# Anti-social Responses to the "Coal to Gas" Regulation:

# An Unintended Consequence of a Residential Energy Policy

Jing Cao[a,*,†] Tracy Xiao Liu[b,*,†] Rong Ma[c,*,†] Ang Sun[d,*,†]

**Abstract**

Policies geared toward environmental and economic improvement could unexpectedly lead to negative consequences in other dimensions. Such cases raise a red flag to economists and policymakers who aim to deliver comprehensive and sensible policy evaluations. This article investigates anti-social behaviors in response to the Clean Winter Heating Policy (CWHP), which seeks to improve outdoor air quality. Our results show that participating villagers are more likely to violate laws to burn agricultural waste and exhibit lower prosociality in incentivized dictator games and public goods games. We further explore treatment heterogeneities and find that two channels are likely to play a part. First, the CWHP was perceived as a negative income shock. Therefore, the villagers would want to reduce their expenditure on straw disposal and behave less generously in the incentivized games. Second, the CWHP could trigger discontent and directly affect social preference. Additional evidence suggests that the anti-social (less prosocial) responses could have been avoided by granting larger upfront subsidies.

Keywords: prosociality, environmental policy, quasi-natural experiment

JEL Classification: K32, L38, P18

[a] Department of Economics, School of Economics and Management, Tsinghua University, Beijing 100084, China.
[b] Department of Economics, School of Economics and Management, National Center for Economic Research at Tsinghua University, Tsinghua University, Beijing 100084, China.
[c] College of Economics and Management, China Agricultural University, Beijing 100083, China.
[d] School of Finance and China Financial Policy Research Center, Renmin University of China, Beijing 100872, China.
[*] J.C., T.X.L., R.M., and A. S. contributed equally to this work.
[†] To whom correspondence may be addressed. Emails: caojing@sem.tsinghua.edu.cn, liuxiao@sem.tsinghua.edu.cn, rongma@cau.edu.cn, ang.sun@ruc.edu.cn.

## 1. Introduction

In recent years, governments worldwide have increasingly turned to mandated regulations to address a wide range of issues and concerns across many different sectors, including public health, energy, finance, and technology. While many see these regulations as essential for ensuring that governments can effectively protect and serve the public interest, some have raised concerns about their potential costs and limitations. One of the major concerns is the potential for unintended consequences.[1]

Air pollution presents a global health hazard, and more than 90% of the world's population suffers drastic and devastating health consequences.[2] Policymakers worldwide have instituted various regulations to improve local air quality. We aim to provide insights on potential unintended consequences by examining a spectrum of behavioral responses to China's recently implemented Clean Winter Heating Policy (CWHP), the "Coal to Gas" regulation. From 2017 to 2019, the CWHP was implemented gradually in certain villages across 28 cities in four provinces. As one of the world's most extensive residential energy transition programs, the CWHP stipulated that more than ten million rural households in northern China replace coal stoves with electricity or natural gas furnaces.

While we demonstrate that the CWHP directly improves indoor and outdoor air quality by transitioning to clean heating fuels, it also induces more anti-social behaviors among participating villagers compared to non-participating villagers. We primarily focus on law-violating straw-burning behaviors, which impose significant negative externalities on air quality, leading to health damages and increased risks of uncontrollable fires. Straw burning is banned by

---

[1] For example, online privacy regulation strengthens data protection but also reduces competition by increasing market concentration (Peukert et al., 2022; Johnson et al., 2023). Director independence mandates on large companies could improve the reliability of financial disclosures but may reduce earnings (Alam et al., 2018).

[2] For example, Chen et al. (2013), Cohen et al. (2017), Ebenstein et al. (2017), and Burnett et al. (2018).

six different laws in China. Using fire spot data collected by the Landsat-8 satellite, an Earth observation satellite designed by the National Aeronautics and Space Administration (NASA), we constructed panel data extending from 2014 to 2019. Based on their latitude and longitude, these fire spots are linked to our sample of 110 villages in nine cities covered by our door-to-door household survey. The temporal and regional variations in the rollout of the CWHP allow for a difference-in-differences (DID) analysis. We find significantly more straw burning in CWHP participant villages following the implementation of CWHP compared to non-participant villages. In addition, we investigate the outcomes of two incentivized games we implemented in the field in 2018. Our causal forests (CF) estimations show that CWHP participant households behaved in a less prosocial manner in both games.

Two forces are likely to play a role in driving anti-social (less prosocial) behavior. First, consistent with the anecdotes in mass media that most subsidies from the central government are spent on infrastructure and installation of new furnaces, our door-to-door survey shows that the subsidies on electricity and gas for heating are minimal and that the CWHP is perceived as a substantial disadvantageous income shock by villagers. Therefore, the incentive to reduce other expenditures drove them to dispose of stalks in the fields and, in the two games, to distribute less money to their anonymous neighbors or public accounts. This mechanism is well supported by two heterogeneity analyses. We find that the pattern of more anti-social and less prosocial behavior is stronger in villages with lower disposable income per capita. The pattern is also more pronounced in years with colder winters or more snowfall when low temperatures and high humidity increase heating costs.

Second, due to the negative income shock and the inconvenience associated with the construction of extended gas pipelines and power grids, the CWHP may trigger stress, dissatisfaction, and resistance among villagers. Such discontent could provoke farmers to violate laws and exhibit less prosocial behavior when distributing money in the two games. We perform another heterogeneity analysis to test this conjecture. In many localities, the construction of natural gas pipelines or power grids failed to keep pace with the fast rollout of the CWHP, or the sudden surge in demand for electric heating has led to an overload and tripping of the power grid, leaving farmers without heating on cold winter nights. We find that heating shortages aggravate farmers' emotional discontent with the CWHP. The DID estimation also indicates a more pronounced increase in straw burning in villages experiencing greater heating shortages compared to other treated villages following the CWHP rollout.

To illustrate the tradeoff between the intended and unintended effects of the CWHP, we show that while the program improves outdoor air quality, it is compromised by increased pollution from straw burning during plowing and harvest seasons, resulting in a benefit-cost ratio slightly less than one. These findings raise concerns for policymakers and contribute to the literature on the unintended consequences of government regulations. For example, energy-saving technologies enhance energy efficiency but may also inadvertently encourage energy consumption (Sorrell et al., 2020), and efforts to reduce energy consumption at home can unintentionally increase energy use elsewhere (Dolan and Galizzi, 2015). Similarly, while smart metering provides better information on energy use, it can unintentionally discourage energy-saving technology investment (McCoy and Lyons, 2017).[3]

Additionally, the negative impacts of the CWHP on prosocial behavior can potentially hinder long-term development and growth. Changes in individual behaviors and social

[3] More broadly, regulations designed to enhance welfare in areas beyond environmental policy have also been found to foster collusion and reduce market competition (Anand and Giraud-Carrier, 2020; Peukert et al., 2022; Johnson et al., 2023) and can even lead to unintended adverse effects (Guan et al., 2019; Li and Wu, 2022; Adepoju et al., 2023).

preferences could result in a series of consequences in socio-economic development in the short and long run (Carpenter and Seki, 2011). These concerns are particularly pertinent for low-income individuals, for whom anti-social and unethical behaviors are more closely correlated with their social identity (Gangadharan et al., 2019; Gsottbauer et al., 2022). Our findings prompt regulators to carefully consider the complex and unintended effects of regulations on individual behavior and social dynamics (Dang et al., 2023).

Our welfare analyses indicate that some unexpected and undesirable policy effects are attributed to inadequate subsidies. This notion is quite sensible, considering it is often costly for individuals to comply with environmental policies that have high externalities (Mobarak et al., 2012; Allcott and Wozny, 2014; Newell and Siikamäki, 2014; Fowlie et al., 2015, 2018; Hanna et al., 2016). De Groote and Verboven (2019) argue that adequate upfront subsidies could be more efficient than production subsidies in promoting solar photovoltaic systems as a cost-effective technology.

The rest of the paper is organized as follows. Section 2 describes the implementation and controversies of the CWHP and briefly discusses straw-burning activities in rural China. Section 3 explains the DID estimation on straw burning and the causal forests approach for analyzing individual-level data. Section 4 describes the satellite data on straw burning, our surveys on winter heating in rural households, and the two incentivized games used to measure prosocial behavior. Section 5 presents the main results. Section 6 discusses potential mechanisms, and Section 7 analyzes welfare changes due to the CWHP. Section 8 concludes the paper and discusses policy recommendations.

## 2. Clean Winter Heating Project

## 2.1 Launch and Implementation

The CWHP was launched as a major measure of the "Blue Sky Protection Campaign" proposed by the central government. Beijing's air condition had attracted significant attention and was intensively criticized by international mass media. In 2015, the CWHP began with a few villages around central Beijing. In 2017, China formally launched the campaign in Beijing, Tianjin, and 26 other cities in the northern provinces. Most cities completed the transformation before 2019. Between 2015 and 2018, the number of households transformed under the CWHP increased from a few thousand to several million. Meng et al. (2019) predict that the contribution of rural coal heating to ambient PM2.5 concentration will decrease by over 50%,[4] and indoor PM2.5 concentration is projected to decline by 40%,[5] assuming that the goal of replacing coal stoves for 60% of rural households within the 28 cities is achieved.

The implementation locality was mainly determined by meteorological and geographical conditions, especially the location relative to Beijing. In addition, feasibility factors were considered, such as the availability of a natural gas infrastructure, the distance to the closest natural gas pipeline, and the gas pipeline's capacity to dispatch.[6]

The actual implementation was conducted by village governments. They were responsible for rigorously enforcing three measures: (1) providing infrastructure upgrades, such

[4] Zhao et al. (2021) find that the transition to clean winter heating can reduce ambient PM2.5 concentration by an average of 2.4 $\mu g/m^3$ in a major Chinese city.

[5] Barrington-Leigh et al. (2019) surveyed 302 households from six villages in Beijing and measured indoor air quality for 55 surveyed households. Their findings indicate that indoor PM2.5 concentrations in CWHP-treated households are about 50% lower than in untreated households in high- and middle-income districts, but there is no significant difference in low-income districts.

[6] According to an interview with an anonymous official from the Department of Atmospheric Environment of Beijing Municipal Ecology and Environmental Bureau, villages that were required to convert from coal to gas were mainly chosen based on the distance to the nearest gas trunk pipelines and the dispatch capacities of these pipelines. In addition, the trunk pipelines in China are sparsely distributed and only connect some major cities to supply gas to their urban residents. Therefore, in rural areas and other cities, the proximity to a trunk pipeline is determined by the major cities where the pipeline begins and ends.

as local gas pipelines, electricity grids, and related constructions; (2) imposing a ban on local coal transactions; and (3) requiring households to install furnaces powered by electricity or gas, along with providing the corresponding subsidies.

The installation cost of a furnace varies by type, averaging around 3,000 RMB. By 2018, the central government had provided 15.8 billion RMB in subsidies to participant villages, while local governments contributed over 75 billion RMB. These subsidies primarily covered furnace installations, with many local governments required to offer up to 5,000 RMB for purchases and installations. In contrast, while gas and electricity are more expensive than coal, villagers receive less subsidy support. Media reports suggest that heating costs with new furnaces were twice as high as those for coal,[7] and some low-income households spent over 50% of their average monthly income on heating following the CWHP implementation.[8]

In addition to insufficient subsidies, the ambitious implementation agenda was also controversial. The program's goal to replace the coal stoves with new furnaces within two years was monitored and evaluated by central government environmental inspections each October, shortly before the heating season. To meet the requirement, some villages had forcefully dismantled coal stoves even before the infrastructure was in place, resulting in villagers being unable to heat their homes. In some regions, this led to schools being unable to operate normally and significant losses in vegetable greenhouse facilities. As a result, the project has slowed down since 2019.[9]

## 2.2 Straw Burning in Rural China

[7]Source: https://www.theguardian.com/world/2017/dec/04/poor-bear-brunt-beijing-coal-cleanup- with-no-heating-at--6c.
[8] Source: https://www.nytimes.com/2018/02/10/world/asia/china-coal-smog-pollution.html.
[9] Source: http://www.nea.gov.cn/2019-07/03/c_138195454.htm.

In rural China, farmers traditionally burn agricultural waste after harvesting to prepare their land for planting the following year. However, this practice generates significant air pollutants (Wang et al., 2009) and subsequent health damage (He et al., 2020; Graff Zivin et al., 2020). Since 1999, straw burning has been prohibited by several laws in China, including the Atmospheric Pollution Prevention and Control Law, the Fire Control Law, and the Public Security Administration Punishment Law of China.

More environmentally friendly methods of handling crop residues include burying them in the fields or recycling them. However, farmers are reluctant to adopt such measures for several reasons. First, extra manpower and resources are required to clear the stalks. For example, an investigation by Henan media Yuji found that collecting corn stalks from 10 mu of land, hauling them home, and stacking them took at least four laborers working hard for three days. China's rural labor shortage is also an indisputable fact. Migrant laborers are obliged to rush back to their villages for the autumn harvest, and it would be unrealistic to expect them to spend more days cleaning up agricultural waste after the harvest.

Second, "returning straw to the field," although considered by the government as the greenest measure, is economically costly and could potentially cause difficulties in future planting. This process involves crushing the stalks, plowing the land deeply, and burying the stalks underground. According to our household survey, the cost is generally around 780 RMB per household each year. In addition, returning crop residues to the field may result in decomposition and pest issues in the subsequent year due to insufficient shredding or burial.

Third, transportation costs restrict farmers from recycling straws. Although crop residues are not heavy, their large volume makes transportation expensive. According to a straw recycler interviewed in Hubei,[10] he can purchase 30 tons of straw per day and sell it to biomass power plants for 7,200 RMB. However, it costs 8,000 RMB to transport the straw to the power plant.

## 3. Empirical Strategies

### 3.1 The Difference-In-Differences Estimation of the Incidence of Straw Burning

In the spring of 2018, we conducted a door-to-door survey of 1,330 households across 110 villages in Beijing, Tianjin, and Hebei to collect information on the CWHP implementation details. Within this sampling frame, we gathered information at the village and household levels. At the village level, we collected data on village coordinates and the timing when each village launched the CWHP. We matched NASA's monthly 30-meter pixel fire point data to villages based on their longitudes and latitudes, constructing a village-month-level fire spot panel dataset from January 2014 to December 2019. We further matched this panel data with the CWHP implementation for each village. The panel construction will be discussed in detail in Section 4.

The nature of the panel enables us to exploit the regional and temporal variations of the CWHP and estimate its effect by applying a DID design. The specification is as follows:

$$Straw_{it} = \beta CWHP_{it} + \alpha Z_{it} + \eta_i + \delta_{pt} + \epsilon_{it} \tag{1}$$

where $Straw_{it}$ represents the measure of fire spots, serving as proxies for straw burning for village *i* at time *t* (precise to the month in the calendar). We investigate the incidence, frequency, and area of the fire spots. $CWHP_{it}$ is a dummy indicator of whether the CWHP has been implemented in village *i* at time *t*. If the CWHP was initiated in a given year, we regard October of that year, i.e., right before the start of the heating season, as the first month when the effects of

[10] Source: https://www.yicai.com/news/4571949.html.

the CWHP begin to manifest.[11] $Z_{it}$ is a set of control variables sourced from the monthly meteorological data, including the average monthly temperature, air pressure, sunlight duration, wind speed, and total precipitation. These weather conditions are likely to affect the feasibility and risks of burning straw and the accuracy of the satellite's remote sensors. $\eta_i$ represents the village-fixed effects, $\delta_{pt}$ denotes the province-specific time-fixed effects, and $\epsilon_{it}$ is the error term clustered at the village level.

Before presenting the data and results, we will briefly discuss a few specifics about the identification strategy. First, we will carefully test the prevailing assumption of parallel trends by incorporating lags and leads in staggered DID regression analysis. Second, we include the province-specific time-fixed effect to prevent potential simultaneous policies from contaminating the effect of the CWHP. In addition, we perform a permutation test to further rule out random concomitant changes as potential major driving forces of the estimated CWHP effect. Finally, the remote sensor identifies hotspots but cannot specify the exact combustible materials. We exploit the seasonality of straw burning as a sensitivity check to address the concern about measurement error. We find that the impact only appears during the few months before the planting season and after the harvest season.

## 3.2 Causal Forests Estimation

To analyze the compliance of CWHP and its impacts on indoor temperature, air pollution, and individual prosociality, we exploit our 2018 cross-sectional survey data and apply the CF method. CF was recently developed by Athey and Imbens (2016), Wager and Athey (2018), and

[11] This is a sensible approach because local government policies explicitly require the completion of the coal-to-gas or electricity transition by the end of October. For example, this mandate is outlined in Beijing's regulations (https://www.beijing.gov.cn/zhengce/zhengcefagui/201905/t20190522_61224.html), Hebei's guidelines (http://www.xiongan.gov.cn/2018-11/02/c_129984560.htm), and Jinnan district of Tianjin's directives (https://www.tjjn.gov.cn/zwgk/zcwj/qjwj/qzfbgs100/202012/t20201208_4678267.html). In addition, the effects of the CWHP can only be observed when households begin using the new equipment during the heating seasons.

Athey et al. (2019). CF offers advantages in dealing with high-dimensional covariates and non-linear relationships between the outcome variable and the treatment assignment. Conventional approaches, such as PSM, often exhibit inferior statistical performance when encountering irrelevant or numerous covariates (Wager and Athey, 2018). In contrast, CFs can adaptively determine weights for each nearby observation, which are then used to generate a counterfactual for the outcome variable. As demonstrated by Wager and Athey (2018), compared with PSM, CFs lead to an increase in statistical power and a reduction in estimation bias.[12]

Furthermore, CFs are specifically designed to capture heterogeneous treatment effects. We employ CFs to estimate heterogeneous treatment effects by income levels and past heating experiences. More importantly, CFs allow us to estimate individualized treatment effects on prosociality for each participant, enabling us to explore how CWHP-induced idiosyncratic changes in prosocial behavior correlate with the incidents of straw burning.

Given these well-designed features, we predominantly focus on CFs for individual-level analyses but also perform sensitivity checks by supplementing the main analyses with OLS and PSM estimations to examine the effects of the CWHP, thereby ensuring the robustness of our findings.

# 4. Data Description and Graphical Evidence

## 4.1 Survey on Winter Heating and Social Preferences of Rural Residents

***Sampling frame.*** We collected the rolling-out information on CWHP in the targeted provinces (Beijing, Tianjin, and Hebei) from local governments and media reports. Villages were randomly selected within each province, resulting in a final sample comprising 25 villages from

[12] Due to these considerations, CFs have been applied in a number of empirical studies, e.g., Davis and Heller (2017, 2020); Carter et al. (2019); Alyakoob et al. (2021); De Neve et al. (2021); Guo et al. (2021); Knaus et al. (2022).

Beijing and 85 villages from the adjacent areas of Tianjin and Hebei. We intentionally oversampled villages in Tianjin to include as many untreated villages as possible. In Beijing, where the CWHP originated, nearly all suburban villages were already part of the program before our survey. Tianjin, a centrally administered municipality with social and economic characteristics similar to Beijing, implemented the CWHP later. We acknowledge that given the higher weighting of Tianjin in our sample, our estimations may not fully represent the entire population of the three provinces. However, our study aims to highlight the unintended consequences of the CWHP rather than provide an estimate of the average treatment effect across the entire targeted area. Appendix Figure A1 plots the research timeline, including the initiation of satellite data collection, key implementation phases of the CWHP, and the schedule for our survey.

Figure 1 maps the scatter points of the villages in our door-to-door survey. Out of the 110 villages surveyed, 11 participated by the end of 2016, 64 participated in 2017, 15 villages participated in 2018, and 9 in 2019, leaving 11 villages untreated when the program concluded in 2019.

[Figure 1 is here]

***Survey variables.*** Based on the village sampling frame, we randomly selected households in each village to participate in our survey. Enumerators interviewed the household member who was at home and best able to provide responses on behalf of the household. The survey included the following five modules:

**A. Basic demographics and socioeconomic statuses of the household**: We collected demographic information such as the gender, age, occupation, and education level of the household head, family size, house area, heating area, and household annual income and consumption.

**B. Costs of winter heating:** The survey gathered detailed information on winter heating, including heating equipment, costs of purchasing and installation, expenditure on fuel during the previous heating season, and government subsidies for heating fuel and equipment.

**C. Willingness to pay for the clean winter heating and actual benefit for CWHP participants.** To assess willingness to pay (WTP) for clean winter heating, all respondents were asked how much they would pay annually for heating with electricity or natural gas. For CWHP participants, enumerators collected data on the timing of CWHP implementation and their attitudes toward the transition to clean heating. Additionally, participants were asked if they experienced any power or natural gas shortages during the heating season. Half of the participants were randomly selected for home visits, during which enumerators measured indoor temperature and indoor PM2.5 concentration.[13]

**D. Straw burning:** Information was collected on cropland, straw disposal methods, and enforcement of bans on straw burning in recent years. For those needing to dispose of straw, enumerators provided information on greener methods if not already applied and inquired about reasons for not using such methods.

**E. Two incentivized games**: Our door-to-door survey included a module measuring the extent of prosociality using two incentivized games. In a dictator game, participants were informed that they had been randomly paired with another survey participant from their village. They were tasked with allocating a sum of 20 RMB between themselves and the other participant,

[13] Due to resource constraints, we equipped one-fourth of enumerators with air quality monitors to measure indoor PM2.5 concentration for the households selected for home visits.

whose identity remained anonymous to them and vice versa. The public goods game introduced subjects to a scenario where they were part of a four-person group from their village. Each was given an endowment of 20 RMB to allocate between private and public accounts. The returns from the public account were contingent on the combined contributions of all four members to this account. With a marginal return rate of 0.4, every RMB contributed to the public account, which resulted in a 0.4 RMB gain for each group member. Thus, while contributions to the public account were costly on an individual level, they were beneficial for the group collectively. Additional information and detailed descriptions of these games can be found in Online Appendix A1.

Finally, we conducted a reamended telephone interview in 2023 with 109 randomly selected households from the initial survey. These interviews investigated the specific costs of straw disposal and probed their specific WTP for air pollution reduction and heating comfort improvement, respectively. In Appendix Table A1, we provide a description of the data structure and coverage for these key outcome variables, as well as the empirical analysis methods used to examine these variables. The coverage of households across these modules varies by subgroup, which, in some cases, results in differing analytical approaches for each module.

***Summary Statistics.*** Table 1 demonstrates household characteristics. The average house and heating areas are 125 and 93 square meters, respectively. These households are slightly below the median of China's overall income distribution, and the annual income per capita is 12,395 RMB (around 1,900 USD). Around 70% of the annual income is spent on consumption, amounting to 8,282 RMB per capita.

[Table 1 is here]

Panel A of Table 2 reports heating expenses for fuel and equipment, along with the respective subsidies, for CWHP participants and non-participants. On average, the annual fuel cost and equipment cost for CWHP participants amount to 2,789 RMB and 2,822 RMB (about 398 and 403 USD), while for non-participants, the expenses are 1,534 RMB (about 219 USD) on fuel and 1,191 RMB (about 170 USD) on equipment. The average fuel and equipment subsidies received by participants are 609 RMB and 1,495 RMB (about 87 and 214 USD), while the average fuel subsidy for non-participants is only 79 RMB (about 11 USD), and their equipment subsidy is 359 RMB (about 51 USD). The average indoor temperature for CWHP participant households is 19.2°C, 0.8°C higher than that of non-participant households. The average indoor PM2.5 concentration for CWHP participant households is 57.4 $\mu g/m^3$, about 4.7 $\mu g/m^3$ lower than that of non-participant households.

[Table 2 is here]

## 4.2 Data on Straw Burning and Weather

***Data Descriptions.*** Straw burning is extracted from Landsat-8 satellite imagery, a NASA Earth observation satellite launched in 2013. It passes over China daily from 9 am to 2 pm, equipped with the Operational Land Imager (OLI), a remote sensing instrument. OLI's advantage over other satellite data, such as the Moderate Resolution Imaging Spectroradiometer (MODIS),[14] lies in its ability to detect thermal anomalies using short-wave infrared radiation (Schroeder et al., 2016). This short-wave technique offers a significantly higher spatial resolution compared to mid-infrared radiation, which is crucial for detecting small-scale events. In North China, where farmers burn straw, the fire sites typically span around 500 square meters. This size

[14] Using the MODIS fire spots, He et al. (2020) and Graff Zivin et al. (2020) consider straw burning at the county level with a radius of 50 km.

is detectable by the 30-meter spatial resolution of the OLI sensor but exceeds the detection capability of MODIS,[15] which usually requires fire areas larger than 1 km under general weather conditions.

The data from the OLI is reanalyzed by the Institute of Remote Sensing and Digital Earth of the Chinese Academy of Sciences (CAS) using a short-wave infrared Normalized Burning Ratio (NBR).[16] This analysis produces a high-resolution fire product for China, including the geographic coordinates, sizes, temperatures, and raw satellite images of detected fire pixels.

In our study, we match pixels with village areas to construct village-level fire spot data. Due to the lack of precise coordinates for village boundaries, we encompass all pixels within a 5-kilometer radius circle centered at each village's midpoint. Given that the average village size in China is around 15 $km^2$, this radius may extend beyond the actual boundaries of some villages. We intentionally chose a larger radius for two reasons. First, central village areas typically host commercial and administrative activities, while agricultural lands are predominantly located on the outskirts. Second, the irregular topography of village landscapes necessitates a broader radius to ensure comprehensive coverage of agricultural areas at the peripheries.[17] Panel B of Table 2 presents the summary statistics on the average probability, frequency, and area of straw burning

[15] There could be instances where simultaneous multiple fires occur within a 30-meter grid. While the probability of such occurrences is low due to the relatively small 30-meter cutoff compared to the land diameter and fire locus of straw burning in North China, we acknowledge that this possibility cannot be entirely ruled out.

[16] The NBR in the burning area is significantly larger than that in the non-burning area. The CAS chooses the threshold for fire pixels by maximizing the increase in the gradient of the NBR. To reduce errors due to the influence of buildings and clouds, the CAS uses the negative correlation between the short-wave spectrum and temperature and compares two short-wave infrared bands (bands 6 and 7 of Landsat-8) to identify misclassified pixels.

[17] In our analysis, we do not employ land cover type data to differentiate agricultural fires from other vegetation fires, such as forest or grassland fires. This decision was informed by the potential for nonclassical measurement errors, as discussed by Alix-Garcia and Millimet (2023), and the specific environmental context of the Beijing-Tianjin-Hebei region, which is characterized by scant forest and grassland coverage. To illustrate the minimal impact of forest and grassland fires on our findings, we employed the MODIS land cover type product (MCD12Q1) to exclude vegetation fires in non-farmland areas. This approach led to a negligible reduction—about 2%—in the average frequency of agricultural fires in our sample, highlighting the limited relevance of non-agricultural vegetation fires to our analysis.

using the 5 km radius for the treated and untreated villages in the pre-treatment period. In Appendix Table A2, we define village fires as those detected within the circles with radii of 4 km and 6 km around the village center as a robustness check for our main findings.

In regression analyses, we use the information on monthly local meteorological conditions that likely affect the detection of straw burning from satellites. The data is sourced from the China Meteorological Administration and constructed by matching villages in our survey to the nearest weather stations based on their geographical locations.

***Graphical Evidence.*** We focus on the time window extending from 2014, the year after Landsat-8 was launched,[18] to 2019, the last year before the outbreak of COVID-19. Figure 2 plots the incidence and intensity of straw burning by month. The blue solid lines stand for the 75 villages enrolled in the CWHP by the end of 2017, shortly before our survey, and the orange dashed lines stand for the 35 non-participant villages by 2017. The left panel shows the seasonality of straw burning before 2017, and the right panel shows the seasonality after 2017. Figure 2 demonstrates two salient features. First, straw burning exhibits strong seasonality, peaking shortly before the spring plowing (March) and after the autumn harvest (November and December), consistent with the agricultural production cycle. Second, before 2017, the blue solid and orange dashed lines closely track each other. However, after 2017, the solid blue line (participants) consistently surpassed the orange dashed line (non-participants) during early spring and late autumn, suggesting more straw burning by participants after the CWHP program.

[Figure 2 is here]

## 5. Empirical Results

[18] Landsat-8 was launched in February 2013, and the straw burning data is available for four months of 2013. We discard the 2013 data to ensure that observations extend over the whole year.

## 5.1 Intended Effects of the CWHP

***Indoor temperature and pollution.*** The data on indoor pollution and temperature are extracted from our cross-section survey in 2018. Building upon the discussion in Section 3.2, we utilize the CF method to optimize the weighting of observations, aiming for a maximum balance between CWHP-participating households and non-participants. We also employ unweighted (OLS) and propensity score matching (PSM) weighted approaches for reference. To assess the performance of CF relative to OLS and PSM, Figure 3 displays the differences in a comprehensive set of covariates between CWHP participants and non-participants, including house area, per-person heating area, household annual income (dummy indicators), the household head's education level (dummy indicators), job type (dummy indicators), gender, age, and family size. The green circles represent the raw difference between participants and non-participants without weighting. Following Stuart (2010), we apply an absolute standardized mean difference cutoff of 0.1 to assess balance. Figure 3 shows that the two groups are, to a large extent, balanced regarding most covariates, except for house size, heating size per person, and education level. Most circles fall to the left of the dashed line of 0.1. This balance alleviates our concern about selection bias and enhances confidence in the principle of assignment, indicating that the CWHP village locality is mainly determined by exogenous factors such as meteorological and geographical conditions. The red diamonds and blue triangles in the figure represent the differences under the weighting strategy of PSM and CF, respectively. Figure 3 shows no noticeable improvement in balancing the sample for PSM. In contrast, CF outperforms both OLS and PSM. Applying CF, all differences decrease to below or around 0.1.[19]

[19] Wager and Athey (2018) show that conventional matching methods, such as nearest neighbor matching, could perform well in the presence of a small set of covariates, but may break down as the number of covariates increases. Hence CFs usually dominate them in terms of bias and variance.

[Figure 3 is here]

Table 3 shows a high compliance rate with the CWHP, along with notable improvements in indoor amenities resulting from the program. Specifically, the results indicate that the CWHP increases indoor temperatures by approximately 1.2 to 1.4 °C and reduces indoor PM2.5 concentrations by 3.6 to 3.8 μg/m³. These findings suggest that the CWHP has been somewhat successful in achieving its objectives of reducing indoor air pollution and enhancing the indoor comfort of rural households.

[Table 3 is here]

***Outdoor air pollution.*** We investigate the effect of CWHP on outdoor air pollution using the high-resolution (0.01° × 0.01°) monthly ground-level PM2.5 concentration data from 2014 to 2019, obtained from van Donkelaar et al. (2021). This dataset is derived from satellite data for aerosol optical depth and a chemical transport model, providing high-resolution data on global PM2.5 concentration across time.

We calculate the village-level PM2.5 concentration by averaging the concentration within the circle of a 5 km radius centered at the latitude and longitude of the central location of each village. Using our DID specification Eq. (1), we estimate the effect of the CWHP on PM2.5 concentration.[20] Table 3 shows that the coefficient on the CWHP is negative and significant at a 10 percent level, indicating that the CWHP leads to an overall reduction in outdoor PM2.5 concentration by 1.06 μg/m³ in the treated villages. It is important to note that this DID estimate likely represents a conservative estimate of the reduction in air pollution. Given the dispersion of

[20] Specifically, we estimate the following equation: $PM2.5_{it} = \beta CWHP_{it} + \alpha Z_{it} + \eta_i + \delta_{pt} + \epsilon_{it}$, where $PM2.5_{it}$ represents the measure of pollution for village $i$ at calendar-specific month $t$. The other variables follow the definitions provided in Eq. (1).

PM2.5 in the atmosphere, the effectiveness of local governance measures at the village level is highly likely underestimated.

## 5.2 Unintended Effect on Straw Burning

***DID Estimators on Straw Burning.*** Table 4 presents the DID estimators of the effect of the CWHP on straw-burning activities. Column (1) reports the impact on the extensive margin, showing that compared to the pre-CWHP period, the likelihood of straw burning in CWHP participant villages is 4.1 percentage points higher than in non-participant villages. Columns (2) and (3) provide estimates for the intensive margin, focusing on the frequency and area of straw burning within a month. Compared to the prior level, the frequency of straw burning is higher by 6.3 percent per month in participant villages than in non-participant villages, and the area of fire scenes is 5.1 percent larger. All the estimands are consistently significant at the 5 percent level.

We split the sample into two groups according to whether a month is in the straw-burning season, i.e., following the autumn harvest (from October to December) or before the spring plowing (from February to April). Fire spots detected during the straw-burning periods are supposed to be more relevant to the practice of straw burning. Columns (4)–(6) present the DID estimators using the subsample composed of months within straw-fire seasons. For all three measures of straw burning, the coefficients on the CWHP are positive and statistically significant at the 1% level. Their magnitudes are also more than twice as much as those yielded from regressions using the full sample. In contrast, columns (7)–(9), which focus on months outside the straw-burning seasons, indicate smaller and less statistically significant changes in straw burning.

[Table 4 is here]

***Tests on the Parallel-Trend Assumption.*** To test the maintaining assumption of parallel trend, we replace the dummy indicator in Eq (1) with a set of dummies of lags and leads. To further address the concern of the internal validity of staggered DID considering potential heterogeneous treatment effects, we employ the method proposed by Borusyak et al. (2024). Compared with other approaches designed for staggered DID, Borusyak et al.'s approach imputes counterfactuals for the treatment group, which is particularly advantageous in our study where few villages were never treated. Moreover, this method allows us to incorporate high-dimensional fixed effects, such as province-calendar-specific-month fixed effects, into our estimation framework.

Figure 4 illustrates the dynamic DID estimates for straw-burning activities in participant villages each year.[21] The connected red solid and blue dashed lines represent the estimated coefficients for the full sample and high straw-burning seasons, respectively, with the vertical bars indicating their 95% confidence intervals. Before the implementation of the CWHP, the DID coefficients for all three outcomes related to both the extensive and intensive margins of straw-burning activities are close to zero, supporting the validity of the parallel trend assumption. Figure 4 also shows substantial and consistent differences in straw-burning activities following the introduction of the CWHP, with most estimates statistically significant at least at the 5% level.

[Figure 4 is here]

***Ruling Out Potential Concomitant Changes.*** Since 2013, China's central government has implemented a series of anti-pollution policies to improve air quality, with some focusing on regulating the emissions of manufacturing industries. We argue that the benchmark analysis

[21] The point estimators depicted in Figure 4 are reported in Appendix Table A3.

results are not likely driven by concomitant industrial regulations, as agriculture is the primary industry in most surveyed villages.

In addition, we perform a permutation test by randomly assigning (falsifying) months of exposure to the CWHP. We construct the corresponding full sample and straw-burning-season sample with falsified exposure, respectively. We re-estimate our benchmark model (Eq. (1)) using these artificial samples. The coefficients of interest should not be significant if the CWHP is indeed the primary driving force. In contrast, if the sample generated from the random-data-generation process yields a similar pattern to that in the benchmark DID analysis, we cannot rule out the effect of omitted concomitant changes. Figure 5 plots the coefficients and p-values obtained from the abovementioned regressions. In all figures examining different outcomes, the coefficient estimated using the observational data (marked by the vertical dashed line) is far outside the distribution of the DID coefficient (represented by the solid blue line).

[Figure 5 is here]

***Addressing the Concern of Measurement Error in Fire Spots.*** The potential measurement error of fire spots as a proxy for straw burning should not be ignored.[22] It is possible that the remote sensor of the satellite captured static fire spots rather than burning stalks, such as chimneys in large factories. To probe the extent to which such potential measurement error could drive our estimation, we exploit the seasonality of straw burning and plot the effects

[22] Heavy smoke can obscure fire detection, and super-heated smoke plumes may lead to false fire detections. For such interference to occur, fires would need to be extremely large, exhibiting explosive growth and rapid plume development, which typically happens in forest fires. Fires in cropland are generally less intense compared to forest fires (Val Martin et al., 2010). Fires in the Beijing-Tianjin-Hebei region are unlikely to reach that level of intensity, given the region's relatively small amount of forest land and the modest average size of each household's cropland parcel. Furthermore, false detections due to super-heated smoke plumes usually happen at night when the satellite fire product is particularly sensitive to heat sources. However, the satellite data used in our study are derived from Landsat-8, which passes over China only during local daytime hours.

of the CWHP on straw burning using the subsample of each of the 12 months.[23] In Figure 6, the red dots are DID coefficients, and the gray area represents the 95% confidence interval. We only observe a statistically significant effect of the CWHP in October and November (after autumn harvests) and in March and April (shortly before spring plowing). The seasonality shown in Figure 6 confirms that the detected effects of the CWHP are mainly associated with burning straws. The point estimates displayed in Figure 6 are reported in Appendix Table A4.[24]

[Figure 6 is here]

## 5.3 Unintended Effect on Prosociality

We employ the responses in the two incentivized experiments from the 2018 household survey to identify the unintended effects on prosociality. Table 5 presents the treatment effects of the CWHP on prosocial behaviors, employing OLS, PSM, and CF. Columns (1)-(3) detail the outcomes from the dictator game, where the dependent variable is the money allocated to an anonymous neighbor from a total of 20 RMB provided for the game. The CF estimation indicates that participation in the CWHP results in a reduction of 0.858 RMB in the amount allocated to an anonymous neighbor. This decrease amounts to over 10% of the sample average (8.361 RMB). Columns (4)-(6) display that, with the CF estimation, CWHP participation leads to a reduction of 0.877 RMB in contributions to the group account. This reduction corresponds to 12% of the average contribution in the sample (7.412 RMB). We note that the effects of the CWHP on prosocial behaviors, although statistically significant, are relatively small in magnitude. Prosocial

[23] Straw burning is linked with plowing and harvesting seasons that may span across a single month. Therefore, to reduce misclassification errors, we assign the month of interest a weight of 0.5, and the month before and the month after each a weight of 0.25 in Figure 5.

[24] Another potential source of measurement error for village fire spots arises from the imperfect overlap between the chosen 5-km radius and exact village boundaries. Appendix Table A2 uses the 4 km and 6 km radii to define village fires and demonstrates the robustness of our analysis against variations in spatial parameters. Given the strong seasonality in straw burning, as revealed in Figure 6, we focus on the effects of CWHP during straw-burning seasons in Appendix Table A2 to reduce contamination from other months.

behaviors, which are deeply rooted in individuals' social preferences, generally evolve slowly and are not highly responsive to short-term interventions like the CWHP. Therefore, while the program may influence behaviors, substantial changes in prosocial tendencies may not manifest immediately.

[Table 5 is here]

To verify if both outcomes are driven by the CWHP and operate through the same channels, as discussed in the following section, we should be able to associate the two outcomes with a negative correlation. That is, the villages where the CWHP induces a greater reduction in prosociality are exactly those with higher instances of law-breaking straw burning. To elucidate this relationship, we compare the results from incentivized games with data on illegal straw-burning behavior.

These graphical results are displayed in Appendix Figure A2, where we aggregate the CF-predicted individualized conditional average treatment effects (CATEs) in the two incentivized games by village and then match these aggregated variables with the measures of straw burning at the village level during straw-burning seasons.[25] In Appendix Figure A2, we outline the association between the two outcomes through a simple regression analysis. We observe strong negative correlations between straw burning and the outcomes of the incentivized games, depicted in Appendix Figure A2(a) and A2(b). All correlations are statistically significant at the 5% level at least. The coefficients and corresponding t-statistics are reported within the graphs.

## 6. Discussions on the Mechanisms

[25] We use the individualized CATEs predicted by CF in the two incentivized games rather than the direct outcomes (money to others or the group account) to control for the influence of other variables (such as income and gender) on their prosocial behavior.

This section discusses the mechanisms of the CWHP impact by examining heterogeneous treatment effects. To sharpen the estimation of straw burning, we focus on the impacts during straw-burning seasons. We consider two possible channels in particular: the negative income effect and emotional discontent. In addition, we endeavor to rule out competing explanations using the comprehensive data collected through our survey.

## 6.1 Re-optimization in response to the adverse income shock

The CWHP is likely perceived as a negative income shock among farmers. As revealed in Table 2, the net heating costs of CWHP participants amount to an average of 3507 RMB, accounting for roughly 10% of participants' annual household income. Additionally, our follow-up telephone interviews reveal that each household's average annual cost of eco-friendly straw disposal is approximately 780 RMB.[26] In contrast, burning straw requires less effort and involves almost no monetary costs. In addition, many farmers in China incorrectly believe that the high temperatures during burning can eliminate pests and that the ash residues can fertilize the soil.[27]

Given that CWHP participation is mandatory, the substantial heating costs imposed by the policy strain villagers' budgets, particularly in financially constrained rural areas. This income shock tightens their budgets, leading them to cut back on other significant expenditures,

[26] In our follow-up telephone interviews, we specifically inquired whether their families operated any cropland. Among the respondents who confirmed that they did (43 interviewed individuals), we further probed into their methods for disposing of crop straw and the associated costs of environmentally friendly straw disposal. We discovered that the average disposal cost per mu of cropland is roughly 100 RMB, and the average size of cropland they operate is about 7.8 mu, indicating that the average annual cost of environmentally friendly straw disposal is approximately 780 RMB.

[27] Although the effects of burning straw on pest elimination and soil fertilization may appear obvious, it cannot eliminate pests buried in the soil, and many nutrients such as nitrogen and sulfur are released into the atmosphere rather than being absorbed by the soil (source: https://www.cenews.com.cn/news.html?aid=1097255).

such as the proper disposal of agricultural waste. Consequently, farmers may find straw burning a more economically attractive option compared to the additional costs of proper disposal.[28]

We propose two heterogeneity analyses to test the negative income effect as a driving force behind the reduction in prosociality. First, we assess how the CWHP-induced budget constraints vary with farmers' original income levels. The income shock is expected to have a more pronounced impact in regions with lower living standards, where the CWHP effects would likely be more significant. We categorize villages exposed to the CWHP into two groups based on the average disposable income per capita in their counties. Each group, along with 11 never-treated villages, forms a subsample for the heterogeneity analysis. We adopt a consistent estimation strategy as in the benchmark DID analyses. To estimate the effect on the two incentivized games, we calculate the average treatment effect on the treated (ATT) within different subgroups based directly on the full-sample CF model trained in Table 5. Table 6 examines all five major outcomes. Columns (1) and (2) report the estimators for the high- and low-income subgroups, respectively. The results show that the coefficients using the sample of the low-income subgroup are of a larger magnitude and more significant, and significant differences in point estimates are observed for both the incidence of straw burning and the amount of money allocated to the public account.

[Table 6 is here]

[28] During our survey, enumerators did not directly inquire whether and why the surveyed households engaged in straw burning, as they might be hesitant to admit to these behaviors honestly. Instead, enumerators informed households with crop straws to dispose of various clean disposal methods, such as selling to factories, feeding livestock, using it in biogas digesters, and using it as a substrate for mushrooms or worms. Then, the enumerators asked why they did not choose these methods for straw disposal. Among the respondents, approximately 60% were aware of these methods but found them impractical due to the significant time and high transportation costs associated with collecting and delivering straws to designated locations. Additionally, 29% lacked knowledge about environmentally friendly disposal methods, and about 10% reported a lack of necessary equipment and technology for environmentally friendly straw disposal, such as pulverizers and biogas digesters.

Second, CWHP-induced heating costs vary by winter temperature and humidity in the locality. Low temperatures and higher humidity during winter imply higher heating costs and greater financial strain. Unlike the income gradient, this heterogeneity analysis exploring variations in weather conditions avoids concerns that villages with different income levels could systematically differ in aspects such as crop types and production methods. In such cases, the results of income heterogeneity could be misinterpreted.

The villages in our sampling frame are located in Beijing, Tianjin, and Hebei, with a maximum distance of roughly 300 kilometers. Therefore, most variations for identification are temporal rather than regional.[29] In this context, we use the intrinsic variations in monthly temperature and snowfall relative to historical levels to examine the heterogeneous effects of the CWHP over different time periods. Specifically, for each heating season (from November to March of the following year), we compare its average temperature and number of snow days with those recorded over the past decade. Subsequently, we divide CWHP participant villages into two groups based on these intrinsic weather differences. Each group, along with the 11 never-treated villages, forms a subsample for our heterogeneity analysis. Table 7, columns (1) and (2) report the CWHP effect in warm and cold winters, respectively. We find that villages encountering lower temperatures are more likely to burn straw. Using a similar method, we examine the heterogeneity in the CWHP effects by snowfall in Table 8, with columns (3) and (4) displaying results for villages experiencing more and less snowfall, respectively. Villages with more snowfall are more inclined to burn straws in response to CWHP.[30]

[Table 7 is here]

[29] For example, there was cold weather in January and February of 2018, but warmer weather was present at the same time in 2017.

[30] Table 7 focuses only on the outcome of straw burning. The heterogeneity is identified mostly from temporal rather than regional variations in weather conditions. Therefore, this strategy cannot be applied to the 2018 cross-sectional data from the incentivized games.

## 6.2 Emotional Discontent

It is not a new notion in social science that when people feel unfairly treated or even wronged, they would possibly behave in a more anti-social (or less prosocial) manner.[31] Anecdotal evidence suggests that in some areas, the CWHP has been perceived as prioritizing Beijing's air quality at the expense of local residents' well-being.[32] This perception, compounded by the negative income effect, the inconvenience of infrastructure changes, stove replacements, and high heating costs, may provoke dissatisfaction among villagers, potentially culminating in instances of illegal straw burning as an extreme expression of their dissatisfaction (Elster, 1998).

The rapid CWHP implementation sometimes led to inadequate infrastructure to support the transition from coal to gas. Coal-burning stoves were usually removed before October to prevent farmers from burning coal during the heating season. However, extending gas pipelines and constructing power grids often take longer, resulting in gas and electricity shortages. Consequently, many villagers were unable to adequately heat their homes, particularly on cold winter nights when heating demand peaked. In addition, as a large number of villages switched from coal to gas and electricity, this exacerbated the shortage of natural gas and placed a strain on electricity grids. Using our door-to-door household survey data, Appendix Table A5 shows that households that experienced power and gas shortages are more likely to rate the CWHP as making them less happy, less warm and comfortable during the heating season, and more insecure. Therefore, we utilize whether villagers had unpleasant experiences with natural gas or electricity shortages in the winter to capture the dissatisfaction induced by the CWHP.

[31] For example, Kuhlken (1999) concludes from fire cases in the US, England, and Algeria that farmers use fire as a weapon of resistance to express their discontent. Kull (2002) notices that half of the grassland and woodland of Madagascar is burned by farmers and herders and points out that peasants burned the land to protest and to manage resources, such as crop field preparation and pest control. Holmes (2007) analyzes the resistance to conservation and asserts that fire is a popular form of protest. Nolte et al. (2013) further highlight that in Brazil, farmers might engage in environmentally degrading activities in protected areas as a protest against protection measures.

[32] https://www.nytimes.com/2018/02/10/world/asia/china-coal-smog-pollution.html

In our sample, 82 of the 681 CWHP participant households have once experienced a shortage of gas or electricity. Generally, shortages do not occur in a single household but rather among the villagers in an entire village using the same type of heating source. To assess the prevalence of shortages across villages, we rank them based on the proportion of households that reported shortages. This proportion provides an indication of the likelihood and frequency of villages experiencing shortages. We then divide the sample into two subgroups based on the median shortage ratio. Each subgroup, along with the 11 never-treated villages, forms a separate subsample for our heterogeneity analysis. Table 8, columns (1) and (2) report the impact of the CWHP on the high shortage and low shortage samples, respectively. Panels A–C show that the effects of the CWHP on straw-burning activities in the villages with a high share of villagers experiencing heating shortages are almost twice as large as the effects for the low shortage group. Panels D and E suggest that the CWHP led subjects in high shortage propensity villages to allocate 4.275 RMB less to neighbors and to contribute 3.673 RMB less to public goods. In contrast, the effects of the CWHP on subjects residing in villages with a low possibility of shortage are much smaller. The differences between the subsamples in the point estimates are statistically significant for four out of the five outcomes. These results in Table 8 and Appendix Table A5 indicate that the experience of energy shortages can exacerbate dissatisfaction with the policy. Our regression results align with the hypothesis that straw burning may be provoked by the desire to express personal dissatisfaction.

Additionally, the heterogeneity analyses on the two games performed in Sections 6.1 and 6.2 are robust to different weighting schemes. Appendix Table A6 shows that heterogeneity remains when using OLS and PSM estimations.[33]

[Table 8 is here]

## 6.3 Excluding Competitive Channels

***Reduced Prosociality due to Air Pollution.*** A growing body of literature has documented the significant impact of air pollution on various aspects of individual behaviors, including cognitive functions (Lai et al., 2021, 2022), violent crimes (Herrnstadt et al., 2021), and risk preferences and prosocial behaviors (Chew et al., 2021). Although the CWHP has overall improved the average air quality throughout the year, the study by Chew et al. (2021) demonstrates that short-term increases in smog concentration can still alter brain function and decision-making. Our survey was conducted in the spring of 2018, coinciding with the peak straw-burning season. According to Chew et al. (2021), short-term pollution may potentially cause villagers to exhibit less prosocial behavior. We conduct two analyses to examine this possibility.

First, we control for PM2.5 concentrations on the survey date, redo the benchmark analyses, and report the results in Table 9. We categorize PM2.5 concentrations into three bins: (0, 35), (35, 65), and above 65 μg/m$^3$ to account for potential non-linear effects of pollution. These bins correspond to PM2.5 levels considered harmless, unhealthy for sensitive groups, and unhealthy for all individuals, respectively, according to the U.S. Environmental Protection Agency guidelines. Controlling for the dummy indicators of PM2.5 concentration falling in the

[33] The heterogeneity analyses for straw burning in Sections 6.1 and 6.2 are also robust when extending the sampling months from the straw burning seasons to the whole year, as outlined in Appendix Tables A7, A8, and A9.

range of (35, 65), and above 65 μg/m$^3$, respectively, the coefficients of the CWHP remain significant, and the magnitudes are similar to those shown in Table 5. This consistency across different specifications suggests that the adverse effects of the CWHP are not primarily driven by different exposures to air pollution between the participants and non-participants. Our findings reveal a notably negative impact of heavy pollution, i.e., concentrations above 65 μg/m$^3$. These results align with the research of Chew et al. (2021), which demonstrates a negative impact of high PM2.5 exposure on prosocial behaviors.

[Table 9 is here]

As another parallel sensitivity check, we confine the sample to villages where the air quality reached the "good" level on the survey date and examine the CWHP's impact on prosocial behaviors. A "good" level is defined by an Air Quality Index under 50, implying minimal air pollution. Appendix Table A10 demonstrates that the impact of the CWHP on prosocial behavior remains significant, and the magnitudes are even larger than the results estimated with the full sample, as shown in Table 5. To sum up, even on days with negligible air pollution, CWHP participants still exhibit significantly lower prosociality than non-participants, implying that the adverse effects of the CWHP on prosociality persist irrespective of individuals' exposure to air pollution.

***CWHP as a Heating Substitution for Burning Straw.*** An alternative explanation for the increase in straw burning due to the CWHP is that villagers may resort to burning straw outdoors after being prohibited from using it indoors. It is argued that farmers might burn straws as agricultural waste immediately after the harvest season and possibly before the next planting season. When straw is used indoors for heating, it is typically burned in small amounts multiple times in hearths, making it less detectable by satellites. In contrast, large-scale straw burning outdoors on agricultural land is more easily detected by remote sensing satellites.

The ostensibly plausible explanation, however, does not hold. Burning straws for heating is inefficient and causes severe indoor pollution. Farmers' living standards have improved during the past few decades, leading many to replace traditional straw-burning hearths with coal-burning stoves. In addition, since 2009, regulations in the study region have prohibited all forms of straw burning, whether for agricultural waste disposal or household heating, without making distinctions between the purposes.

We perform a heterogeneity analysis to further rule out this mechanism. In our survey, we asked about the traditional heating method in the regions, especially the burning of straw in the hearth. Based on this information, we divide the sample into two groups. Table 10 shows that the CWHP effects are similar based on whether straw burning was historically adopted as a heating method. Columns (1), (3), and (5) report the impact of the CWHP in villages that used to burn straw for indoor heating, and the other three columns report the impact in villages that did not. Although the CWHP effects are slightly larger for the group that historically used straw as heating fuel, they are less statistically significant. The coefficient differences between the two groups are modest and lack statistical significance.[34]

[Table 10 is here]

***Change in Enforcement of the Ban on Straw Burning.*** When the CWHP is vigorously promoted, one concern is whether local authorities are more lenient to inter-seasonal substitution of air pollution as long as the average pollution levels remain the same over the year.

[34] Additionally, estimation results using observations from all 12 months are presented in Appendix Table A11, which closely align with the findings in Table 10.

Figure 7 shows that this is not the case.[35] Our survey collected detailed information on punishment for straw burning from subjects with farmland and straws to dispose of. We conduct *t*-tests on penalties regarding straw burning across the CWHP participants and non-participants in 2018. In particular, we examine dummy indicators such as whether the village received government inspections, whether bans were strictly enforced, whether fines or confiscations were imposed for straw burning, and whether those responsible were detained. The t-tests show no statistical differences in any indicators of the strictness of straw-burning bans. These results indicate that the CWHP implementation is unlikely to result in any relaxation of straw-burning prohibitions.

[Figure 7 is here]

Additionally, Table 3 illustrates a marginally significant reduction in PM2.5 concentration, indicating that local authorities are motivated to maintain consistent reductions in air pollution throughout the year. Despite improvements in winter pollution levels, local authorities are unlikely to condone straw burning. The CWHP and bans on straw burning are direct responses to China's Air Pollution Prevention and Control Action Plan, initiated by the State Council in 2013 (Greenstone et al., 2021). This national initiative mandates Beijing, Tianjin, and Hebei to achieve a 25% reduction in PM2.5 concentrations by the end of 2017, with ongoing improvements thereafter subject to annual evaluation by the State Council. Local officials are held accountable for meeting these stringent targets, as any relaxation of pollution control measures could jeopardize achieving these ambitious goals and potentially impact their future career prospects.

[35] It is worth noting that the potential for misreporting of inspections and penalties related to straw burning by the surveyed households cannot be entirely excluded. If villages that have implemented the policy are more inclined to over-report such activities, this could attenuate the significance of the findings presented in Figure 7.

# 7. Welfare Analyses

On the one hand, the CWHP inevitably raises heating costs despite substantial subsidies from the local government. On the other hand, as detailed in Section 5.1, the CWHP has reduced indoor pollution, potentially leading to substantial health benefits. Estimating the overall health impact and that on household well-being based on willingness to pay is crucial from an economic perspective.

## 7.1 Costs of the CWHP

Although CWHP is heavily subsidized by the government, the subsidies cannot fully cover the costs, especially the costs of the fuels, as reviewed in Section 2.1. Additionally, the subsidies are inconvenient to claim for most villages. In our sample, only 9.1% of treated households received the fuel subsidy as an immediate deduction from the fuel purchase. Around 40% of treated households received their fuel subsidy on their payment card, while another 41% had to seek reimbursement by approaching the responsible person with an invoice after paying the full fuel fee. Another 10% were unsure how to claim the subsidy. Apart from those who received the subsidy through immediate price reduction, 30% of treated households received the subsidy 1–2 months after paying the fuel bill, 40% received it between 2–4 months, and about 20% indicated that they had not received the subsidy at the time of our survey. Apart from those who received the subsidy through immediate price reduction, 30% of treated households received the subsidy 1–2 months after paying the fuel bill, 40% received it between 2–4 months, and about 20% indicated that they had not received the subsidy at the time of our survey. Keeping these figures in mind, it is not hard to understand that heating costs could consume a large proportion of villagers' disposable income with overdue or not fully materialized subsidies. This is particularly challenging for villagers experiencing credit constraints.

Table 11 formally quantifies the CWHP's impact on heating expenditure and subsidies using causal forests. Columns (1) and (2) show that the CWHP increases heating fuel costs by 1,200 RMB (about 171 USD), with about 50% (595 RMB) of this amount being covered by a government subsidy. In addition, columns (3) and (4) indicate that the increase in heating equipment expenditure is 1,732 RMB, 80% of which is offset by a government subsidy (1,399 RMB). Following Zhao et al. (2021), we assume a 10-year lifespan for the heating equipment and a discount rate of 6%. In this context, the average increase in the net annualized heating expenditure of CWHP participants is about 649 RMB (93 USD). This figure is close to, but slightly smaller than, the average annual cost of clean straw disposal (780 RMB) revealed in our follow-up telephone interviews.

[Table 11 is here]

## 7.2 The Impact on Wellbeing

To analyze the impact of the CWHP on well-being, we follow Allcott and Kessler (2019) in weighing the willingness to pay and the costs induced by the CWHP, as presented in Figure 8. To measure a household's WTP, the survey asked how much they would pay for clean heating fuel and equipment without any subsidy in place. CWHP participants reported an average WTP of 1962 RMB (280 USD) for heating with clean fuel. Additionally, follow-up telephone interviews offer more detailed information, showing an average WTP of 913 RMB (130 USD) for air quality improvements and 875 RMB (125 USD) for enhanced heating comfort. However, after deducting subsidies, their average net expenditure was 2107 RMB (301 USD), resulting in an expenditure gap of 145 RMB (21 USD) for fuel. Similarly, participants expressed an average WTP of 1171 RMB (167 USD) for heating equipment, with a net expenditure of 1267 RMB (181 USD), indicating an expenditure gap of 96 RMB (14 USD). Combining these gaps for fuel and equipment yields an average welfare loss of 241 RMB (34 USD) per CWHP participant.

[Figure 8 is here]

## 7.3 Health Impacts of the CWHP

The reduced indoor air pollution brought about by the CWHP implies direct health benefits to participants. In addition, the transition to clean heating has substantial potential to decrease outdoor air pollution, as indicated by studies conducted by Meng et al. (2019) and Zhao et al. (2021). This reduction in outdoor pollution would also yield health benefits for non-CWHP participants. However, the increase in straw burning can generate considerable air pollutants and potentially offset the direct impact of the CWHP on reducing outdoor pollution. In this section, we aim to quantify the health benefits and costs associated with the direct and indirect effects of the CWHP.

The DID estimator in Table 3 reports the overall effect on air pollution, which is the net of the pollution reduction from clean heating conversion and the pollution increase because of more straw burning. To isolate the latter, we employ the Hybrid Single-Particle Lagrangian Integrated Trajectory model to simulate the transport trajectories of air pollutants caused by straw burning. The simulation procedures and estimation results are presented in Online Appendix A2 and Appendix Table A12. In summary, our findings indicate that the increase in PM2.5 concentration caused by straw burning due to the CWHP amounts to 0.30 $\mu g/m^3$. This estimation implies that the direct effect of the CWHP on reducing monthly ambient PM2.5

concentration is approximately 1.36 μg/m$^3$, and the indirect effect offsets about 22% of the intended outdoor air quality benefits.

Based on the estimated effects of the CWHP on indoor and outdoor pollution, we follow the approach of Zhao et al. (2021) and utilize the integrated exposure-response function to evaluate the health impacts. Details of this estimation process are provided in Online Appendix A3. Assuming that the CWHP were to cover all rural households in the Beijing-Tianjin-Hebei region, Panel A of Table 12 outlines the direct health effects of the CWHP. The CWHP reduces 4,285 annual rural mortalities and 2,619 urban mortalities in this region.[36] However, as indicated in Panel B, the increased straw burning results in a 12% offset of the total health benefits achieved through the clean heating conversion. Panel C, using the value of a statistical life from He et al. (2021), calculates the net health benefits at 17,333 million RMB, with 11,353 million attributed to rural residents and 5,980 million to urban residents. Panel D further compares these monetized health benefits to expenditure costs. The net health benefits for rural residents significantly outweigh their expenditure costs (8,173 million RMB). However, due to the substantial subsidy borne by the government (9,809 million RMB), the total cost amounts to 17,982 million RMB, slightly higher than the net health benefits, resulting in a benefit-cost ratio of 0.96. This ratio, which is less than one and considerably lower than that of other widely implemented environmental policies (e.g., Jacobsen et al., 2023; Barwick et al., 2024), suggests a limited net welfare gain from the CWHP regarding health benefits.

[Table 12 is here]

---

[36] According to the Exposure Factors Handbook of Chinese Population, complied by the Ministry of Environmental Protection of China, rural residents typically spend approximately 20% of their time on outdoor activities. Therefore, the reduction of average PM2.5 exposure for rural residents amounts to 1.36 * 0.2 + 3.60 * 0.8 = 3.15 μg/m$^3$. For urban residents, we assume that the change in their indoor PM2.5 concentration equals the change in outdoor concentration, resulting in a reduction of their average PM2.5 exposure by 1.36 μg/m$^3$.

## 8. Conclusions

In this paper, we leverage the context of the CWHP in northern China as a quasi-natural experiment to examine the effect of environmental policy implementation on individual anti- and pro-social behavior. Our findings indicate that treated villages exhibit higher incidences of anti-social behavior, measured by straw-burning activities, after enrolling in the CWHP. This trend is particularly pronounced in villages with lower living standards and those experiencing heating shortages due to inadequate power or natural gas supply. These heterogeneities suggest that both advantageous income shocks and discontent triggered by the CWHP likely contribute to the incidence of law-violating straw burning. In addition, employing two incentivized games, we observe that participants in the CWHP demonstrate lower levels of prosocial behavior, as evidenced by their allocation decisions in dictator and public goods games, compared to non-participants.

Finally, we show that the CWHP enhances the indoor amenities of participant households. However, the substantial increase in fuel expenditure, coupled with inadequate subsidies, diminishes the improvement in subjective well-being. In addition to subjective utility, we also quantify the objective impact of the CWHP on urban and rural mortalities. The overall benefit-cost ratio of the program in terms of reducing air pollution is slightly lower than one due to the high expenditure cost and unintended impacts on straw burning.

Despite its aim to improve air quality in northern China, the CWHP has been fraught with controversy since its inception due to increased heating costs for households as well as shortages in the natural gas and electricity supply. The government attempted to mitigate supply issues by slowing down the original expansion timetable. However, two critical lessons regarding the CWHP necessitate urgent attention and are of serious concern. First, the indirect consequences may be far more complex than initially anticipated by authorities. Economists and policymakers need to adopt a more comprehensive approach to policy evaluation, particularly in planning surveys. Our findings indicate that the additional costs stemming from the CWHP can influence not only energy consumption but also production-related behaviors. More significantly, anti-social behaviors could lead to negative and enduring consequences for social capital.

Secondly, millions of households have already transitioned from coal to natural gas and electricity. A critical issue for these households is how future subsidies should be structured. The CWHP prioritized equipment replacement, effectively eliminating coal burning and contributing to achieving air quality goals. However, inadequate and delayed fuel subsidies could result in unintended consequences, including increased pollution during planting and harvest seasons.

**Acknowledgments**

The authors thank Jianuo Gai for her excellent research assistance. Cao gratefully acknowledges financial support from the National Natural Science Foundation of China (72243007, 72250064) and Ministry of Science and Technology of the People's Republic of China (2023YFE0112900). Liu gratefully acknowledges financial support by the National Natural Science Foundation of China (72222005, 72342032). Ma gratefully acknowledges financial support from the National Natural Science Foundation of China (72134006, 72304272). Sun gratefully acknowledges financial support from the National Natural Science Foundation of China (72373157), Fundamental Research Funds for the Central Universities, and the Research Funds of Renmin University of China (22XNA003). The authors are also thankful for the support by the Energy Foundation, China Southern Power Grid Co., Ltd, Research Center for Green Economy and Sustainable Development and Institute for Global Development of Tsinghua University, and Harvard-China Project on Energy, Economy and Environment.

## Reference


Adepoju, T., Carson, A. L., Jin, H. S., & Manasseh, C. S. (2023). Hospital Boarding Crises: The Impact of Urgent vs. Prevention Responses on Length of Stay. Management Science, 69(10), 5948-5963.

Alam, Z. S., Chen, M. A., Ciccotello, C. S., & Ryan Jr, H. E. (2018). Board structure mandates: Consequences for director location and financial reporting. Management Science, 64(10), 4735-4754.

Alix-Garcia, J., & Millimet, D. L. (2023). Remotely incorrect? Accounting for nonclassical measurement error in satellite data on deforestation. Journal of the Association of Environmental and Resource Economists, 10(5), 1335-1367.

Allcott, H., & Kessler, J. B. (2019). The welfare effects of nudges: A case study of energy use social comparisons. American Economic Journal: Applied Economics, 11(1), 236-276.

Allcott, H., & Wozny, N. (2014). Gasoline prices, fuel economy, and the energy paradox. Review of Economics and Statistics, 96(5), 779-795.

Alyakoob, M., Rahman, M. S., & Wei, Z. (2021). Where you live matters: Local bank competition, online marketplace lending, and disparity in borrower benefits. Information Systems Research, 32(4), 1390-1411.

Anand, K. S., & Giraud-Carrier, F. C. (2020). Pollution regulation of competitive markets. Management Science, 66(9), 4193-4206.

Athey, S., & Imbens, G. (2016). Recursive partitioning for heterogeneous causal effects. Proceedings of the National Academy of Sciences, 113(27), 7353-7360.

Athey, S., Tibshirani, J., & Wager, S. (2019). Generalized random forests. Annals of Statistics, 47(2), 1148-1178.

Barrington-Leigh, C., Baumgartner, J., Carter, E., Robinson, B. E., Tao, S., & Zhang, Y. (2019). An evaluation of air quality, home heating and wellbeing under Beijing's programme to eliminate household coal use. Nature Energy, 4(5), 416-423.

Barwick, P. J., Li, S., Lin, L. & Zou, E. (2024). From Fog to Smog: the Value of Pollution Information. American Economic Review, 114(5), 1338-1381.

Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event-study designs: robust and efficient estimation. Review of Economic Studies.

Burnett, R., Chen, H., Szyszkowicz, M., Fann, N., Hubbell, B., Pope III, C.A., Apte, J.S., Brauer, M., Cohen, A., Weichenthal, S. and Coggins, J. (2018). Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. Proceedings of the National Academy of Sciences, 115(38), 9592-9597.

Carter, M. R., Tjernström, E., & Toledo, P. (2019). Heterogeneous impact dynamics of a rural business development program in Nicaragua. Journal of Development Economics, 138, 77-98.

Carpenter, J., & Seki, E. (2011). Do social preferences increase productivity? Field experimental evidence from fishermen in Toyama Bay. Economic Inquiry, 49(2), 612-630.

Chen, Y., Ebenstein, A., Greenstone, M., & Li, H. (2013). Evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River policy. Proceedings of the National Academy of Sciences, 110(32), 12936-12941.

Chew, S. H., Huang, W., & Li, X. (2021). Does haze cloud decision making? A natural laboratory experiment. Journal of Economic Behavior & Organization, 182, 132-161.

Cohen, A. J., Brauer, M., Burnett, R., Anderson, H. R., Frostad, J., Estep, K., ... & Forouzanfar, M. H. (2017). Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. The lancet, 389(10082), 1907-1918.

Dang, V. A., Gao, N., & Yu, T. (2023). Climate policy risk and corporate financial decisions: Evidence from the NOx budget trading program. Management Science, 69(12), 7517-7539.

Davis, J. M., & Heller, S. B. (2017). Using causal forests to predict treatment heterogeneity: An application to summer jobs. American Economic Review, 107(5), 546-550.

Davis, J. M., & Heller, S. B. (2020). Rethinking the benefits of youth employment programs: The heterogeneous effects of summer jobs. Review of economics and statistics, 102(4), 664-677.

De Groote, O., & Verboven, F. (2019). Subsidies and time discounting in new technology adoption: Evidence from solar photovoltaic systems. American Economic Review, 109(6), 2137-2172.

De Neve, J. E., Imbert, C., Spinnewijn, J., Tsankova, T., & Luts, M. (2021). How to improve tax compliance? Evidence from population-wide experiments in Belgium. Journal of Political Economy, 129(5), 1425-1463.

Dolan, P., & Galizzi, M. M. (2015). Like ripples on a pond: Behavioral spillovers and their implications for research and policy. Journal of Economic Psychology, 47, 1-16.

Ebenstein, A., Fan, M., Greenstone, M., He, G. and Zhou, M., 2017. New evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River Policy. Proceedings of the National Academy of Sciences, 114(39), 10384-10389.

Elster, J. (1998). Emotions and economic theory. Journal of Economic Literature, 36(1), 47-74.

Fowlie, M., Greenstone, M., & Wolfram, C. (2015). Are the non-monetary costs of energy efficiency investments large? Understanding low take-up of a free energy efficiency program. American Economic Review, 105(5), 201-204.

Fowlie, M., Greenstone, M., & Wolfram, C. (2018). Do energy efficiency investments deliver? Evidence from the weatherization assistance program. The Quarterly Journal of Economics, 133(3), 1597-1644.

Gangadharan, L., Grossman, P. J., Molle, M. K., & Vecci, J. (2019). Impact of social identity and inequality on anti-social behaviour. European Economic Review, 119, 199-215.

Graff Zivin, J., Liu, T., Song, Y., Tang, Q. & Zhang, P. (2020). The unintended impacts of agricultural fires: Human capital in China. Journal of Development Economics, 147, 102560.

Gsottbauer, E., Müller, D., Müller, S., Trautmann, S. T., & Zudenkova, G. (2022). Social class and (un) ethical behaviour: Causal and correlational evidence. The Economic Journal, 132(647), 2392-2411.

Guan, Y., Li, C., Lu, H., & Wong, M. F. (2019). Regulations and brain drain: Evidence from Wall Street star analysts' career choices. Management Science, 65(12), 5766-5784.

Guo, T., Sriram, S., & Manchanda, P. (2021). The effect of information disclosure on industry payments to physicians. Journal of Marketing Research, 58(1), 115-140.

Hanna, R., Duflo, E., & Greenstone, M. (2016). Up in smoke: the influence of household behavior on the long-run impact of improved cooking stoves. American Economic Journal: Economic Policy, 8(1), 80-114.

He, G., Liu, T. & Zhou, M., (2020). Straw burning, PM2.5, and death: evidence from China. Journal of Development Economics, 145, 102468.

Herrnstadt, E., Heyes, A., Muehlegger, E., & Saberian, S. (2021). Air pollution and criminal activity: Microgeographic evidence from Chicago. American Economic Journal: Applied Economics, 13(4), 70-100.

Holmes, G. (2007). Protection, politics and protest: understanding resistance to conservation. Conservation and society, 5(2), 184-201.

Ito, K., & 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, M. R., Sallee, J. M., Shapiro, J. S. Regulating untaxable externalities: Are vehicle air pollution standards effective and efficient? Quarterly Journal of Economics, 138(3), 1907-1976.

Johnson, G. A., Shriver, S. K., & Goldberg, S. G. (2023). Privacy and market concentration: intended and unintended consequences of the GDPR. Management Science, 69(10), 5695-5721.

Knaus, M. C., Lechner, M., & Strittmatter, A. (2022). Heterogeneous employment effects of job search programs: A machine learning approach. Journal of Human Resources, 57(2), 597-636.

Kuhlken, R. (1999). Settin 'the woods on fire: rural incendiarism as protest. Geographical Review, 89(3), 343-363.

Kull, C. A. (2002). Madagascar aflame: landscape burning as peasant protest, resistance, or a resource management tool?. Political Geography, 21(7), 927-953.

Lai, W., Li, S., Li, Y., & Tian, X. (2022). Air pollution and cognitive functions: evidence from straw burning in China. American Journal of Agricultural Economics, 104(1), 190-208.

Lai, W., Song, H., Wang, C., & Wang, H. (2021). Air pollution and brain drain: Evidence from college graduates in China. China Economic Review, 68, 101624.

Li, J., & Wu, D. (2022). The price effect of drug price ceilings: Intended and unintended consequences. Management Science, 68(8), 5758-5777.

McCoy, D., & Lyons, S. (2017). Unintended outcomes of electricity smart-metering: trading-off consumption and investment behaviour. Energy Efficiency, 10(2), 299-318.

Meng, W., Zhong, Q., Chen, Y., Shen, H., Yun, X., Smith, K. R., ... & Tao, S. (2019). Energy and air pollution benefits of household fuel policies in northern China. Proceedings of the National Academy of Sciences, 116(34), 16773-16780.

Mobarak, A. M., Dwivedi, P., Bailis, R., Hildemann, L., & Miller, G. (2012). Low demand for nontraditional cookstove technologies. Proceedings of the National Academy of Sciences, 109(27), 10815-10820.

Newell, R. G., & Siikamäki, J. (2014). Nudging energy efficiency behavior: The role of information labels. Journal of the Association of Environmental and Resource Economists, 1(4), 555-598.

Nolte, C., Agrawal, A., Silvius, K. M., & Soares-Filho, B. S. (2013). Governance regime and location influence avoided deforestation success of protected areas in the Brazilian Amazon. Proceedings of the National Academy of Sciences, 110(13), 4956-4961.

Ottoni-Wilhelm, M., Vesterlund, L., & Xie, H. (2017). Why do people give? Testing pure and impure altruism. American Economic Review, 107(11), 3617-3633.

Peukert, C., Bechtold, S., Batikas, M., & Kretschmer, T. (2022). Regulatory spillovers and data governance: Evidence from the GDPR. Marketing Science, 41(4), 746-768.

Schroeder, W., Oliva, P., Giglio, L., Quayle, B., Lorenz, E., & Morelli, F. (2016). Active fire detection using Landsat-8/OLI data. Remote sensing of environment, 185, 210-220.

Sorrell, S., Gatersleben, B., & Druckman, A. (2020). The limits of energy sufficiency: A review of the evidence for rebound effects and negative spillovers from behavioural change. Energy Research & Social Science, 64, 101439.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. Statistical science: a review journal of the Institute of Mathematical Statistics, 25(1), 1.

Val Martin, M., Logan, J.A., Kahn, R.A., Leung, F.Y., Nelson, D.L. and Diner, D.J. (2010). Smoke injection heights from fires in North America: analysis of 5 years of satellite observations. Atmospheric Chemistry and Physics, 10(4), 1491-1510.

Van Donkelaar, A., Hammer, M.S., Bindle, L., Brauer, M., Brook, J.R., Garay, M.J., Hsu, N.C., Kalashnikova, O.V., Kahn, R.A., Lee, C. and Levy, R.C. (2021). Monthly global estimates of fine particulate matter and their uncertainty. Environmental Science & Technology, 55(22), 15287-15300.

Wang, G., Kawamura, K., Xie, M., Hu, S., Cao, J., An, Z., ... & Chow, J. C. (2009). Organic molecular compositions and size distributions of Chinese summer and autumn aerosols from Nanjing: Characteristic haze event caused by wheat straw burning. Environmental Science & Technology, 43(17), 6493-6499.

Wager, S., & Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. Journal of the American Statistical Association, 113(523), 1228-1242.

Zhao, B., Zhao, J., Zha, H., Hu, R., Liu, Y., Liang, C., Shi, H., Chen, S., Guo, Y., Zhang, D., & Aunan, K. (2021). Health benefits and costs of clean heating renovation: an integrated assessment in a major Chinese city. Environmental Science & Technology, 55(14), 10046-10055.

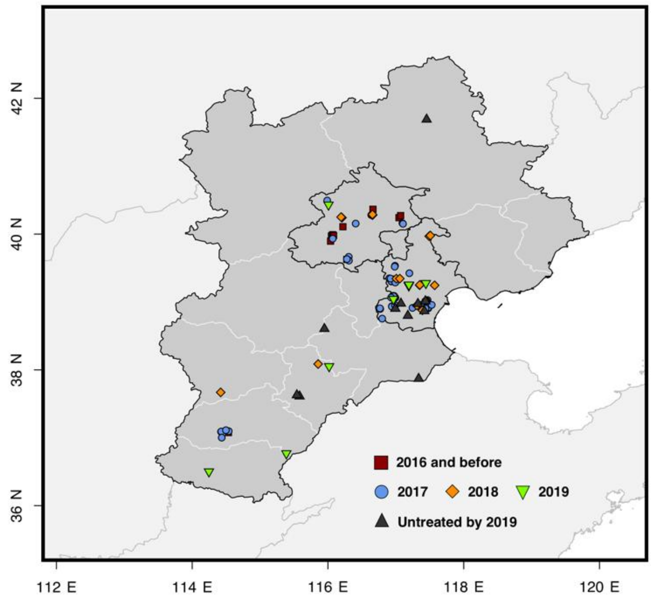


Figure 1 Spatial distribution of treated and untreated villages

Notes: Red squares represent the 11 villages treated in 2016 and before. Blue circles represent the 64 villages treated in 2017. Orange diamonds represent the 15 villages treated in 2018. Green down-pointing triangles represent the 9 villages treated in 2019. Black up-pointing triangles represent the 11 untreated villages by the end of 2019.

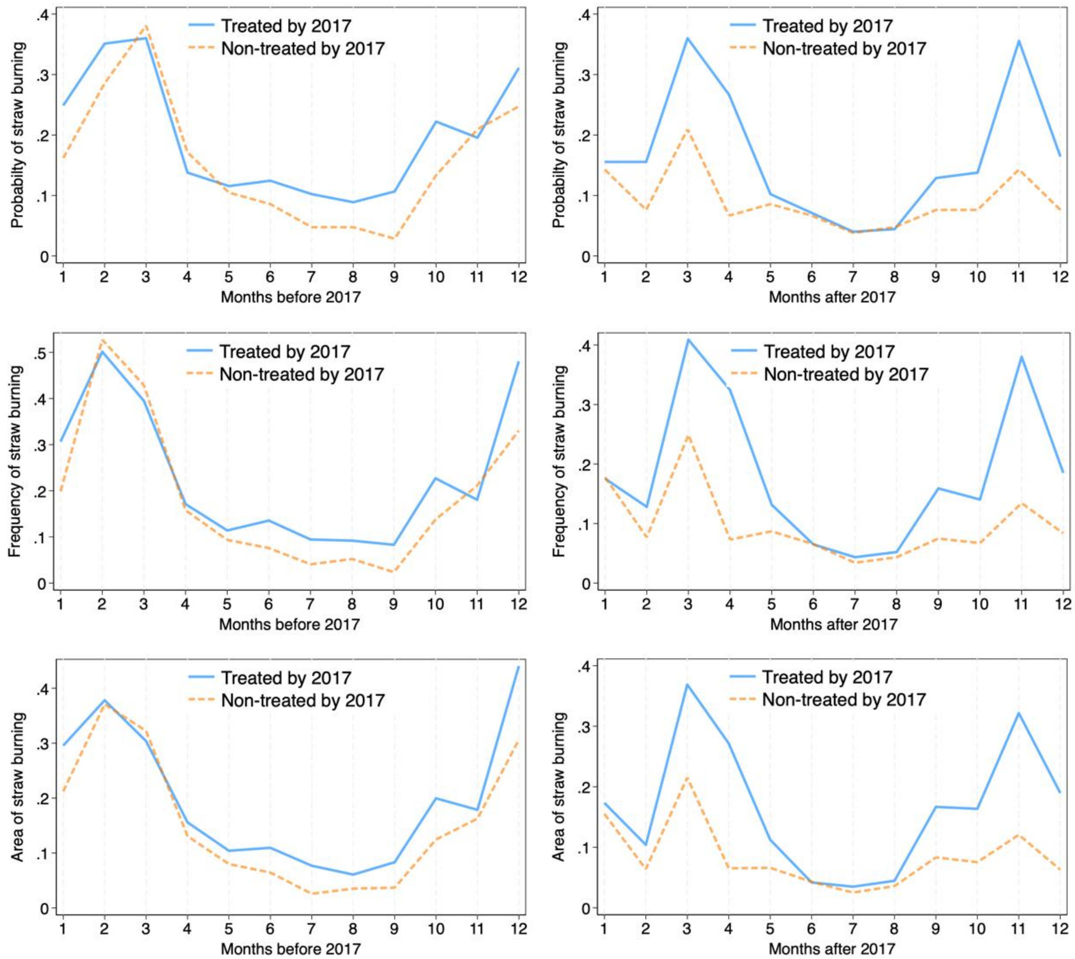



Figure 2 The probability, frequency, and area of straw burning by month before and after 2017

Notes: The first, second, and third rows show the probability, frequency, and area of straw burning, respectively, for the years before 2017 (left panel) and after 2017 (right panel). The solid blue lines represent the villages enrolled in CWHP by the end of 2017, and the orange dashed lines represent other villages.

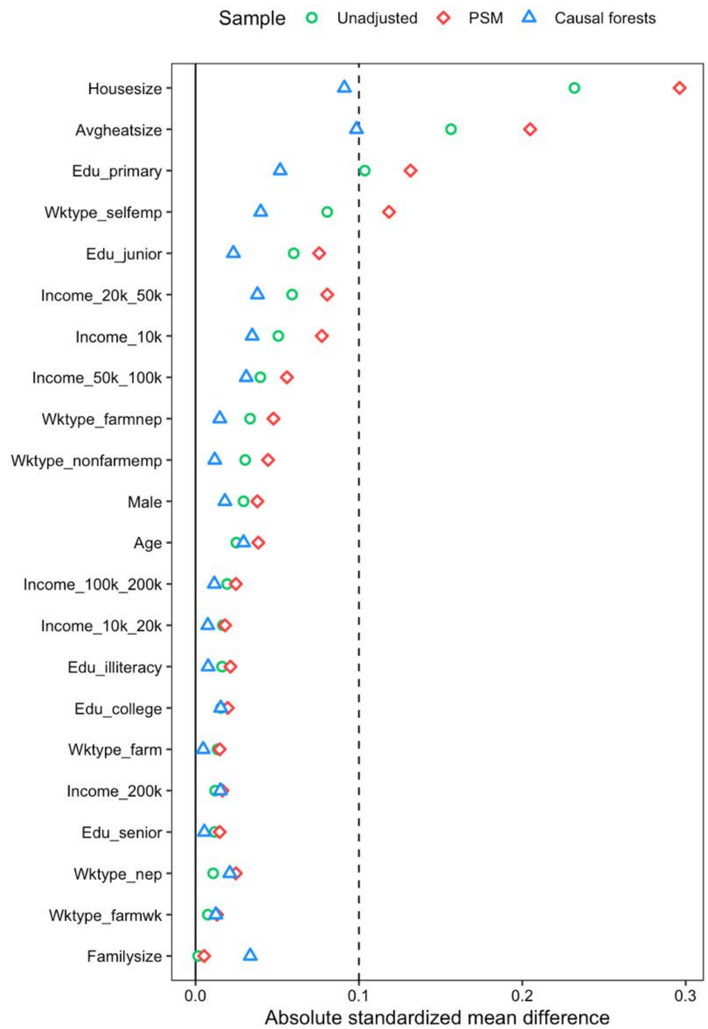



Figure 3 Differences in covariates

Notes: Housesize represents the house area of the survey respondents. Avgheatsize represents the per-person heating area of the survey respondents. Edu_illiteracy, Edu_primary, Edu_junior, Edu_senior, Edu_college are dummy variables for household heads with education levels of illiteracy, primary school, junior high school, senior high school, college, or higher. Wktype_farmwk is a dummy variable that equals 1 if the household head owns a farm business. Wktype_farmnep is a dummy variable that equals 1 if the household head works on others' farms. Wktype_selfemp is a dummy variable that equals 1 if the household head is self-employed. Wktype_nonfarmemp is a dummy variable that equals 1 if the household head works in non-agricultural industries. Wktype_nep is a dummy variable that equals 1 if the household head is unemployed. Income_10k, Income_10k–20k, Income_20k–50k, Income_50k–100k, Income_100k–200k, Income_200k are dummy variables for survey respondents with household annual incomes of less than 10k, 10k–20k, 20k–50k, 50k–100k, 100k–200k, 200k, or higher. Male is a dummy variable for the household head's gender. Age represents the household head's age. Familysize is the number of household members. The 0.1 cutoff suggested by Stuart (2010) is shown by the vertical dashed line.

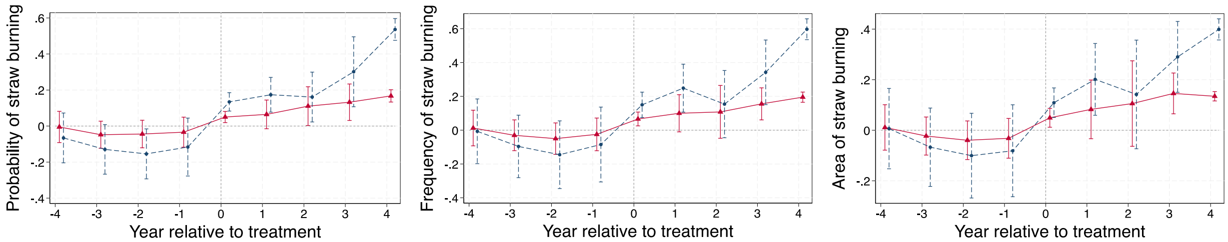
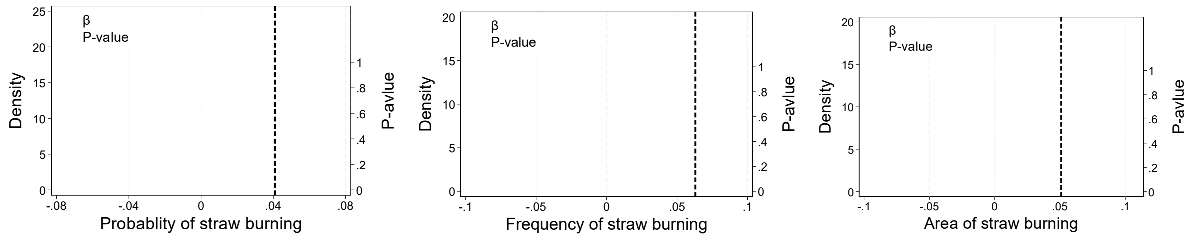
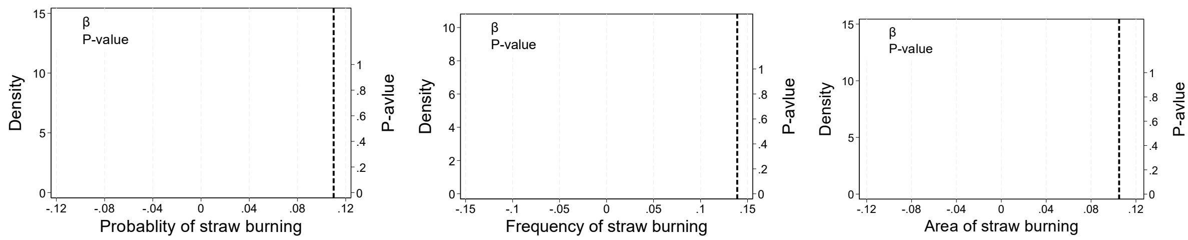



Figure 4 Time trends of the effects of the CWHP on straw burning

Notes: The dependent variables in the left, middle, and right figures are the probability, frequency, and area of straw burning. The red triangles and solid vertical lines represent the point estimates and their 95% confidence intervals for the full sample period. The blue circles and dashed vertical lines represent the point estimates and their 95% confidence intervals for the high straw-burning seasons. The method proposed by Borusyak et al. (2024) is used.




(a) Full sample




(b) Straw-burning seasons

Figure 5 Permutation tests for the effects of the CWHP on straw burning

Notes: The dependent variables in the left, middle, and right figures are the probability, frequency, and area of straw burning. The blue lines represent the distribution of the CWHP coefficients. The red dots represent the p-values. The vertical dashed lines mark the DID coefficients estimated using the observational data.

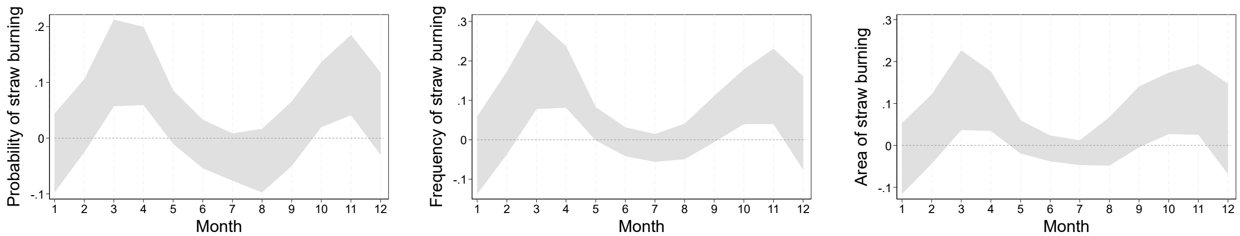



Figure 6 Seasonality in the effects of the CWHP on straw burning

Notes: The dependent variables in the left, middle, and right figures are the probability, frequency, and area of straw burning. The shaded area represents the 95% confidence interval. To account for the possibility that the plowing and harvesting seasons may span across a single month and reduce misclassification errors, the month of interest is assigned a weight of 0.5, while the month before and the month after are each assigned a weight of 0.25.

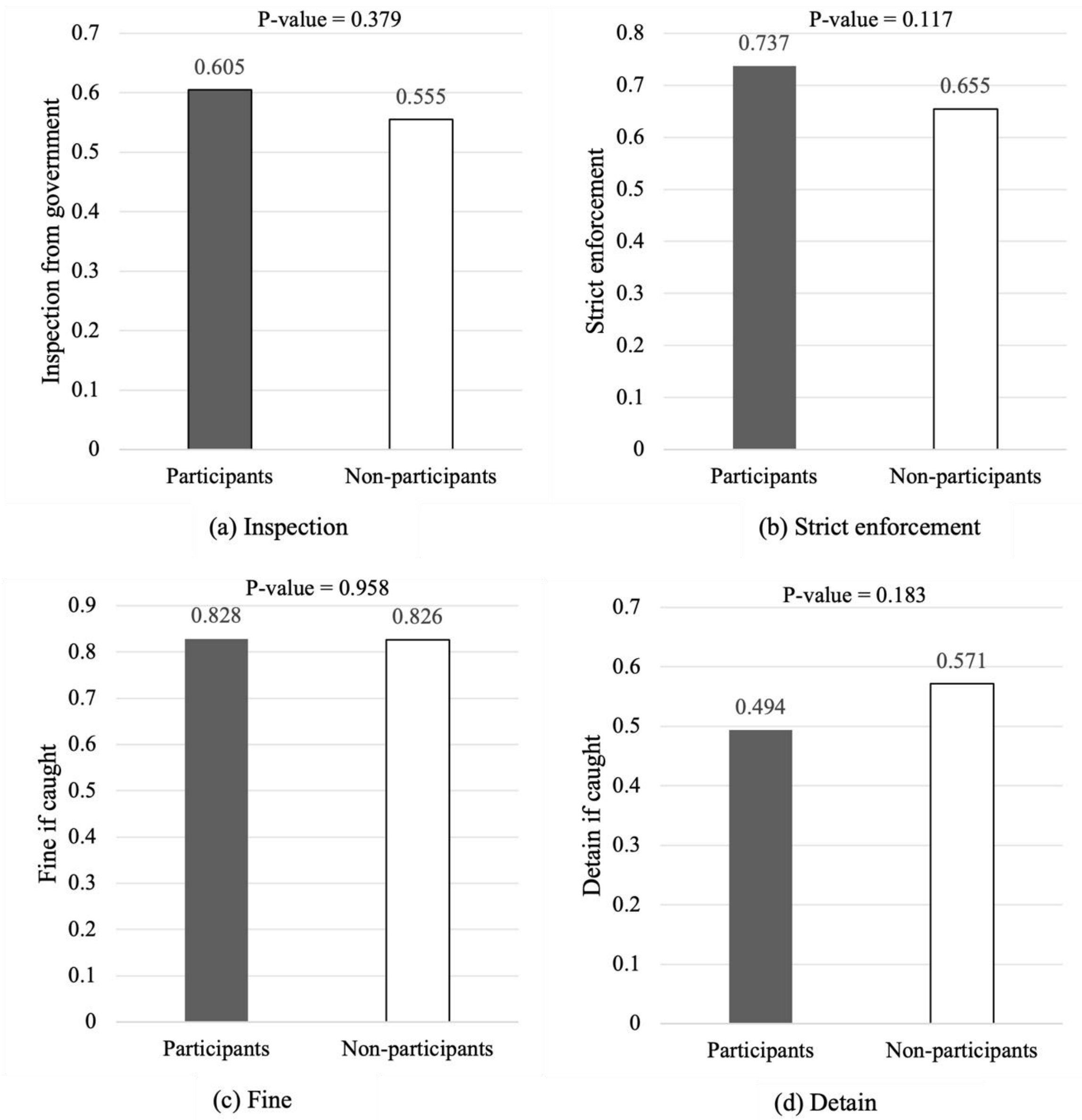



Figure 7 Differences in straw-burning ban stringency between the CWHP participants and non-participants

Notes: The figure presents the response of surveyed households to the following questions: (a) whether there is an inspection from government officials for straw burning, (b) whether the enforcement of the straw burning ban is strict in the village, (c) whether a person will pay a fine if caught burning straw, and (d) whether a person will be detained if caught burning straw.

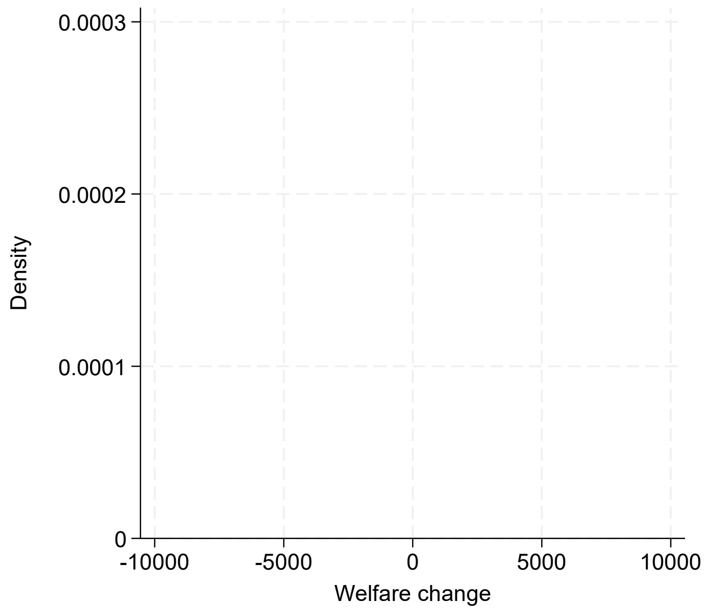



Figure 8 Subjective welfare change

Notes: The welfare change is measured by the difference between the willingness to pay for clean heating and the cost of heating fuel and equipment, plus the heating subsidy.

Table 1 Summary statistics of household characteristics

| Variable | Obs | Mean | Median | SD | Min | Max |
| --- | --- | --- | --- | --- | --- | --- |
| Family size | 1291 | 3.668 | 3 | 1.590 | 1 | 13 |
| House area | 1292 | 124.591 | 110 | 76.476 | 20 | 1000 |
| Heating area | 1292 | 93.343 | 80 | 65.711 | 3 | 1000 |
| Household income | 1292 | 42051 | 35000 | 37158 | 5000 | 200000 |
| Household consumption | 1292 | 27872 | 15000 | 22966 | 5000 | 150000 |
| Income per capita | 1291 | 12395 | 8750 | 11919 | 500 | 150000 |
| Consumption per capita | 1291 | 8282 | 7000 | 7164 | 625 | 75000 |
| Agricultural household | 1292 | 0.548 | 1 | 0.498 | 0 | 1 |
| Education | 1292 | 2.665 | 3 | 0.932 | 1 | 5 |

Notes: Family size is the number of family members. House area and heating area are the house size and heating size in square meters. Household income and household consumption are household-level annual income and consumption. Income and consumption per capita are the annual income per capita and consumption per capita. Agricultural household is a dummy variable that equals 1 if the household head does not have non-agricultural jobs. Education is a categorical variable for education levels of illiteracy, primary school, junior high school, senior high school, and college or higher.

Table 2 Summary statistics for the CWHP participants and non-participants

| Variable | CWHP participants | | Non-participants | | Difference |
|---|---|---|---|---|---|
| | Obs | Mean (SD) | Obs | Mean (SD) | |
| | | | Panel A: Household level | | |
| Fuel cost | 615 | 2789 (1274) | 527 | 1534 (902) | 1255*** |
| Equipment cost | 681 | 2822 (4151) | 611 | 1191 (1724) | 1631*** |
| Fuel subsidy | 681 | 609 (1192) | 611 | 79 (258) | 530*** |
| Equipment subsidy | 681 | 1495 (3978) | 611 | 359 (766) | 1136*** |
| Indoor temperature | 367 | 19.219 (4.753) | 307 | 18.383 (3.733) | 0.836** |
| Indoor PM2.5 | 98 | 57.357 (17.897) | 69 | 62.101 (16.195) | -4.744* |
| | | | Panel B: Village level | | |
| Fire occurrence | 99 | 0.174 (0.161) | 11 | 0.144 (0.178) | 0.030 |
| Fire frequency | 99 | 0.513 (0.626) | 11 | 0.496 (0.639) | 0.017 |
| Fire area | 99 | 0.388 (0.607) | 11 | 0.361 (0.644) | 0.027 |

Notes: Panel A presents summary statistics for household-level heating cost, subsidy, indoor temperature, and indoor PM2.5 concentration. Fuel cost is the heating fuel expenditure (in RMB) of the surveyed households in the 2017 heating season. Fuel subsidy is the subsidy that households receive for their heating fuel in the 2017 heating season. Equipment cost is the expenditure on heating equipment of the surveyed households from 2015 to 2017. Equipment subsidy is the subsidy given to surveyed households for heating equipment from 2015 to 2017. Indoor temperature represents the measured indoor temperature (Celsius degree). Indoor PM2.5 represents the measured indoor PM2.5 concentration (μg/m$^3$). The numbers of observations for indoor temperature and PM2.5 are smaller because only 674 surveyed subjects were measured for indoor temperature, and 167 were measured for indoor PM2.5 concentration. Panel B presents the average fire probability, frequency, and area in the pre-treatment period. Fire occurrence is a dummy variable that equals 1 if the village is detected with straw burning. Fire frequency is the frequency of straw burning. Fire area is the area of straw fires (square meters). ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 3 Intended effects of the CWHP

| Variable \ Method | (1) OLS | (2) PSM | (3) Causal forests | (4) DID | (5) Observations |
|---|---|---|---|---|---|
| Use of clean fuel | 0.653***<br>(0.019) | 0.658***<br>(0.020) | 0.672***<br>(0.018) | / | 1,285 |
| Use of clean equipment | 0.635***<br>(0.020) | 0.639***<br>(0.021) | 0.648***<br>(0.018) | / | 1,285 |
| Indoor temperature | 1.229***<br>(0.326) | 1.392***<br>(0.326) | 1.187***<br>(0.316) | / | 671 |
| Indoor PM2.5 concentration | -3.597**<br>(1.659) | -3.804**<br>(1.748) | -3.676*<br>(2.185) | / | 166 |
| Outdoor PM2.5 concentration | / | / | / | -1.064*<br>(0.599) | 7,920 |

Notes: The dependent variables are a dummy variable that equals 1 if the subject uses clean heating fuel (electricity and natural gas), a dummy variable that equals 1 if the subject uses clean heating equipment (such as air conditioners and gas heaters), the measured indoor temperature (Celsius degree), the measured indoor PM2.5 concentration ($\mu g/m^3$), and the monthly outdoor PM2.5 concentration in a village ($\mu g/m^3$). Control variables for the first four dependent variables include the household head's gender, age, job type (dummy variables), education level (dummy variables), family size, house size, per-person heating size, and household annual income (dummy variables). The number of observations for the use of clean fuel and equipment (1,285) is slightly smaller than that of Table 1 (1,291 or 1,292) because of missing values in the control variables. Control variables for the outdoor PM2.5 concentration include monthly average temperature, air pressure, wind speed, total precipitation, and sunlight duration, with fixed effects at the village level and province-year-month level. Standard errors for Panel E are clustered at the village level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 4 DID estimations: Effects of the CWHP on straw burning

| | Full sample | | | Straw-burning seasons | | | Non-straw-burning seasons | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| | Probability | Frequency | Area | Probability | Frequency | Area | Probability | Frequency | Area |
| CWHP | 0.041** | 0.063** | 0.051** | 0.110*** | 0.139*** | 0.105*** | -0.036* | -0.022 | -0.011 |
| | (0.019) | (0.025) | (0.024) | (0.028) | (0.040) | (0.035) | (0.018) | (0.018) | (0.020) |
| Village FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Province-Year-Month FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Mean of Dep Var | 0.163 | 0.187 | 0.163 | 0.227 | 0.266 | 0.228 | 0.100 | 0.108 | 0.097 |
| Observations | 7,920 | 7,920 | 7,920 | 3,960 | 3,960 | 3,960 | 3,960 | 3,960 | 3,960 |
| R-squared | 0.351 | 0.413 | 0.413 | 0.350 | 0.414 | 0.405 | 0.353 | 0.405 | 0.428 |

Notes: The dependent variables in columns (1), (4), and (7) are the dummy variable that equal 1 if straw burning is detected in the village. The dependent variables in columns (2), (5), and (8) are the logarithm of the frequency of straw burning (plus 1). The dependent variables in columns (3), (6), and (9) are the logarithm of the area of straw fires (plus 1). Control variables include monthly average temperature, air pressure, wind speed, total precipitation, and sunlight duration. Fixed effects are at the village level and province-year-month level. Standard errors are clustered at the village level. Standard errors are clustered at the village level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 5 Effects of the CWHP on prosocial behavior

| | Dictator | | | Public goods | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | OLS | PSM | Causal forests | OLS | PSM | Causal forests |
| CWHP | -0.849*** | -0.678** | -0.858*** | -0.944*** | -0.922*** | -0.877*** |
| | (0.320) | (0.326) | (0.318) | (0.330) | (0.347) | (0.322) |
| Mean (SD) of Dep Var | | 8.361 (5.152) | | | 7.412 (5.132) | |
| Observations | 1,072 | 1,072 | 1,072 | 993 | 993 | 993 |

Notes: The dependent variable in columns (1)–(3) is the cash amount allocated to others. The dependent variable in columns (4)–(6) is the cash amount contributed to public goods. Control variables include the household head's gender, age, job type (dummy variables), education level (dummy variables), family size, house size, per-person heating size, and household annual income (dummy variables). ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 6 Heterogeneous effects of the CWHP by disposable income per capita

| | High income | Low income |
|---|---|---|
| | (1) | (2) |
| Panel A: probability of straw burning | | |
| CWHP | 0.073* | 0.142*** |
| | (0.038) | (0.036) |
| Difference | -0.069* (0.039) | |
| Observations | 2,196 | 2,160 |
| R-squared | 0.439 | 0.326 |
| Panel B: frequency of straw burning | | |
| CWHP | 0.127** | 0.171*** |
| | (0.055) | (0.052) |
| Difference | -0.044 (0.053) | |
| Observations | 2,196 | 2,160 |
| R-squared | 0.482 | 0.418 |
| Panel C: area of straw burning | | |
| CWHP | 0.095** | 0.115** |
| | (0.046) | (0.048) |
| Difference | -0.020 (0.045) | |
| Observations | 2,196 | 2,160 |
| R-squared | 0.496 | 0.382 |
| Panel D: Dictator | | |
| CWHP | -0.607 | -1.363* |
| | (0.481) | (0.713) |
| Difference | 0.756 (1.009) | |
| Treated observations | 348 | 225 |
| Panel E: Public goods | | |
| CWHP | 0.111 | -2.443*** |
| | (0.515) | (0.753) |
| Difference | 2.554*** (0.809) | |
| Treated observations | 322 | 195 |

Notes: The dependent variables in Panels A–C are the probability, frequency, and area of straw burning. The control variables are the same as in Table 4. Fixed effects are at the village level and province-year-month level. Standard errors are clustered at the village level. The dependent variables in D and E are the cash amount allocated to others and the cash amount contributed to public goods. The control variables are the same as in Table 5. Standard errors for coefficient differences are obtained through 1000 bootstrap iterations. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 7 DID estimations by temperature and snow variations

| | Temperature | | Snowfall | |
|---|---|---|---|---|
| | High | Low | High | Low |
| | (1) | (2) | (3) | (4) |
| | Panel A: probability of straw burning | | | |
| CWHP | 0.024 | 0.150*** | 0.137*** | 0.066** |
| | (0.047) | (0.033) | (0.051) | (0.033) |
| Difference | -0.126*** (0.037) | | 0.071* (0.039) | |
| Observations | 2,186 | 2,170 | 2,336 | 2,020 |
| R-squared | 0.363 | 0.404 | 0.376 | 0.400 |
| | Panel B: frequency of straw burning | | | |
| CWHP | 0.030 | 0.207*** | 0.199*** | 0.091** |
| | (0.074) | (0.049) | (0.073) | (0.045) |
| Difference | -0.177*** (0.045) | | 0.108** (0.048) | |
| Observations | 2,186 | 2,170 | 2,336 | 2,020 |
| R-squared | 0.455 | 0.465 | 0.450 | 0.458 |
| | Panel C: area of straw burning | | | |
| CWHP | 0.018 | 0.154*** | 0.164** | 0.058 |
| | (0.071) | (0.036) | (0.071) | (0.040) |
| Difference | -0.136*** (0.041) | | 0.106** (0.043) | |
| Observations | 2,186 | 2,170 | 2,336 | 2,020 |
| R-squared | 0.463 | 0.427 | 0.442 | 0.445 |

Notes: The dependent variables in Panels A–C are the probability, frequency, and area of straw burning. The control variables are the same as in Table 4. Fixed effects are at the village level and province-year-month level. Standard errors for coefficients are clustered at the village level. Standard errors for coefficient differences are obtained through 1000 bootstrap iterations. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 8 Heterogeneous effects of the CWHP by winter heating supply shortage

| | High shortage | | Low shortage |
|---|---|---|---|
| | (1) | | (2) |
| Panel A: probability of straw burning | | | |
| CWHP | 0.176** | | 0.101*** |
| | (0.070) | | (0.030) |
| Difference | | 0.075 (0.058) | |
| Observations | 756 | | 3,600 |
| R-squared | 0.466 | | 0.342 |
| Panel B: frequency of straw burning | | | |
| CWHP | 0.279** | | 0.127*** |
| | (0.101) | | (0.043) |
| Difference | | 0.152** (0.076) | |
| Observations | 756 | | 3,600 |
| R-squared | 0.583 | | 0.404 |
| Panel C: area of straw burning | | | |
| CWHP | 0.208** | | 0.096** |
| | (0.078) | | (0.038) |
| Difference | | 0.112** (0.056) | |
| Observations | 756 | | 3,600 |
| R-squared | 0.534 | | 0.397 |
| Panel D: Dictator | | | |
| CWHP | -4.275*** | | -0.076 |
| | (0.827) | | (0.453) |
| Difference | | -4.199** (1.677) | |
| Treated observations | 113 | | 460 |
| Panel E: Public goods | | | |
| CWHP | -3.673*** | | -0.167 |
| | (1.103) | | (0.459) |
| Difference | | -3.506** (1.468) | |
| Treated observations | 101 | | 416 |

Notes: The dependent variables in Panels A–C are the probability, frequency, and area of straw burning. The control variables are the same as in Table 4. Fixed effects are at the village level and province-year-month level. Standard errors are clustered at the village level. The dependent variables in D and E are the cash amount allocated to others and the cash amount contributed to public goods. The control variables are the same as in Table 5. Standard errors for coefficient differences are obtained through 1000 bootstrap iterations. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 9 Effects of the CWHP on prosocial behavior with PM2.5 as controls

| | Dictator | | | Public goods | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | OLS | PSM | Causal forests | OLS | PSM | Causal forests |
| CWHP | -0.902*** | -0.715** | -0.941*** | -0.977*** | -0.954*** | -0.864*** |
| | (0.330) | (0.336) | (0.317) | (0.346) | (0.362) | (0.333) |
| 35≤PM2.5<65 | -0.203 | -0.104 | Included | -0.069 | -0.074 | Included |
| | (0.369) | (0.371) | | (0.411) | (0.446) | |
| PM2.5≥65 | -0.864** | -0.859* | Included | -0.814* | -0.865* | Included |
| | (0.436) | (0.463) | | (0.418) | (0.445) | |
| Observations | 1,071 | 1,071 | 1,071 | 992 | 992 | 992 |

Notes: The dependent variable in columns (1)–(3) is the cash amount allocated to others. The dependent variable in columns (4)–(6) is the cash amount contributed to public goods. PM2.5 concentration below 35 μg/m$^3$ on the survey day is used as the reference. 35≤PM2.5<65 is a dummy variable for PM2.5 concentration between 35 and 65 μg/m$^3$ . PM2.5≥65 is a dummy variable for PM2.5 concentration above 65 μg/m$^3$. Both variables are incorporated as covariates in the estimations of causal forests, and therefore, no coefficients are reported for them. Other control variables are the same as in Table 5. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 10 DID estimations by prior heating method

| | Probability | | Frequency | | Area | |
|---|---|---|---|---|---|---|
| | Yes | No | Yes | No | Yes | No |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| CWHP | 0.107* | 0.090*** | 0.149* | 0.119*** | 0.101 | 0.085** |
| | (0.056) | (0.028) | (0.085) | (0.039) | (0.070) | (0.034) |
| Difference | 0.017 (0.049) | | 0.030 (0.069) | | 0.016 (0.055) | |
| Observations | 1,044 | 3,312 | 1,044 | 3,312 | 1,044 | 3,312 |
| R-squared | 0.394 | 0.356 | 0.449 | 0.424 | 0.440 | 0.411 |

Notes: The dependent variable in columns (1) and (2), columns (3) and (4), and columns (5) and (6) are the probability, frequency, and area of straw burning, respectively. Columns (1), (3), and (5) present the impact of the CWHP in participant villages that used to burn straw for indoor heating, and the other three columns present the impact in participant villages that did not. The control variables are the same as in Table 4. Fixed effects are at the village level and province-year-month level. Standard errors for coefficients are clustered at the village level. Standard errors for coefficient differences are obtained through 1000 bootstrap iterations. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 11 Effects of the CWHP on heating costs and subsidies

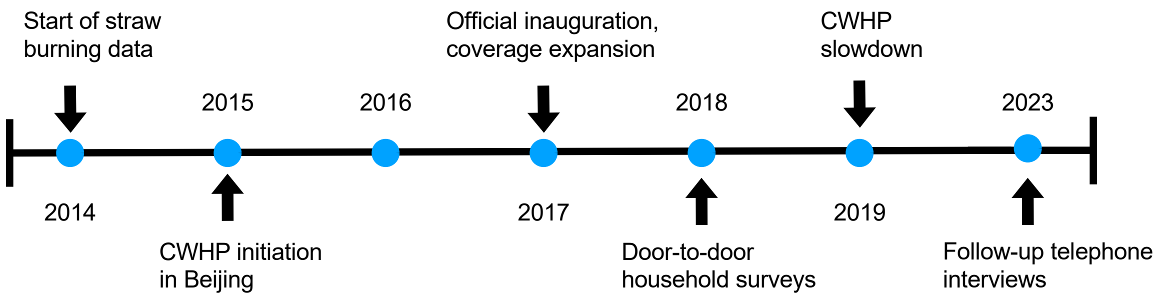
| | Heating fuel | | Heating equipment | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | Fuel costs | Fuel subsidies | Eq. costs | Eq. subsidies |
| CWHP | 1200.025*** | 594.616*** | 1731.747*** | 1398.600*** |
| | (89.644) | (56.062) | (234.208) | (226.664) |
| Observations | 1,136 | 1,285 | 1,285 | 1,285 |

Notes: The dependent variables in columns (1) and (3) are the expenditure on heating fuel and equipment. The dependent variables in columns (2) and (4) are the subsidies for heating fuel and equipment. The causal forests method is used for estimation. Control variables are the same as in Table 5. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

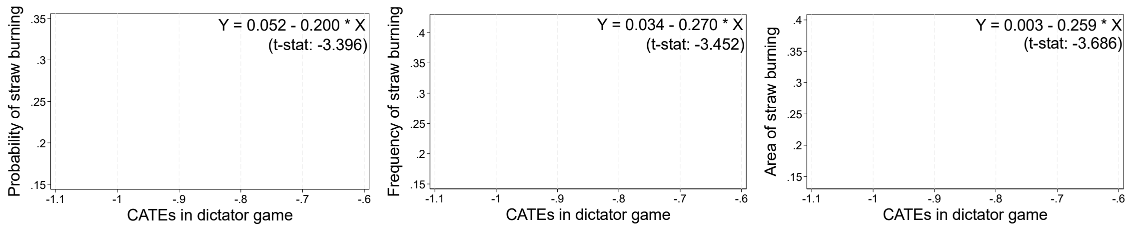
Table 12 Health benefits and cost of full CWHP implementation in Beijing-Tianjin-Hebei rural areas

| | |
|---|---|
| Panel A. Direct effects of the CWHP on PM2.5 | |
| 1. Outdoor PM2.5 change (μg/m$^3$) | -1.36 |
| 2. Indoor PM2.5 change (μg/m$^3$) | -3.60 |
| 3. PM2.5 exposure change for rural residents (μg/m$^3$) | -3.15 |
| 4. PM2.5 exposure change for urban residents (μg/m$^3$) | -1.36 |
| 5. Annual reduction in rural mortality (persons) | 4285 |
| 6. Annual reduction in urban mortality (persons) | 2619 |
| 7. Health benefits of rural residents (million RMB) | 12512 |
| 8. Health benefits of urban residents (million RMB) | 7647 |
| | |
| Panel B. Indirect effects of the CWHP on PM2.5 | |
| 9. Outdoor PM2.5 change (μg/m$^3$) | 0.30 |
| 10. Annual increase in rural mortality (persons) | 397 |
| 11. Annual increase in urban mortality (persons) | 571 |
| 12. Health costs of rural residents (million RMB) | 1159 |
| 13. Health costs of urban residents (million RMB) | 1167 |
| | |
| Panel C. Net health benefits | |
| 14. Net health benefits of rural residents (million RMB) | 11353 |
| 15. Net health benefits of urban residents (million RMB) | 5980 |
| 16. Total net benefits (million RMB) | 17333 |
| | |
| Panel D. Private and government expenditures | |
| 17. Private cost per rural household (RMB) | 649 |
| 18. Government cost per rural household (RMB) | 779 |
| 19. Total private cost of rural residents (million RMB) | 8173 |
| 20. Total government cost (million RMB) | 9809 |
| 21. Total cost (million RMB) | 17982 |
| 22. Benefit-to-cost ratio | 0.96 |

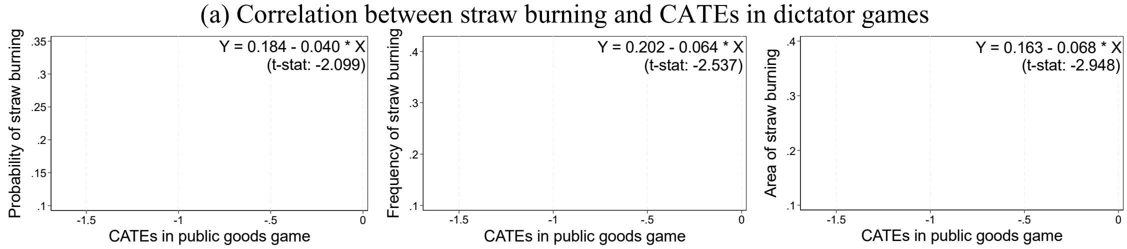
## Appendix Figures and Tables




Appendix Figure A1 Research timeline




(a) Correlation between straw burning and CATEs in dictator games




(b) Correlation between straw burning and CATEs in public goods games

Appendix Figure A2 Correlation between straw burning and the individualized CATEs from the two games

Notes: The variables in the vertical axis in the left, middle, and right panels are the probability, frequency, and area of straw burning during straw-burning seasons. The variable in the horizontal axis in (a) is the individualized conditional average treatments from the dictator game, averaged at the village level. The variable in the horizontal axis in (b) is the individualized conditional average treatment effects from the public goods game, averaged at the village level. The t-statistics are in brackets.

Appendix Table A1 Data description of key variables

| Variable | Source | Structure | Coverage | Analysis |
|---|---|---|---|---|
| Fire occurrence | Satellite | Panel | Treated & Control | DID regression |
| Fire frequency | Satellite | Panel | Treated & Control | DID regression |
| Fire area | Satellite | Panel | Treated & Control | DID regression |
| Outdoor PM2.5 concentration | Satellite | Panel | Treated & Control | DID regression |
| Fuel cost | Survey | Cross section | Treated & Control | OLS<br>PSM<br>Causal forests |
| Fuel subsidy | Survey | Cross section | Treated & Control | OLS<br>PSM<br>Causal forests |
| Equipment cost | Survey | Cross section | Treated & Control | OLS<br>PSM<br>Causal forests |
| Equipment subsidy | Survey | Cross section | Treated & Control | OLS<br>PSM<br>Causal forests |
| Clean heating fuel switch | Survey | Cross section | Treated & Control | OLS<br>PSM<br>Causal forests |
| Clean heating equipment switch | Survey | Cross section | Treated & Control | OLS<br>PSM<br>Causal forests |
| Indoor temperature | Survey | Cross section | Treated & Control | OLS<br>PSM<br>Causal forests |
| Indoor PM2.5 concentration | Survey | Cross section | Treated & Control | OLS<br>PSM<br>Causal forests |
| Money allocated to others | Survey | Cross section | Treated & Control | OLS<br>PSM<br>Causal forests |
| Money contributed to public goods | Survey | Cross section | Treated & Control | OLS<br>PSM<br>Causal forests |
| Emotion change (happier) | Survey | Cross section | Treated only | Summary statistics |
| Emotion change (warmer) | Survey | Cross section | Treated only | Summary statistics |
| Emotion change (more comfortable) | Survey | Cross section | Treated only | Summary statistics |
| Emotion change (safer) | Survey | Cross section | Treated only | Summary statistics |

Appendix Table A2 DID estimations: Alternative radii for village fire measurement

| | 4 km radius | | | 6 km radius | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | Probability | Frequency | Area | Probability | Frequency | Area |
| CWHP | 0.064*** | 0.060** | 0.041** | 0.071** | 0.110*** | 0.084** |
| | (0.020) | (0.025) | (0.019) | (0.032) | (0.042) | (0.034) |
| Village FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Province-Year-Month FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Control | Yes | Yes | Yes | Yes | Yes | Yes |
| Mean of Dep Var | 0.151 | 0.168 | 0.135 | 0.294 | 0.366 | 0.321 |
| Observations | 3,960 | 3,960 | 3,960 | 3,960 | 3,960 | 3,960 |
| R-squared | 0.322 | 0.373 | 0.371 | 0.394 | 0.468 | 0.460 |

Notes: The dependent variables in columns (1) and (4), columns (2) and (5), and columns (3) and (6) are the probability, frequency, and area of straw burning in straw-burning seasons, respectively. Straw burning is measured with the 4 km radius in columns (1)–(3) and with the 6 km radius in columns (4)–(6). Control variables are the same as in Table 4. Fixed effects are at the village level and province-year-month level. Standard errors are clustered at the village level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix Table A3 Pre-trend tests for DID estimations on straw burning

| | Full sample | | | Straw-burning seasons | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | Probability | Frequency | Area | Probability | Frequency | Area |
| $CWHP_{-4}$ | -0.005 | 0.013 | 0.011 | -0.066 | -0.007 | 0.006 |
| | (0.044) | (0.054) | (0.046) | (0.071) | (0.098) | (0.081) |
| $CWHP_{-3}$ | -0.048 | -0.030 | -0.023 | -0.130* | -0.096 | -0.067 |
| | (0.038) | (0.046) | (0.038) | (0.070) | (0.094) | (0.079) |
| $CWHP_{-2}$ | -0.044 | -0.050 | -0.040 | -0.155** | -0.145 | -0.100 |
| | (0.039) | (0.047) | (0.039) | (0.071) | (0.102) | (0.085) |
| $CWHP_{-1}$ | -0.034 | -0.025 | -0.032 | -0.116 | -0.085 | -0.081 |
| | (0.043) | (0.049) | (0.040) | (0.082) | (0.113) | (0.093) |
| $CWHP_{0}$ | 0.050*** | 0.066*** | 0.049*** | 0.134*** | 0.150*** | 0.109*** |
| | (0.016) | (0.021) | (0.019) | (0.026) | (0.038) | (0.030) |
| $CWHP_{+1}$ | 0.065 | 0.100* | 0.083 | 0.174*** | 0.249*** | 0.201*** |
| | (0.041) | (0.056) | (0.059) | (0.049) | (0.072) | (0.072) |
| $CWHP_{+2}$ | 0.111** | 0.108 | 0.106 | 0.161** | 0.154 | 0.142 |
| | (0.055) | (0.080) | (0.086) | (0.071) | (0.101) | (0.110) |
| $CWHP_{+3}$ | 0.132** | 0.155*** | 0.146*** | 0.301*** | 0.342*** | 0.290*** |
| | (0.052) | (0.048) | (0.041) | (0.099) | (0.097) | (0.072) |
| $CWHP_{+4}$ | 0.167*** | 0.195*** | 0.134*** | 0.536*** | 0.597*** | 0.399*** |
| | (0.017) | (0.015) | (0.009) | (0.031) | (0.032) | (0.021) |
| Observations | 7,830 | 7,830 | 7,830 | 3,870 | 3,870 | 3,870 |

Notes: The dependent variables in columns (1) and (4), columns (2) and (5), and columns (3) and (6) are the probability, frequency, and area of straw burning, respectively. Control variables are the same as in Table 4. Fixed effects are at the village-month level and province-year-month level. Standard errors are clustered at the village level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix Table A4 Effects of the CWHP on straw burning by month

| | (1)<br>Probability | (2)<br>Frequency | (3)<br>Area |
|---|---|---|---|
| January | -0.027 | -0.039 | -0.031 |
| | (0.036) | (0.049) | (0.043) |
| February | 0.040 | 0.068 | 0.039 |
| | (0.033) | (0.053) | (0.042) |
| March | 0.135*** | 0.191*** | 0.132*** |
| | (0.039) | (0.057) | (0.048) |
| April | 0.129*** | 0.159*** | 0.106*** |
| | (0.035) | (0.039) | (0.036) |
| May | 0.038 | 0.041* | 0.020 |
| | (0.024) | (0.021) | (0.020) |
| June | -0.010 | -0.005 | -0.007 |
| | (0.022) | (0.018) | (0.016) |
| July | -0.034 | -0.021 | -0.018 |
| | (0.021) | (0.018) | (0.015) |
| August | -0.040 | -0.004 | 0.009 |
| | (0.029) | (0.023) | (0.029) |
| September | 0.008 | 0.053* | 0.068* |
| | (0.029) | (0.030) | (0.037) |
| October | 0.078*** | 0.109*** | 0.100*** |
| | (0.029) | (0.035) | (0.037) |
| November | 0.113*** | 0.135*** | 0.110** |
| | (0.036) | (0.048) | (0.043) |
| December | 0.044 | 0.043 | 0.039 |
| | (0.037) | (0.060) | (0.055) |

Notes: The dependent variables are the probability, frequency, and area of straw burning, respectively. Control variables are the same as in Table 4. Fixed effects are at the village level and province-year-month level. Standard errors are clustered at the village level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix Table A5 Heating outage and emotion changes

| Emotion change | Heating shortage | | | No shortage | | | Difference |
|---|---|---|---|---|---|---|---|
| | N | Mean | Std | N | Mean | Std | |
| Feeling happier | 82 | 0.402 | 0.493 | 588 | 0.616 | 0.487 | -0.213*** |
| Feeling warmer | 82 | 0.598 | 0.493 | 589 | 0.696 | 0.460 | -0.099* |
| Feeling more comfortable | 82 | 0.622 | 0.488 | 588 | 0.738 | 0.440 | -0.116** |
| Feeling safer | 82 | 0.537 | 0.502 | 586 | 0.667 | 0.472 | -0.131** |

Notes: Emotion changes are reported by the CWHP participants. They were asked if they felt happier, warmer, more comfortable, and safer after switching to clean heating with electricity or gas.

Appendix Table A6 Robustness checks for the heterogenous effects of the CWHP on prosocial behavior

| | Disposable income | | Heating supply | |
|---|---|---|---|---|
| | High | Low | High shortage | Low shortage |
| | (1) | (2) | (3) | (4) |
| | Panel A: OLS, Dictator | | | |
| CWHP | -0.664* | -1.023** | -2.815*** | -0.443 |
| | (0.351) | (0.496) | (0.591) | (0.333) |
| Difference | 0.359 (0.473) | | -2.372*** (0.572) | |
| Observations | 847 | 724 | 594 | 977 |
| R-squared | 0.024 | 0.038 | 0.076 | 0.022 |
| | Panel B: OLS, Public goods | | | |
| CWHP | -0.513 | -1.768*** | -2.563*** | -0.625* |
| | (0.367) | (0.497) | (0.732) | (0.333) |
| Difference | 1.255*** (0.450) | | -1.938*** (0.709) | |
| Observations | 798 | 671 | 561 | 908 |
| R-squared | 0.043 | 0.088 | 0.073 | 0.064 |
| | Panel C: PSM, Dictator | | | |
| CWHP | -0.589 | -0.706 | -2.737*** | -0.239 |
| | (0.364) | (0.487) | (0.612) | (0.339) |
| Difference | 0.117 (0.471) | | -2.498*** (0.575) | |
| Observations | 847 | 724 | 594 | 977 |
| R-squared | 0.029 | 0.040 | 0.080 | 0.026 |
| | Panel D: PSM, Public goods | | | |
| CWHP | -0.603 | -1.568*** | -2.567*** | -0.578 |
| | (0.387) | (0.517) | (0.769) | (0.351) |
| Difference | 0.965** (0.448) | | -1.989*** (0.721) | |
| Observations | 798 | 671 | 561 | 908 |
| R-squared | 0.043 | 0.087 | 0.076 | 0.064 |

Notes: The dependent variable in Panels A and C is the cash amount allocated to others. The dependent variable in Panels B and D is the cash amount contributed to public goods. Panels A and B are estimated by OLS, and Panels C and D by PSM. Control variables are the same as in Table 5. Standard errors for coefficient differences are obtained through 1000 bootstrap iterations. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix Table A7 DID estimations by disposable income per capita for the full sample

| | High income | Low income |
|---|---|---|
| | (1) | (2) |
| Panel A: probability of straw burning | | |
| CWHP | 0.004 | 0.064** |
| | (0.027) | (0.025) |
| Difference | -0.060*** (0.023) | |
| Observations | 4,392 | 4,320 |
| R-squared | 0.440 | 0.323 |
| Panel B: frequency of straw burning | | |
| CWHP | 0.033 | 0.084** |
| | (0.036) | (0.033) |
| Difference | -0.051* (0.030) | |
| Observations | 4,392 | 4,320 |
| R-squared | 0.487 | 0.401 |
| Panel C: area of straw burning | | |
| CWHP | 0.019 | 0.062* |
| | (0.035) | (0.033) |
| Difference | -0.043 (0.027) | |
| Observations | 4,392 | 4,320 |
| R-squared | 0.496 | 0.378 |

Notes: The dependent variables in Panels A–C are the probability, frequency, and area of straw burning. Control variables are the same as in Table 4. Fixed effects are at the village level and province-year-month level. Standard errors for coefficients are clustered at the village level. Standard errors for coefficient differences are obtained through 1000 bootstrap iterations. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix Table A8 DID estimations by temperature and snow variations for the full sample

| | Temperature | | Snowfall | |
|---|---|---|---|---|
| | High | Low | High | Low |
| | (1) | (2) | (3) | (4) |
| | Panel A: probability of straw burning | | | |
| CWHP | 0.022 | 0.061*** | 0.063 | 0.018 |
| | (0.044) | (0.021) | (0.043) | (0.022) |
| Difference | -0.039 (0.025) | | 0.045* (0.025) | |
| Observations | 4,466 | 4,246 | 4,664 | 4,048 |
| R-squared | 0.363 | 0.382 | 0.376 | 0.389 |
| | Panel B: frequency of straw burning | | | |
| CWHP | 0.022 | 0.104*** | 0.091 | 0.047 |
| | (0.060) | (0.028) | (0.057) | (0.029) |
| Difference | -0.082*** (0.030) | | 0.044 (0.031) | |
| Observations | 4,466 | 4,246 | 4,664 | 4,048 |
| R-squared | 0.440 | 0.440 | 0.440 | 0.443 |
| | Panel C: area of straw burning | | | |
| CWHP | 0.013 | 0.084*** | 0.070 | 0.041 |
| | (0.063) | (0.024) | (0.060) | (0.029) |
| Difference | -0.071** (0.027) | | 0.029 (0.028) | |
| Observations | 4,466 | 4,246 | 4,664 | 4,048 |
| R-squared | 0.456 | 0.418 | 0.438 | 0.443 |

Notes: The dependent variables in Panels A–C are the probability, frequency, and area of straw burning. Control variables are the same as in Table 4. Fixed effects are at the village level and province-year-month level. Standard errors for coefficients are clustered at the village level. Standard errors for coefficient differences are obtained through 1000 bootstrap iterations. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix Table A9 DID estimations by heating supply shortage for the full sample

| | High shortage | Low shortage |
|---|---|---|
| | (1) | (2) |
| Panel A: probability of straw burning | | |
| CWHP | 0.066 | 0.040* |
| | (0.044) | (0.020) |
| Difference | 0.026 (0.026) | |
| Observations | 1,512 | 7,200 |
| R-squared | 0.441 | 0.351 |
| Panel B: frequency of straw burning | | |
| CWHP | 0.113* | 0.059** |
| | (0.060) | (0.027) |
| Difference | 0.054 (0.036) | |
| Observations | 1,512 | 7,200 |
| R-squared | 0.546 | 0.408 |
| Panel C: area of straw burning | | |
| CWHP | 0.079 | 0.046* |
| | (0.058) | (0.026) |
| Difference | 0.033 (0.029) | |
| Observations | 1,512 | 7,200 |
| R-squared | 0.507 | 0.409 |

Notes: The dependent variables in Panels A–C are the probability, frequency, and area of straw burning. Control variables are the same as in Table 4. Fixed effects are at the village level and province-year-month level. Standard errors for coefficients are clustered at the village level. Standard errors for coefficient differences are obtained through 1000 bootstrap iterations. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix Table A10 Effects of the CWHP on prosocial behavior on days with “Good” air quality

| | Dictator | | | Public goods | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | OLS | PSM | Causal forests | OLS | PSM | Causal forests |
| CWHP | -3.310*** | -3.073*** | -2.794*** | -3.398*** | -3.628*** | -3.048*** |
| | (0.878) | (0.853) | (0.955) | (0.961) | (0.959) | (0.965) |
| Mean of Dep Var | 7.961 | 7.961 | 7.961 | 6.580 | 6.580 | 6.604 |
| Observations | 152 | 152 | 152 | 143 | 143 | 143 |

Notes: The dependent variable in columns (1)–(3) is the cash amount allocated to others. The dependent variable in columns (4)–(6) is the cash amount contributed to public goods. Control variables are the same as in Table 5. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix Table A11 DID estimations by prior heating method for the full sample

| | Probability | | Frequency | | Area | |
|---|---|---|---|---|---|---|
| | Yes | No | Yes | No | Yes | No |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| CWHP | 0.040 | 0.030 | 0.066 | 0.051** | 0.045 | 0.038 |
| | (0.038) | (0.019) | (0.055) | (0.024) | (0.054) | (0.025) |
| Difference | 0.010 (0.028) | | 0.014 (0.040) | | 0.007 (0.036) | |
| Observations | 2,088 | 6,624 | 2,088 | 6,624 | 2,088 | 6,624 |
| R-squared | 0.401 | 0.354 | 0.447 | 0.419 | 0.443 | 0.416 |

Notes: The dependent variable in columns (1) and (2), columns (3) and (4), and columns (5) and (6) are the probability, frequency, and area of straw burning, respectively. Columns (1), (3), and (5) present the impact of the CWHP in participant villages that used to burn straw for indoor heating, and the other three columns present the impact in participant villages that did not. Control variables are the same as in Table 4. Fixed effects are at the village level and province-year-month level. Standard errors for coefficients are clustered at the village level. Standard errors for coefficient differences are obtained through 1000 bootstrap iterations. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix Table A12 Effects of straw burning on PM2.5 concentration

| | PM2.5 concentration | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Fire incidence | 0.227*** | 0.129*** | 0.103*** |
| | (0.016) | (0.013) | (0.013) |
| County FE | Yes | Yes | Yes |
| Year-Month FE | Yes | | |
| Region-Year-Month FE | | Yes | Yes |
| Control | | | Yes |
| Mean of Dep Var | 40.529 | 40.529 | 40.525 |
| Observations | 204,984 | 204,984 | 204,712 |
| R-squared | 0.759 | 0.821 | 0.830 |

Notes: The dependent variable is the monthly PM2.5 concentration at the county level. Fire incidence represents the number of trajectories passing through the counties. Control variables include the monthly average temperature, air pressure, wind speed, total precipitation, and sunlight duration. Fixed effects are at the county level and year-month level (column 1) or the region-year-month level (columns 2 and 3). Regions are portioned as North China, Northeast China, East China, Central China, South China, and Northwest China. Standard errors are clustered at the county level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.