How Do Household Energy Transitions Work?

Jill Baumgartner (Co-PI)   Sam Harper (Co-PI)

2024-01-15

Table of contents

Abstract . . . 1
Introduction . . . 2
Specific Aims and Overarching Approach . . . 2
Study Design and Methods . . . 2
Data Analysis . . . 4
  Total Effect . . . 4
  Mediation Analysis . . . 4
Results . . . 5
  Description of study sample (Table) . . . 5
  Aim 1: Policy impacts and potential mediation . . . 5
  Aim 2: Source contributions . . . 6
  Aim 3 . . . 6
Discussion and Conclusions . . . 6
Implications of Findings . . . 6
Data Availability Statement . . . 6
Acknowledgements . . . 6
References . . . 6
Appendices . . . 7
About the authors . . . 7
Other publications . . . 7

Abstract

Introduction

Methods

Results

Conclusions

Brief summary of what we did.

Introduction

China is deploying an ambitious plan to transition up to 70% of all households in northern China to clean space heating, including Beijing. To meet this target the Beijing municipal government announced a two-pronged program that designates coal-restricted areas and simultaneously offers subsidies to night-time electricity rates and for the purchase and installation of electric-powered, air-source heat pumps to replace traditional coal-heating stoves. The program is being rolled out on a village-by-village basis; however there is uncertainty as to when villages will receive the program. The variability in when the policy is applied to each village allows us to treat the roll-out of the program as a quasi-randomized intervention. Households may also be differentially affected by this program due to factors such as financial constraints, preferences and social capital, and there is uncertainty about whether and how this intervention may affect indoor and outdoor air pollution, as well as health behaviors and health outcomes.

Specific Aims and Overarching Approach

This study builds on three data collection campaigns in winter 2018/19, winter 2019/20, and winter 2021/22, as well as a partial campaign in winter 2020/21 (CIHR-funded) with the following specific aims:

1. Estimate how much of the policy's overall effect on health, including respiratory symptoms and cardiovascular outcomes (blood pressure, central hemodynamics, blood inflammatory and oxidative stress markers), can be attributed to its impact on changes in PM2.5;
2. Quantify the impact of the policy on outdoor air quality and personal air pollution exposures, and specifically the source contribution from household coal burning (Previously Aim 3);
3. Quantify the contribution of changes in the chemical composition of PM2.5 from different sources to the overall effect on health outcomes (Previously Aim 2).

Study Design and Methods

- Field equipment
- DiD schematic
- Mediation DAG

To understand how Beijing's policy works we used a difference-in-differences (DiD) design (Callaway 2020), leveraging the staggered rollout of the policy across multiple villages to estimate its impact on health outcomes and understand the mechanisms through which it works. Simple comparisons of treated and untreated (i.e., control) villages after the CBHP policy has been implemented are likely to be biased by unmeasured village-level characteristics (e.g., migration, average winter temperature) that are associated with health outcomes. Similarly, comparisons of only treated villages before and after exposure to the program are susceptible to bias by other factors associated with changes in outcomes over time (i.e., secular trends, impacts of the COVID-19 pandemic). By comparing *changes* in outcomes among treated villages to *changes* in outcomes among untreated villages, we can control for any unmeasured time-invariant characteristics of villages as well as any general secular trends affecting all villages that are unrelated to the policy

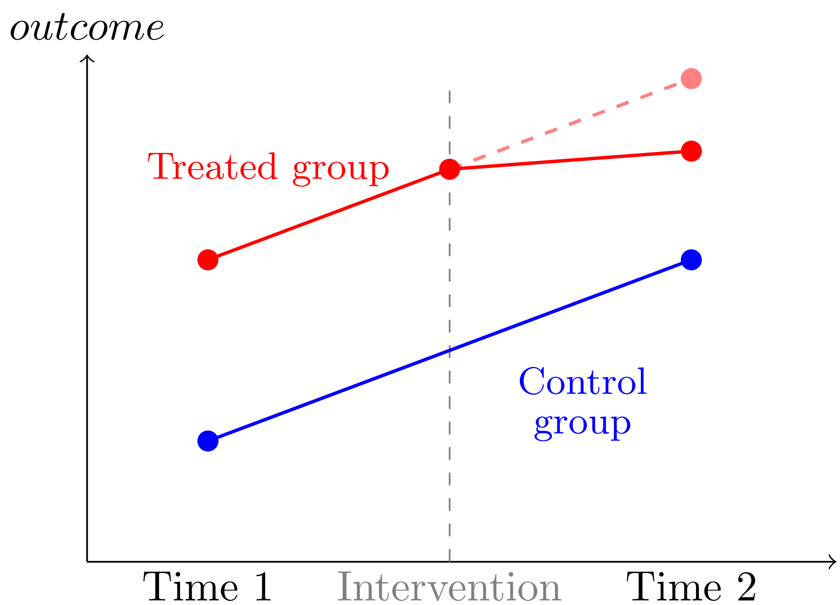

Figure 1: Stylized example of difference-in-differences

The DiD design compares outcomes before and after an intervention in a treated group relative to the same outcomes measured in a control group. The control group trend provides the crucial "counterfactual" estimate of what would have happened in the treated group had it not been treated. By comparing each group to itself, this approach helps to control for both measured and unmeasured fixed differences between the treated and control groups. By measuring changes over time in outcomes in the control group unaffected by the treatment, this approach also controls for any unmeasured factors affecting outcome trends in both treated and control groups. This is important since there are often many potential factors affecting outcome trends that cannot be disentangled from the policy if one only studies the treated group (as in a traditional pre-post design).

The canonical DiD design (Card and Krueger 1994) compares two groups (treated and control) at two different time periods (pre- and post-intervention, Figure X). In the first time period both groups are untreated, and in the second time period one group is exposed to the intervention. If we assume that the differences between the groups would have remained constant in the absence of the intervention (parallel trends assumption), then an unbiased estimate of the impact of the intervention in the post period can be calculated by subtracting the pre-post difference in the untreated group from the pre-post difference in the treated group.

However, when multiple groups are treated at different time periods, the most common approach has been to use a two-way fixed effects model to estimate the impact of the intervention which controls for secular trends and differences between districts. However, recent evidence suggests that the traditional two-way fixed effects estimation of the treatment effect may be biased in the context of heterogeneous treatment effects (Callaway and Sant'Anna 2021; Goodman-Bacon 2021)

Data Analysis

Total Effect

To estimate the total effect of the policy we used a DiD analysis that accommodates staggered treatment rollout. To allow for heterogeneity in the context of staggered rollout we used 'extended' two-way fixed effects (ETWFE) models (Wooldridge 2021) to estimate the total effect of the CBHP policy. The mean outcome (replaced by a suitable link function $g(\cdot)$ for binary or count outcomes) was defined using a set of linear predictors:

$$Y_{ijt} = g(\mu_{ijt}) = \alpha + \sum_{r=q}^{T} \beta_r d_r + \sum_{s=r}^{T} \gamma_s fs_t + \sum_{r=q}^{T} \sum_{s=r}^{T} \tau_{rt}(d_r \times fs_t) + \varepsilon_{ijt} \tag{1}$$

where Y_{ijt} is the outcome for individual i in village j at time t, d_r represent treatment cohort dummies, i.e., fixed effects for cohorts of villages that were first exposed to the policy at the same time q (e.g., in 2019, 2020, or 2021), fs_t are time fixed effects corresponding to different winter data collection campaigns (2018-19, 2019-20, or 2021-22), and τ_{rt} are the cohort-time $ATTs$.

Mediation Analysis

As noted above, with respect to the mediation analysis we are chiefly interested in the CDE, which can be derived by adding relevant mediators M to this model. If we also allow for exposure-mediator interaction and potentially allow for adjustment for confounders W of the mediator-outcome effect, we can extend equation Equation 1 as follows:

$$
\begin{aligned}
Y_{ijt} = g(\mu_{ijt}) = \alpha + \sum_{r=q}^{T} \beta_r d_r + \sum_{s=r}^{T} \gamma_s fs_t + \sum_{r=q}^{T}\sum_{s=r}^{T} \tau_{rt}(d_r \times fs_t) \\
+\delta M_{it} + \sum_{r=q}^{T}\sum_{s=r}^{T} \eta_{rt}(d_r \times fs_t \times M_{it}) + \zeta \mathbf{W} + \varepsilon_{ijt}
\end{aligned}
\tag{2}
$$

where now δ is the conditional effect of the mediator M at the reference level of the treatment (again, represented via the series of group-time interaction terms), and the collection of η terms are coefficients for the product terms allowing for mediator-treatment interaction. Finally, ζ is a vector of coefficients for the set of confounders contained within $\mathbf{W}$.

As noted above, in the staggered DiD framework that allows for heterogeneity we do not have a single treatment effect but a collection of group-time treatment effects that may be averaged in different ways. This extends to the estimation of the CDE, in which case we will also have several CDEs that can be averaged to make inferences about the extent to which the policy's impact is mediated by $PM_{2.5}$. Based on the setup in Equation 2 the CDE is estimated as: $\delta + \eta_{rt}MT$. In the absence of interaction between the exposure and the mediator (i.e., $\eta_{rt} = 0$) the CDE will simply be the estimated treatment effects $\sum_{r=q}^{T}\sum_{s=r}^{T}\tau_{rt}$, i.e., the effect of the policy holding M constant. For a valid estimate of the CDE we must account for confounding of the mediator-outcome effect, represented by W in the equation above. Baseline measures of both the outcome and the proposed mediators inherent in our DiD strategy will help to reduce the potential for unmeasured confounding of the mediator-outcome effect (Keele et al. 2015).

Results

Description of study sample (Table)

- Study flowchart of participants (Figure)
- Description of PM measurements (Figure)
- Uptake of the policy (Sankey energy use Figure)
- Impact of 'treatment assignment' on coal use (Figure? Table?)

Aim 1: Policy impacts and potential mediation

- Impact of policy on PM mass (Figure)
- Table of CDEs (Central SBP, Central DBP, FeNO, Respiratory outcomes, inflammatory markers), mediated by indoor PM (CDEs for personal and outdoor in SI)
- Table for multiple mediation analysis for BP

Aim 2: Source contributions

- Figure of source contributions (6 or fewer components)
- Source contributions by treatment status
- DiD for source contributions to PM

Aim 3

- Table of mediated health effects by source contribution (coal and biomass)

Discussion and Conclusions

Other relevant results (Tables or figures in SI)

Policy impacts on other relevant outcomes:

- Temperature
- Heating room
- Well-being

Implications of Findings

Data Availability Statement

To come…

Acknowledgements

To come…

References

Callaway B. 2020. Difference-in-Differences for Policy Evaluation. In: *Handbook of Labor, Human Resources and Population Economics* (K.F. Zimmermann, ed). Springer International Publishing:Cham. 1–61.

Callaway B, Sant'Anna PHC. 2021. Difference-in-Differences with multiple time periods. Journal of Econometrics 225:200–230; doi:10.1016/j.jeconom.2020.12.001.

Card D, Krueger AB. 1994. Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. American Economic Review 84: 772–93.

Goodman-Bacon A. 2021. Difference-in-differences with variation in treatment timing. Journal of Econometrics 225:254–277; doi:10.1016/j.jeconom.2021.03.014.

Keele L, Tingley D, Yamamoto T. 2015. Identifying mechanisms behind policy interventions via causal mediation analysis. Journal of Policy Analysis and Management 34: 937–963.

Wooldridge JM. 2021. Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators.; doi:10.2139/ssrn.3906345.

Appendices

About the authors

Other publications