Research Report 235, *How Do Household Energy Transitions Work?*, by J. Baumgartner and S. Harper et al.

INTRODUCTION

Protecting environmental quality and human health through air pollution actions or interventions typically incurs an economic cost. It is therefore important to understand whether environmental policies result in the intended improvements. This area of study, known as environmental accountability research (sometimes also referred to as intervention studies), evaluates the extent to which air quality actions or interventions have reduced air pollutant emissions and concentrations and improved public health. A major challenge in this research field is isolating changes that can be attributed to the actions in question from improvements that might be due to other actions or long-term trends. This challenge is of particular concern for actions that target numerous pollutant sources over large geographic regions and are implemented over several years.

Over the past two decades, HEI has played a prominent role as a leader in accountability research, contributing to research design, funding, study oversight, and evaluation of such research (see Preface). Through a series of Requests for Applications (RFAs),* HEI has now funded many studies that have assessed a wide variety of actions that have targeted both point and mobile sources of air pollution. Earlier studies tended to focus on local-level actions that were implemented over a relatively short time frame. HEI later solicited research that evaluated actions with a larger geographical scope or that were implemented over longer time frames.

In its 2018 research solicitation, *RFA 18-1* Assessing Improved Air Quality and Health from National, Regional, and Local Air Quality Actions, HEI aimed to fund empirical studies to assess the health effects of air quality actions or to develop methods required for, and specifically suited to, conducting such research and make them accessible and available to other researchers. Areas of interest included national- or regional-scale actions implemented over multiple years, local actions targeted at improving air quality in urban areas with well-documented air quality problems, and programs to improve air quality around major ports and transportation hubs and corridors.

In response, Baumgartner, Harper, and colleagues proposed to assess the effects of a household clean heating policy that mandated and subsidized 3,700 villages in the Beijing region of China to switch from highly polluting residential heaters fueled by coal to efficient electric- or gas-powered heat pumps between 2017 and 2021. They also proposed to assess potential behavioral, environmental, and chemical mechanisms that might explain how the policy affects health outcomes. The HEI Research Committee recommended the proposal for funding because it applies a strong study design to address an important policy and because it builds on an ongoing study.

This Commentary provides the HEI Review Committee's independent evaluation of the study. It is intended to aid the sponsors of HEI and the public by highlighting both the strengths and limitations of the study and by placing the results presented in the Investigators' Report into a broader scientific and regulatory context.

SCIENTIFIC AND REGULATORY BACKGROUND

EVALUATING THE EFFECTIVENESS OF HOUSEHOLD AIR POLLUTION INTERVENTIONS

It is well established that household air pollution from the combustion of solid fuels (e.g., coal and biomass) contributes to cardiorespiratory diseases and other adverse health effects (Chowdhury et al. 2023; HEI 2025; HEI Household Air Pollution Working Group 2018; Lai et al. 2024; WHO 2014). However, the extent to which household air pollution levels and health are improved by relevant interventions (e.g., replacing solid fuel stoves with more efficient electric or gas heaters) is unclear (HEI Household Air Pollution Working Group 2018; Lai et al. 2024). Given that as recently as 2022, about a quarter of the world's population (1.8 billion) relied on high-emission fuel sources for heating, there is a pressing need to understand how to effectively reduce exposures and improve health (IEA et al. 2025).

The effectiveness of household air pollution interventions is typically evaluated by intervention studies that compare measurements of air quality and health before and after an intervention, using randomized controlled trials or observational study designs. Most intervention studies of household

Dr. Jill Baumgartner's and Dr. Sam Harper's 3½-year study, "How Do Household Energy Transitions Work?" began in May 2020. Total expenditures were $1,094,118. The draft Investigators' Report was received for review in May 2024. A revised report, received in October 2024, was accepted for publication in November 2024. During the review process, the HEI Review Committee and the investigators had the opportunity to exchange comments and clarify issues in the Investigators' Report and its Commentary. Note: Review Committee members Frank Kelly, Jennifer Peel, and John Volckens did not partake in the review of the report due to conflicts of interest.

* A list of abbreviations and other terms appears at the end of this volume.

air pollution to date have focused on more efficient biomass cookstoves or changes to fuels, including liquified petroleum gas for cooking. These studies often measure fine particulate matter ($PM_{2.5}$; sometimes supplemented with measurements of black carbon or other major chemical constituents of $PM_{2.5}$) or carbon monoxide, and sometimes measure nitrogen dioxide (HEI Household Air Pollution Working Group 2018; Lai et al. 2024; Lee et al. 2020). The mixed results in the published literature about the effectiveness of these interventions have been attributed to study design choices, such as which health outcomes are assessed, the length of follow-up, and small sample sizes in some studies. Other issues have included that the adaptation and use of the interventions were context dependent and often limited (e.g., because the interventions were hard to use or broke), and that reductions in emissions from the intervention might be small relative to other sources of air pollution (HEI Household Air Pollution Working Group 2018; Lai et al. 2024).

Given the heterogeneous results of cookstove interventions and the more limited information on exposure to household air pollution from heating, studies are needed to provide credible and convincing estimates of the potential health benefits and the costs of interventions intended to accelerate transitions to clean household energy sources (HEI Household Air Pollution Working Group 2018). Such studies should address some of the limitations of the earlier cookstove studies by including other sources of household air pollution (e.g., home heating), evaluating the extent to which the new household energy technology has been adopted, and using prospective study designs to capture important pre-policy information (HEI Household Air Pollution Working Group 2018; Lai et al. 2024). Additionally, evidence on the mechanisms of how interventions produce changes in outcomes — a research question that can be assessed using a statistical approach known as mediation analysis — could be important to evaluate past interventions and design future interventions that are more effective (Keele et al. 2015).

HOUSEHOLD COAL COMBUSTION IN BEIJING

In suburban and rural parts of Beijing, fuel for heating during the cold winter season has historically consisted predominantly of coal and biomass. Households traditionally burned these fuels in standalone heating stoves and traditional kangs (i.e., raised platforms for both heating and cooking). Residential space heating in northern China required more than 200 million tons of coal in 2017 (Dispersed Coal Management Research Group 2023). The Global Burden of Disease — Major Air Pollution Sources (GBD MAPS) Working Group found that coal burning from residential, industrial, and power-generating sources contributed just over one-third of the total ambient wintertime $PM_{2.5}$ across China in 2013 (GBD MAPS Working Group 2016). According to GBD MAPS, emissions from residential coal combustion in Beijing alone were responsible for about 4.4 μg/m^3 of the annual average $PM_{2.5}$ concentrations in Beijing, accounting for 5.9% of the overall contributions to population-weighted $PM_{2.5}$ exposures and about 770 annual deaths related to air pollution in 2013. Separately, it has been estimated that residential coal combustion across the Beijing-Tianjin-Hebei region was responsible for about half of the wintertime $PM_{2.5}$ concentrations in Beijing before implementation of the clean heating policy, with lower contributions in other seasons (Zhang et al. 2013, 2017).

CHINA'S CLEAN HEATING POLICY

To reduce ambient air pollution concentrations, the Chinese government issued an *Air Pollution Prevention and Control Action Plan* in 2013 (China State Council 2013). Among other things, the Plan mandated (and subsidized) that over 1.5 million residents (up to 70% of all households) in rural areas of northern China, including the Beijing region, switch from highly polluting coal heaters to efficient electric or gas-powered heat pumps between 2017 and 2021. The government simultaneously designated coal-restricted areas and offered subsidies for electricity and electric-powered heaters, including electric heat pumps, to replace traditional coal-heating stoves. The clean heating policy included a pilot phase in 2015 and was rolled out on a village-by-village basis based on various factors related to policy priorities and local capacity starting in 2016.

EARLIER STUDIES EVALUATING CHINA'S CLEAN HEATING POLICY

Air quality and health improvements have been reported following the sweeping air pollution control policies that began in 2013, including the transition of cooking and heating fuels in the residential sector (e.g., Ding et al. 2019; Li et al. 2024; Zhang et al. 2019). The causal links between these air pollution regulations, emissions, ambient air pollution, and mortality have been assessed in a study led by Dr. Patrick Kinney with funding from HEI RFA 18-1 (Kinney et al. in press).

Because reducing the use of household coal was a major part of the air pollution control plan, researchers quickly initiated studies to evaluate the effectiveness of the clean heating policy in northern China. In a cross-sectional pilot study conducted in the winter of 2017, the investigators of the current study enrolled 302 households from three Beijing districts. Half of the households were located in three villages that had participated in the first wave of the clean heating policy, and the other half were located in three geographically contiguous villages in the same district that had not yet participated (Barrington-Leigh et al. 2019). They found that clean heating technologies completely replaced coal heaters in two of the villages with the policy, and partially replaced them in one of the villages with the policy. Indoor $PM_{2.5}$ concentrations were 67% lower and indoor temperatures were 1.4°C higher in villages participating in the policy compared with neighboring villages that were not participating. Additionally, in parallel with the current study, the investigators conducted a study that showed reductions in acute myocardial infarction in Beijing townships after roll-out of the clean heating policy (Lee et al. 2024). A separate study conducted by other investigators at the same time evaluated the clean heating policy in a multicity cohort of Chinese adults and found small decreases in the risk of chronic lung diseases (1.1% to 3.0%), but no change in cardiovascular disease risk after a one-year post-policy evaluation period (Wen et al. 2023).

MOTIVATION FOR THE CURRENT STUDY

The current HEI study builds on the ongoing Beijing Household Energy Transitions (BHET) study, which was intended to evaluate the effect of the clean heating policy on health. The observed differences in $PM_{2.5}$ concentrations and indoor wintertime temperatures in their earlier cross-sectional study motivated the investigators to delve deeper with difference-in-differences analyses that could account for stable characteristics that differ between the villages with and without the clean heating policy, as well as for shared trends in air pollution, health, and related factors over time. They therefore identified 50 suburban and rural Beijing villages that were eligible for the clean heating policy, but where the policy had not yet been implemented (**Commentary Figure 1**). The investigators expected about half the villages to implement the policy over the course of the BHET study. In December 2018, they enrolled 1,003 participants (about 20 per village) in 977 different households in the 50 participating villages. Trained staff collected information on socio-demographic characteristics, indoor temperature, outdoor and personal $PM_{2.5}$ exposures, and health starting in winter 2018–2019. They observed early indications of improved wintertime indoor $PM_{2.5}$ concentrations and indoor temperature after implementation of the policy.

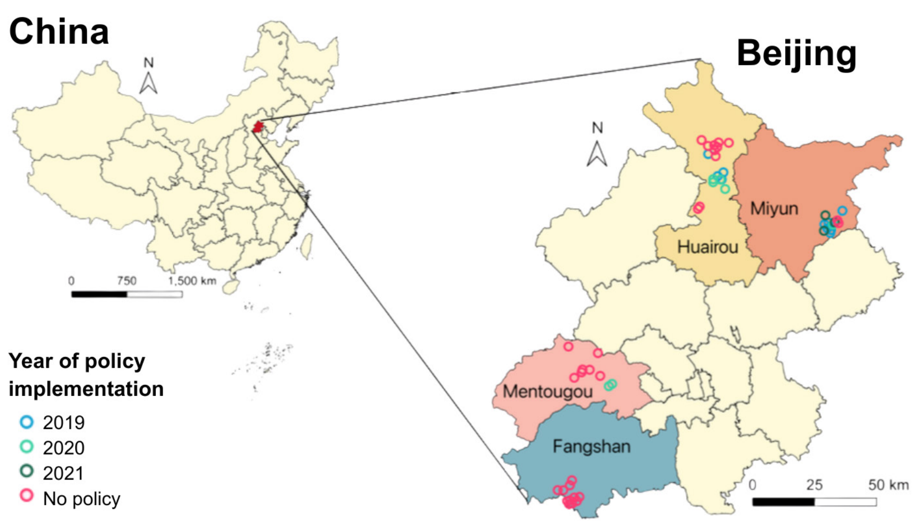
STUDY OBJECTIVES AND APPROACH

Baumgartner, Harper, and colleagues used HEI funding to extend the BHET study to assess the effects of the clean heating policy on outdoor and indoor air quality, personal exposures to air pollution, and cardiovascular and respiratory health. They also evaluated whether any changes in health outcomes observed after the implementation of the policy could be explained (i.e., were mediated) by changes in $PM_{2.5}$ concentrations and indoor temperature. Specific study aims were as follows:

1. Estimate the effect of the policy on outdoor, indoor, and personal $PM_{2.5}$ concentrations, and specifically estimate the source contribution from household coal burning to indoor and outdoor $PM_{2.5}$.
2. Estimate the overall effect of the policy on health, including respiratory symptoms and cardiovascular outcomes.
3. Assess the proportion of the observed effect of the policy on health that can be attributed to changes in $PM_{2.5}$, indoor temperature, and household coal burning.

To address these aims, the investigators measured many parameters related to air quality, temperature, and health in four consecutive winters from 2018 to 2022 (**Commentary Table 1**). They also collected information on household fuel use and self-reported respiratory health through participant questionnaires. To identify the contribution of household coal combustion to the $PM_{2.5}$ exposures, they conducted source apportionment of outdoor and personal measurements of $PM_{2.5}$ chemical composition.

To estimate the effects of the clean heating policy, the investigators conducted difference-in-differences analyses,



Commentary Figure 1. Map of village implementation of the clean heating policy. Each circle represents one recruited village. The colors of the circles indicate the year the clean heating policy was implemented in the villages. (Source: Adapted from Investigators' Report Figure 1.)

Commentary Table 1. Data Collection Across Waves in Four Consecutive Winters[a]

	Wave 1	Wave 2	Wave 3	Wave 4
Winter	2018–2019	2019–2020	2020–2021	2021–2022
Enrolled villages	50	50	50	50
Cumulative villages where the policy had been implemented	0	10	17	20
Household questionnaire	✓	✓		✓
Health measurements (blood pressure and cardiovascular biomarkers)	(✓)	(✓)		(✓)
Outdoor village-level $PM_{2.5}$	✓	✓	✓	✓
Indoor household $PM_{2.5}$		(✓)	(✓)	(✓)
Personal exposure to $PM_{2.5}$ and black carbon	(✓)	(✓)		(✓)
Chemical speciation of outdoor and personal $PM_{2.5}$	(✓)	(✓)		(✓)
Indoor temperature	✓	✓	✓	✓
Outdoor temperature	✓	✓	✓	✓
Confirmation of heating device use		(✓)	(✓)	(✓)

[a] ✓ indicates that the data were collected for the full population, and (✓) indicates that the data were collected for a subsample of the population as part of the study design. Numbers of measurements for each parameter and wave are included in the Investigators' Report. HEI funded Waves 3 and 4, and parts of Wave 2 (measurement of indoor $PM_{2.5}$, chemical speciation of outdoor and personal $PM_{2.5}$, and confirmation of heating device use using sensors).

a statistical approach used to compare changes over time in villages that had implemented the policy versus those that might implement it in the future. For health outcomes where they detected changes associated with the policy, the investigators then used mediation analysis to assess what proportion of the observed changes in health were caused by changes in air quality or temperature inside the home.

The investigators were forced to make some changes during the study because of the COVID-19 pandemic. Among the main changes were a partial campaign in winter 2020–2021 and the addition of a fourth full winter data collection campaign (including surveys, personal exposure, and health measurements) in winter 2021–2022.

SUMMARY OF METHODS AND STUDY DESIGN

STUDY AREA AND POPULATION

Local guides developed rosters of households that could be recruited into the study and recruited participants from about 20 randomly selected households in each of the 50 BHET villages for each study wave. About three-quarters of the recruited households remained in the study across all waves. Forty-three percent of individuals from the recruited households contributed observations across all three waves, and 31% of individuals participated in two waves. Individuals were eligible to participate if they were over 40 years old, lived in the study villages, were not planning to move out of the village in the next year, and were not currently receiving immunotherapy or corticosteroid treatments.

ENVIRONMENTAL EXPOSURE MEASUREMENTS

The study team performed detailed measurements for outdoor, indoor, and personal exposure to $PM_{2.5}$, along with outdoor and indoor temperatures. They measured $PM_{2.5}$ using both real-time sensors (one measurement per minute; Plantower) and gravimetric analysis of filters collected with Ultrasonic Personal Aerosol Samplers or Personal Exposure Monitors. All air pollution measurements followed established standard operating procedures with detailed quality assurance.

Community-level outdoor measurements of $PM_{2.5}$ were made using co-located sensors and filter samplers at one location near the center of each village and one or two locations at least 500 m away from the first measurement location. All outdoor measurement locations were distant from visible sources of $PM_{2.5}$. Co-located weekly filter samples were used to calibrate sensors and to measure the chemical composition of particles on the filters (i.e., elemental carbon, organic carbon, individual elements, and water-soluble ions).

Indoor measurements of $PM_{2.5}$ were made in six randomly selected households from each village starting with Wave 2. Sensors were deployed between late November and mid-January in the rooms where participants reported spending most of their time. The sensors made continuous measurements until they were collected in late April. Filter-based samplers were co-located with a subset of the indoor sensors during the first 24 hours of sensor-based measurements to aid in sensor calibration.

Personal exposures to $PM_{2.5}$ and black carbon were assessed for about 500 participants in each wave (10 randomly selected study participants in each village) who each carried a filter sampler for 24 hours. Outdoor temperature and relative humidity were obtained from the Beijing meteorological network. Additional meteorological data (e.g., boundary layer height) were obtained from the fifth generation of the European Center for Medium-Range Weather Forecasting Forecasts global climate reanalysis dataset (ERA5).

The investigators also measured indoor air temperature in a centrally located room in each house for the 5 minutes before blood pressure measurements and, for 75% of households, every 125 minutes for up to 4 months in the winter and early spring to check which heating and cooking devices were in use.

$PM_{2.5}$ SOURCE APPORTIONMENT

The investigators used a source apportionment approach to determine the contribution of household coal burning to the combined set of outdoor and personal $PM_{2.5}$ samples from which both $PM_{2.5}$ mass and chemical components were quantified. The investigators used the United States Environmental Protection Agency's source apportionment model, positive matrix factorization (PMF) 5.0, which has been widely used for air pollution analyses in China. They accounted for differences in wind speed and boundary layer height in the chemical analysis data prior to their use in the PMF model.

They conducted sensitivity analyses using source apportionment models with different measurement subsets and three to six source factors and evaluated the fit of the final PMF model using standard PMF model diagnostics. Factors identified from the PMF analyses were interpreted and named based on the field observations, investigators' knowledge of local sources, and relevant previous studies.

HEALTH OUTCOMES

The investigators focused on various cardiovascular outcomes and respiratory symptoms that have well-established links to $PM_{2.5}$. Blood pressure (brachial systolic, brachial diastolic, central systolic, and central diastolic) and chronic airway symptoms were assessed at the participants' homes at the time of the questionnaire visits (see below). Blood pressure was measured by trained staff using factory-calibrated devices with standard protocols and following detailed quality control procedures that included an appropriately sized cuff, correct positioning of the arm, and having both feet on the ground. Self-reported airway symptoms were assessed using a questionnaire (details below) with standard validated questions about chronic airway symptoms, including cough, phlegm, wheeze, and tightness in the chest. Additionally, the study team measured airway inflammation (as the fractional concentration of exhaled nitric oxide [FeNO]) in about one-quarter of participants.

Within 1 month of the researchers' visit, participants visited a village clinic where the study team collected fasting blood samples. Samples were tested for glucose, a complete lipid profile, and biomarkers of systemic inflammation or oxidative stress (C-reactive protein, interleukin-6, tumor necrosis factor alpha, and malondialdehyde) that are associated with both exposure to air pollution and the development of cardiovascular disease and events. Body weight, height, and waist circumference were also measured.

QUESTIONNAIRES

In addition to the respiratory symptoms described above, the study team collected information on households and individual participants by administering structured questionnaires in Mandarin Chinese. They gathered information on the types of fuels and stoves present, patterns of fuel and stove use, and the amount of fuel used for space heating in winter. They also collected supporting information on socio-demographic factors, house structure, and individual lifestyle factors (e.g., smoking). One representative from each village completed a survey about the policy implementation and the level of village interest in the policy, as well as information about any other policies being implemented simultaneously.

STATISTICAL ANALYSES

Measuring Effects of the Clean Heating Policy

To estimate the effects of the policy, the investigators conducted difference-in-differences analyses using multivariable "extended" two-way fixed effects models to account for the staggered roll-out of the policy, as described by Wooldridge (2021). This design compared changes in outcomes among villages that had the policy implemented in the same year, different years, or not at all. The villages where the policy was not implemented by the end of the study period provided an estimate of what changes in the health and exposure outcomes would have been expected to happen in the other villages if they had not implemented the policy (a "counterfactual"). By comparing changes in outcomes among villages with and without the policy, the difference-in-differences approach accounts for both time-invariant characteristics of each village as well as trends in air quality and health over time that are unrelated to the policy. As a result, these analyses produce estimates of the total effect of the policy on air quality and health (**Commentary Figure 2, top**).

Comparing Direct and Indirect Effects of the Clean Heating Policy

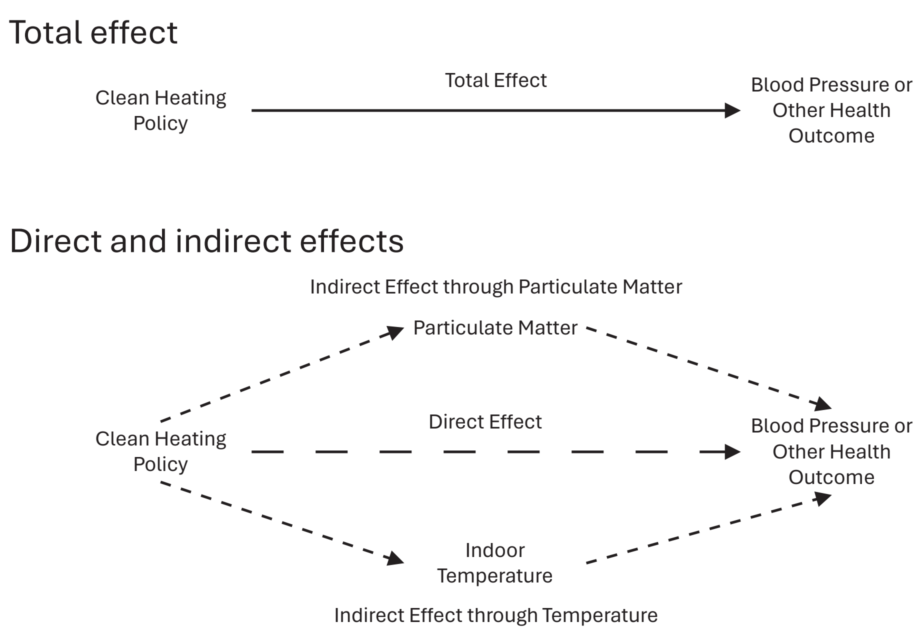
An important part of the study was to evaluate whether the clean heating policy influenced health directly, or indirectly through its effects on other factors (known as mediators), which then affect health. To do this, the investigators applied an approach for mediation analysis in the causal inference framework (Hernán and Robins 2020; Keele et al. 2015).

For those outcomes where they found a total effect, the investigators estimated the contributions of potential mediators of those associations (**Commentary Figure 2, bottom**). The potential mediators they considered were indoor or personal $PM_{2.5}$ concentrations and indoor temperature. The investigators tested the influence of these potential mediators by repeating the initial models, but with the potential mediator being tested held constant for the entire study population. This analysis produced the "direct effect" of the policy to describe the relationship between the policy and health outcomes independent of the potential mediators. Finding that the estimated direct effect was smaller than the total effect would suggest that the effect of the clean heating policy on health was at least partly because of its effect on the mediating factor being tested (in other words, that there was mediation). Conversely, finding that the direct and total effects were similar would suggest that the health benefits of the clean heating policy were related to factors other than changes in measured indoor $PM_{2.5}$ concentrations or measured temperatures, such as residents going outside less frequently to refuel the heaters or more even heating within the homes.

Adjustment for Potential Confounders and Imputation of Missing Data

The investigators identified potential confounders using directed acyclic graphs (DAGs) and included different sets of adjustments for potential confounding in the models for each outcome. To obtain valid results, they adjusted the total effect and direct effect models for confounding by factors such as participant age, sex, exposure to tobacco smoke, height, and weight. They also considered potential confounders related to the size of the villages and homes, and related to other factors that might vary over time and differ among villages.

Several key covariates for blood pressure models (e.g., waist circumference, height, and weight) were missing from 15% of participants in Waves 1 and 2 because those individuals did not participate in the clinic visits. Also, indoor $PM_{2.5}$ concentrations were measured at only about one-fifth of the homes in each wave by design, so 80% of the participants had no matching indoor $PM_{2.5}$ concentrations. There was little overlap between participants with blood pressure measurements and participants whose homes had been sampled for indoor $PM_{2.5}$. The investigators therefore used multiple



Commentary Figure 2. Simplified diagrams of the main models in the study. Top: The total effect of the policy, including all mechanisms by which the policy might work. Bottom: Separating indirect effects through changes in two potential mediators (particulate matter and indoor temperature; dotted lines) from the direct effect of the policy (dashed line), which includes all other possible mechanisms.

imputation with chained equations to replace the missing covariates and indoor $PM_{2.5}$ concentrations so that they could make full use of the blood pressure data.

SUMMARY OF KEY RESULTS

STUDY POPULATION

The investigators enrolled 1,438 participants over the study period (**Box 1: A Comprehensive Study**). Demographic characteristics, health, $PM_{2.5}$ concentrations, and indoor temperature were similar at baseline for participants in villages where the policy was implemented by the end of the study and where it was not. The average age of study participants at baseline (i.e., in Wave 1) was 60 years. About 60% of participants were female and about 80% were current or former smokers or lived with someone who smoked. On average, the participants had slightly elevated blood pressure (130 mm Hg systolic blood pressure and 82 mm Hg diastolic blood pressure), and just over half of the participants reported at least one respiratory problem, usually shortness of breath or phlegm. There was little missing data, and the investigators did not find a relationship between the implementation of the policy and missing data.

> **Box 1: A Comprehensive Study**
>
> **Large Sample Size**
>
> - 1,438 participants from 1,236 households in 50 villages visited over four winters
>
> **Detailed Characterization of Environmental Exposures**
>
> - 11,174 outdoor, 399 indoor, and 1,270 personal sensor-based 24-hour $PM_{2.5}$ measurements
> - 968 outdoor, 288 indoor, and 1,295 personal $PM_{2.5}$ filter samples (a subset of which were also used to quantify chemical constituents of $PM_{2.5}$)
> - 1,161 personal measurements of black carbon on filters
> - 717 outdoor and 1,158 personal estimates of exposure to “mixed combustion” from source apportionment
> - 2,999 measurements of 5-minute indoor temperature
>
> **Individual Health Outcomes**
>
> - 3,082 measurements of blood pressure
> - 3,076 measurements of self-reported respiratory effects (coughing, phlegm, wheezing attacks, trouble breathing, and chest trouble)
> - 793 measurements of exhaled nitric oxide (FeNO)
> - 1,603 measurements of inflammatory markers (C-reactive protein [CRP], interleukin-6 [IL-6], tumor necrosis factor-alpha [TNF-α]), and an oxidative stress marker (malondialdehyde [MDA]) in blood

POLICY UPTAKE

During the study, villages where the policy was implemented had almost complete uptake and adherence to the policy, as shown by the shift to heat pumps as the primary heating device (**Commentary Figure 3, top**). About one-third of households from treated villages that previously used coal stoves replaced the coal stoves with exclusive heat pump usage, and the rest used a mixture of heat pumps with biomass kangs.

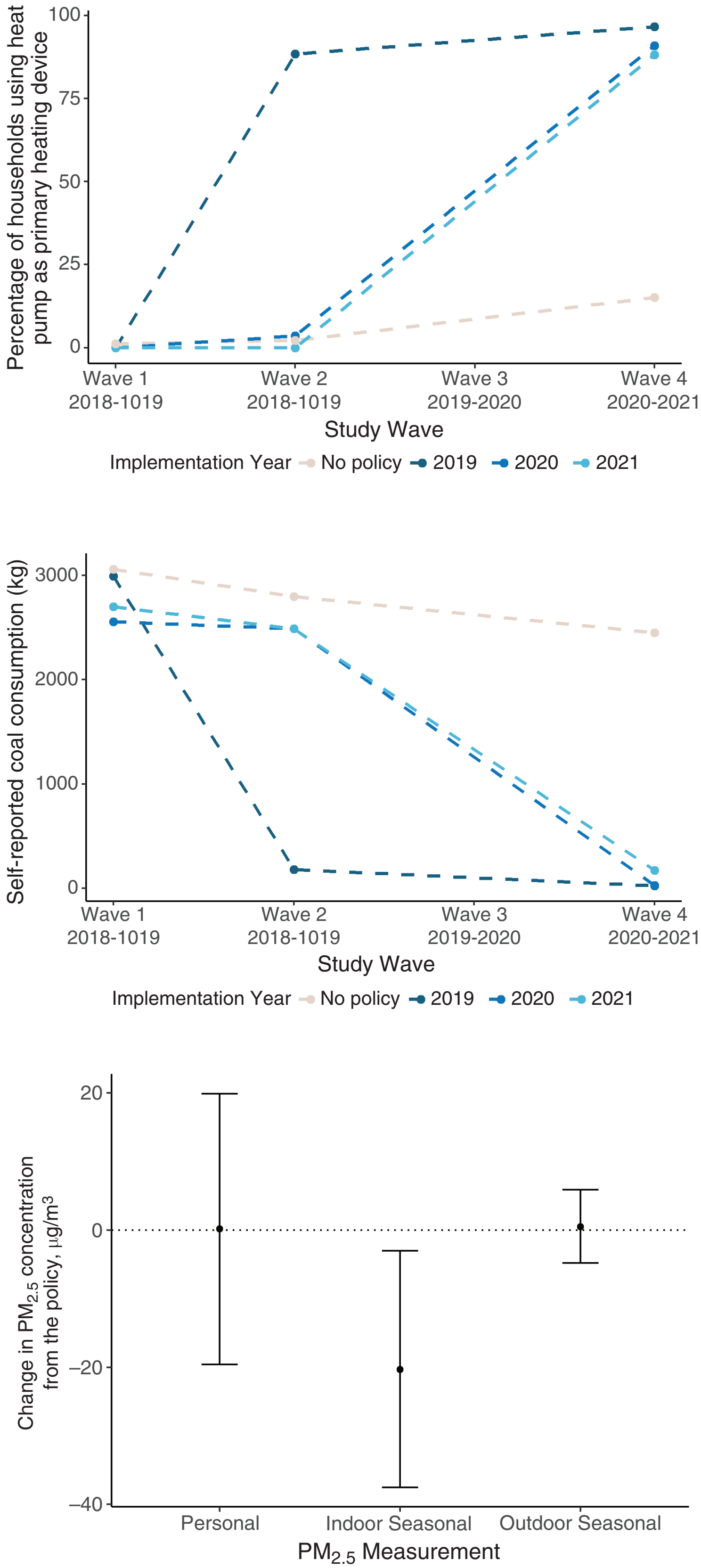
Among villages where the policy was not implemented by the end of the study, the proportion of households that reported using electric heat pumps increased from 3% at the start of the study to 16% at the end of the study, and most of those households that used heat pumps reported using them exclusively.

Implementation of the policy corresponded with steep declines in the use of self-reported coal use (**Commentary Figure 3, center**). Self-reported electricity expenditures increased over time, regardless of whether the villages had implemented the policy. At the same time, self-reported use of biomass — which was not targeted by the policy — remained mostly stable.

EFFECTS OF THE POLICY ON AIR QUALITY AND TEMPERATURE

Measured personal concentrations of $PM_{2.5}$ and black carbon were higher than measured indoor concentrations, which were higher than measured outdoor concentrations. This was the case both for villages where the clean heating policy was implemented and where it was not. Generally, personal, indoor, and outdoor concentrations decreased over the four winters of the study period, regardless of policy implementation. For example, seasonal average $PM_{2.5}$ concentrations over all villages decreased from 38 µg/m³ in Wave 1 to 26 µg/m³ in Wave 4. Baseline personal exposures to $PM_{2.5}$ and black carbon were 20% and 2% lower, respectively, in villages where the policy was eventually implemented than in other villages, which could mask any effects of the policy. However, the investigators were able to account for this difference at baseline using the difference-in-differences approach.

Implementation of the policy reduced indoor $PM_{2.5}$ concentrations by about 20 µg/m³ but had no effect on outdoor and personal $PM_{2.5}$ concentrations (**Commentary Figure 3, bottom**). The effect on indoor $PM_{2.5}$ concentrations was roughly one quarter of the baseline Wave 1 concentrations in 2018–2019 and about the same as the average change in indoor $PM_{2.5}$ concentrations across all residences between Wave 2 (2019–2020) and Wave 4 (2021–2022).



Commentary Figure 3. Effects of the policy on heat pump availability, coal consumption, and air quality. Top: Percentage of households using heat pumps as primary heating device. Center: Self-reported household coal consumption in kg. Darker colors represent villages where the clean heating policy was implemented earlier. Points indicate when the questionnaires were completed; effects on villages where the policy was implemented in 2019 or 2020 were first observed in the 2021 data collection because no questionnaires were collected in Wave 3. Bottom: Effect of the clean heating policy on personal, indoor, and outdoor $PM_{2.5}$ concentrations with 95% confidence intervals, adjusted for household size, tobacco smoking category, outdoor temperature, and outdoor dew point (negative values mean that the policy reduced $PM_{2.5}$).

From their source apportionment analyses, the investigators identified a factor that they called "mixed combustion" because it included chemical species typical of direct emissions from coal combustion, biomass burning, and possibly tobacco smoking. Over the study period, "mixed combustion" accounted for about 5–10 µg/m^3 of the outdoor $PM_{2.5}$ and about 20–30 µg/m^3 of the personal $PM_{2.5}$. Implementation of the policy had statistically insignificant and imprecise effects on "mixed combustion" exposures.

Households in villages that had implemented the policy had about 2°C higher mean and 4°C higher minimum winter indoor temperatures than households in villages without the policy, even though both groups of households had similar winter indoor temperatures at baseline. The investigators did not report any systematic heterogeneity by the year of policy implementation in the effect of the policy on $PM_{2.5}$ concentrations or temperature.

TOTAL EFFECTS OF THE POLICY ON HEALTH OUTCOMES

Individuals who lived in households in villages where the policy had been implemented had improved blood pressure and self-reported chronic respiratory symptoms. In adjusted models, both systolic and diastolic blood pressure decreased by about 1.5 mm Hg after villages implemented the policy, although there was no meaningful change in pulse pressure or blood pressure amplification. For comparison, the literature suggests that each 5 mm Hg reduction in systolic blood pressure reduces the risk of developing cardiovascular events by 10% (Canoy et al. 2022).

Implementation of the policy was associated with a 7.5% reduction in self-reported respiratory symptoms. This finding was largely driven by reductions in reports of coughing, having chest trouble, or difficulty breathing on several or most days of the week. The investigators did not find evidence that exposure to the policy affected phlegm, wheezing symptoms, levels of biomarkers in blood, or levels of FeNO in exhaled air.

Effects of the policy on blood pressure and phlegm were strongest in the villages where the policy was implemented earliest and were weaker in later years. The investigators did not report systematic patterns related to when the policy was implemented for other respiratory symptoms.

DIRECT AND INDIRECT EFFECTS OF THE POLICY ON HEALTH OUTCOMES

Because the overall models provided evidence that the policy had affected blood pressure and self-reported respiratory symptoms, the investigators conducted mediation analyses on whether changes to these health outcomes could be attributed to changes in $PM_{2.5}$ and temperature (**Commentary Table 2**). In the mediation analyses, the estimates of direct effects of the policy on all four measures of blood pressure decreased, suggesting that the policy had its effects on health mediated through changes in air quality and temperature. When the mediation of the effect of the policy on blood pressure by $PM_{2.5}$ and indoor temperature together was tested, the direct effect of the policy almost completely disappeared, suggesting that the effect of the policy on blood pressure could be explained almost completely by the policy-related reductions in indoor $PM_{2.5}$ concentrations and indoor temperature.

The direct effects for the self-reported respiratory outcomes were statistically indistinguishable from the total effects, so the investigators concluded that changes in $PM_{2.5}$ and temperature were not contributing factors to the policy-related changes in respiratory outcomes.

HEI REVIEW COMMITTEE EVALUATION

Baumgartner, Harper, and colleagues evaluated the air quality and health effects of China's policy to replace coal heaters with heat pumps in 50 villages that did not have the clean heating policy at the start of the study. By the end of the study, the policy had been implemented in 20 of the villages, and there was nearly complete compliance with the policy. Implementation of the policy reduced indoor $PM_{2.5}$ by about 20 µg/m^3, with smaller or negligible changes in outdoor $PM_{2.5}$, personal $PM_{2.5}$, and personal black carbon. At the same time, indoor temperatures during winter increased after implementation of the policy. The policy also slightly improved blood pressure and respiratory symptoms. Almost all the effects of the policy on blood pressure could be explained by the imputed improvements in indoor or personal $PM_{2.5}$ concentrations and indoor temperature, and the improvements in respiratory symptoms were most likely because of other factors.

In its independent evaluation of the study, the Review Committee thought that Baumgartner, Harper, and colleagues had completed an important study to evaluate the benefits of a clean energy policy on air quality and health. The study design and the causal framework in which they applied DAGs, staggered difference-in-differences analyses, and mediation analyses were strong. The Committee was impressed with the efforts to adapt the study design to the least extent necessary during the COVID-19 pandemic. Other strengths of the study included the use of sensors alongside filter samplers to measure $PM_{2.5}$ and long-term sampling in rural areas. Additionally, the statistical analysis methods were well described, and the extended two-way fixed effects model was well justified with details on the model and the DAG. The Committee agreed with the overall conclusion that the clean heating policy achieved its intended goals by dramatically reducing residential coal burning and improving indoor environmental quality, which provided some benefits to blood pressure and respiratory symptoms.

DIFFERENCE-IN-DIFFERENCES MODELING

The Review Committee highlighted that the staggered difference-in-differences modeling was generally a strong approach because it accounted for several key sources of potential confounding. First, this design removed all temporally fixed differences in the study populations in villages where the policy was implemented and where it was not. It

Commentary Table 2. Summary of Mediation Analysis Results[a]

	Was this factor found to substantially mediate the effect of the policy? (Percentage of the total effect that could be explained by mediation, ranges for different health outcomes in this group)			
Health Outcome or Effect	**Indoor or Personal $PM_{2.5}$[b]**	**Indoor Temperature[c]**	**Indoor or Personal $PM_{2.5}$ and Indoor Temperature[b,c]**	**"Mixed Combustion"**
Measured brachial blood pressure (systolic or diastolic)	YES (31% to 43%)	YES (31% to 79%)	YES (63% to 121%)	Not tested
Measured central blood pressure (systolic or diastolic)	YES (31% to 43%)	YES (25% to 71%)	YES (56% to 114%)	NO (–44% to 6%)
Self-reported respiratory outcomes (any symptom, coughing, phlegm, wheezing attacks, trouble breathing, or chest trouble)	NO (–13% to 33%)	NO (–143% to 8%)	NO (–40% to 15%)	Not tested

[a] Only those outcomes where mediation analyses were conducted are included in the table. No mediation analyses were conducted for pulse pressure, blood pressure amplification, FeNO, or inflammatory blood markers because there was no overall effect of the policy on these outcomes.

[b] Indoor $PM_{2.5}$ concentrations from about one-third of the participants in Waves 2–4 were used for blood pressure, and personal $PM_{2.5}$ concentrations from half of the participants in Waves 1, 2, and 4 were used for respiratory effects.

[c] Seasonal average temperatures were used for blood pressure, and 5-minute average temperatures were used for respiratory outcomes.

also accounted for background trends in air pollution and health during the study that might have occurred across the area regardless of whether the policy had been implemented. Nonetheless, some limitations in the difference-in-differences models and how they were used by the investigators bear discussion.

The key assumption for the difference-in-differences approach was that the time trends in the villages without the policy are an accurate reflection of the time trends that would have occurred in the villages with the policy had they, counterfactually, not implemented the policy. Because this policy was implemented over a period spanning years before, during, and after the COVID-19 pandemic, one must consider if the study conduct, the intervention implementation, or the outcomes could have been affected differently in the villages with and without the policy. There are several assurances that this is not the case. First, the investigators held to the original study plan wherever possible while also adjusting as needed. For example, they continued village-level tracking of the policies even when they were unable to recruit new participants or collect individual-level data. Second, the investigators also confirmed, to the extent possible, that all villages experienced similar COVID-19 burdens. In fact, according to the best available data, there were no COVID-19 cases in any of the suburban and rural villages of the investigation during the study period. Additionally, the investigators report that the villages all had similar requirements to prevent the spread of COVID-19. Collectively, these factors provided confidence to the investigators and the Committee that there should be similar trends in the health outcomes and exposures across villages with and without interventions during the pandemic period.

The investigators also explored the possibility of other unexplained differences in the trends of air pollution exposure and health outcomes by comparing pre-policy trends in the villages that implemented the policy at different times during the study period versus those that did not implement the policy. Largely, they observed no differences in trends between the villages before the policy implementation. The only exposure with different trends between villages with and without the policy was personal $PM_{2.5}$, which was found to have higher baseline concentrations and to be initially declining more steeply in villages where the policy was never implemented, as compared to those where the policy was eventually implemented. Similarly, there was a greater development of the probability of cough and chest complaints in the villages that received the intervention in 2021 as compared to other groups. For all situations, however, these trends suggest that the effect of the policy, if anything, would be underestimated.

Overall, the Committee thought that the investigators had carefully applied difference-in-differences modeling, that they addressed the limitations of the approach in a real-world accountability study, and that the results were robust.

CHALLENGES FOR MEDIATION ANALYSIS USING IMPUTED DATA

The use of causal mediation analysis was also felt to be a constructive addition to the study. At the same time, a major challenge for the mediation analysis was that data were needed on parameters that had little overlap in availability, thus leading the investigators to impute data for parameters that were not measured. In particular, it should be noted that imputation was used to account for missing indoor $PM_{2.5}$ measurements for the majority of participants, which adds uncertainty to analyses that involved indoor $PM_{2.5}$. The use of imputation might have made it difficult to detect true mediation or introduced measurement error or noise to the analyses. As a result, the Committee thought that the results of analyses using imputed data, especially the mediation analyses for blood pressure, should be treated as exploratory and interpreted with caution.

OPPORTUNITIES FOR FUTURE ANALYSES

Some of the surprising results in this study suggest the opportunity for future analyses that were out of the scope of the current study but might provide additional insight into the effects of the clean heating policy and the mediating factors through which it works.

Surprisingly, the investigators did not find that the clean heating policy had significant effects on personal or outdoor seasonal $PM_{2.5}$. It is possible that some of this difference had to do with averaging across an entire season when the intensity of heating (and thus fuel combustion) might have varied throughout the season. Therefore, it is possible that the policy did affect personal and outdoor $PM_{2.5}$ even though no effect was found over the entire winter season. Future analyses making use of heating-degree-days might be able to test for potential effects of the policy on $PM_{2.5}$ on the days where the policy would be expected to affect emissions the most.

Because less than one-third of the total effect of the policy on respiratory symptoms was mediated by indoor $PM_{2.5}$, a more complete understanding of the effect of the policy on respiratory health might be obtained if different potential mediators were tested, such as other pollutants, time spent cooking, heating days, and thermal comfort.

PUTTING THE AIR POLLUTANT EXPOSURES INTO GLOBAL CONTEXT

To put the air pollution exposures measured in the study into context, the investigators compared the measured baseline personal $PM_{2.5}$ concentrations to WHO air quality guidelines, which are recommendations for health-protective concentrations of ambient air pollution (WHO 2021). Assuming that the measured wintertime personal concentrations were similar to annual average ambient concentrations, which is a large assumption, the investigators stated that the observed personal $PM_{2.5}$ concentrations were aligned with WHO's Interim Target 1. The Committee observed that the seasonal average outdoor $PM_{2.5}$ concentrations were also of similar magnitude. They agreed with the investigators that $PM_{2.5}$ concentrations — whether personal, indoor, or outdoor — in this region were high. It was therefore valuable to evaluate the effect of China's clean heating policy on air pollution and health, and reassuring to see that the policy had some effects.

SUMMARY AND CONCLUSIONS

Baumgartner, Harper, and colleagues conducted a thorough accountability study of the air quality and health effects of China's clean heating policy. They evaluated more than 1,400 participants from more than 1,200 households in 50 suburban and rural villages of Beijing over four consecutive winters between 2018 and 2022. By the end of the study period, 20 villages had implemented the policy, which led to improved air quality and some health improvements in individuals living in those villages. The improvements in blood pressure were linked to improvements in air quality and indoor temperature.

The HEI Review Committee members were impressed with the strong study design and causal framework, showing that the policy was linked to some improvements in air quality that were in turn linked to some health improvements. They appreciated the difference-in-differences approach, which allowed the study team to account for many factors other than the clean heating policy that could have resulted in differences between the villages or changes in air pollution and health over time. They also acknowledged the effort involved for the investigators to follow the original study plan wherever possible, including through the COVID-19 pandemic.

Overall, the study demonstrated that the clean heating policy achieved its intended goals to electrify household heating in the villages where it was implemented and that it dramatically reduced residential coal burning and improved indoor environmental quality in the first years after implementation. The policy provided some benefits to heart and lung health, some of which (systolic and diastolic blood pressure) were related to decreases in air pollution exposure.

Although there is an abundance of evidence on associations of negative health outcomes with exposure to $PM_{2.5}$, it is important to quantify (rather than assume) the effects of specific actions on air quality and health. This study provides evidence that replacing coal-fueled heaters with heat pumps reduces indoor air pollution and improves health. These results are encouraging for other countries seeking to implement policies to replace highly polluting residential heating sources.

ACKNOWLEDGMENTS

The HEI Review Committee thanks the ad hoc reviewers for their help in evaluating the scientific merit of the Investigators' Report. The Committee is also grateful to Hanna Boogaard and Eleane van Vliet for oversight of the study, to Allison Patton for assistance with review of the Investigators' Report and preparation of its Commentary, to Progressive Publishing Services and Tom Zaczkiewicz for editing the Investigators' Report and its Commentary, and to Kristin Eckles for her role in preparing this Research Report for publication.

REFERENCES

Barrington-Leigh C, Baumgartner J, Carter E, Robinson BE, Tao S, Zhang Y. 2019. An evaluation of air quality, home heating, and well-being under Beijing's programme to eliminate household coal use. Nat Energy 4:416–423; *https://doi.org/10.1097/01.EE9.0000606212.67011.88*.

Canoy D, Nazarzadeh M, Copland E, Bidel Z, Rao S, Li Y, et al. 2022. How much lowering of blood pressure is required to prevent cardiovascular disease in patients with and without previous cardiovascular disease? Curr Cardiol Rep 24:851–860; *https://doi.org/10.1007/s11886-022-01706-4*.

China State Council. 2013. Air Pollution Prevention and Control Action Plan: China Clean Air Updates (English Translation). Clean Air Alliance of China. Available: *https://policy.asiapacificenergy.org/sites/default/files/Air%20Pollution%20Prevention%20and%20Control%20Action%20Plan%20%28EN%29.pdf* [accessed 8 August 2025].

Chowdhury S, Pillarisetti A, Oberholzer A, Jetter J, Mitchell J, Cappuccilli E, et al. 2023. A global review of the state of the evidence of household air pollution's contribution to ambient fine particulate matter and their related health impacts. Environment International 173:107835; *https://doi.org/10.1016/j.envint.2023.107835*.

Ding D, Xing J, Wang S, Liu K, Hao J. 2019. Estimated contributions of emissions controls, meteorological factors, population growth, and changes in baseline mortality to reductions in ambient $PM_{2.5}$ and $PM_{2.5}$-related mortality in China, 2013–2017. Environ Health Perspect 127:067009; *https://doi.org/10.1289/ehp4157*.

Dispersed Coal Management Research Group. 2023. 中国散煤综合治理研究报告 (China Dispersed Coal Governance Report). 北京大学能源研究院气候变化与能源转型项目 (Climate Change and Energy Transition Project, Institute of Energy, Peking University). Available: *https://www.ditan.com/static/upload/file/20231106/1699241777630902.pdf* [accessed 23 July 2025].

GBD MAPS Working Group. 2016. Burden of Disease Attributable to Coal-Burning and Other Air Pollution Sources in China. Health Effects Institute. Available: *https://www.healtheffects.org/publication/burden-disease-attributable-coal-burning-and-other-air-pollution-sources-china* [accessed 22 July 2025].

Health Effects Institute (HEI). 2025. State of Global Air 2025: A Report on Air Pollution and Its Role in the World's Leading Causes of Death. Boston, MA: Health Effects Institute.

HEI Household Air Pollution Working Group. 2018. Household Air Pollution and Noncommunicable Disease. Communication 18. Boston, MA: Health Effects Institute.

Hernán M, Robins J. 2020. Causal Inference: What If. Boca Raton, FL: Chapman & Hall/CRC.

IEA, IRENA, UNSD, World Bank, WHO. 2025. Tracking SDG 7: The Energy Progress Report. Washington, DC: World Bank. Available: *https://www.irena.org/Publications/2024/Jun/Tracking-SDG-7-The-Energy-Progress-Report-2024*.

Keele L, Tingley D, Yamamoto T. 2015. Identifying mechanisms behind policy interventions via causal mediation analysis: Methods for policy analysis. J Pol Anal Manage 34:937–963; *https://doi.org/10.1002/pam.21853?urlappend=%3Futm_source%3Dresearchgate.net%26medium%3Darticle*.

Kinney P, Beizhan Yan, Xiaoming Shi, Maigeng Zhou, Shuxiao Wang, Haotian Zheng, et al. In press. Accounting for the Health Benefits of Air Pollution Regulations in China, 2008–2019. Research Report. Health Effects Institute: Boston, MA.

Lai PS, Lam NL, Gallery B, Lee AG, Adair-Rohani H, Alexander D, et al. 2024. Household air pollution interventions to improve health in low- and middle-income countries: An official American Thoracic Society Research statement. Am J Respir Crit Care Med 209:909–927; *https://doi.org/10.1164/rccm.202402-0398st*.

Lee KK, Bing R, Kiang J, Bashir S, Spath N, Stelzle D, et al. 2020. Adverse health effects associated with household air pollution: A systematic review, meta-analysis, and burden estimation study. Lancet Glob Health 8:e1427–e1434; *https://doi.org/10.1016/s2214-109x(20)30343-0*.

Lee M, Chang J, Deng Q, Hu P, Bixby H, Harper S, et al. 2024. Effects of a coal-to-clean heating policy on acute myocardial infarction in Beijing: A difference-in-differences analysis. Lancet Planet Health 8:e924–e932; *https://doi.org/10.1016/s2542-5196(24)00243-2*.

Li T, Chen C, Zhang M, Zhao L, Liu Y, Guo Y, et al. 2024. Accountability analysis of health benefits related to National Action Plan on Air Pollution Prevention and Control in China. PNAS Nexus 3:pgae142; *https://doi.org/10.1093/pnasnexus/pgae142*.

Wen H, Nie P, Liu M, Peng R, Guo T, Wang C, et al. 2023. Multi-health effects of clean residential heating: Evidences from rural China's coal-to-gas/electricity project. Energy for Sustainable Development 73:66–75; *https://doi.org/10.1016/j.esd.2023.01.013*.

Wooldridge JM. 2021. Two-way fixed effects, the two-way Mundlak regression, and difference-in-differences estimators. SSRN (August 17, 2021); *https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3906345*.

World Health Organization (WHO). 2014. WHO Guidelines for Indoor Air Quality: Household Fuel Combustion. Geneva, Switzerland: WHO. Available: *https://www.who.int/publications/i/item/9789241548885*.

World Health Organization (WHO). 2021. WHO Global Air Quality Guidelines: Particulate Matter ($PM_{2.5}$ and PM_{10}), Ozone, Nitrogen Dioxide, Sulfur Dioxide, and Carbon Monoxide. Available: *https://www.who.int/publications/i/item/9789240034228* [accessed 21 July 2025].

Zhang R, Jing J, Tao J, Hsu S-C, Wang G, Cao J, et al. 2013. Chemical characterization and source apportionment of $PM_{2.5}$ in Beijing: Seasonal perspective. Atmos Chem Phys 13:7053–7074; *https://doi.org/10.5194/acpd-13-9953-2013*.

Zhang Z, Wang W, Cheng M, Liu S, Xu J, He Y, et al. 2017. The contribution of residential coal combustion to $PM_{2.5}$ pollution over China's Beijing-Tianjin-Hebei region in winter. Atmos Environ 159:147–161; *https://doi.org/10.1016/j.atmosenv.2017.03.054*.

Zhang Q, Zheng Y, Tong D, Shao M, Wang S, Zhang Y, et al. 2019. Drivers of improved $PM_{2.5}$ air quality in China from 2013 to 2017. Proc Natl Acad Sci USA 116:24463–24469; *https://doi.org/10.1073/pnas.1907956116*.