assumption of stable unit treatment value assignment (SUTVA) is apparently violated. Within our framework, the ATE corresponds to the effect of reform on the potential market size (the "income effect"), which is determined by how much gasoline consumption in total. In comparison, β corresponds to the effect of reform on relative utility between two products (the"substitution effect"), which is determined where to purchase gasoline. The former corresponds to the consumption elasticity, while the latter corresponds to the purchasing elasticity. Theoretically, the overall effects are the combination of two.

To better understand the difference between the two estimands, think about an extreme case that prohibits consumers from crossing the boundary, which means the purchasing elasticity is zero. In this case, the relative utility between H and L no longer plays a role and the consumption elasticity dominates.[60] Since the SUTVA assumption is no longer violated, the estimand now is the ATE of the enforcement of high-quality gasoline standards in Hebei. On the contrary, this paper consider the other extreme case, namely, consumption elasticity is zero. In this case, only the relative utility between H and L matters and the purchasing elasticity dominates. Therefore, consumers' choices between two gasoline standards are directly informative to the preference parameters. Our assumption is supported by existing papers about the retail gasoline market which document that the purchasing elasticity is much larger than the consumption elasticity (Houde, 2012; Huntington et al., 2019) in the short run. In the empirical analysis, we also formally test this assumption by exploiting the unique feature of our context: for some adjacent regions, cross-boundary purchasing is not likely to happen since gasoline sold on two sides of the boundary can not be used interchangeably, which shuts down the relative utility channel.

D.4 Random Coefficient

We assume homogeneous α and β in our random utility model of gasoline demand. There are several reasons for this modelling choice. First, homogeneous logit model can be estimated by a linear least squares model and allows us to perform the two-step estimator. Besides, it provide a direct mapping from reduced-form coefficient to preference parameters. Second, the advantage of the well-known BLP estimator is to allow for a more flexible substitution pattern across products by relaxing the independence of irrelevant alternatives (IIA) assumptions. Since, in our setting, there are only two products to choose from, Beijing gasoline and Hebei gasoline. The heterogeneity in substitution pattern is not what we focus on in this paper. Third, compared to other studies using BLP estimator with Chinese data (Ito and Zhang, 2020; Li, 2018), our sample is more granular in the sense that we use gas station-by-day-level data for one market instead of products-by-city-by-year-level data. Therefore, we lack of enough variation in demographic distribution to identify the random coefficients associated with demographic.[61] Fourth, due to the regulated nature of retail gasoline price, it is difficult to find non-weak instrument variables for price within the event window to precisely estimate the random coefficients.

Although the random coefficient model is not implemented in this paper, we briefly discuss how to interpret our results when including random component in the utility function. First, we allow α and β to change with household income. For expositional purpose, consider the case that WTP on environmental value for household i is perfectly sorting on income y_i, such that $\beta_i = \beta_0 + \beta_1 y_i$, y_i is assumed to be drawn from log normal distribution $F_{\sigma_y}(i)$. Assuming log normal distribution is to make

[60] Actually, this is the exact setting in papers to study the electricity market at regional boundaries. See Ito (2014) and Deryugina et al. (2020) for examples.

[61] In paper with similar data structural, (Houde, 2012) use income and other demographic distribution calculated from household travel surveys at traffic area zone (TAZ) level and monthly labor force surveys. Unfortunately, we cannot access to data including demographic information at TAZ level or road segement level. Without demographic component in random coefficients, additional information that the random coefficient model can buy us does not worth its computational complexity.

sure positive βs.

$$s_{jkt} = \exp(\alpha\, p_{jkt} + \mathbf{\Lambda_{jt}} + G_k \cdot \gamma + \xi_{jkt}) \int_i \frac{\exp(\beta_1 y_i\, D_j\, \mathbb{1}(t \geq T^*))}{\sum_{j',k'} \exp(\alpha\, p_{j'k't} + \mathbf{\Lambda_{ij't}} + G_{k'} \cdot \gamma + \xi_{j'k't})}\, dF_{\sigma_y}(i)$$

Existing literature have documented a positive relationship between WTP and income (Ito and Zhang, 2020). Since Beijing implements higher emission standard during our sample period, consumer have higher probability to shop in Beijing stations if they are richer. By assuming extreme value distribution of individual error term, all consumers will increase their purchasing probability or frequency of purchasing in Hebei stations. And the richest consumers in Beijing would adjust their purchasing behavior most. Therefore, When considering heterogeneous income and sorting, our estimated WTP should be interpreted as a nonlinear weighted average of consumers across all income distribution.

Second, we allow for unobserved heterogeneity term in price coefficient such that $\alpha_i = \alpha_0 + e_i$, where e_i is drawn from distribution F_{σ_a}. Then the market share equation becomes,

$$\begin{aligned} s_{jkt} &= \int_i \frac{\exp(\alpha_i\, p_{jkt} + \mathbf{\Lambda_{jt}} + G_k \cdot \gamma + \xi_{jkt})}{\sum_{j',k'} \exp(\alpha_i\, p_{j'k't} + \mathbf{\Lambda_{j't}} + G_{k'} \cdot \gamma + \xi_{jk't})}\, dF_{\sigma_a}(i) \\ &= \exp(\mathbf{\Lambda_{jt}} + G_k \cdot \gamma + \xi_{jkt}) \int_i \frac{\exp(\alpha_i\, p_{jkt})}{\sum_{j',k'} \exp(\alpha_i\, p_{j'k't} + \mathbf{\Lambda_{j't}} + G_{k'} \cdot \gamma + \xi_{jk't})}\, dF_{\sigma_a}(i) \end{aligned}$$

and,

$$log(s_{jkt}) = f(p_{jkt}) + \mathbf{\Lambda_{jt}} + G_k \cdot \gamma + \xi_{jkt} \tag{C.4}$$

In equation C.4, the log of market share now depend on a nonlinear function of price $f(p_{jkt})$. Therefore, our two-step estimator is no longer valid. To alleviate this problem, we re-estimate main results using one-step estimation and include a polynomial function of price of order three. The results are shown in Table A.12. The sign and statistical significance of estimation do not change much, which implies the nonlinearity in price are not suppose to alter our main findings substantially.

Appendix E Additional Welfare Analysis

In this section, we use estimated WTP to discuss the potential distortion effects of regulated price and its implication on optimal gasoline price.

E.1 Distortion Effects

Recall that under our preferred specification, the WTP of the gasoline III to IV reform for the regular gasoline is 0.345 Yuan/liter, which account for about the 4.65% of its total price. The estimated relative price increasing associated with the reform as about 2.48%. A simple back-of-the-envelope calculation illustrates that the actual price change is about 1.42% lower than equivalent price change that does not alter consumers relative utility and conditional choice probability between H and L. This being said, the reform "distort" consumers' choice toward the Hebei side of the border, as it improves the "price-adjusted-quality" of the fuel L. This distortion on consumer's behavior might generate unexpected welfare effect since the reform essentially provides a de facto subsidy to the less-environmental friendly fuel type.

To study these effects formally, we construct the (per capita) social welfare function as,

$$W(\delta_j, p_j, \Lambda_j) = V(\delta_j) - E(p_j, \Lambda_j) - C(p_j, \Lambda_j), \quad j \in \{H, L\} \tag{D.1}$$

Note that for simplicity, we ignore the gas station characteristics in the welfare function and thus suppress the subscript k. Following the notation in section 3, δ_j is the mean utility of the fuel type j, which is identical for all consumers; p_j is the regulated price, Λ_j is the emission standard level; The social welfare W is composed of three parts: the expected value of private utility from the optimal choice $V(\delta_{jk})$, the social cost of the environmental externality from the emission $E(p_j, \Lambda_j)$, and other social costs for implementing the price and emission level regulation $C(p_j, \Lambda_j)$.[62] Exploiting the properties of the distribution assumption on ϵ_{ij}, $V(\delta_j)$ takes the form of,

$$V(\delta_j) = E[\max\{\delta_j + \epsilon_{ij}\}] = \hat{\gamma} + log\Big(\sum_j exp\{\delta_j\}\Big) \tag{D.2}$$

where $\hat{\gamma} \approx 0.52277$ is the Euler constant. The expected consumer surplus from the most preferred gasoline and gas station equals private utility $V(\delta_{jk})$ adjusted by marginal utility,

$$E[CS] = \frac{1}{MU} V(\delta_j) \tag{D.3}$$

where the marginal utility MU is assumed to be $-\alpha$.

$E(p_j, \Lambda_j)$ is defined as the expected pollution level for each consumer's choice, which is,

$$E(p_j, \Lambda_j) = \sum_j \pi_j(p_j, \Lambda_j)\, m_j \tag{D.4}$$

where $\pi_j = exp\{\delta_j\}/\Big(exp\{\delta_H\} + exp\{\delta_L\}\Big)$ is the probability of choosing j, and m_j is the per unit pollution level of j.

Combining equation D.1,D.2, and D.4, we could calculate the welfare effects of the III to IV reform by taking the total derivative,

$$\begin{aligned} \triangle W &= \Big[\partial V/\partial p_L \cdot \triangle p_L + \partial V/\partial \Lambda_L \cdot \triangle \Lambda_L\Big] - \Big[\partial E/\partial p_L \cdot \triangle p_L + \partial E/\partial \Lambda_L \cdot \triangle \Lambda_L\Big] - c \\ &= \pi_L \cdot \Bigg\{ \underbrace{\Big[\hat{\alpha}\,\triangle p_L + \hat{\beta}\Big]}_{\text{private utility changes}} - \underbrace{\Big[\pi_H \cdot (\hat{\alpha}\,\triangle p_L + \hat{\beta}) \cdot (m_L - m_H)\Big]}_{\text{distortion effects}} - \underbrace{\triangle m_L}_{\text{quality improvement}} \Bigg\} - c \end{aligned} \tag{D.5}$$

Formula D.5 decomposes the welfare effects of the reform into four parts. The first part is the private utility changes as a result of improving the "price-adjusted-quality". According to the estimates in section 5, $\hat{\alpha}\,\triangle p_L + \hat{\beta}$ approximately equals to 0.051, or equivalently 1.42 % of the gasoline price. The second part is the distortion effects on the environmental externality, which being said, the reform induces the (marginal) consumers who would have chosen H to purchase the more polluting fuel L. The magnitude depends on the proportion of original consumer of H, π_H; the private utility changes, or called, distortion scales; and the relative polluting gaps between L and$H$, $m_L - m_H > 0$. The third term $\triangle m_L < 0$ is the traditional pollution alleviation effects from improving emission standards that documented by the previous literature (Li et al., 2020). The last term c is the social cost of implementing the III to IV reform. The central parameter in this paper β, consumer's private value of higher emission standard enters the first two terms of the welfare effect, which together be named the "behavioral response". To evaluate the relative importance of behavioral response, we conduct a back-of-the envelope calculate. Take the value of $\pi_H \approx 0.5$, $\hat{\beta} = 0.14$. Borrowing from the estimation in Li et al. (2020), $\triangle m_L \approx -12.9\% \cdot m_L$. Based on International Council on Clean Transportation (2014), set $m_L - m_H \approx 30\% \cdot m_L$. Therefore, $\pi_H \cdot \hat{\beta} \cdot (m_L - m_H)$ approximately equals to 16.3% of the

[62] For example, the government have to subsidy to the state-owned retailers to maintain the regulated price. Readers can think of $C(p_j, \Lambda_j)$ as reduced-form measurement of the general equilibrium effects of the price and quality regulation.

polluting alleviation effects and the total distortion effects account for about 6% of the polluting alleviation effects.[63] These numbers imply a non-negligible effect from behavioral response when designing environmental regulation.

E.2 Deriving the Equation D.5

To calculate the total derivative of the welfare function $\triangle W$, we calculate the total derivative of each part of its composition. Note that the formula of V is derived based on the standard logit model results. See Small and Rosen (1981) among many others. Since $V = \hat{\gamma} + log\Big(\sum_j exp\{\delta_j\}\Big)$ and $\delta_j = \delta(p_L, \Lambda_L)$

$$\begin{aligned}\triangle V &= \partial V/\partial p_L \cdot \triangle p_L \,+\, \partial V/\partial \Lambda_L \cdot \triangle \Lambda_L \\ &= \partial V/\partial \delta_j \cdot \partial \delta_j/\partial p_L \cdot \triangle p_L \,+\, \partial V/\partial \delta_j \cdot \partial \delta_j/\partial \Lambda_L \cdot \triangle \Lambda_L \\ &= exp\{\delta_L\}/\Big(\sum_j exp\{\delta_j\}\Big) \cdot \Big[\hat{\alpha} \cdot \triangle p_L + \hat{\beta}\Big] \\ &= \pi_L \cdot \Big[\hat{\alpha} \cdot \triangle p_L + \hat{\beta}\Big]\end{aligned} \tag{D.6}$$

Note from equation D.3, the changes of expected consumer surplus,

$$\triangle E[CS] = \pi_L \cdot [WTP - \triangle p_L] \tag{D.7}$$

As for the environment externality part, recall that $E = \sum_j \pi_j(p_j, \Lambda_j) \cdot m_j$,

$$\begin{aligned}\triangle E &= \partial E/\partial p_L \cdot \triangle p_L \,+\, \partial E/\partial \Lambda_L \cdot \triangle \Lambda_L \\ &= \Big[\sum_j \frac{\partial \pi_j}{\partial p_L} \cdot \triangle p_L \cdot m_j\Big] \,+\, \Big[\sum_j \frac{\partial \pi_j}{\partial \Lambda_L} \cdot \triangle \Lambda_L \cdot m_j + \pi_L \cdot \frac{\partial m_L}{\partial \Lambda_L} \cdot \triangle \Lambda_L\Big]\end{aligned} \tag{D.8}$$

where $\pi_j = exp\{\delta_L\}/\Big(\sum_j exp\{\delta_j\}\Big)$, and,

$$\begin{aligned}\frac{\partial \pi_L}{\partial p_L} \cdot \triangle p_L &= \pi_L \frac{\partial log(\pi_L)}{\partial p_L} \cdot \triangle p_L \\ &= \pi_L \left(\hat{\alpha} - \hat{\alpha}\,\pi_L\right) \cdot \triangle p_L = \pi_L\,\pi_H\,\hat{\alpha} \cdot \triangle p_L\end{aligned} \tag{D.9}$$

and by construction, $\frac{\partial \pi_H}{\partial p_L}$ = - $\frac{\partial \pi_L}{\partial p_L}$. Similarly,

$$\begin{aligned}\frac{\partial \pi_L}{\partial \Lambda_L} \cdot \triangle \Lambda_L &= \pi_L \frac{\partial log(\pi_L)}{\partial \Lambda_L} \cdot \triangle \Lambda_L \\ &= \pi_L\,\pi_H\,\hat{\beta}\end{aligned} \tag{D.10}$$

and $\frac{\partial \pi_H}{\partial \Lambda_L}$ = - $\frac{\partial \pi_L}{\partial \Lambda_L}$. Denote $\frac{\partial m_L}{\partial \Lambda_L} \cdot \triangle \Lambda_L$ as $\triangle m_L$. Therefore, equation D.8 becomes,

$$\triangle E = \pi_L\,\pi_H \left(\hat{\alpha} \cdot \triangle p_L + \hat{\beta}\right) \left(m_L - m_H\right) + \pi_L \cdot \triangle m_L \tag{D.11}$$

Combine equation D.6, D.11, and D.1, we could derive the equation D.5,

$$\triangle W = \pi_L \cdot \Bigg\{\Big[\hat{\alpha}\,\triangle p_L + \hat{\beta}\Big] \,-\, \Big[\pi_H \cdot \left(\hat{\alpha}\,\triangle p_L + \hat{\beta}\right) \cdot (m_L - m_H)\Big] \,-\, \triangle m_L \,-\, c\Bigg\} \tag{D.12}$$

[63] In the calculation, 16.3% = (0.5 × 0.14 × 30%)/12.9% and 6.0% = (0.5 × 0.051 × 30%)/12.9%

■