Three Essays in Health and Energy Economics

Wenmei Tu

Department of Economics

McGill University, Montreal

December 2024

A thesis submitted to McGill University in partial fulfillment of the requirements

for the degree of PhD in Economics

Acknowledgements

I am profoundly grateful to everyone who made it possible for me to complete this thesis. My deepest appreciation goes to my supervisors, Prof. Erin Strumpf and Prof. Franque Grimard, whose invaluable suggestions and unwavering support were fundamental to my research. Their expertise deftly guided me through complex challenges, and their continuous engagement and encouragement were crucial in helping me make consistent progress throughout this journey. Without their intellectual generosity and personal commitment, this thesis would not have come to fruition.

Special thanks are also owed to Prof. Jill Baumgartner for allowing me to participate in the BHET project, which significantly contributed to the completion of two pivotal chapters of my thesis. The opportunity to access vital data and resources was a privilege that enabled me to conduct intriguing and valuable research.

I am deeply indebted to all members of the BHET team, whose openness in sharing their work and data, coupled with insightful feedback during our weekly discussions, profoundly impacted my growth as a researcher. I extend my heartfelt thanks to Xiang and Talia, whose readiness for discussion and consistent assistance were instrumental in refining my work.

The unwavering support and continuous encouragement from my parents have been the cornerstone of my academic pursuits. Their steadfast belief in my potential has been a constant source of strength and motivation throughout my years of study and the rigorous process of researching and writing this thesis.

Abstract

This dissertation comprises three essays that collectively explore themes of energy transition, public health, and human capital investment. The first essay investigates the dynamics of fuel transition within rural communities. The second essay evaluates the effects of a clean heating policy on the sleep patterns of rural residents in Beijing. The third essay delves into the role of later human capital investment as a means to mitigate the impacts of early-life adversities.

The first essay delves into the dynamics of heating fuel choices among rural households during the mandatory coal to electricity transition program in Beijing. Utilizing three rounds of panel data, the essay employs a difference-in-differences (DID) methodology to assess the impact of the coal ban policy. Findings reveal a significant, yet incomplete, transition from coal to electricity in affected households, alongside a spontaneous transition in regions not directly impacted by the policy. This prompts a deeper investigation into the primary factors influencing households' fuel choices and behaviors in areas without a coal ban policy. A correlated random effects generalized ordered probit model is utilized to control for unobserved individual heterogeneity and dissect the determinants of fuel choices. Key variables such as fuel prices, household income, education, marital status, and house area influence these choices. Notable differences are observed in the determinants affecting fuel choices and usage patterns, particularly in the roles of coal prices and income. These findings suggest that the selection of fuel types and determination of usage levels might be governed by separate household decision-making processes.

The second essay leverages data from Beijing's Coal to Electricity transition program, applying a DID methodology to explore the coal ban policy's effects on sleep patterns. This analysis recognizes that sleep is not always more is better, and thus, investigates the intervention's

heterogeneous impacts across different sleep durations. By examining various sleep outcomes and defining categories such as short sleepers and long sleepers, alongside deviations from recommended sleep durations, this study clarifies which groups are most influenced by the policy. The findings indicate that the policy significantly affects specific demographics: short sleepers, long sleepers, males, and individuals aged 60-70, particularly in regions undergoing substantial shifts from polluting to clean heating fuels. Although the policy successfully changed the way homes are heated from coal to electricity, which should have improved air quality and maintained more consistent temperatures indoors, these changes did not have a big impact on improving sleep quality as might have been expected. Despite cleaner air and better temperature control, these factors were not to play a major role in enhancing the sleep experiences of individuals.

The third essay utilizes two distinct exogenous variables that influence human capital accumulation in rural Indonesia—early-life rainfall shocks and access to free Early Childhood Education and Development (ECED) services—to test the Critical Programming Period hypothesis, which suggests that shocks experienced at birth could have long lasting consequences. The study also explores whether access to ECED services can mitigate the long-term consequences of these negative shocks experienced during the critical programming period. Our analysis reveals that adverse rainfall shocks during the year of birth negatively impact children's physical health, emotional maturity, and language and cognitive skills. Contrarily, the provision of free ECED services yields mixed and generally minor effects on these developmental dimensions, providing inconclusive evidence of their capacity to mitigate early-life adversities.

Abrégé

Cette thèse comprend trois essais explorant collectivement la transition énergétique, la santé publique et l'investissement en capital humain. Le premier essai étudie la dynamique de transition des combustibles dans les communautés rurales. Le second évalue les effets d'une politique de chauffage propre sur les habitudes de sommeil des résidents ruraux de Pékin. Le troisième explore le rôle de l'investissement en capital humain tardif comme moyen d'atténuer les impacts des adversités vécues en début de vie.

Le premier essai explore les choix de combustibles de chauffage des ménages ruraux lors du programme de transition obligatoire charbon-électricité à Pékin, utilisant trois vagues de données longitudinales et la méthode de différences de différences (DID) pour évaluer l'impact de la politique d'interdiction du charbon. Les résultats révèlent une transition importante mais incomplète du charbon à l'électricité, ainsi qu'une transition spontanée dans les zones non directement affectées par la politique. Ceci incite à une enquête plus approfondie des facteurs influençant les choix et comportements énergétiques des ménages dans les régions sans interdiction du charbon, utilisant un modèle probit ordonné à effets aléatoires corrélés pour contrôler l'hétérogénéité individuelle non observée et identifier les déterminants des choix de combustibles. Des variables clés telles que les prix des combustibles, le revenu des ménages, l'éducation, le statut matrimonial et la superficie de la maison influencent ces choix. Des différences notables sont observées dans les déterminants affectant les choix de combustibles et les modèles d'utilisation, particulièrement en ce qui concerne les rôles des prix du charbon et du revenu. Ces résultats suggèrent que la sélection des types de combustibles et la détermination des niveaux d'utilisation pourraient être régies par des processus de décision distincts au sein des ménages.

Le deuxième essai utilise des données du programme de transition charbon-électricité de Pékin, en appliquant une méthode DID pour étudier les effets de la politique sur les habitudes de sommeil. L'analyse reconnait qu'une augmentation du sommeil n'est pas toujours bénéfique et explore les impacts hétérogènes de l'intervention selon différentes durées de sommeil. En examinant divers résultats liés au sommeil et en définissant des catégories telles que les petits et gros dormeurs, ainsi que les écarts par rapport aux durées de sommeil recommandées, cette étude clarifie quels groupes sont les plus affectés par la politique. Les résultats montrent que la politique affecte significativement certains groupes démographiques, notamment les petits et gros dormeurs, les hommes, et les individus âgés de 60 à 70 ans, en particulier dans les régions passant de combustibles polluants à propres. Bien que la politique ait modifié le chauffage des maisons, améliorant théoriquement la qualité de l'air et la constance des températures intérieures, ces changements n'ont pas significativement impacté de sommeil. Malgré un air plus propre et un meilleur contrôle de la température, ces facteurs n'ont pas grandement amélioré l'expérience de sommeil.

Le troisième essai utilise deux variables exogènes influençant l'accumulation de capital humain en Indonésie rurale: les chocs pluviométriques précoces et l'accès aux services gratuits d'Éducation et de Développement de la Petite Enfance (ECED), pour tester l'hypothèse de la période critique, qui suggère que les des conséquences à long-terme, voir même permanentes. L'étude examine également si l'accès aux services ECED atténue les effets à long terme de ces chocs. L'analyse révèle que les chocs pluviométriques défavorables pendant l'année de naissance nuisent à la santé physique, la maturité émotionnelle et les compétences linguistiques et cognitives des enfants. En revanche, l'accès aux services ECED gratuits donne des résultats mitigés sur ces dimensions développementales, n'offrant pas de preuves concluantes sur leur capacité à atténuer les adversités durant la petite enfance.

Contribution of Authors

This thesis comprises three distinct studies. I am the sole author of each chapter, responsible for formulating the research ideas, analyzing the data, estimating the models, and writing the manuscripts. My supervisors, Prof. Erin Strumpf and Prof. Franque Grimard, alongside all members of the BHET group, provided invaluable advice and insightful feedback that significantly enhanced the quality and depth of the research presented. An AI tool was employed to refine the language of this thesis; however, the content is based entirely on my own work.

Contribution to Original Knowledge

The first essay contributes to the literature by providing an evaluation of clean heating policies on both the fuel choices and the broader heating behaviors at the household level. This analysis distinctly focuses on heating energy, an aspect often overshadowed by cooking energy in previous studies. I employ a random effects generalized ordered probit model with Mundlak correction to effectively control for unobserved individual heterogeneity, allowing me to dissect the determinants influencing household fuel choices and heating behaviors separately. This research sheds light on the complex yet distinct nature of household decision-making regarding fuel use, a subject that has been minimally explored. Furthermore, by categorizing energy types into polluting, mixed, and clean fuels and examining Fuel-stacking behaviors, which involve the simultaneous use of multiple fuels, my study uncovers factors driving these behaviors, addressing a gap in the existing research on fuel-stacking determinants.

The second essay evaluates the impact of Beijing's clean heating policy, which transitions households from coal-based heating systems to heat pumps, on sleep patterns in rural areas. I am the first to investigate the unintended health benefits of this environmental policy on sleep, exploring how changes in heating affect sleep duration across various demographic groups. A key strength of this analysis is the recognition that more sleep is not always beneficial, prompting an examination of the policy's heterogeneous impacts on different initial sleep durations. By categorizing individuals as short or long sleepers and examining deviations from recommended sleep durations, this study identifies which groups are most affected by the policy. Additionally, this essay explores potential mechanisms through which clean heating might influence sleep, focusing on anticipated environmental improvements such as reductions in indoor air pollution and enhancements in temperature level and stability, which could account for the observed improvements in sleep.

The third essay tests the Critical Programming Period hypothesis (Heckman, 2007), which posits that shocks experienced at birth can have long-lasting consequences. The study also explores whether access to early childhood education and development (ECED) services can mitigate the medium-term consequences of these adverse shocks experienced during the critical programming period. This research not only adds to the evaluation of policies aimed at assisting disadvantaged groups but also enhances our understanding of how early-life conditions influence childhood development outcomes, filling a gap in the literature on the medium-term effects of early adversities. By assessing diverse indicators of early childhood development, the study identifies which developmental dimensions are most affected by early shocks and how they can potentially be mitigated by subsequent investments, offering valuable insights for future policy and program design targeting disadvantaged populations.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

In low- and middle-income countries (LMICs), the interplay between public policy and household decisions is a critical area of study that sheds light on the broader implications of governmental initiatives on individual and family behaviors. Public policies across healthcare, education, energy, and housing aim to enhance welfare and drive sustainable development by influencing household behaviors. Assessing how these policies shape household choices and their outcomes is crucial for evaluating their effectiveness and achieving intended goals. This interaction is particularly vital in LMICs, where it significantly impacts resource allocation, vulnerability management, and key health and educational outcomes within rapidly urbanizing societies with deep-rooted cultural norms. By analyzing how households respond to policy shifts and the consequent benefits—both expected and unintended—this research offers insights into development dynamics and the scope for policy-driven enhancement of well-being.

Environmental policy plays an important role in enhancing well-being and promoting sustainable development. Globally, the reliance on biomass and coal for daily cooking and heating remains a persistent challenge, significantly contributing to indoor and regional air pollution. In 2020 alone, such pollution was estimated to be responsible for approximately 3.2 million deaths (WHO, 2024; Cheng et al., 2017). This pressing issue not only poses major health concerns but also contributes to climate change, underscoring the urgent need to transition to modern energy services as emphasized by the United Nations Sustainable Development Goals (UN, 2016). My dissertation explores the dynamics of fuel transition within households, particularly in rural areas where dependence on polluting fuels exacerbates public health risks and environmental degradation.

The second chapter of my thesis assesses the effectiveness of Beijing's "Coal to Electricity" program or "Clean Heating Policy," which incentivizes the adoption of heat pumps and electricity over coal through financial support. This policy, initiated in 2017, aims to reduce the reliance on polluting fuels, thus mitigating air pollution in rural settings. Utilizing a quasi-experimental design with three rounds of panel data, I investigate how households adjust their fuel selection and heating hours in response to the policy. My findings indicate substantial shifts towards cleaner energy, with a 63% decrease in the likelihood of using polluting fuels only and a 73% increase in the probability of using electricity as the primary heating fuel. However, the complete elimination of coal has not yet been accomplished. This trend towards cleaner energy is also evident in households not directly impacted by the policy. This suggests that additional factors may influence fuel choices. This variability in responses prompted me to further investigate household decision-making processes and the persistent use of traditional fuels alongside cleaner alternatives. Using a random effects generalized ordered probit model with Mundlak correction, I assessed the determinants of fuel choices and fuel usage separately, acknowledging that these decisions might be influenced by distinct processes. The analysis revealed that fuel prices primarily influence the type of fuels chosen but not the quantity consumed. Conversely, household income mainly impacts the amount of energy consumed rather than the type of fuel.

In rural northern China, traditional heating methods such as biomass or coal-fueled kangs are prevalent. While effective at providing warmth, they produce considerable indoor air pollution, primarily from particulate matter, significantly affecting health (Guo, 2002; Liu et al., 2016; Lawrence et al., 2018). These systems also necessitate overnight refueling and cause temperature fluctuations that can disrupt sleep, potentially impacting physical and cognitive health (Lan et al., 2017; Medic et al., 2017). Sleep, a fundamental pillar of health essential for cognitive, emotional, and physical functioning (Cappuccio et al., 2010; Lim and Dinges, 2010; Baglioni et al., 2011), is the focus of the third chapter of this dissertation. Employing a difference-in-differences approach with self-reported sleep data, this chapter examines the unintended public health impacts of the Beijing's clean heating policy, on sleep patterns among rural populations transitioning from traditional coal-based to cleaner heating systems. Research indicates a U-shaped relationship between sleep duration and health risks (Wang et al., 2016; Yin et al., 2017), suggesting that the effects of the clean heating intervention might vary among individuals with different initial sleep durations. This study explores the heterogeneous effects of the policy on sleep, by examining various sleep outcomes and deviations from recommended sleep durations, identifying which groups are most affected by the policy. The results indicate that the policy primarily benefits specific subgroups, including initial short sleepers (with an increase of approximately 1 hour in sleep duration), long sleepers (around a 0.5-hour decrease), men (about a 0.5-hour increase), individuals aged 60–70 (with over a 0.5-hour increase), and those in regions with significant transitions from polluting to clean heating fuels (exceeding a 0.5-hour increase). Additionally, I investigate the potential mechanisms through which clean heating might influence sleep, focusing on the anticipated benefits of reduced indoor air pollution and enhanced temperature stability. However, the findings reveal that these factors account for less than 10% of the observed improvements in sleep duration, suggesting complex interactions between environmental changes and health outcomes.

Lastly, the fourth chapter focuses on the impact evaluation of policy, with a particular emphasis on interventions aimed at children. This chapter examines how initiatives targeting early childhood can help those who experienced adversities during their birth year to catch up at later stages. Extensive literature supports the Critical Programming Period hypothesis, which posits that early-life shocks have long-lasting effects on human development (Almond, 2006; Lavy et al., 2016; Lee, 2014). I begin by testing this hypothesis using the example of rainfall shocks in rural Indonesia, which impact family income and initial investments in children's human capital. This confirms the effects of early-life shocks, demonstrating improvements such as a 0.15 standard deviation increase in the Height-for-age Z-score and a reduction of 15% or more in the likelihood of poor development in emotional maturity, conduct problems, and pro-social behavior problems. A subsequent question arises concerning whether disadvantaged children can catch up through later human capital investments, such as the Early Childhood Education and Development (ECED) program in rural Indonesia. By analyzing the interplay between two exogenous shocks—the early-life rainfall shocks and the introduction of the ECED program—I assess how these shocks and subsequent investments in human capital affect childhood developmental outcomes. The results indicate limited evidence of catching-up effects from the ECED program. This analysis contributes to the understanding of the complex dynamics between early adversities and the effectiveness of subsequent educational interventions in mitigating their long-term impacts.

Overall, my dissertation contributes to enriches our understanding of how environmental policies influence household energy decisions and public health and broadens our perspective on the dynamic interplay between human capital investments and the conditions of early life. By offering comprehensive impact evaluation of policy interventions aimed at mitigating environmental challenges and bolstering human capital development, this research serves as an invaluable resource for policymakers and stakeholders committed to fostering sustainable development and enhancing public welfare through informed policy actions. Through its detailed analysis and implications, this work encourages a deeper consideration of how policies can be optimally designed and implemented to address the multifaceted challenges faced in low- and middle-income countries.

Chapter 2

Household Heating Fuel Choice and Behaviours: Evidence from Rural China

Wenmei Tu

Department of Economics, McGill University, Montreal, Quebec, Canada

Abstract

The Clean Heating Policy, aimed at encouraging households in Northern China to transition from coal to electricity or gas, marks a pivotal effort in combating air pollution. However, the effectiveness of household fuel substitution policies in fuel choices and heating behaviors has seldom been thoroughly assessed. This paper utilizes data from three rounds of a detailed panel dataset collected during the Coal to Electricity program in rural areas of Beijing to explore the dynamics of heating fuel choice and behavior among rural households. By using a difference-in-differences (DID) approach, my analysis identifies a significant, though incomplete, shift from coal to electricity among affected households. There was a 63% decrease in the likelihood of exclusively using polluting fuels and a 73% increase in the probability of using electricity as the primary heating fuel. Additionally, a spontaneous transition was observed in areas not directly targeted by the policy. This observation prompts a deeper exploration into the determinants influencing households' fuel choices and behaviors in the absence of coal ban policy. Employing the correlated random effects generalized ordered probit model to account for unobserved individual heterogeneity, I identify key variables such as fuel prices, household income, education, marital status, and house area that markedly influence these choices. Notable differences are observed in the determinants affecting fuel choices and usage patterns, particularly in the roles of coal prices and income. These findings indicate that choosing fuel types and determining usage levels might be governed by separate household decision-making processes. The study highlights the necessity to integrate both the energy ladder and fuel stacking theories into the models of household fuel consumption to effectively capture the nuanced dynamics of energy transition in rural China.

2.1 Introduction

Currently, approximately 2.1 billion people in developing countries rely heavily on biomass and coal for their daily cooking and heating needs, contributing significantly to both household indoor and regional air pollution. This dependence poses a substantial health risk, associated with approximately 3.2 million deaths in 2020 alone, and exacerbates the challenges of climate change (WHO, 2024; Cheng et al., 2017). Consequently, the United Nations Sustainable Development Goals (SDGs) have recognized the access to affordable, reliable, sustainable, and modern energy services as pivotal for catalyzing economic growth and enhancing quality of life. Improved access to modern energy services directly supports advancements in health, education, gender equality, while mitigating environmental degradation and global warming (UN, 2016). Evidently, residential fuel switching and fuel substitution stand out as essential strategies to achieve these broad development goals, particularly due to their significant role in reducing solid fuel combustion.

There have been numerous efforts to promote cleaner fuels through inter-fuel substitution globally. Governments and organizations worldwide have launched initiatives to replace polluting fuels with cleaner alternatives such as natural gas, liquefied petroleum gas (LPG), and electricity. These initiatives often involve financial incentives, infrastructural improvements, and educational campaigns to facilitate the transition. For instance, programs in low and middle income countries have successfully reduced the dependency on kerosene and charcoal by promoting solar energy and improved cooking stoves, contributing to enhanced public health and environmental preservation (Rehfuess et al., 2014).

In China, reliance on solid fuels such as biomass and coal remains widespread, especially in rural areas, significantly exacerbating local air pollution and contributing to global climate change (Pachauri and Jiang, 2008; Q. Chen et al., 2024). This dependence on traditional energy sources has led to a concerning situation where, as of 2013, an astonishing 99.6% of the Chinese population was exposed to air pollution levels that exceeded WHO guidelines. Moreover, only a mere 1% of the urban population lived in cities that met EU air quality standards, as reported by MEP, 2013 and Zheng and Kahn, 2013. These conditions underscore the ongoing challenges in air quality management, with particulate matters such as PM10 and PM2.5 posing major public health risks. In response to this critical issue, the Chinese government has enforced stringent policies aimed at reducing air pollution, with coal control as a central strategy. Initiated with the Air Pollution Prevention and Control Action Plan of 2013, extensive household fuel substitution programs have been deployed, particularly focusing on coal-reliant rural areas[1]. The Clean Heating Program, launched during the heating seasons in residential sectors, is pivotal in diminishing air pollution levels. In 2015, it was estimated that rural China utilized approximately 200 million tons of scattered coal for heating (NRDC, 2017). The policy's main strategies include transitioning from coal to gas or electricity, supported by financial subsidies for households to adopt cleaner heating technologies and fuels, thereby fostering a significant reduction in pollution and enhancing public health.

Xu et al., 2024 have estimated that the clean heating policy led to substantial emission reductions-1.83 million metric tons of carbon dioxide equivalent in the aggregate in 2018. Song et al., 2023 further demonstrated that from 2015 to 2021, such policies reduced PM2.5 by 41.3 % Beijing and surrounding areas, significantly more than in other northern cities, and decreased China's annual PM2.5 by $1.9\ \mu\mathrm{g/m}^3$, preventing roughly 23,556 premature deaths in 2021. Despite these benefits, Li, 2018 observed a significant natural gas shortage in winter 2017, which inflated prices and heating costs, making them unaffordable for many residents despite subsidies. Wu et al., 2020 noted that the coal-to-electricity shift, while cleaning the air, did not ensure adequate warmth due to decreased energy delivery.

Research has extensively examined the health and environmental effects of clean heating policies, yet there remains a limited focus on quantifying their impact on household fuel use and heating behaviors, which are critical to residents' welfare. This study aims to address this gap by analyzing the effects of clean heating policies on household fuel choices and consumption patterns at the micro level. Utilizing data from the Beijing Household Energy Transition Project, this study leverages the staggered and quasi-random enrollment into the clean heating policy. This design allows for a robust evaluation of fuel substitution impacts and changes in heating behaviors, utilizing

[1] A comprehensive description of the series of policies regarding household fuel substitution is detailed by Wu et al., 2020.

the Difference-in-Differences (DID) method and supported by detailed questionnaire data on the usage of heating devices.

The clean heating policy evaluation in this study revealed a significant reduction in polluting fuel use with electricity becoming the primary heating source, leading to an increase in total heating hours due to more clean/electric heating despite a reduction in polluting fuel use. The policy successfully maintained overall heating hours and showed more pronounced benefits in groups treated earlier, suggesting that early adoption boosts efficacy. However, coal use was not completely eliminated, even with the coal ban, and a similar shift from coal to electricity was observed in non-treated households, albeit to a lesser degree. This indicates other factors may influence fuel choices, underscoring the need for further investigation into why some households transition to cleaner fuels while others do not.

Despite extensive research on household energy choices in China, most studies use cross-sectional data, which may not effectively capture the dynamic and complex nature of these decisions and fail to control for unobserved individual heterogeneity. And there is also an emphasis on cooking energy, often neglecting heating energy, which a major component of northern households' energy use. Additionally, the phenomenon of fuel-stacking, where households adopt cleaner fuels as partial substitutes without fully abandoning traditional energy sources, has not been thoroughly examined during transitions to cleaner energy. Evidence on the determinants of fuel stacking is quite limited. Furthermore, previous empirical studies on fuel consumption often use simplistic models focusing on limited socioeconomic factors. However, household decisions in rural areas are complex due to market failures for biomass, commercial fuels, and labor. These failures mean that consumption decisions are intertwined with production decisions, including fuel production, food supply, and labor allocation, suggesting that consumption and production decisions are interconnected. Thus, I adopt a comprehensive model that integrates a broader range of socioeconomic factors affecting both consumption and production to better understand household responses to market failures. Furthermore, I emphasize the difference between the determinants of fuel choices and fuel usage, which could reveal distinct aspects of household decision-making processes.

This study bridges existing research gaps by using longitudinal data from rural Beijing to focus on heating energy and explore the dynamics of energy transition. I classify energy types into polluting, mixed, and clean fuels to provide evidence for fuel stacking behaviors and separately analyze the determinants of fuel choices and usage. Employing a random effects generalized ordered probit model with Mundlak correction helps us control for unobserved heterogeneity, thereby reducing the bias of the results. The study aims to understand why the use of polluting fuels persists in rural households, the factors driving their transition towards clean fuels, whether households ascend a fuel quality ladder with increasing income, and if the determinants of fuel choices differ from those of fuel usage. These insights are crucial for crafting effective policies to promote a fuel transition.

The analysis of heating fuel choices and behaviors highlights their interconnected yet distinct nature. Notably, while coal prices significantly affect fuel choices—driving a shift towards cleaner options as prices rise—they do not impact total heating hours, indicating that price hikes alone may not reduce pollution. Conversely, while income influences energy consumption levels, it has less impact on the type of fuels chosen. Higher income typically leads to increased electricity use for heating, aligning with the energy ladder theory, which suggests that as household income rises, families transition to higher-ranked energy sources while phasing out less advanced alternatives; however, it does not eliminate the use of polluting fuels, supporting the fuel stacking theory where households maintain a diverse energy mix. This suggests that fully transitioning to clean energy involves more than economic incentives; it also requires changes in social norms and behaviors, enhanced energy education, and better integration of rural households into broader energy policy frameworks. These findings underscore the need to combine energy ladder and fuel stacking theories to create comprehensive policies that effectively promote energy transitions.

The paper is structured as follows: Section 2 presents a detailed literature review; Section 3 describes the data used; Section 4 evaluates the clean heating policy in terms of fuel choices and heating behaviors; Section 5 examines the determinants of two processes in heating fuel consumption: the selection of fuels and the amount of heating for each fuel or device; Section 6 provides a discussion; and Section 7 offers the conclusion.

2.2 Literature Review

In this section, I begin by thoroughly reviewing evaluations of household energy interventions globally, with a specific focus on China, identifying key limitations and gaps in existing research. I then present the two predominant theories of fuel choices—the energy ladder and fuel stacking—and provide a summary of empirical evidence concerning various factors that influence household fuel decisions. Additionally, I investigate the existing research on household fuel choices in China, pinpointing methodological shortcomings and emphasizing the contributions of this study to the field.

2.2.1 Evaluation of Household Energy Interventions

Extensive fuel intervention programs have been launched worldwide to reduce air pollution and enhance public health. Numerous studies have explored the impacts of these initiatives globally, including specific analyses focused on China. A primary strategy in these interventions is the implementation of fuel bans, which have been demonstrated to significantly decrease air pollution and its associated health risks. For instance, the coal ban in Dublin led to marked reductions in air pollution levels, correlating with lowered incidences of respiratory and cardiovascular diseases (Dockery et al., 2013; Clancy et al., 2002). Similarly, bans on residential wood-burning were shown to decrease cardiovascular hospitalizations and mortality rates due to reduced particulate matter exposure (Yap and Garcia, 2015; Johnston et al., 2013).

Another strategy for energy switching programs involves substituting clean energy by providing improved appliances or stoves. Budya and Arofat, 2011 examined Indonesia's nationwide initiative to transition from kerosene to LPG, observing significant reductions in household air pollution. Pachauri et al., 2018 highlighted the necessity of effective subsidy mechanisms to facilitate such transitions, especially in economically disadvantaged communities where cost barriers could restrict access to cleaner fuels. Scott and Scarrott, 2011 analyzed air quality interventions in New Zealand, noting marked improvements in PM10 concentrations resulting from stringent regulations and heater replacements.

China has dedicated significant efforts to mitigate its severe air pollution over recent decades. A major focus of these initiatives is the transition from residential coal to cleaner fuels, recognizing the heavy dependence on coal in the residential sector. As China's coal replacement policies evolve, scholarly attention on residential coal substitution within the country is increasing. Many studies focus on evaluating the economic, health, and environmental impacts via cost-benefit analysis. For instance, X. Zhang et al., 2019 developed an integrated model to evaluate the health impacts and economic costs of cleaner heating within the Beijing-Tianjin-Hebei (BTH) region, concluding that the overall public health improvements from enhanced air quality yield net social benefits, including spillover effects. Similarly, Xiaolin et al., 2019 applied cooperative game theory to devise cost-effective strategies that maximize environmental benefits in the BTH region, suggesting substantial potential savings if investments and support are optimally distributed.

Another focus of these evaluations is the effect of such programs on air pollution reduction and the resultant health benefits. Tian et al., 2018 highlighted the efficacy of high-quality coal replacements in reducing emissions from coal stoves. Studies such as those by Niu et al., 2024, Song et al., 2023, and Yu et al., 2021 noted modest reductions in outdoor PM2.5 levels in areas where the coal ban with heat pump subsidy was implemented, in contrast to neighboring regions without such measures. Meng et al., 2019 reported a significant 36% decrease in personal PM2.5 exposure following changes in household fuel use. However, Wen et al., 2023 observed only slight reductions in chronic lung diseases, with no significant changes in physician-diagnosed cardiovascular diseases following the coal ban policy.

In the rural areas of Beijing—the focus of this study—seasonal indoor PM2.5 levels experienced a substantial decrease of 30.9 $\mu g/m^3$ as a result of the Clean Heating policy, significantly enhancing indoor air quality across all villages. However, personal PM2.5 exposures did not similarly decrease, which may be due to sporadic high-pollution activities such as smoking (Brehmer and Li, 2023). Health-wise, the policy was linked to notable reductions in blood pressure, with systolic BP falling by 1.4 mmHg and diastolic BP by 1.6 mmHg (Sternbach et al., 2023).

Despite the proliferation of studies at the macro level, there is a notable gap in micro-level research on the impacts of residential coal replacement programs in China, especially on the fuel consumption and well-being. A few studies like those by Barrington-Leigh et al., 2019 and Wu et al., 2020 have begun to fill this void. Barrington-Leigh et al., 2019 assessed the effects of a program that subsidized electric heating devices and banned coal, analyzing its impact on household energy use, expenditure, and indoor environmental quality. Wu et al., 2020 found that the coal-to-electricity policy effectively reduced pollution but resulted in decreased energy delivery, adversely affecting winter warmth. Conversely, the high-quality coal replacement policy maintained energy delivery but failed to enhance indoor air quality.

This paper addresses this gap by evaluating the impact of coal replacement policies on household heating fuel choices and behaviors at the micro-level. It investigates how households distribute their heating needs among various fuels and devices in the context of coal replacement policies, and assesses the evidence for both the ”energy ladder” and ”fuel stacking” theories in understanding the effects of these interventions.

2.2.2 Determinants of Household Fuel Choices

After reviewing the evaluation of household energy interventions, I now explore another critical aspect of the literature concerning the determinants of household fuel choices. The energy ladder and fuel stacking theories prominently guide discussions on how households make fuel decisions. Empirical studies have extensively identified various determinants influencing these decisions. Additionally, I delve into the research on fuel choice dynamics within China, highlighting its limitations and underscoring the contributions of this study in addressing these gaps.

Energy Ladder Theory

The energy ladder theory illustrates a hierarchical relationship between a household’s economic status and the types of fuel they use for cooking and heating. As household income increases, families progressively transition to higher-ranked fuels while phasing out alternatives(Hosier and Dowd, 1987; Heltberg, 2005), where the ranking of fuels is determined by cleanliness, ease of use, cooking or heating speed, and efficiency(Hiemstra-Van der Horst and Hovorka, 2008). Initially, families rely on biomass, then transition to fuels like kerosene and coal, and ultimately adopt cleaner

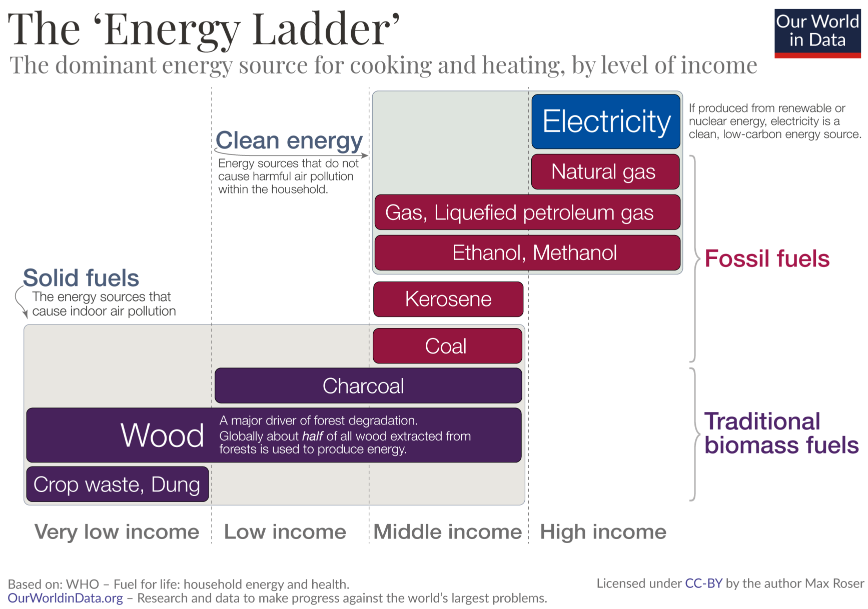


Figure 2.1: Energy Ladder

Source: from WHO. https://ourworldindata.org/energy-ladder

energy sources such as LPG and electricity(See Figure 2.1). This transition is driven not only by the pursuit of fuel efficiency and reduced pollution exposure but also by a desire to reflect improved socio-economic status (Masera et al., 2000).

While this correlation between income and fuel choice has been observed at both the country and individual levels (Farsi et al., 2007; Davis, 1998; Gupta and Köhlin, 2006), the model's simplicity is often challenged by empirical evidence. Research across various developing countries indicates that fuel wood remains a vital energy source for households at all income levels despite the availability of modern fuels(Hiemstra-Van der Horst and Hovorka, 2008; Hosier and Kipondya, 1993; Bhagavan and Giriappa, 1995; Brouwer and Falcão, 2004). Furthermore, the use of higher-ranked fuels such as electricity and LPG by low-income households (Campbell, 2003) demonstrates that energy choices are influenced by a variety of factors beyond income. This indicates a more complex interplay of factors influencing energy usage than what is depicted by the energy ladder, necessitating a broader understanding of the impact of various socio-economic variables on energy consumption behaviors.

Fuel Stacking Theory

Numerous studies found that the energy transition is not a simple linear progression but involves households simultaneously using multiple fuels, a concept known as fuel stacking (Leach, 1992; Campbell, 2003; Arnold et al., 2006; Karekezi and Majoro, 2002). This model posits that households adopt new fuels as partial substitutes without abandoning traditional energy sources. Often, fuel switching process is not unidirectional; households may revert to using traditional fuels even after adopting modern energy sources due to factors like price fluctuations (Masera et al., 2000; Arnold et al., 2006; Maconachie et al., 2009; Wickramasinghe, 2011). Masera et al., 2000 proposed the multiple fuel model, which suggests households do not entirely switch from one fuel to another but maintain a diverse portfolio of energy options for common needs such as cooking. Studies support the prevalence of fuel stacking across both urban and rural households in developing countries(Heltberg, 2005; Hiemstra-Van der Horst and Hovorka, 2008; Mekonnen and Köhlin, 2009; Mirza and Kemp, 2011), serving as a strategy to ensure energy security (Davis, 1998), cope with fluctuating fuel prices or incomes (Hosier and Kipondya, 1993), manage unreliable fuel supplies (Masera et al., 2000; Hosier and Kipondya, 1993), and adhere to cultural traditions (Masera et al., 2000).

Determinants of Fuel Use

I present empirical evidence on the principal determinants of household fuel decisions. For a comprehensive review of the factors influencing fuel use, readers are encouraged to refer to the detailed studies by Muller and Yan, 2018 and Van der Kroon et al., 2013.

Income The energy ladder theory has emphasized income as a important determinant of fuel choices, however, empirical studies offer mixed results. The relationship between income and fuel choice is complex and varies by context, as shown in studies like Ouédraogo, 2010 in Burkina Faso and Gupta and Köhlin, 2006 in urban India, where higher incomes or expenditures led to the adoption of gases like natural gas and LPG. Notably, the income effect on fuel choice is not always linear or positive; for instance, Démurger and Fournier, 2011 found that Chinese rural households substitute coal for firewood as wealth increases, challenging the simple linear model of the energy ladder. This complexity is further highlighted by the varied income elasticities across fuel types and contexts, suggesting that while some fuels behave as necessities, others are luxury goods, depending on the economic and social setting (Muller and Yan, 2018; Akpalu et al., 2011; Hughes-Cromwick, 1985).

Fuel Prices Extensive empirical research illustrates the significant role of fuel prices on household fuel choices, noting considerable variability in price sensitivity across fuels and regions. Numerous studies, such as those by Farsi et al., 2007 in India and Jingchao and Kotani, 2012 in Beijing, demonstrate that increases in the prices of fuels like LPG significantly reduce their consumption and likelihood of selection by households. Similarly, research by Gupta and Köhlin, 2006 and Akpalu et al., 2011 indicates that the relationship between firewood consumption and its price is predominantly negative. However, the effects vary; while some studies find strong negative price elasticities, others, like those on coal, show no significant price impact, emphasizing the complexity of fuel price effects across different contexts. Additionally, cross-price effects, which suggest potential substitutability or complementarity between fuels, have yielded mixed results in the literature. For example, Heltberg, 2005 and Pitt, 1985 observed substitution effects between various fuels, whereas others like Gupta and Köhlin, 2006 report negative cross-price elasticities, challenging simple interpretations of fuel substitutability. This indicates that fuel choices are influenced by a complex interplay of price dynamics and other socioeconomic factors, necessitating nuanced economic models to accurately capture these relationships.

Age Empirical evidence on the impact of household head age on fuel choice presents mixed findings. Older household heads tend to favor traditional fuels; studies like Baiyegunhi and Hassan, 2014 in Nigeria and Edwards and Langpap, 2012 in Guatemala report a positive association between age and the use of wood or fuel wood. Similarly, Démurger and Fournier, 2011 find a significant preference for firewood in older rural Chinese households, while Gebreegziabher and Kooten, 2013 note older Ethiopian heads prefer charcoal. Conversely, some research suggests older individuals prefer modern fuels; Farsi et al., 2007 and Gupta and Köhlin, 2006 observed a preference for LNG over wood in older Indian household heads, and Özcan et al., 2013 report a shift from wood to cleaner fuels like natural gas and electricity in older Turkish household heads. These divergent findings could reflect a life cycle effect, where older individuals, potentially with fewer liquidity constraints, can afford cleaner fuels, and at older ages, it may become too difficult to gather firewood or haul coal. However, several studies like those by Abebaw, 2007 and Israel, 2002 argue that age has no significant impact on fuel use, adding complexity to the understanding of age dynamics in fuel choice behavior.

Gender Gender also influences household fuel choices, with mixed evidence on its impact. Studies like Farsi et al., 2007, Rao and Reddy, 2007, and Rahut et al., 2020 suggest that female-headed households often prefer modern fuels over traditional ones, likely due to women's central role in cooking and their direct exposure to harmful pollutants from traditional fuels. Conversely, some research finds no significant gender effect; Abebaw, 2007, and Ouédraogo, 2010 report an insignificant gender coefficient in varying contexts. In Nepal, Link et al., 2012 observed that households with a higher proportion of female members tended to use more fuel wood, attributing this to women's roles in fuel wood gathering. However, contrasting findings from Heltberg, 2005 in Guatemala and Israel, 2002 in urban Bolivia suggest that the proportion of females does not influence fuel wood use significantly, with Israel noting that women earning a larger share of family income correlates with less firewood use, possibly reflecting the higher opportunity costs of women's time. Gupta and Köhlin, 2006 also find that the employment status of women in India does not impact fuel choice. These disparate findings indicate that gender's role in fuel choice may be shaped by a blend of preferences, time costs, and intra-household dynamics.

Education Education plays a pivotal role in influencing household fuel choices, often leading to a shift from traditional to modern fuels due to increased awareness of health impacts and the efficiencies offered by alternative energy sources. Studies such as Abebaw, 2007 and Démurger and Fournier, 2011 show a negative correlation between education level and firewood consumption, partly due to higher opportunity costs associated with fuel collection time. Further, research from Nigeria and India by Baiyegunhi and Hassan, 2014 and Gupta and Köhlin, 2006 supports that increased educational attainment encourages a transition towards kerosene and LPG. Similarly, in Ethiopia and Kenya, Gebreegziabher and Kooten, 2013 and Lay et al., 2013) found that higher education levels correlate with a preference for electricity over wood. These findings highlight education's role not only in enhancing income but also in promoting energy literacy, which can drive changes in fuel consumption patterns.

Household Size Household size influences energy consumption patterns, often reflecting both economic scale effects and income constraints. Studies like Abebaw, 2007 and Jingchao and Kotani, 2012 highlight a negative correlation between household size and per capita energy use, suggesting economies of scale. However, this association might also mask an underlying income effect, as larger households, which are often poorer, might not afford modern fuels, thus relying on cheaper alternatives like firewood. On the contrary, some research, such as by Baiyegunhi and Hassan, 2014 and Gupta and Köhlin, 2006, indicates that larger households are more inclined towards clean fuels, possibly due to greater aggregate resources. The relationship between household size and fuel choice remains mixed, with some studies noting no significant impact on fuel switching (Hosier and Dowd, 1987; Guta, 2014), while others observe a preference for traditional fuels in larger families due to their need for larger fuel quantities and the lower unit costs of traditional energy sources (Rao and Reddy, 2007). These contrasting findings suggest that the effects of household composition on fuel usage are complex and warrant further investigation to uncover the underlying dynamics.

Availability and Accessibility The availability and accessibility of fuels is another determinant of household energy choices. Traditional fuels' availability often relates to geographical factors like proximity to fuel wood sources and community perceptions of fuel wood availability (Hosier and Dowd, 1987; Hosier and Dowd, 1987). Conversely, the accessibility of modern fuels involves factors such as community access to electricity and modern renewable energy technologies (Heltberg, 2005; Gupta and Köhlin, 2006; Lay et al., 2013). Studies consistently find that limited access to traditional fuel sources often pushes households toward modern alternatives due to the high opportunity costs associated with fuel collection. For instance, households with better access to electricity and other modern energy sources are more likely to transition away from traditional fuels like wood and kerosene (Lay et al., 2013; Heltberg, 2005). This shift is also affected by infrastructure reliability, as frequent power outages can hinder the adoption of electricity (Kemmler, 2007).

Production Characteristics The conrrelation between agricultural production characteristics and household fuel choice is particularly pronounced in rural areas, where the decisions surrounding consumption and production are deeply interconnected. Authors like L. Chen et al., 2006, Démurger and Fournier, 2011, and Pandey and Chaubal, 2011 demonstrate varying impacts of farmland size and livestock on fuel choices across different regions in China and India, with some findings indicating that larger farm sizes and the presence of livestock increase the use of biomass fuels due to the availability of agricultural by-products for burning. Conversely, some studies note a decrease in firewood consumption with an increase in livestock, suggesting a complex relationship between farm assets and fuel usage that transcends simple income effects. Additionally, the labor availability within households, often associated with larger family sizes, can reduce the opportunity costs related to fuel wood collection and management, influencing a continued preference for traditional fuels (L. Chen et al., 2006; Baland et al., 2017). This correlation suggests that agricultural activities and labor resources within households play crucial roles in shaping fuel consumption patterns, yet these dynamics are often overlooked in energy source studies.

Other Factors Lifestyle factors impact fuel choices, particularly when traditional cooking habits demand prolonged use of a fuel, influencing the type of energy used. Baiyegunhi and Hassan, 2014 note that Nigerian households hesitate to transition from fuel wood to natural gas or electricity due to prolonged cooking times. Similarly, Ouédraogo, 2010 finds that frequent traditional cooking increases firewood use in Ouagadougou. The choice of fuel is also intertwined with the type of cooking and heating appliances available, which can limit or promote the use of certain fuels. For instance, the presence of modern appliances is often necessary for adopting cleaner fuels (Hughes-Cromwick, 1985; Manning and Taylor, 2014). However, high appliance costs can impede this transition, a factor not yet empirically explored in depth regarding its direct financial impact on fuel choice (Edwards and Langpap, 2012; Louw et al., 2008). Additionally, household structure and regional factors can influence fuel preferences, with married heads of households and specific ethnic or regional groups showing distinct fuel usage patterns (Paudel et al., 2018; Liao et al., 2019).

Literature Review on Fuel Choices in China

The existing literature on household energy choices in China has predominantly relied on either aggregated statistics or localized surveys, with a limited number of studies utilizing nationwide data. Studies like those by Cai and Jiang, 2008 utilized aggregate statistics to validate the energy ladder hypothesis, showing a distinct contrast between urban and rural energy consumption patterns, with urban households favoring cleaner and more efficient fuels compared to the traditional biomass and coal used in rural settings. This hypothesis was similarly explored by Peng et al., 2010, who analyzed fuel switching at the household level within rural Hubei and noted significant geographic variations in energy source availability and preferences, highlighting a gradual shift from biomass to commercial energy sources like coal.

Further contributions to this field include Jingchao and Kotani, 2012, who, through survey data from rural Beijing, identified that price factors for coal and LPG did not support substitution effects between these fuels. At a national level, Jiang and O'Neill, 2004 expanded the scope of analysis by incorporating a nationally representative rural household survey combined with aggregate statistics to explore broader patterns of residential energy use across rural China.

More recent studies have applied more sophisticated econometric models to dissect the nuances of energy choice determinants. For instance, utilizing panel data from the China Health and Nutrition Survey (CHNS), Muller and Yan, 2018 employed random effects multinomial logit models to confirm that fuel choices are significantly influenced by a complex interplay of socioeconomic and market factors. Similarly, X.-B. Zhang and Hassen, 2017 employed generalized ordered models with random effects to scrutinize urban households' cooking energy preferences, finding that policies enhancing income could lessen coal's competitive edge over liquefied natural gas. J. Zhang et al., 2020 employ the Seemingly Unrelated Regression (SUR) model to analyze urban household data in China, identifying building type, income, and fuel prices as key determinants of energy choice behavior.

Additionally, Zhu et al., 2022 integrate the least absolute shrinkage and selection operator (LASSO) algorithm with a multinomial Logit model to analyze the determinants of energy choice for household cooking in China, utilizing data from the Chinese Family Panel Studies (CFPS), finding household expenditure, off-farm employment, health levels, the presence of children, the education level of the household head, living conditions, and energy accessibility significantly influence the transition towards cleaner cooking fuels. Wu and Zheng, 2022 further studied the energy ladder using data from the Chinese Residential Energy Consumption Survey (CRECS), revealing an inverted U-shaped relationship between income and fuel diversity and noting that the ”Coal-to-Electricity” policy might hinder natural income-driven transitions towards cleaner energy.

Despite the depth of research on household energy choices in China, the majority of studies suffer from methodological limitations, particularly the use of cross-sectional data which may not fully account for unobserved individual heterogeneity, thus possibly confounding the true effects of explanatory variables. Additionally, there is a pronounced focus on cooking energy within the existing literature, often overlooking heating energy, which constitutes a 32% of building energy use in Chinese households(International Energy Agency, 2015). Moreover, few studies have explored the distinction between the determinants of fuel choices and fuel usage, which may represent two separate decision-making processes. Wu and Zheng, 2022 briefly touch on this concept but do not provide detailed insights into how these decision processes differ.

Based on the above literature review on the determinants of fuel choices, the influence of factors such as age, income, and others largely depends on the context. To better understand how these determinants impact heating fuel choices under the Beijing clean heating policy, I will need to conduct an analysis within the relevant context, as the previous literature from other settings does not provide a reliable basis for prediction.

This study addresses these gaps in the literature by implementing several methodological enhancements to provide more accurate and comprehensive insights into rural heating energy decisions. Firstly, I control for unobserved heterogeneity by utilizing longitudinal data, which enables us to offer more precise estimates of the effects being studied. Secondly, I specifically focus on heating—a critical but often under-examined aspect of household energy use—to enrich the empirical evidence regarding rural energy decisions. Lastly, I differentiate between the determinants of heating fuel choices and fuel usage, examining these two decision processes separately to highlight potential variations and to better understand the underlying mechanisms influencing each.

2.3 Data

2.3.1 Clean Heating Policy

In response to the severe smog and escalating particulate matter pollution across various regions, the Chinese government has implemented significant air pollution reduction measures over recent decades. These measures include a series of policies mandating household fuel substitutions, such as replacing coal with natural gas or electricity, especially during the heating seasons. This initiative is crucial since residential space heating is one of the largest sources of wintertime air pollution (Dispersed Coal Management Research Group, 2023). The policies primarily target the ”2+26” cities[2] in the northern region, emphasizing mandatory coal-to-gas and coal-to-electricity transitions, with financial subsidies provided to households for purchasing cleaner heating solutions. These strategies aim to ensure a clean and warm winter while reducing emissions from the residential sector, thereby enhancing regional air quality.

In the context of Beijing Municipality, where the data for this study was collected, most peri-urban and rural households have traditionally relied on traditional coal heaters and biomass kangs[3]. In line with national environmental policies, the Beijing municipal government has launched an ambitious program–the Clean Heating Program– aimed at reducing pollution and modernizing heating methods.

The program's design incorporates two primary strategies: prohibiting the use of coal and providing financial incentives for cleaner alternatives. Before the implementation of the current

[2]The ”2” refers to Beijing and Tianjin municipalities. The ”26” refers to 26 cities located in Shanxi, Henan, and Hebei provinces. The Ministry of Environmental Protection has identified these ”2+26” cities as the key transmission channel cities for air pollution in the Beijing-Tianjin-Hebei region, and they are considered a top priority in the prevention and control of air pollution.

[3]A kang is an integrated traditional Chinese heating technology that serves multiple home functions including cooking, sleeping, space heating, and ventilation.

clean heating policy analyzed in this study, Beijing had already initiated another coal replacement program. Under this earlier initiative, rural households dependent on coal were provided with subsidized low-sulfur coal briquettes (National Energy Administration, 2013). These briquettes were cheaper and emitted less pollution. Each autumn, village heads would collect household coal needs and place bulk orders with government-designated suppliers, ensuring controlled distribution directly to households. By 2017, this program had reached all villages in Beijing, placing all coal supply under governmental control. Upon a village's entry into the Clean Heating Policy, the government immediately ceased coal supply, removing the option for households to purchase coal independently.

Regarding the subsidies, financial support for electric heat pumps is calculated based on house size, set at 200 RMB per square meter, with a cap of 24,000 RMB ("People's Government of Beijing Municipality", 2017a). This arrangement generally covers the full cost of a heat pump adequate for each household's needs. Once a village enrolls in the Clean Heating Policy, selected heat pumps are installed and configured at no cost during the summer, ensuring they are easy to use with no additional operation required beyond turning the device on. And Additionally, households benefit from electricity subsidies provided by both municipal and district governments. Beijing operates a tiered pricing system for electricity, starting at 0.488 RMB/kWh for the initial 240 kWh per day, with rates increasing with higher usage ("People's Government of Beijing Municipality", 2024). An off-peak rate of 0.3 RMB/kWh applies for electricity used from 8 PM to 8 AM during the heating season (Nov 1st to March 31st) ("People's Government of Beijing Municipality", 2017b). Both the municipal and district governments further subsidize this rate by 0.1 RMB/kWh, reducing the actual electricity cost to 0.1 RMB/kWh. The subsidy electricity limit is up to 10,000 kWh per household for each heating season(("People's Government of Beijing Municipality", 2022a).

In addition to these financial incentives, the State Grid Corporation of China enhances the electrical infrastructure in villages prior to the winter of their enrollment year, ensuring a reliable power supply.

As a result of these comprehensive measures, by the end of 2020, 86.4% of rural villages and 1.38 million rural households in Beijing had successfully transitioned to clean heating ("People's Government of Beijing Municipality", 2022b).

2.3.2 Beijing Household Energy Transition Project

To assess the impact of the Clean Heating policy on air quality and public health, the Beijing Household Energy Transition Project[4] conducted four consecutive seasons of data collection during the winters of 2018-19, 2019-20, 2020-21 , and 2021-22 (referred to as Season 1 to Season 4, respectively) in rural Beijing. The study encompassed 50 villages across four administrative districts—Fangshan, Huairou, Mentougou, and Miyun. Initially, these villages predominantly utilized coal for heating and were eligible for, but not yet participating in, the clean heating policy.

The research project implemented a staggered treatment approach, as illustrated in Figure 2.2 and Table 2.1. By the conclusion of the study, 20 villages had adopted the clean heating policy, while 30 remained as control groups. The treatment villages were equipped with subsidized heat pumps during the preceding summer, and households began paying for electricity at reduced rates starting the subsequent winter. For the study, approximately 20 households per village were randomly selected to participate, and these households were followed up in subsequent years to complete detailed surveys on health and energy use. In addition, data on indoor air quality and temperature were collected to assess the environmental impacts. If a household drops out of the study, a new household is enrolled to maintain the sample size. The majority of the households consist of two people or fewer, with an average respondent age of 60. Most participants have received only primary education, are married, and work in agriculture. 40% of the respondents are male.

During Season 3, coinciding with the peak of the COVID-19 pandemic, data collection was confined to indoor air quality, temperature, and stove use due to travel restrictions, limiting

[4]The project is primarily led by Prof. Jill Baumgartner and Sam Harper at McGill University, with funding from the Health Effects Institute (HEI), the Canadian Institutes of Health Research (CIHR), the Social Sciences and Humanities Research Council (SSHRC), and the Government of Canada.

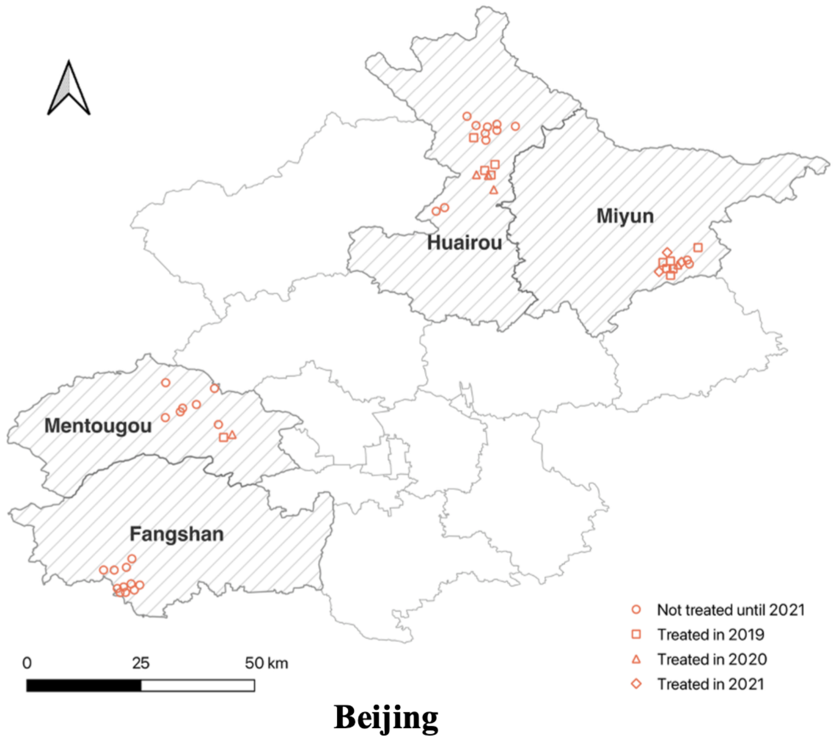


Figure 2.2: Treatment Map

District	No Ban	Treated in S2	Treated in S3	Treated in S4	Total
Miyun	2	6	1	3	12
Huairou	10	4	4	0	18
Mentougou	7	0	2	0	9
Fangshan	11	0	0	0	11
Total	30	10	7	3	50

Table 2.1: Treatment Schedule of clean heating policy

comprehensive data on other fuel types and heating devices. Consequently, for the purposes of heating analysis, I will focus on data from Seasons 1, 2, and 4 only.

2.3.3 Data Description

The study collected comprehensive information on household energy choices and living conditions, including income, expenditure, demographics, health, educational status, occupation, farming activities, asset ownership, and other variables at both the household and individual levels. A strength of this dataset is its detailed recording of energy choices for all purposes, including cooking, heating, and boiling hot drinking water. Additionally, it captured detailed heating hours for various heating devices in each room of the households (see Table 2.2). In this study, I will utilize this information to delve into household heating energy choices.

Heat source by room		
Room	Q. 1 How is the room heated in winter?	Q. 2 How many hours, on average, is the heat source on each day in the winter?
(1) Common room (2) Bedroom (3) Kitchen (4) Bathroom (5) Other (6) Storage	(1) Not heated (2) Hot water wall radiator (3) Heated floor (by water) (4) Kang - wood (5) Kang - coal (6) Coal stove (directly) (7) Wood stove (directly) (8) ATA heat pump (9) Thermal storage (electricity) (10) Mobile electric heater (11) Electric blanket (12) Air conditioner (13) Heated floor (by electricity) (90) Other	(1) Seldom or only on special occasions (2) < 2 hours (3) $2 - 4$ hours (4) $4 - 6$ hours (5) $6 - 8$ hours (6) $8 - 10$ hours (7) $10 - 12$ hours (8) $12 - 16$ hours (9) All day and night

Table 2.2: Survey Questions on heat sources by room during winter

Dependent Variables

To comprehensively analyze household heating dynamics, this study examines two pivotal categories of outcomes: heating fuel choices and heating behaviors. Heating fuel choices represent the energy sources selected by households for heating, which directly influence environmental and health outcomes due to emissions. Heating behaviors, on the other hand, encompass the frequency and intensity of fuel usage, as well as the distribution of heating needs across various fuels and heating devices. All dependent variables may be time-varying.

Heating Fuel Pattern The first set of dependent variables in this study are binary indicators representing the types of energy households choose for heating[5]. To align with the theories of fuel stacking and energy ladder, households in this study are categorized into three groups based on their

[5] In this study, I concentrate on fuels and devices used regularly for daily heating, excluding those used infrequently. If a household indicates that a device is used ”seldom or only on special occasions” for heating, it is classified as not being used for routine heating purposes.

energy usage: clean, mixed, and polluting. Households exclusively using electricity, natural gas, or LPG for heating fall into the ”clean energy” category. Those relying on traditional biomass fuels like firewood, charcoal, and crop residues are categorized as ”polluting fuel” users. The ”mixed fuels” category encompasses households that use a combination of both clean and polluting fuels, employing modern energy sources like electricity and gas for certain heating tasks while continuing to depend on traditional biomass fuels for others.

Primary Heating Fuel The primary heating fuels form the second set of dependent variables in this study, represented as binary variables. Although fuel stacking occurs, the analysis of primary heating fuel remains crucial for a comprehensive understanding of household fuel choices, as most households, at least within this study, predominantly rely on a single heating device to meet the majority of their heating needs. I utilize detailed usage data for various heating devices in each room to compute the aggregate heating hours for each device and fuel type. The primary heating device is identified as the one with the maximum aggregate heating hours across all rooms[6].Consequently, the primary heating fuel is the fuel used by the primary heating device. The vast majority of the households in the study use electricity, coal, and wood as their primary heating fuels.

Total Heating Hours across all rooms The first set of heating behaviours variables captures the total heating hours for all rooms on a typical day, organized by fuel pattern, fuel type, or specific heating device. This approach enables an exploration of how households allocate their heating hours among different types of fuels—be it clean, polluting, or more specifically, options like electricity, coal, and wood. The study also delves into the heating behaviors associated with various electric heating devices, hypothesizing that the adoption of more efficient heat pumps might reduce reliance on less efficient alternatives. However, the availability of electricity subsidies might encourage the continued use of other electric heating devices, complicating this transition. Thus, the analysis examines how heating hours are allocated among three distinct categories: subsidized heat pumps,

[6]In cases where two or more devices have identical heating duration, the one with the highest heating efficiency is designated as the primary device. The hierarchy of efficiency levels is as following: Heat Pump > Radiator(Coal) > Heated floor by electricity > AC > Kang(Wood) = Kang(Coal) > Mobile electric heater > Coal stove

other high-efficiency electric devices such as heated floors and air conditioners, and less efficient portable electric heaters.

Average Heating Hours per Room The second set of heating behaviours variables measures the average heating hours, defined as the total heating hours divided by the number of rooms heated daily. For this study, 'daily heating' is characterized by rooms consistently heated for a specified duration each day, excluding those heated only occasionally.

The independent variables include fuel prices, household income, age, sex, education, marital status, occupation of the respondent, household size, agricultural assets, and other factors. A summary of these variables will be presented in Table 2.3 in a later section.

2.4 Clean Heating Policy Evaluation

This section addresses the first research question of the paper: how does the clean heating policy impact the heating fuel choices and behaviors of households? I begin with a descriptive summary of the energy dynamics over the sample period, followed by an outline of the empirical strategy. Subsequently, I present the estimated treatment effects on the heating fuel choices and behaviors of households.

2.4.1 Descriptive Energy Dynamics

Figure 2.3 illustrates the current usage rates of electricity, coal, and biomass. For the treated villages, coal use decreases drastically following the implementation of the clean heating policy, although a few households continue to use coal. Meanwhile, the electricity usage rate for heating in these villages increases to nearly 100%. In the control villages, there is also an observed trend of increasing electricity use and decreasing coal use over time, even without a coal ban. This indicates that households are gradually transitioning from coal to electricity on their own. Understanding the determinants behind this spontaneous transition will be valuable for future policy design. Regarding wood usage, there is a slow decreasing trend in both the control and treated groups.

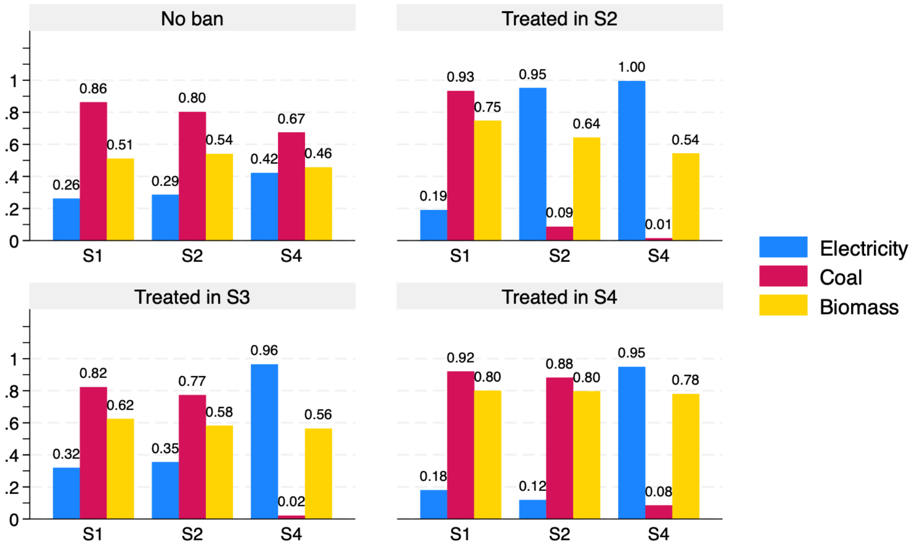


Figure 2.3: Current Use of Heating Fuel

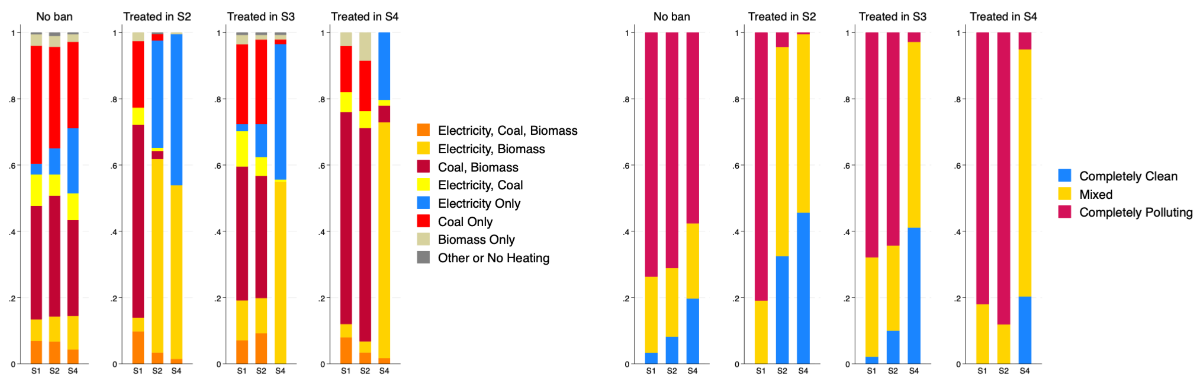


Figure 2.4: Heating Fuel Combination

Figure 2.4 illustrates the heating fuel combinations for households, highlighting how they change their heating fuel mix in response to the clean heating policy. There are two main scenarios for treated households: some completely switch to electricity heating, while others do not fully switch and instead partially use biomass as a substitute for coal. Consequently, more households adopt either a completely clean or a mixed heating style. This trend is also observed among households in the control villages. The Sankey diagrams in Figures 2.A.1 and 2.A.2 in the appendix visually illustrate the transitions households make between different fuel patterns.

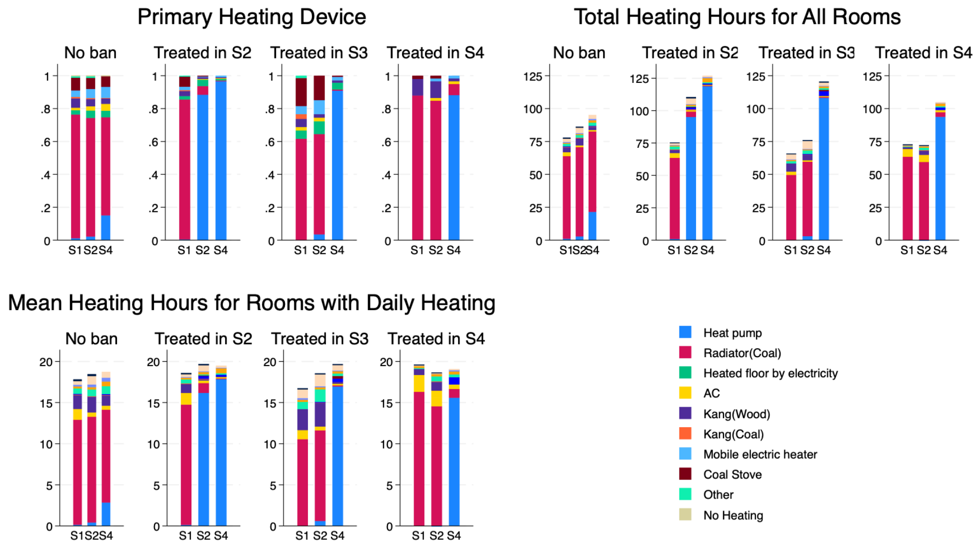


Figure 2.5: Heating Behaviours By Device

Figures 2.5 and 2.6 present the heating behaviors by device and fuel, respectively. Before the coal ban, radiators fueled by coal were the main heating devices for most households. After the coal ban, the subsidized heat pump became the primary device for most treated households. Even in the control villages, an increasing number of households are adopting electric heating. Figure 2.A.3 in appendix illustrates the transitions in primary heating fuel.

Figures 2.5 and 2.6 present the total daily heating hours across all rooms by heating device and fuel type, respectively, as well as the mean heating hours across rooms with daily heating in the treated group. Notably, there is a significant shift in heating hours from coal-powered devices to subsidized heat pumps for the treated households. Furthermore, there has been a marked increase in the total heating hours for the treated post-clean heating policy implementation, while the mean heating hours for rooms with daily heating have remained relatively stable. This suggests that households are choosing to heat more rooms (see Figure 2.A.4) rather than increasing the heating duration per room. There is a notable reduction in heating hours using polluting fuels and devices, with a corresponding increase in the use of clean energy, particularly electricity, which supports the policy's objectives.

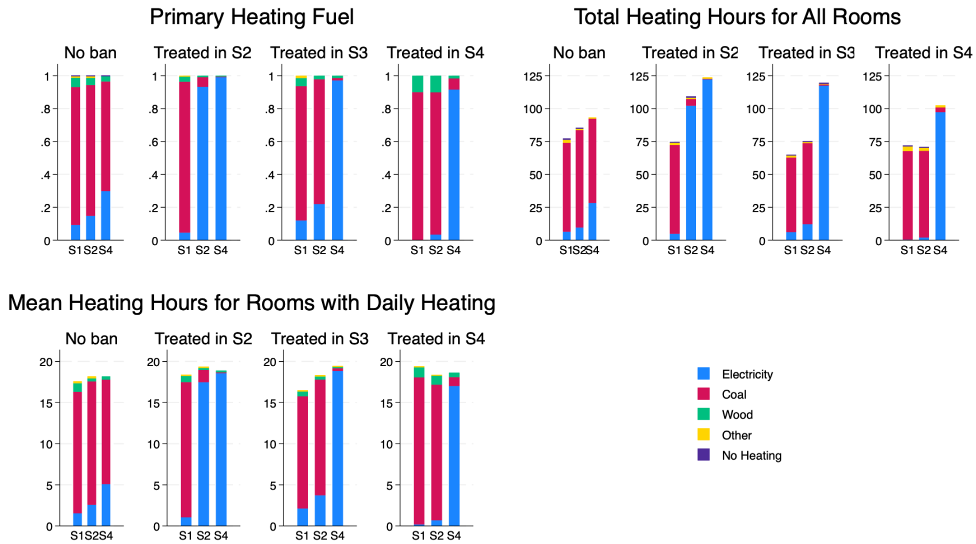


Figure 2.6: Heating Behaviours By Fuel Type

To summarize, household heating dynamics have changed significantly following the policy implementation, as evidenced by an increased adoption of electricity and heat pumps for heating. This shift is characterized by a substantial reduction in coal usage, though coal was not completely phased out, and wood usage remained unchanged. Following the ban, there was a notable rise in total heating hours across all rooms utilizing electricity and subsidized heat pumps. However, the average heating hours per room saw only a minimal increase, indicating that households opted to heat more rooms rather than extend the heating duration in any single room. Additionally, a similar, albeit less pronounced, transition from coal to electricity was observed in households not directly affected by the ban, suggesting broader underlying trends in energy preferences.

2.4.2 Empirical Strategy

Two Way Fixed Effects(TWFE) Model

Based on the study design of the Beijing Household Energy Transition Project, I utilized the Difference-in-Difference (DID) method to assess the effect of the clean heating policy by comparing changes in the treatment group against the control group before and after policy implementation. I formulated the following two-way fixed effect regression equation to estimate the policy's treatment effect:

$$Y_{it} = \alpha_i + \gamma_t + \beta \times D_{it} + X_{it} + \varepsilon_{it} \tag{2.1}$$

where Y_{it} represents the heating outcome for household i in period t; D_{it} indicates treatment status, with 1 if unit i is treated in period t and 0 otherwise; X_{it} are control variables; γ_t is the time fixed effect for period t; α_i is the household fixed effect for household i; ε_{it} is the error term; and β measures the average treatment effect.

A critical assumption of the DID method is the parallel trend assumption, which posits that in the absence of treatment, the average outcome for both treated and untreated groups would have followed a similar trend. I test for pre-existing trend differences as a check on this assumption by estimating the following dynamic TWFE model:

$$Y_{it} = \alpha_i + \gamma_t + \sum_{r \neq 0} 1[R_{it} = r]\beta_r \times D_{it} + X_{it} + \varepsilon_{it} \tag{2.2}$$

where $R_{it} = t - T_{treat,i} + 1$ denotes the time relative to treatment. A non-violation of the parallel trends is indicated by $\beta_r = 0$ for all $r < 0$.

Advanced DID Estimators for Staggered Treatment

Recent studies have highlighted potential issues with the two-way fixed effects model when dealing with staggered treatment adoption(Goodman-Bacon, 2021; Callaway and Sant'Anna, 2021; Gardner, 2022; Borusyak et al., 2024; Sun and Abraham, 2021; De Chaisemartin and d'Haultfoeuille, 2020). This model typically assumes homogeneity in treatment effects—implying that all units are impacted uniformly by the treatment regardless of the timing of its reception or the duration since its initiation. This can be problematic as it includes early-treated units as controls for those treated later, potentially distorting the estimated effects if the uniform effect assumption does not hold.

To address these issues, various improved Difference-in-Differences (DID) estimators have been developed that relax the homogeneity assumption, allowing for the treatment effects to vary across units and over time. A particularly notable approach among these new methodologies is proposed by Callaway and Sant'Anna, 2021, which I will employ as a robustness check in my analysis. This method avoids using already-treated groups as controls and allows for the estimation of treatment effects that are specific to each group and period. These effects can then be aggregated as needed—whether as an overall effect, or segmented by treatment group, calendar time, or duration since treatment initiation. By using this method, we can address the potential biases inherent in traditional DID estimators and more accurately capture the varied impacts of the treatment across different contexts and times.

Advanced DID Estimators for Non-parallel Trend

In some cases, a significant pre-treatment trend is detected, suggesting that the parallel trends assumption may not hold. However, researchers may still wish to draw conclusions about the treatment effects, particularly if the violation of parallel trends is relatively small in magnitude. The conventional approach does not provide clear guidance on how to proceed in such cases (Roth et al., 2023).

Rambachan and Roth, 2023 introduce an innovative DID estimator that incorporates pre-treatment trends to assess and adjust for potential violations of the parallel trends assumption in DID analyses. Their approach formalizes two key methods: Bounds on relative magnitudes and Smoothness restrictions. The former sets limits on the extent of post-treatment trend deviations relative to pre-treatment, with parameters like $\bar{M}$ determining the allowable magnitude of these deviations. The latter imposes constraints on the rate of change in pre-treatment trends, ensuring that post-treatment deviations do not stray significantly from a linear extrapolation of earlier trends. These innovative methods enhance the robustness of DID estimators, allowing for more reliable causal inference in complex scenarios where traditional assumptions may not hold. For further details, please refer to Rambachan and Roth, 2023.

2.4.3 Empirical Results

In this section, I analyze the impacts of the clean heating policy on two primary aspects: heating fuel choice and heating behaviors. My primary objective is to understand how this policy influences household decisions regarding which fuel portfolios to adopt. Recognizing that households frequently use multiple fuels simultaneously, it is crucial to identify not only the overall fuel patterns, but also the primary heating fuel they use. To this end, I utilize two sets of binary dependent variables: one set categorizes heating fuel patterns as clean, mixed, or polluting, and another identifies the primary heating fuel, which may be electricity, coal, or wood.

After identifying the fuel choices, I delve into the related yet different decisions about heating behaviors. This includes how much households choose to heat and how they allocate heating hours among the selected fuels. I assess the total heating hours per fuel source across all rooms on a typical day. For a detailed analysis of these behaviors, I also include in the appendix the average heating hours per room, broken down by each fuel source, for rooms that are heated daily.

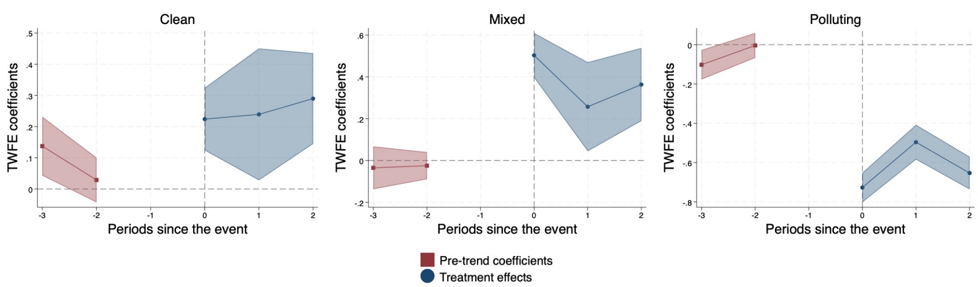
Table 2.3 presents the descriptive statistics for the socioeconomic variables of households at baseline, clearly indicating that there are no significant differences between the treated and control groups. This provides a robust foundation for further analysis of the impacts of the clean heating policy.

	Control					Ever Treated				
	Mean	SD	Min	Max	N	Mean	SD	Min	Max	N
Covariate										
Price of Coal (RMB/tonne)	541	118	350	1000	581	516	67	0	1000	385
Income per Capita(RMB/year)	16310	11168	3000	86000	581	16631	13522	3000	135000	385
Household Size	2.48	1.23	1	8	581	2.30	0.99	1	6	385
Age	60.34	9.36	33.75	88.95	581	60.88	9.21	37.24	85.98	385
Male	0.41	0.49	0.00	1.00	581	0.41	0.49	0.00	1.00	385
Number of Children Under 5	0.10	0.35	0.00	3.00	581	0.08	0.30	0.00	2.00	385
Education										
Primary School	0.68	0.47	0.00	1.00	581	0.73	0.44	0.00	1.00	385
Secondary or High School	0.21	0.41	0.00	1.00	581	0.14	0.35	0.00	1.00	385
Higher Education	0.08	0.27	0.00	1.00	581	0.07	0.25	0.00	1.00	385
Work in Agriculture	0.62	0.49	0.00	1.00	581	0.67	0.47	0.00	1.00	385
Married	0.90	0.30	0.00	1.00	581	0.87	0.33	0.00	1.00	385
Agriculture Land(mu[7]	0.88	1.56	0	15	581	0.77	1.31	0	10	385
Forest Land(mu)	3.04	21.53	0	500	581	6.03	31.93	0	500	385

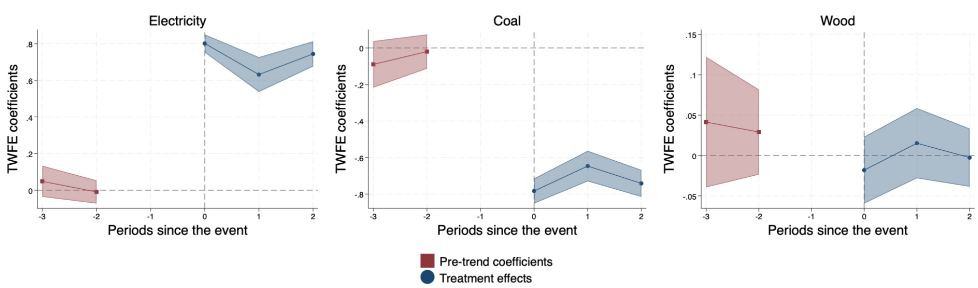
Table 2.3: Summary Statistics at Baseline

Heating Fuel Choices

Parallel Trend Figure 2.7 presents the results of the pre-trend testing for the binary variables representing heating fuel choices, with the -1 period omitted as the reference period. Panel A examines heating fuel patterns, while Panel B focuses on primary heating fuel. Significant pre-treatment effects were observed for the 'Clean' and 'Polluting' heating fuel patterns three periods before the implementation of the clean heating policy, potentially suggesting violation of the parallel trends assumption essential. However, the direction of potential pre-trends is opposite to that of the post-trends, mitigating concerns about overestimating treatment effects. To further enhance the robustness of my findings, I will employ advanced DiD estimators that can accommodate non-parallel trends in later section. For all other variables, no significant pre-trends were detected.



Panel A: Heating Fuel Pattern



Panel B: Primary Heating Fuel

Figure 2.7: Pre-trend Test for Heating Fuel Choices

Note: Dynamic treatment effects estimated by the TWFE Model are displayed alongside their 95% confidence intervals. The -1 period is omitted as the reference period. Units for all variables are shown as probabilities.

Estimation Results Table 2.4 outlines the estimated results for binary heating fuel choices by TWFE model. Panel A details the results for heating fuel patterns, while Panel B focuses on primary heating fuels. The average treatment effect is noted at the top of each panel, with the heterogeneous effects by cohort, treated in different waves, displayed in the subsequent rows.

From Panel A, it is evident that the clean heating policy resulted in a 63% decrease in the probability of households choosing entirely polluting fuel portfolios, a 40% increase in adopting a mixed fuel strategy, and a 23% increase in choosing entirely clean fuels. These results suggest that the clean heating policy significantly accelerated the transition to cleaner heating patterns. However, the policy did not completely eliminate the use of polluting fuels. Instead, a substantial proportion of households opted for a mixed fuel strategy, incorporating both clean and polluting fuels. This indicates that while the policy led to a shift toward cleaner fuels, the transition was not absolute, and many households continue to rely on a combination of both clean and traditional energy sources.

The transition from polluting to clean fuels was most marked in the group treated during season 2, and less pronounced in the group treated during season 3, possibly because this latter group had already adopted cleaner fuels before the clean heating policy (see Figure 2.4). The effects for transitioning fully to clean fuels were not statistically significant at the 5% level for the groups treated later in season 3 and season 4, and the impact was minimal for the group treated in season 4. This indicates that the shift to a completely clean fuel portfolio primarily occurred in the group treated early in season 2, underscoring that such a transition requires time.

The findings support the fuel stacking hypothesis, indicating that households commonly utilize multiple fuels concurrently. Despite this, most heating hours are primarily concentrated across one or two fuels or devices. As detailed in Table 2.5, 87% of households predominantly use their primary heating device for more than 80% of their heating duration. Furthermore, for over 99% of households, the primary and secondary heating devices combined account for 80% of total heating hours. This pattern shows that while fuel stacking is common, typically one or two main heating fuels dominate, making the analysis of primary heating fuel crucial to understanding overall household fuel choices.

Panel A: Heating fuel pattern						
	Clean		Mixed		Polluting	
ATT	0.227	***	0.403	***	-0.630	***
	(0.054)		(0.059)		(0.037)	
G2	0.288	***	0.416	***	-0.704	***
	(0.062)		(0.075)		(0.044)	
G3	0.219	*	0.274	**	-0.493	***
	(0.113)		(0.114)		(0.052)	
G4	0.073		0.583	***	-0.656	***
	(0.046)		(0.050)		(0.027)	
Panel B: Primary heating fuel						
	Electricity		Coal		Wood	
ATT	0.733	***	-0.721	***	-0.012	
	(0.030)		(0.032)		(0.016)	
G2	0.791	***	-0.791	***	0.001	
	(0.028)		(0.028)		(0.015)	
G3	0.632	***	-0.631	***	-0.001	
	(0.048)		(0.041)		(0.020)	
G4	0.735	***	-0.662	***	-0.072	
	(0.038)		(0.100)		(0.075)	

Table 2.4: Treatment Effects on Binary Heating Fuel Choices

[a] *** p<.01, ** p<.05, * p<.1
[b] G# refers to group treated in season #
[c] All treatment effects are estimated by TWFE Model.

Heating Device Combination	Freq.	Percent	Cum.
Primary Only	222	7.49	7.49
Primary > 80%	2357	79.55	87.04
Primary+Secondary	99	3.34	90.38
Primary+Secondary > 80%	270	9.11	99.49
Primary+Secondary+Tertiary	6	0.20	99.70
Primary+Secondary+Tertiary > 80%	9	0.30	100.00
Total	2963	100.00	

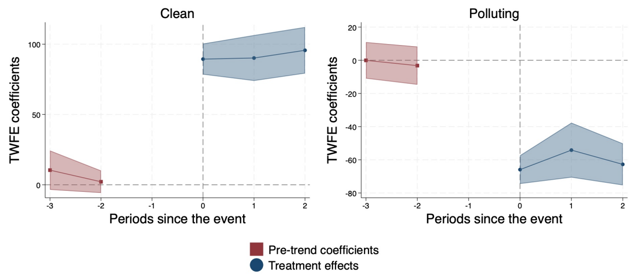
Table 2.5: Heating Device Combination

In Panel B, there is a 73% increase in the probability of using electricity as the primary heating fuel following the implementation of the clean heating policy, accompanied by a nearly equal decrease in coal use, with no significant change observed for wood. This suggests that while a fully clean heating portfolio may not be widely adopted, a substantial number of households previously reliant on coal have shifted to using electricity as their primary heating fuel. This transition is consistent with the objectives of the clean heating policy, which aims to reduce the reliance on polluting fuels and encourage the use of cleaner alternatives. Many households have adopted a mixed fuel portfolio, combining electricity and wood, as indicated by earlier panel analysis. However, the results from this panel clarify that wood is not used as a primary heating fuel but rather as a supplementary source. This is less detrimental to regional air quality and climate due to the significant reduction in the use of coal. The transition towards a completely clean heating portfolio would involve phasing out even this supplementary use of wood, a step that, while challenging, raises fewer environmental concerns.

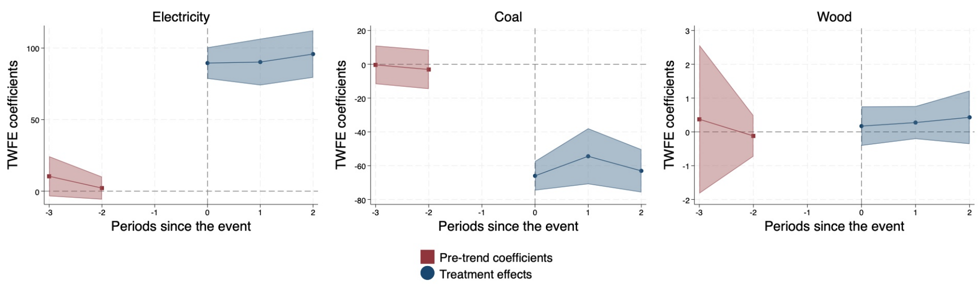
Effects were slightly less marked in households from the season 3 cohort, which already exhibited lower primary reliance on coal before the policy implementation.

Heating Behaviours Having examined the heating fuel choices of households, it is crucial to extend my analysis to heating behaviors, that is, how households allocate heating hours among the fuels or devices they have chosen. Heating behaviors—specifically, the duration of fuel use—offer additional insights into the broader impacts of clean heating policy, including efficiency, cost implications, health, environmental and welfare effects. Analyzing heating behaviors is crucial for developing more precise and effective interventions. These interventions should not only influence the choice of fuel but also enhance how these fuels are used to maximize benefits and minimize risks. By exploring heating behaviors, I bridge the gap between identifying fuel choices and understanding their broader implications, thus offering a comprehensive view of household energy consumption patterns following the clean heating policy.

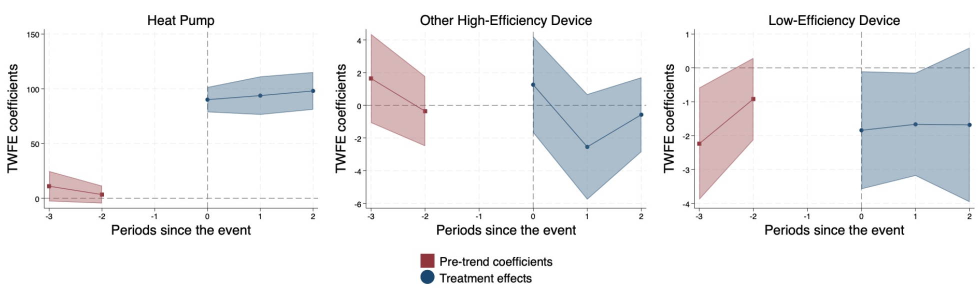
Parallel Trend Figure 2.8 displays pre-trends test results for total daily heating hours across all rooms, categorized by fuel type and heating device, while Figure 2.9 shows the average daily heating hours for rooms that are regularly heated, also broken down by fuel type and device. For most variables, there is no evidence of pre-trends. However, significant pre-treatment effects are observed for both total and mean heating hours for low-efficiency electric devices. To address this, I utilize an advanced difference-in-differences (DID) estimator later in the analysis that relaxes the parallel trends assumption, thereby enhancing the robustness of the results.



Panel A: Total Heating Hours by Fuel Pattern



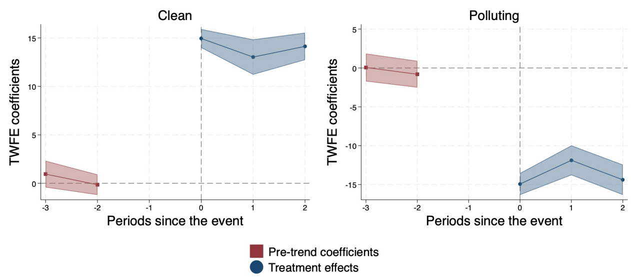
Panel B: Total Heating Hours by Fuel Type



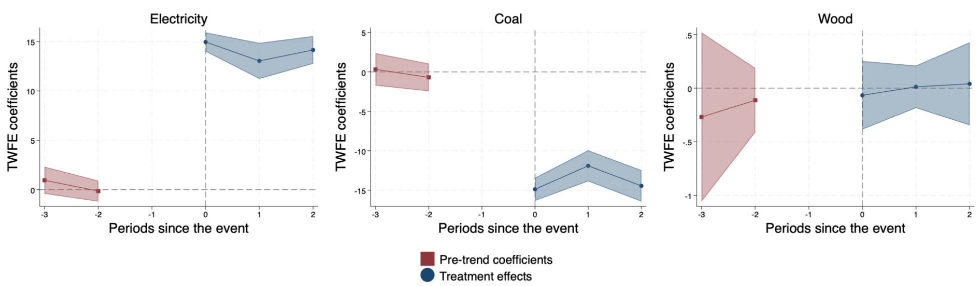
Panel C: Total Heating Hours by Heating Device

Figure 2.8: Pre-trend Test for Heating Behaviors

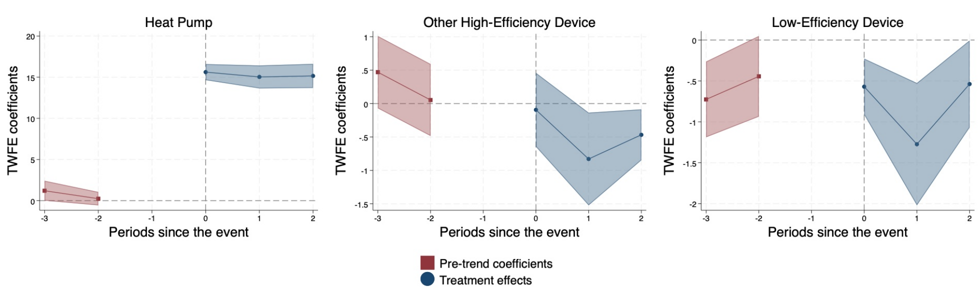
Note: Dynamic treatment effects estimated by the TWFE Model are displayed alongside their 95% confidence intervals. The -1 period is omitted as the reference period. Units for all variables are in hours.



Panel A: Mean Heating Hours by Fuel Pattern



Panel B: Mean Heating Hours by Fuel Type



Panel C: Mean Heating Hours by Heating Device

Figure 2.9: Pre-trends Test for Heating Behaviors

Note: Dynamic treatment effects estimated by the TWFE Model are displayed alongside their 95% confidence intervals. The -1 period is omitted as the reference period. Units for all variables are in hours.

Estimation Results Table 2.6 details the total heating hours across all rooms on a typical day by fuel sources. Panel A presents the daily total heating hours across all rooms, categorized by clean or polluting fuels. Panel B shows the breakdown by fuel types, including electricity, coal, and wood. Panel C details the total heating hours by types of electric heating devices.

In Panel A, there is an increase of 90 hours in heating by clean energy sources and a decrease of 61 hours in polluting heating hours, resulting in a net increase of approximately 29 room hours by all fuels. The effects on clean heating hours are notably stronger for the group treated earlier. Conversely, the effects on polluting heating fuels are more pronounced for the groups treated in seasons 2 and 4, and less so for the group treated in season 3. A combined analysis reveals that the group treated in season 3 experienced the most significant increase in total heating hours by all fuels, approximately 36 room hours.

Panel B analyzes specific fuel types. Heating hours from electricity align with those from clean sources, as electricity is the primary clean fuel used by the majority of households. Similarly, coal heating hours closely match polluting hours. There are no significant shifts in heating hours by wood, suggesting that the clean heating policy did not significantly affect other polluting fuels and that households did not substantially substitute coal with wood following the ban.

In Panel C, the impact on heating hours by heat pumps is slightly larger in magnitude compared to those by electricity, attributed to a reduction in the use of low-efficiency electric devices. High-efficiency electric devices other than heat pumps such as heated floors and air conditioners show no significant changes. As found in earlier results, the more pronounced effects for increases in heat pump heating hours occur in earlier treated groups. The results indicate that households have substituted lower-efficiency devices with more efficient heat pumps and have not exploited electricity subsidies to increase heating hours by other electric heating devices.

From the analysis of total room heating hours, it is evident that treated households have not only substituted coal heating with heat pump heating, but also extended their total heating hours following the clean heating policy. This raises an important question about the allocation of these additional heating hours: Are households heating more rooms or simply extending the heating duration per room? This question is crucial as it helps to understand whether the increase

Panel A: By Fuel Pattern						
	Clean		Polluting			
ATT	89.675	***	-60.618	***		
	(5.138)		(4.830)			
G2	95.380	***	-66.499	***		
	(6.688)		(5.235)			
G3	88.544	***	-52.175	***		
	(7.992)		(7.160)			
G4	76.846	***	-58.705	***		
	(10.893)		(7.574)			
Panel B: By Fuel Type						
	Electricity		Coal		Wood	
ATT	89.772	***	-60.815	***	0.255	
	(5.131)		(4.902)		(0.277)	
G2	95.505	***	-66.912	***	0.557	
	(6.673)		(5.237)		(0.365)	
G3	88.592	***	-52.429	***	0.259	
	(7.986)		(7.209)		(0.283)	
G4	76.902	***	-57.929	***	-0.778	*
	(10.888)		(7.931)		(0.410)	
Panel C: By Electric Heating Device						
	Heat Pump		Other High -efficiency Ele-Device		Low-efficiency Ele-Device	
ATT	91.350	***	-0.329		-1.485	**
	(5.377)		(1.072)		(0.654)	
G2	96.889	***	0.548		-1.932	*
	(6.881)		(1.128)		(1.151)	
G3	91.548	***	-2.391	*	-1.225	*
	(8.674)		(1.411)		(0.663)	
G4	77.069	***	0.230		-0.396	
	(11.304)		(1.029)		(0.625)	

Table 2.6: Treatment Effects on Total Heating Hours

[a] *** p<.01, ** p<.05, * p<.1
[b] *G#* refers to group treated in season #
[c] All treatment effects are estimated by TWFE Model.
[d]Other high-efficiency electric heating devices includes AC, heated floor by electricity; low efficiency electric heating devices includes mobile electric heater.

in heating hours is spread across more spaces or concentrated in the same number of rooms with prolonged heating periods. The answer to this will provide further insights into the behavioral changes in household heating practices post-clean heating policy and help assess the efficiency and effectiveness of the policy in real household settings.

Table 2.7 reports the findings regarding mean heating hours. In the Panel A and B, there is an observed increase of 14 hours in clean or electricity heating, while polluting or coal heating hours decrease by 13.6 hours. This close parity in magnitude indicates no significant change in the average heating hours per room. Consistent with previous results, there are no notable changes in wood heating hours. This suggests that the rise in total heating hours likely results from an increase in the number of rooms heated, rather than extended heating periods per room.

Interestingly, the effects on average heating hours are milder for the group treated in season 3, which also shows the largest difference—approximately 1.5 hours—between average clean and polluting heating hours. This group also exhibits the largest increase in total heating hours, indicating they have reaped the most substantial thermal benefits, with both an increase in the number of rooms heated and longer heating duration. In contrast, the other groups primarily increased the number of rooms heated without significantly extending heating duration. This could be attributed to the fact that households in season 3 previously had shorter heating duration before the clean heating policy, whereas those in other groups might already have been heating their rooms continuously throughout the day and night.

In Panel C, which examines electric device usage, a decrease is observed in both high-efficiency and low-efficiency electric devices. The most significant reductions, around an hour per day for both device types, are seen in the group treated in season 3. This reduction highlights another thermal benefit for this group: they not only have more rooms heated and longer heating duration per room, but they also transition from lower to higher efficiency heating devices, specifically to subsidized heat pumps. Notably, the increase in heating hours attributed to subsidized heat pumps is more pronounced in groups treated earlier, emphasizing the benefits of early adoption in transitioning to more efficient energy usage.

Panel A: By Fuel Pattern						
	Clean		Polluting			
ATT	14.148	***	-13.603	***		
	(0.522)		(0.711)			
G2	14.834	***	-14.832	***		
	(0.517)		(0.814)			
G3	13.037	***	-11.470	***		
	(0.879)		(1.064)			
G4	13.795	***	-14.277	***		
	(0.864)		(1.492)			
Panel B: By Fuel Type						
	Electricity		Coal		Wood	
ATT	14.158	***	-13.623	***	0.010	
	(0.519)		(0.740)		(0.131)	
G2	14.843	***	-14.939	***	0.093	
	(0.514)		(0.828)		(0.195)	
G3	13.045	***	-11.502	***	0.032	
	(0.878)		(1.095)		(0.106)	
G4	13.805	***	-13.963	***	-0.314	
	(0.861)		(1.661)		(0.208)	
Panel C: By Electric Heating Device						
	Heat Pump		Other High -efficiency Ele-Device		Low-efficiency Ele-Device	
ATT	15.250	***	-0.434	**	-0.696	***
	(0.473)		(0.204)		(0.199)	
G2	15.755	***	-0.330		-0.581	***
	(0.457)		(0.240)		(0.209)	
G3	14.826	***	-0.827	***	-1.063	***
	(0.728)		(0.303)		(0.326)	
G4	14.010	***	-0.068		-0.137	
	(0.981)		(0.209)		(0.274)	

Table 2.7: Treatment Effects on Mean Heating Hours

[a] *** p<.01, ** p<.05, * p<.1
[b] *G#* refers to group treated in season #
[c] All treatment effects are estimated by TWFE Model.
[d]Other high-efficiency electric heating devices includes AC, heated floor by electricity; low efficiency electric heating devices includes mobile electric heater.

To summarize the findings, the clean heating policy significantly reduced coal usage, establishing electricity and subsidized heat pumps as the predominant heating sources, though it did not completely eliminate coal use, and wood usage remained relatively stable. It led households away from entirely polluting fuel portfolios towards mixed fuel use, with a limited transition to fully clean fuel portfolios.

Households benefited from the clean heating policy, seeing a substitution of clean heating hours for polluting ones and an increase in total room heating hours due to both more rooms being heated and extended heating duration per room. Additionally, there was a noticeable shift from lower to higher efficiency heating devices.

The effects of the coal ban were most pronounced in groups treated earlier and those with previously less efficient heating behaviors, highlighting a temporal sensitivity to the policy's implementation and its varied impact on different household groups.

2.4.4 Robustness Check

In this section, I perform multiple robustness checks on the estimators of the treatment effects of the clean heating policy. I conduct a placebo test to validate the estimated effects. Additionally, I employ advanced difference-in-differences (DID) methodologies tailored for staggered treatment adoption and potential violations of the parallel trend assumption, to ensure the robustness of my findings.

Placebo Tests

A placebo test is a widely used method in empirical research to assess the validity of causal inferences, particularly in studies that involve interventions or treatments. This test involves creating a scenario where no actual intervention is administered but the procedures of the study are otherwise followed as if it were. The goal is to establish whether the observed effects in the study can be attributed solely to the treatment or if they could be plausibly explained by other factors, such as the placebo effect, data mining biases, or external influences.

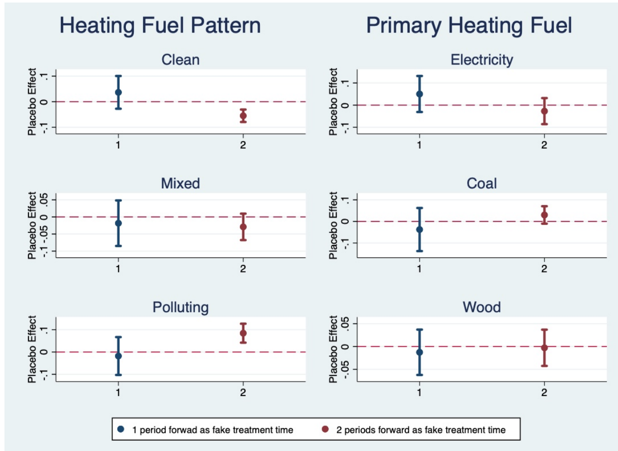
In-time Placebo Test In the context of policy evaluation, a in-time placebo test typically entails applying the analytical method to a time when the treatment is known not to have been applied. The analysis is conducted using data before the policy was implemented. By artificially advancing the treatment timing in the dataset to a period before the policy's actual implementation, the test assesses whether the effects attributed to the policy could plausibly arise in the absence of any intervention, simply due to pre-existing trends or random data fluctuations

Figure 2.10 presents the placebo effects. Temporal placebo effects are evident when the treatment timing is artificially advanced by two periods, indicating significant pre-existing differences between the treatment and control groups prior to the implementation of the coal ban policy. This observation underscores the importance of considering these differences when interpreting the treatment effects. Notably, the placebo effects exhibit the opposite sign to the treatment effects, suggesting a potential underestimation of the actual treatment effects.

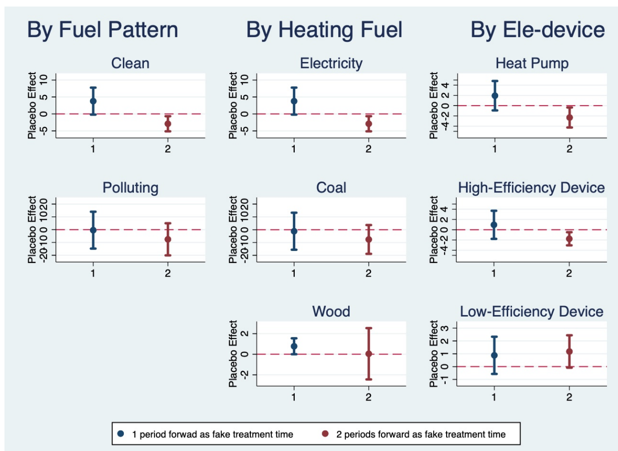
Space Placebo Test A space placebo test is a method used in research to assess the validity of causal inferences when studying the effects of spatially-based interventions. This test works by randomly assigning the treatment to regions where no intervention is actually applied, while keeping the treatment timing unchanged. The goal is to estimate ’fake’ treatment effects in these areas and compare them to the effects observed in the actual treatment region. This helps ensure that any observed effects in the treatment area are genuinely due to the intervention, rather than resulting from broader, non-local trends or coincidental spatial patterns. By conducting a space placebo test, I can control for spatial heterogeneity and potential confounding factors, which might otherwise lead to biased or erroneous conclusions about the effectiveness of a policy.

Figure 2.11 illustrates the results of the placebo test, depicted by the distribution of placebo effects shown as grey bars, with the estimated actual treatment effects marked by a red solid vertical line. The analysis demonstrates that for most of the estimated treatment effects, they are distinctly separate from or located at the tail of the placebo distribution, indicating that these effects are unlikely due to chance or spatial heterogeneity. For the few estimates that align more closely with the placebo distribution, the treatment effects themselves are statistically insignificant. This suggests that there is no evidence of confounding spatial placebo effects impacting the validity of the results.

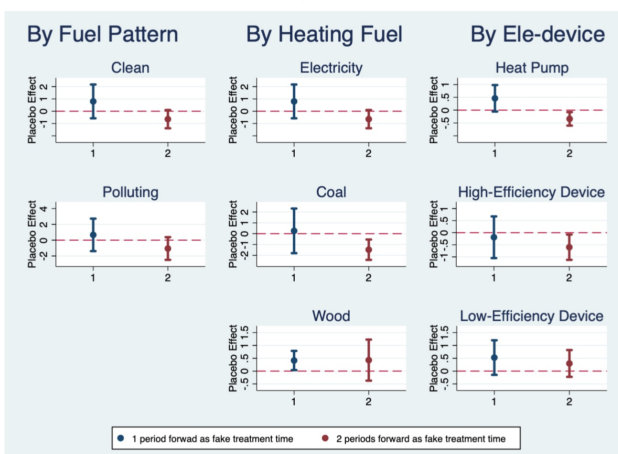
Mixed Placebo Test A mixed placebo test is a combination of temporal and spatial elements to strengthen the robustness of causal inferences, particularly in studies involving interventions or policies with both time and location variables. This test involves applying the original analysis to different time periods or locations where the intervention is known not to have been implemented, thereby mixing aspects of both time-based and space-based placebo tests. By doing so, I aim to



Panel A: Binary Heating Fuel Pattern



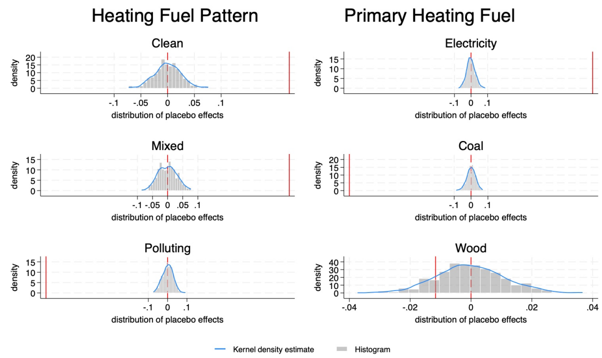
Panel B: Total Heating Hours by Fuel Sources



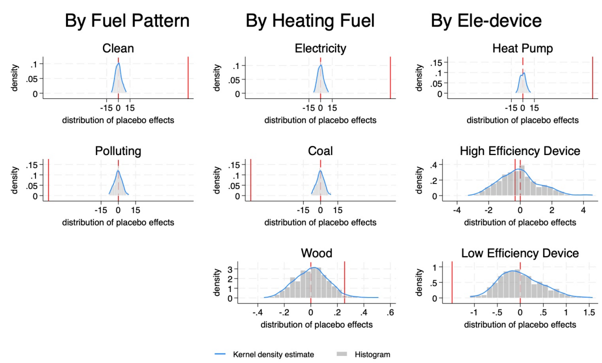
Panel C: Mean Heating Hours by Fuel Sources

Figure 2.10: In-Time Placebo Test

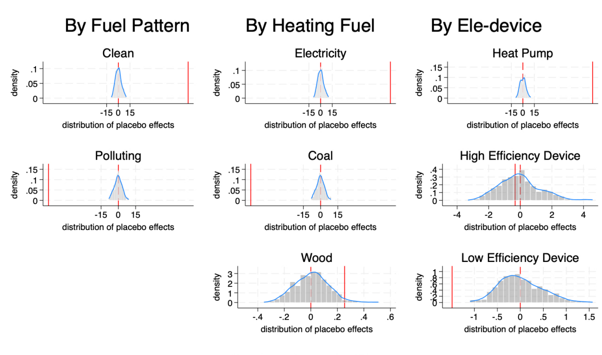
Note: Placebo treatment effects estimated by the TWFE Model are displayed alongside their 95% confidence intervals. Units for variables in Panel A are shown as probabilities , and the units for Panel B and C are presented in hours.



Panel A: Binary Heating Fuel Pattern



Panel B: Total Heating Hours by Fuel Sources



Panel C: Mean Heating Hours by Fuel Sources

Figure 2.11: Space Placebo Test

Note: The distribution of placebo effects are shown as grey bars, with the estimated actual treatment effects marked by a red solid vertical line.

detect any confounding influences that might falsely suggest an effect due to the time or location rather than the intervention itself.

Figure 2.14 displays the results concerning mixed placebo effects. Analogous to the findings on spatial placebo effects, this analysis also indicates that there are no significant mixed placebo effects.

Advanced DID Estimators for Staggered Treatment

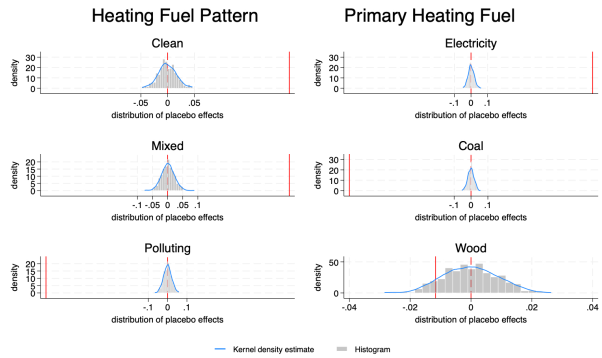
Figure 2.13 presents the results obtained using the Callaway and Sant'Anna, 2021 method designed for the staggered treatment compared with those derived from traditional Two-Way Fixed Effects (TWFE) estimators. The alignment between the two sets of results indicates a robust consistency in the treatment effects measured, reinforcing the reliability of the findings.

Advanced DID Estimators for Non-parallel Trend

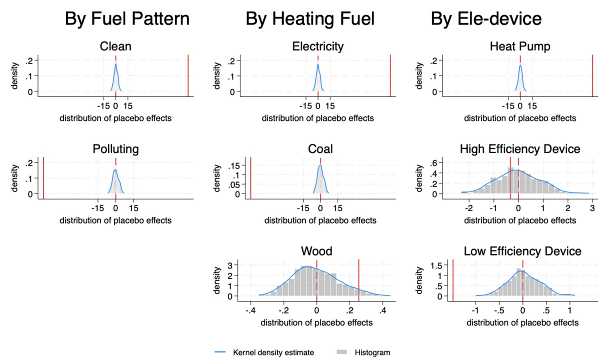
Figure 2.14 illustrates the application of Rambachan and Roth, 2023's first method to the data, which imposes limits on the extent of post-treatment trend deviations relative to pre-treatment trends. Parameters such as $\bar{M}$ determine the allowable magnitude of these deviations. The analysis demonstrates that the treatment effects remain robust, even when deviations are allowed up to twice the magnitude of pre-treatment trends for most variables with initially significant effects. However, it is important to note the broader confidence intervals, which reflect increased uncertainty around these estimates under the new methodological constraints.

2.5 Determinants of Fuel Choices and Heating Behaviors

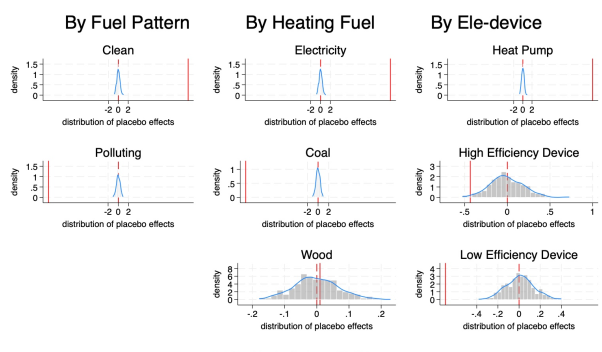
The analysis from the previous section shows that the clean heating policy has effectively facilitated the transition towards clean heating fuels, significantly reducing coal usage, however, it did not completely eradicate the use of polluting fuels, as evidenced by some households' continued coal use and the unchanged wood consumption levels. Concurrently, a spontaneous transition toward cleaner heating portfolios is observed in untreated villages, albeit without achieving a complete shift to clean



Panel A: Binary Heating Fuel Pattern



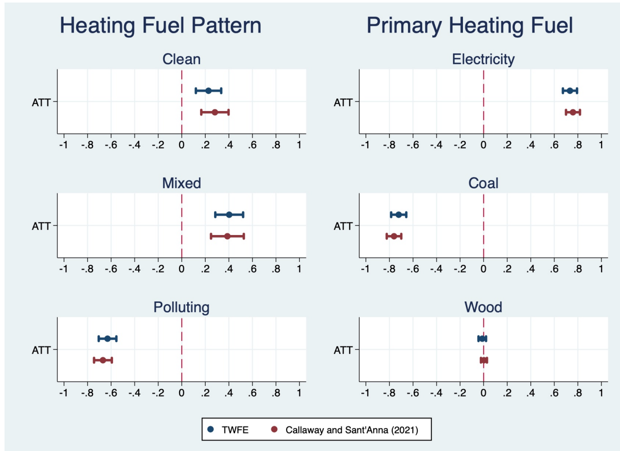
Panel B: Total Heating Hours by Fuel Sources



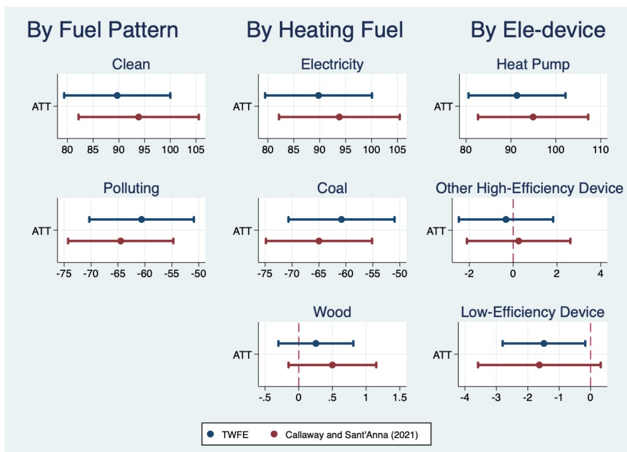
Panel C: Mean Heating Hours by Fuel Sources

Figure 2.12: Mixed Placebo Test

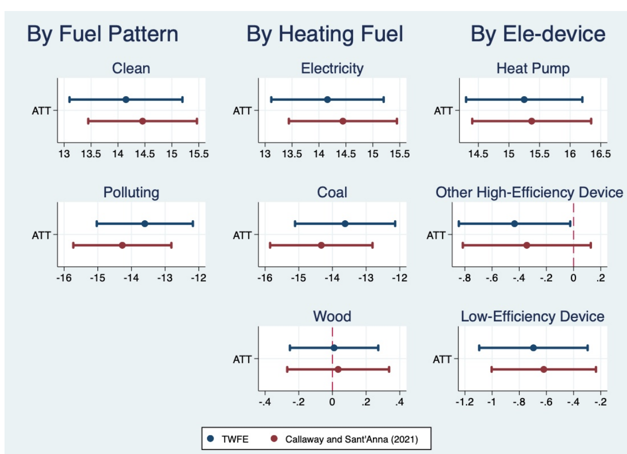
Note: The distribution of placebo effects are shown as grey bars, with the estimated actual treatment effects marked by a red solid vertical line.



Panel A: Binary Heating Fuel Pattern

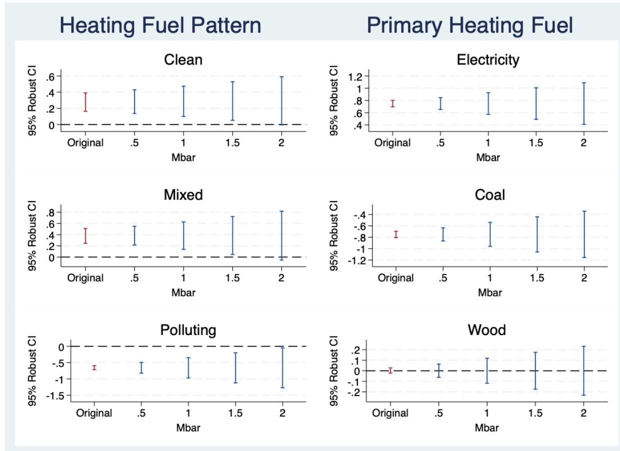


Panel B: Total Heating Hours by Fuel Sources

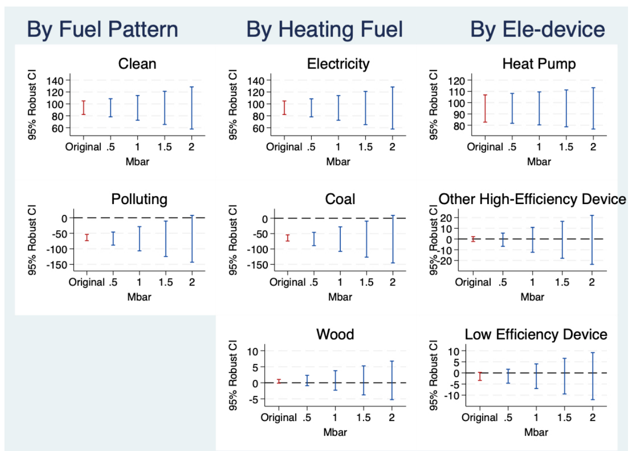


Panel C: Mean Heating Hours by Fuel Sources

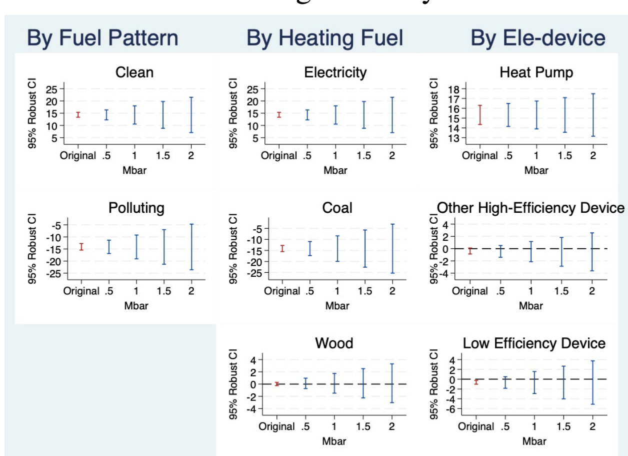
Figure 2.13: Comparison between TWFE and Advanced DID Estimator
Note: Treatment effects are presented with their 95% confidence intervals. In Panel B and C, variables are measured in hours, while probabilities are used as units for variables in Panel A.



Panel A: Binary Heating Fuel Pattern



Panel B: Total Heating Hours by Fuel Sources



Panel C: Mean Heating Hours by Fuel Sources

Figure 2.14: Advanced DID Estimators By Roth et al., 2023

Note: Treatment effects estimated by Roth et al., 2023 are presented with their 95% confidence intervals. In Panel B and C, variables are measured in hours, while probabilities are used as units for variables in Panel A.

fuels. These observations indicate that neither spontaneous transitions nor policy-driven bans have led households to fully embrace clean fuel portfolios, which necessitates a deeper investigation into the underlying factors that govern fuel usage decisions. Consequently, this analysis will now pivot to a detailed examination of the determinants influencing households' fuel choices and behaviors, to better understand the complexities of energy consumption dynamics.

This section is structured as follows: Initially, I introduce the theoretical model for heating fuel choices. Subsequently, I outline the Correlated Random Effects Generalized Ordered Probit (CORE-GOP) Model as the empirical strategy. Lastly, I explore the determinants influencing households' fuel choices and behaviors separately, acknowledging that these decisions may be driven by different processes. This analysis focuses exclusively on households not yet affected by the clean heating policy, to prevent confounding effects from the clean heating policy.

2.5.1 Theory Model

Economists have developed various fuel consumption models that conceptualize a household as a profit-maximizing producer and subsequently as a utility-maximizing consumer, based on the profits earned (Singh et al., 1986; Edwards and Langpap, 2005; Gupta and Köhlin, 2006). However, such models assume the absence of market failures and may not be fully applicable to rural households in developing countries. These households often face incomplete or absent markets, not only for fuels such as firewood, crop residues, and animal dung but also for agricultural products, labor, and credit (L. Chen et al., 2006; Shi et al., 2009; Bowlus and Sicular, 2003; Heltberg et al., 2000). Under these conditions of market failure, the decisions regarding production and consumption cannot be distinctly separated but are instead made jointly in a non-separable fashion, reflecting the complex inter-dependencies within the household's economic activities.

Building upon the insights of Muller and Yan, 2018 and recognizing the incomplete market for firewood in the sampled villages, I have developed an agricultural household fuel consumption model. This model posits that households maximize their utility over fuel use (F) , other good consumption(excluding fuels) C, and leisure (l):

$$U = U(C, F, l; Z)$$

Here, Z represents a vector of household characteristics that influence their preferences. The model assumes that the preferences for other consumption goods are sufficiently independent, allowing us to concentrate specifically on the decision-making processes related to fuel use. The fuel function integrates choices between polluting and clean fuels:

$$F = \left(F^P(F^{wood}, F^{coal}), F^C(F^{ele}); V\right)$$

where F^P and F^C denote polluting fuels (wood and coal) and clean fuels (electricity), respectively, with V encompassing variables pertinent to fuel choice.

The market for firewood is assumed to be missing: firewood consumption, F^{wood}, is tied directly to the amount gathered: $q^w(L_{wood})$, where L_{wood} is the labor time dedicated to collecting firewood.

$$F^{wood} = q^{wood}(L_{wood})$$

Households also engage in agricultural activities, producing crops Q_c based on labor L_c and fixed farm inputs A_c, influenced by land endowments ϕ

$$Q_c = Q_c(L_c, A_c; \phi)$$

This model's budget constraint reflects the integration of farm income, labor income, and other exogenous income sources:

$$p^C C + p^{coal} F^{coal} + p^{ele} F^{ele} = \left(p^H Q_c - p^{ac} A_c\right) + w L_{off} + Y_0$$

where L_{off} denotes the household labour time allocated to off-farm work; w is the wage rate; p^H is the crops price and Y_0 denotes the other exogenous income.

And the time constraint for the household balances time spent on various activities against the total available time:

$$L_{wood} + L_c + L_l + L_{off} + l = T$$

Finally, the Lagrangian of the constrained optimization problem is:

$$\begin{aligned}
\Gamma = {} & U\left[C, F\left(F^P\left(F^{wood}, F^{coal}\right), F^C\left(F^{ele}\right); V\right), l; Z\right] \\
& - \lambda\left[p^C C + p^{coal} F^{coal} + p^{ele} F^{ele}\right. \\
& \left. - \left(p^H Q_c(L_c, A_c; \phi) - p^{ac} A_c\right) - w L_{off} - Y_0\right] \\
& - \eta\left(L_{wood} + L_c + L_l + L_{off} + l - T\right) \\
& - \mu\left(F^{wood} - q^{wood}(L_{wood})\right)
\end{aligned}$$

where $\lambda, \eta$and μ are Lagrange multipliers. Assuming convexity in both preferences and technology sets, and focusing on interior solutions, the constrained optimization problem within the agricultural household fuel consumption model can be solved as

$$\left.\begin{array}{l} F^{wood} \\ F^{coal} \\ F^{ele} \end{array}\right\} = f\left(p^C, p^{coal}, p^{ele}, Y, Z, V, \phi\right)$$

In this formulation, the fuel demands are influenced by a comprehensive set of factors: market prices (p^C, p^{coal}, p^{ele}), household income (Y), household preferences and characteristics (Z), variables pertinent to fuel choice (V), and agricultural endowments (ϕ). These elements collectively define the household-specific shadow prices of fuel, underscoring how production-side factors shape energy consumption patterns in environments marked by market imperfections, beyond individual preferences and broader socio-economic influences.

2.5.2 Empirical Strategy

Having outlined the theoretical model of agricultural fuel consumption, I now transition to the empirical strategy that will enable us to estimate and test the derived hypotheses. The subsequent analysis employs a correlated random effects generalized ordered probit model, an approach specifically chosen for its robustness in handling the ordinal nature of fuel consumption categories and its ability to account for unobserved heterogeneity across observational units. This model is adept at capturing the latent variables that influence the probability of a household transitioning between different ranked fuel pattern. By integrating the theoretical constructs into this empirical framework, we can rigorously assess the factors that drive fuel consumption patterns in the agricultural households, thus providing a comprehensive understanding of the energy transition dynamics.

Correlated Random Effects Generalized Ordered Probit (CORE-GOP) Model Research on household fuel choices often employs multinomial logit (MNL) and multinomial probit (MNP) models. The MNL model is commonly chosen for its computational simplicity but requires adherence to the Independence of Irrelevant Alternatives (IIA) assumption. Failure to meet this assumption can lead to biased and inconsistent results. On the other hand, the MNP model does not require the IIA assumption, offering a more flexible estimation approach, but it overlooks the ordinal nature of fuel preferences, which is essential for capturing the preference hierarchy among different fuels.

To address these limitations and accurately reflect the hierarchical preferences for different fuel types, I will employ the random effect ordered probit model in this study. This model allows us to incorporate the ordinal ranking of fuel types based on convenience, comfort, and efficiency. In the analysis, electricity is deemed the most efficient and convenient, followed by coal, with firewood ranked as the least. This ranking aligns with the energy ladder theory, illustrating the transition of households in rural China from traditional solid fuels to cleaner and more efficient energy sources. Furthermore, to acknowledge the fact of fuel stacking, I also categorize fuel patterns, ranking completely clean portfolios as superior to mixed ones, and mixed ones as better than entirely polluting ones.

Household heating fuel choices are posited to depend on a latent variable Y_{it}^*, formulated as:

$$Y_{it}^* = x_{it}'\beta + \alpha_i + \mu_{it} \tag{2.3}$$

where

$$(\mu_{it}|x_{it}) \sim N(0,1) \tag{2.4}$$

$$(\alpha_i|x_{it}) \sim N(0,\sigma_\alpha^2) \tag{2.5}$$

$$corr(\alpha_i + \mu_{it}, \alpha_i + \mu_{i\tau}) = \rho = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\mu^2} \quad \text{for any} \quad t \neq \tau \tag{2.6}$$

where x_{it} represents the vector of explanatory variables of household i in period t; β is the vector of coefficients to be estimated; α_i denotes time-invariant unobserved heterogeneity, assumed to follow a normal distribution conditioned on x_{it} ; the error term μ_i is also assumed to be normally distributed. In this study, heating fuel choices are categorized into three patterns: completely polluting, mixed, and completely clean. The dependent variable is thus treated as an ordered variable with these three categories. The observed heating fuel choice Y_{it} is defined by thresholds:

$$Y_{it} = j \Leftrightarrow k_{j-1} < Y_{it}^* \leq k_j \Leftrightarrow Y_{it} = \begin{cases} 1, \text{ if } -\infty \leq Y_{it}^* \leq k_1 & \text{(Completely Polluting)} \\ 2, \text{ if } k_1 < Y_{it}^* \leq k_2 & \text{(Mixed)} \\ 3, \text{ if } k_2 < Y_{it}^* \leq \infty & \text{(Completely Clean)} \end{cases} \tag{2.7}$$

where $-\infty = k_0 < k_1 < k_2 < k_3 = \infty$ defining the unknown intervals for each choice, $j \in \{1,2,3\}$. The probability of choosing heating fuel pattern j is defined as

$$Pr(Y_{it} = j|x_{it}) = Pr(k_{j-1} < Y_{it}^* \leq k_j|x_{it}) \tag{2.8}$$

$$= Pr(k_{j-1} < x_{it}'\beta + \alpha_i + \mu_{it}|x_{it}) \tag{2.9}$$

$$= \Phi(k_j - x_{it}'\beta - \alpha_i|x_{it}) - \Phi(k_{j-1} - x_{it}'\beta - \alpha_i|x_{it}) \tag{2.10}$$

where Φ is the standard normal cumulative distribution function.

The random effects ordered probit model offers a nuanced approach to handling unobserved heterogeneity, allowing us to account for household-specific traits that standard ordered probit analysis might miss. Nonetheless, this model enforces the parallel regression assumption, implying constant effects of explanatory variables across different choices (Maddala, 1983; Boes and Winkelmann, 2006). To address this limitation, I extend to a generalized ordered probit model, which allows for varying impacts of explanatory variables across outcomes by modeling the thresholds k_j as a linear functions of x_{it}(Ierza, 1985): $k_{ij} = a_j + x'_{it}\lambda_j$, where λ_j represents the influence parameter of the covariates on the thresholds; k_{ij} denotes the threshold values for household i for each heating fuel choice j; a_j is a constant term. Therefore, the probability of choosing alternative j will be

$$Pr(Y_{it} = j|x_{it}) = \Phi(a_j - x'_{it}(\beta - \lambda_j) - \alpha_i|x_{it}) - \Phi(a_{j-1} - x'_{it}(\beta - \lambda_{j-1}) - \alpha_i|x_{it}) \tag{2.11}$$

where the estimated coefficients $\beta_j = \beta - \lambda_j$ are specific to each fuel alternative.

While the generalized random effects ordered probit(RE-GOP) model relaxes the parallel regression assumption, it traditionally assumes that the unobserved heterogeneity α_i is independent of the explanatory variables x_{it}. This assumption can be restrictive as it neglects potential correlations between the unobserved heterogeneity and the regressors, which may introduce bias into the coefficient estimates. To address this, I incorporate the Mundlak approach, following the methodologies of Boes and Winkelmann, 2010, Mentzakis and Moro, 2009 , Chamberlain, 1982, and Mundlak, 1978, which involves integrating the averages of time-varying regressors within households into the model. This transformation posits that the unobserved effects are linearly related to the explanatory variables:

$$\alpha_i = \bar{x}_i\gamma + \epsilon_i \tag{2.12}$$

where $\bar{x}_i$ represents the average of x_{it} over time, γ is the parameter, and ϵ_i is an orthogonal error term with $\epsilon_i|\bar{x}_i \sim N(0, \sigma^2_\epsilon)$. Incorporating the Mundlak approach transforms the equation (2.5.2) as follows:

$$Pr(Y_{it} = j|x_{it}, \bar{x}, \epsilon_i; \gamma, \beta_j) = \Phi(a_j - x'_{it}(\beta - \lambda_j) - \bar{x}_i\gamma - \epsilon_i|x_{it}) \tag{2.13}$$

$$- \Phi(a_{j-1} - x'_{it}(\beta - \lambda_{j-1}) - \bar{x}_i\gamma - \epsilon_i|x_{it}) \tag{2.14}$$

Equation (2.5.2) describes the correlated random-effects generalized ordered probit(CORE-GOP) model.

Therefore, the CORE-GOP model not only controls for unobserved heterogeneity but also allows for heterogeneous effects of explanatory variables across different choices. Additionally, it accounts for potential correlations between unobserved heterogeneity and the regressors. Unobserved factors such as an individual's family background or socio-economic upbringing may be correlated with observable explanatory variables like education and income. By incorporating these correlations—for instance, acknowledging that individuals from more affluent or educated families are likely to have higher education and income levels—the CORE-GOP model provides more accurate results than the standard ordered probit model. This is particularly important when such correlations exist, as accounting for them helps to avoid biased estimates and results in a more reliable analysis. The CORE-GOP model can be estimated via maximum likelihood as detailed in Boes and Winkelmann, 2010.

Previous studies, such as Wu and Zheng, 2022, propose that household fuel consumption decisions unfold in a two-stage process: initially deciding which fuels to use, followed by determining the quantity of fuel to consume. These stages are potentially influenced by distinct factors. To further explore this hypothesis, I investigate the determinants of heating behaviors—specifically, how households allocate their heating hours among chosen fuels. For this purpose, I employ fixed effects models, using log-transformed total heating hours across all rooms by fuel source as the dependent variable. Like the analysis of fuel choices, this segment of the study focuses exclusively on households not yet treated by the policy.

2.5.3 Results

Determinants of Fuel Choices

Variables	Clean		Mixed		Polluting	
Log Price of Coal	0.087	***	-0.055		-0.033	
	(0.029)		(0.034)		(0.021)	
Log Income per Capita	0.003		0.004		-0.007	
	(0.007)		(0.010)		(0.017)	
Age	0.000		0.001		-0.001	
	(0.000)		(0.001)		(0.001)	
Male	-0.018	**	-0.027	**	0.045	**
	(0.008)		(0.012)		(0.021)	
Married	-0.024	*	-0.037	*	0.061	*
	(0.013)		(0.020)		(0.033)	
Secondary or Higher Education	0.017	**	0.026	**	-0.044	**
	(0.009)		(0.013)		(0.022)	
Log Household Size	0.012		0.103	***	-0.115	***
	(0.018)		(0.029)		(0.041)	
Number of Children Under 5	-0.001		-0.001		0.002	
	(0.017)		(0.026)		(0.043)	
Log House Size	0.064	***	0.095	***	-0.159	***
	(0.011)		(0.016)		(0.027)	
Agricultural Assets						
Log Agriculture Land	0.003	*	0.005	*	-0.009	*
	(0.002)		(0.003)		(0.005)	
Log Forest Land	0.000		-0.012	***	0.012	***
	(0.003)		(0.004)		(0.005)	
Region						
Fangshan	0.136	***	0.104	***	-0.240	***
	(0.027)		(0.032)		(0.040)	
Huairou	0.087	***	0.130	***	-0.217	***
	(0.016)		(0.022)		(0.037)	
Mentougou	0.159	***	0.105	***	-0.265	***
	(0.023)		(0.030)		(0.041)	
Time						
Wave 2	0.066	***	-0.052	***	-0.014	
	(0.014)		(0.020)		(0.018)	
Wave 4	0.095	***	-0.000		-0.095	***
	(0.016)		(0.022)		(0.020)	
Number of observations	2332		2332		2332	

Table 2.8: Determinants of Household Heating Fuel Pattern estimated by CORE-GOP Model

[a] *** p<.01, ** p<.05, * p<.1

Table 2.8 reports the marginal effects of CORE-GOP model on fuel choices. It is evident that coal price, gender, education, marital status, agricultural assets, house area, household size, region, and season dummies significantly influence household heating fuel choices in rural China.

Fuel prices play a role in shaping heating energy patterns in rural China, influencing household decisions to transition towards cleaner energy sources. The findings indicate that an increase in coal prices encourages the adoption of clean heating patterns. Specifically, a one-unit increase in the log of coal prices enhances the likelihood of households opting for completely clean fuels by 8.8%. This highlights the potential of fuel price control as an effective policy tool to promote a transition to clean heating among rural households.

The impact of fuel prices is well illustrated in the Fangshan district during season 4, a period marked by a near doubling of coal prices. This significant price increase led to a 23% rise in the adoption of clean heating portfolios among households, indicating a direct correlation between fuel pricing strategies and the shift towards cleaner heating options. This observation underscores the importance of strategic fuel pricing in driving significant and sustainable changes in heating energy patterns, particularly in regions that are traditionally dependent on polluting fuels.

The marginal effects of income per capita on heating choices are found to be insignificant, contradicting the energy ladder hypothesis, which suggests that households should transition to cleaner fuels as their income increases. This finding supports the fuel stacking theory, indicating that households continue to use traditional solid fuels alongside more advanced fuels, rather than completely transitioning away. This observation implies that the persistence of polluting fuels is not merely a result of financial constraints or the affordability of clean energy. Consequently, providing subsidies for electricity or other clean energies may have limited effectiveness in encouraging households to fully transition to clean energy solutions.

Gender significantly influences household fuel choices. The analysis reveals that men are 4.5% more likely than women to opt for completely polluting heating fuels and are less likely—by 1.8% and 2.7%—to select clean or mixed fuels, respectively. This trend aligns with findings from other studies suggesting that women are more inclined to choose cleaner fuels. This preference may be due to women typically handling labor-intensive tasks like collecting biomass fuels and managing traditional heating sources, such as kangs or stoves, including the continual addition of fuel. This gender difference highlights the importance of incorporating social and behavioral factors into the development of energy policies and interventions to promote cleaner heating solutions.

Additionally, marital status plays a significant role in heating fuel decisions. Married couples are 6.1% more likely than single individuals (including divorced, separated, widowed, and never married) to choose completely polluting heating fuels. This trend could be attributed to married couples having more household labor available to gather solid fuels, such as firewood, which requires considerable effort and time.

According to existing literature, education is an important policy tool to raise households' awareness about the benefits of clean energy sources and the risks associated with dirty fuel sources. As anticipated, higher educational levels within a household tend to increase the likelihood of opting for cleaner energy sources. The findings support this perspective, albeit with a modest effect. Households with secondary education or higher are 1.7% more likely to choose clean fuels and 2.6% more likely to opt for mixed fuels compared to those with no or only primary education, and are 4.3% less likely to choose entirely polluting fuels.

Regarding household size, existing studies provide mixed evidence on its impact on fuel choice and fuel switching behaviors. The analysis reveals that larger household sizes negatively influence the use of completely polluting fuels and positively affect the adoption of mixed fuels, while their impact on clean fuel use is positive but not statistically significant. Larger households typically face greater heating demands, which drives a preference for more efficient heating methods compared to smaller households.

House size significantly impacts household heating fuel choices. A larger house size is associated with a preference for clean or mixed heating fuels. A one-unit increase in house size (log) reduces the likelihood of selecting clean fuels by 4% and mixed fuels by 9.5%, while the probability of opting for polluting fuels decreases by 11.5%. House size often correlates closely with household wealth, with wealthier households more likely to select higher-ranked fuels as suggested by the energy ladder theory. Additionally, a larger house size typically necessitates greater heating needs, prompting households to choose more efficient heating methods to meet these demands more effectively than smaller households.

Agricultural assets significantly influence household heating pattern choices, aligning with the agricultural household consumption model where fuel choices are affected by production-side factors due to incomplete markets in rural areas. However, the effects of agricultural land and forest land on fuel choices are divergent, likely due to the interplay of two opposing influences. On one hand, households with more agricultural and forest land have easier access to freely available but polluting fuels such as crop residues and firewood, encouraging the use of these fuels. This is particularly evident with forest land, which positively correlates with the choice of polluting fuels. On the other hand, crop residues are scarcely used in rural Beijing, with less than 1% of households reporting their use as fuel in this study, suggesting that agricultural land may influence fuel choices predominantly through an income effect. Households with larger expanses of agricultural and forest land are generally wealthier and can afford cleaner and more convenient energy sources. The contrasting impacts of these assets highlight the complex dynamics influencing rural household energy decisions.

The marginal effects of the seasonal dummies illustrate a clear trend of households increasingly transitioning from polluting to clean heating fuels over time. Regional differences also significantly influence heating fuel choices, which can be attributed to geographic and climatic variations across different areas.

Contrary to much existing literature, this study finds that age does not significantly impact household heating fuel patterns. A potential explanation for this discrepancy lies in the method of data collection concerning age. While most studies collect age data specifically for the household head, the dataset captures this information from a randomly selected household member who responded to the survey. Although it is reasonable to assume that the household head often dictates heating decisions and is likely the survey respondent, this methodology may introduce substantial measurement bias if other household members, who might not represent the head's perspective, provide the data.

The results confirm the agricultural household fuel consumption model's predictions that a variety of factors influence household fuel choices. These include market prices, household characteristics such as gender, marital status, education, household size, and area, as well as variables tied to production decisions, notably agricultural assets. These elements collectively establish the household-specific shadow price of fuel. Notably, my findings deviate from the energy ladder theory by demonstrating that income does not significantly predict fuel choices, highlighting the fuel stacking behavior in rural settings.

Determinants of Heating Behaviors

Variables	Electricity		Coal		Wood		All Fuels	
Log Price of Coal	0.232		-0.273		-0.100		0.062	
	(0.218)		(0.273)		(0.067)		(0.052)	
Log Income per Capita	0.259	**	-0.020		-0.022		0.084	**
	(0.105)		(0.106)		(0.067)		(0.038)	
Age	0.002		-0.011		0.022	***	-0.006	*
	(0.010)		(0.014)		(0.006)		(0.003)	
Male	-0.310	*	0.321	*	-0.002		0.046	
	(0.171)		(0.191)		(0.095)		(0.041)	
Married	-0.387	*	0.179		0.449	***	0.017	
	(0.211)		(0.279)		(0.165)		(0.085)	
Secondary or Higher Education	0.481	**	-0.202		-0.311	**	0.075	
	(0.191)		(0.179)		(0.120)		(0.055)	
Log Household Size	0.058		1.070	***	0.355	**	0.453	***
	(0.257)		(0.256)		(0.162)		(0.066)	
Number of Children Under 5	-0.082		0.126		-0.415	***	-0.007	
	(0.255)		(0.219)		(0.144)		(0.048)	
Log Household Area	1.266	***	0.097		-0.324	*	0.557	***
	(0.235)		(0.275)		(0.172)		(0.058)	
Agricultural Assets								
Log Agriculture Land	0.017		-0.027		0.024		-0.008	
	(0.034)		(0.032)		(0.026)		(0.012)	
Log Forest Land	-0.011		0.050	*	-0.003		0.015	*
	(0.024)		(0.027)		(0.021)		(0.008)	
Fixed Effects								
Village	Yes		Yes		Yes		Yes	
Time	Yes		Yes		Yes		Yes	
Number of observations	2346		2346		2345		2345	

Table 2.9: Determinants of Household Heating Behaviours

[a] *** p<.01, ** p<.05, * p<.1
[b] Dependent variable are log-transformed total heating hours across all rooms by fuel sources
[c] Estimated by Fixed Effect Model.

Having explored the determinants of heating fuel choices, I shift my focus to the determinants of heating behaviors. While fuel choices pertain to the selection of energy sources, heating behaviors examine how extensively and in what manner these fuels are utilized. This transition in focus enables us to investigate the factors that influence the actual usage of fuels, providing insights into household heating practices and their broader energy consumption patterns.

Table 2.9 outlines the determinants influencing heating behaviors. Key factors include income, age, gender, marital status, education, household size, the number of children under five, household area, and forest land. These determinants largely align with those influencing fuel choices, although there are variations in their effects between choosing fuels and determining heating behaviors, which implies differences between these two decision-making processes.

Interestingly, while coal prices influence fuel choices, they do not significantly impact heating hours. This distinction underscores the separate decision-making processes for choosing and utilizing fuels. Although price fluctuations may encourage a transition to cleaner fuels, they do not necessarily impact the volume of heating. On the other hand, although income does not significantly affect fuel choice, it markedly influences heating hours. With every unit increase in log income per capita, households increase their electricity heating hours by 26% and their total heating hours by 8.4%. These findings align with both the energy ladder and the fuel stacking hypothesis regarding heating hours allocations, suggesting that as income rises, households tend to increase their use of clean fuels without completely phasing out polluting fuels.

In my previous analyses, age did not emerge as a significant determinant of fuel choices. However, it significantly influences wood usage, with each additional year in age increasing wood heating hours by 2.22%. This finding aligns with existing literature, suggesting that older individuals tend to use more polluting fuels.

Gender differences are also notable in heating behaviors. Males use 31% less electricity and 32% more coal for heating than females, which corroborates both my earlier findings on fuel choices and broader research indicating that females prefer cleaner fuels.

Similarly, marital status impacts heating behaviors. Married couples use 39% less electricity and 45% more wood for heating compared to singles, echoing trends observed in fuel choice preferences.

Educational attainment significantly affects heating behaviors as well. Households with a secondary education or higher use 48% more electricity and 31% less wood, consistent with their tendency to choose cleaner fuels, as noted in the fuel choice analysis.

Interestingly, household size influences heating behaviors differently than fuel choices. Although larger households tend to opt for cleaner fuel portfolios, the findings indicate they use significantly more coal and wood for heating as household size increases. For example, a doubling in household size correlates with a doubling in coal heating hours and a 35.5% increase in wood heating hours. This suggests that while electricity might be adopted as a supplementary heating method in larger households, the predominant fuels used remain polluting.

The presence of young children also significantly impacts heating behaviors. Households with an additional child under five use 41.5% fewer wood heating hours. This could be due to reduced time for fuel collection due to care giving responsibilities and concerns over the health and safety risks posed by traditional wood stoves to young children.

House area also affects heating hours, with a 100% increases in area leading to a 126% increase in electricity heating hours and a 32.4% decrease in wood heating hours. This change suggests a shift towards more efficient heating methods as house size increases, which aligns with the broader patterns observed in fuel choice dynamics.

Lastly, agricultural assets show a minor influence on heating behaviors. Unlike expectations, ownership of forest land does not lead to an increase in wood heating hours but slightly raises coal heating hours and total heating hours. This suggests that the availability of agricultural and forest lands as resources for fuel does not straightforwardly translate into increased use due to potential logistical constraints, such as the distance of land from the household or other uses for the land. Thus, the impact of these assets is likely mediated through income effects rather than direct fuel access.

Heating hours are significantly influenced by a variety of factors such as income, age, marital status, education, household size, area, and the presence of children under five, with each factor exerting a relatively substantial impact. Notably, the determinants of how heating hours are allocated differ from those influencing fuel choices for some factors. For instance, income and agricultural assets demonstrate different effects on these two decision-making processes, highlighting distinct dynamics in heating behavior compared to fuel selection.

2.6 Discussion

The findings contribute to provide more evidence on household fuel choices, particularly in the context of rural Beijing. Contrary to the singular progression predicted by the energy ladder theory—where households move from less to more efficient fuels as income increases—the results reveal a more complex behavior consistent with fuel stacking. Households in this study continue to utilize multiple heating fuels and retain older heating devices alongside newer technologies like heat pumps. However, the models focus primarily on intra-household comparisons over time. The apparent inconsistency with the energy ladder theory may be due to the short duration of the survey period during which household incomes may not change significantly enough to observe a transition up the energy ladder. If data were available over a longer period with greater variation in income, we might observe clearer evidence supporting the energy ladder theory.

Furthermore, the analysis confirms an inverted U-shaped relationship between income and the diversity of fuel types used, as also identified by Wu and Zheng, 2022. This pattern, illustrated in Figure 2.15, suggests that households initially increase the variety of fuels they use as their income grows but later consolidate to more efficient fuels once a certain income threshold is reached. Interestingly, I observe a concurrent increase in the use of electricity for heating with rising income, aligning with the broader principles of the energy ladder that suggest a shift towards more efficient energy sources at higher income levels. These observations underline the importance of integrating the energy ladder and fuel stacking theories to provide a more comprehensive framework for understanding household fuel choices.

However, despite the apparent diversity in fuel use, the data indicate that most households rely predominantly on one primary fuel type that satisfies the majority of their heating needs. This finding highlights the critical role of dominant fuels in shaping overall household heating strategies and suggests that the primary fuel choice is a pivotal factor in understanding household energy dynamics. Thus, identifying primary fuel sources is also crucial in depicting household energy needs. This dual perspective on household fuel use—recognizing both the diversity of fuels

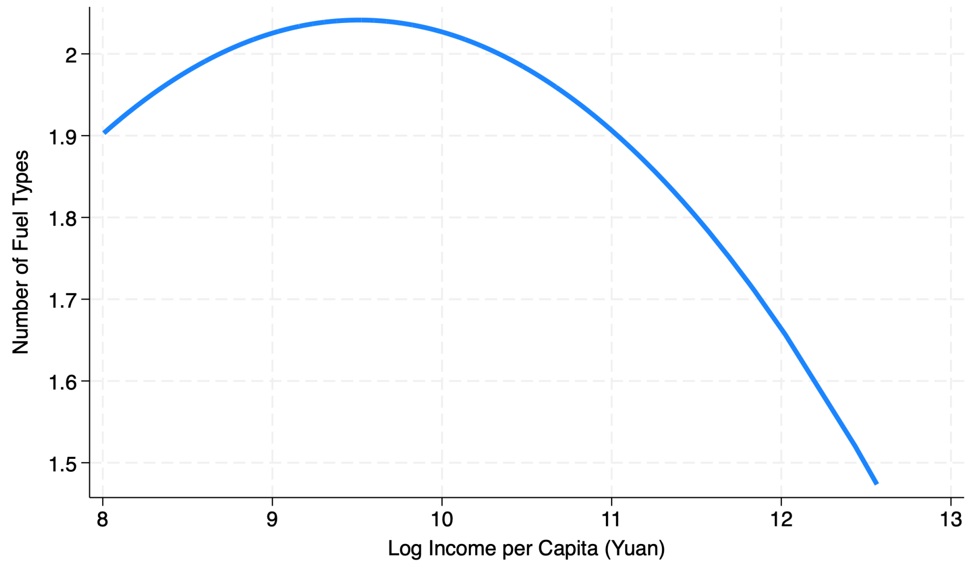


Figure 2.15: Relationship between Income and Number of Fuel Types

employed and the dominance of primary fuels—provides a more detailed understanding of the complexities involved in transitioning to cleaner energy sources.

The findings from this study demonstrate that the clean heating policy effectively encouraged the transition from coal to electricity, with a notable increase in the use of electricity and subsidized heat pumps, alongside a significant reduction in coal use. Additionally, the intervention led households to increase their heating hours by warming more rooms and extending heating durations, contributing to warmer winters in treated regions. More pronounced effects were observed in regions that adopted the policy earlier or had more polluting initial heating setups. However, the transition to completely clean heating was not fully realized, as coal use was not entirely phased out and the usage of other polluting fuels, like wood, remained unchanged. Concurrently, a spontaneous but incomplete shift towards cleaner heating was observed in areas not directly affected by the policy.

Further investigation into the determinants of fuel choices and behaviors reveals multiple factors that could enhance the effectiveness of future policies. For instance, controlling fuel prices may serve as an effective tool to facilitate the transition to cleaner fuels. Household income also emerged as a crucial factor, suggesting that ensuring the affordability of new technologies through subsidies for purchase and installation is essential for sustaining thermal benefits without imposing additional financial burdens on households.

Moreover, the analysis indicates that other demographic factors such as gender, education, and age significantly impact heating fuel choices and behaviors. These insights are valuable for designing effective policy. For instance, as households adopt heat pumps, targeted guidance on their use could be specifically directed towards women and younger adults, who are more likely to adopt cleaner heating solutions. Emphasizing the health and environmental advantages of using heat pumps might also enhance compliance and satisfaction with the policy. Such tailored approaches in policy implementation could maximize the benefits of clean heating interventions, promoting broader and more effective adoption of cleaner heating solutions.

This study acknowledges several limitations that suggest avenues for future research. Firstly, the data collection was confined to rural regions in Beijing, which may limit the generalizability of the findings to other regions. I examined the households' self-reported heating fuel choices and behaviors rather than the actual use, which could introduce bias. However, the study attempted to mitigate this by installing temperature sensors on heating devices such as kangs and heat pumps to capture actual usage. The analysis of these direct measurements yielded results consistent with the self-reported data (Brehmer et al., 2023), reinforcing the reliability of the findings.

Methodologically, there are pre-existing differences between the outcomes of the treatment and control groups, which potentially violate the parallel trends assumption. Although the data includes detailed records of quantity and expenditure for each type of fuel used for all purposes, I lack specific data on the quantity of fuel or expenditure solely for heating. I used heating hours as a proxy for fuel consumption, which, while informative, is not a perfect measure. Additionally, my approach to separately estimate the determinants of fuel choices and fuel consumption aimed to clarify whether these decisions are driven by different processes. However, this method does not account for the potential interconnections between these two decisions.

For future research, a more comprehensive model that integrates the dual aspects of fuel choice and usage would be beneficial. Such a model could explore the two-stage decision-making process in greater detail. Both qualitative analyses and the development of sophisticated econometric models will be demanded for this topic. Additionally, I didn't explore the interactions between various factors that influence fuel choices and the intervention, there might exist heterogeneous effects of policies across different demographics, which could yield insights that enhance the effectiveness and targeting of similar interventions.

2.7 Conclusion

This paper explores the micro-level effects of a mandatory residential coal-to-electricity transition in rural Beijing, shedding light on the dynamics of household heating behaviors under the clean heating policy. This policy aims to reduce reliance on polluting fuels and foster a transition towards cleaner energy alternatives, which has been implemented in the context of China's broader environmental goals. Utilizing three rounds of panel data captured through a quasi-experimental design, this study evaluates the causal impacts of this policy on household heating fuel choices and behaviors. I delve into how households adjust their fuel selection and allocate heating hours across different fuels and devices in response to the policy intervention.

The analysis, leveraging a difference-in-differences methodology, indicates that the clean heating policy has effectively decreased coal use and facilitated a significant shift towards electricity and subsidized heat pumps. However, it did not entirely eliminate the use of coal, and other polluting fuels such as firewood remained in use. The policy also resulted in increased total heating hours and a shift towards more efficient heating devices. Interestingly, the policy encouraged a mixed-fuel usage pattern, incorporating both clean and polluting fuels, which reflects the nuances of both the energy ladder and fuel stacking theories. This suggests that while households have progressed up the energy ladder by adopting cleaner fuels like electricity, they have not fully transitioned away from traditional fuels such as wood, supporting the idea of fuel stacking where multiple fuel types continue to meet diverse household needs.

This partial transition and the continued reliance on mixed fuel portfolios highlight the complexity of achieving a full shift to clean energy in rural settings, influenced by economic, cultural, and logistical factors. Furthermore, spontaneous transitions towards cleaner fuels were observed in regions not directly targeted by the policy, indicating broader shifts in energy preferences and practices.

To address these dynamics further, the latter part of the paper investigates the determinants influencing households' fuel choices and heating behaviors in regions unaffected by the policy. Employing a correlated random effects model, I uncover that factors such as fuel prices, household income, and agricultural assets play significant roles in these decisions, emphasizing that the choice of fuels and the volume of energy consumed are influenced by distinct factors. For example, while fuel prices affect the type of fuels chosen—promoting cleaner options as prices of polluting fuels rise—they do not significantly influence the total heating hours. Conversely, income predominantly affects energy consumption volumes without altering fuel type preferences, underscoring the limitations of price mechanisms alone in driving complete transitions to cleaner energy.

This study highlights the importance of integrating both energy ladder and fuel stacking theories to fully understand and address the complexities of household energy transitions. It suggests that policies aimed at encouraging clean heating should consider the diverse economic and socio-demographic factors affecting household energy decisions.

References

Abebaw, D. (2007). Household determinants of fuelwood choice in urban Ethiopia: A case study of Jimma Town. The Journal of Developing Areas, 41(1), 117–126.

Akpalu, W., Dasmani, I., & Aglobitse, P. B. (2011). Demand for cooking fuels in a developing country: To what extent do taste and preferences matter? Energy Policy, 39(10), 6525–6531.

Arnold, J. M., Köhlin, G., & Persson, R. (2006). Woodfuels, livelihoods, and policy interventions: Changing perspectives. World development, 34(3), 596–611.

Baiyegunhi, L. J., & Hassan, M. (2014). Rural household fuel energy transition: Evidence from giwa lga kaduna state, nigeria. Energy for sustainable development, 20, 30–35.

Baland, J., Libois, F., & Mookherjee, D. (2017). Forest degradation and economic growth in rural Nepal 2003–2010.

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

Bhagavan, M., & Giriappa, S. (1995). Biomass, energy and economic and natural resource differentiation in rural southern India. Biomass and Bioenergy, 8(3), 181–190.

Boes, S., & Winkelmann, R. (2006). Ordered response models. Allgemeines Statistisches Archiv, 90, 167–181.

Boes, S., & Winkelmann, R. (2010). The effect of income on general life satisfaction and dissatisfaction. Social Indicators Research, 95, 111–128.

Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event-study designs: Robust and efficient estimation. Review of Economic Studies, rdae007.

Bowlus, A. J., & Sicular, T. (2003). Moving toward markets? Labor allocation in rural China. Journal of Development Economics, 71(2), 561–583.

Brehmer, C., et al. (2023). Impacts of a clean heating program on stacked heating stove usage in northern china (Working Paper).

Brehmer, C., & Li, X. (2023). Multi-year, field-based assessment on the impacts of a Coal-to-Clean Energy policy on indoor, outdoor, and personal exposure to PM2.5 in Beijing, China (Working Paper).

Brouwer, R., & Falcão, M. P. (2004). Wood fuel consumption in Maputo, Mozambique. Biomass and Bioenergy, 27(3), 233–245.

Budya, H., & Arofat, M. Y. (2011). Providing cleaner energy access in Indonesia through the megaproject of kerosene conversion to LPG. Energy Policy, 39(12), 7575–7586.

Cai, J., & Jiang, Z. (2008). Changing of energy consumption patterns from rural households to urban households in China: An example from Shaanxi Province, China. Renewable and Sustainable Energy Reviews, 12(6), 1667–1680.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. Journal of econometrics, 225(2), 200–230.

Campbell, D. (2003). Intra-and intersectoral effects in environmental disclosures: Evidence for legitimacy theory? Business Strategy and the Environment, 12(6), 357–371.

Chamberlain, G. (1982). Multivariate regression models for panel data. Journal of econometrics, 18(1), 5–46.

Chen, L., Heerink, N., & van den Berg, M. (2006). Energy consumption in rural China: A household model for three villages in Jiangxi Province. Ecological Economics, 58(2), 407–420.

Chen, Q., Yang, H., Su, L., & Pei, H. (2024). Household residential energy choices in green transition: insights from a household survey in rural China. Environmental Science and Pollution Research, 31(50), 60539–60554.

Cheng, M., Zhi, G., Tang, W., Liu, S., Dang, H., Guo, Z., Du, J., Du, X., Zhang, W., Zhang, Y., et al. (2017). Air pollutant emission from the underestimated households' coal consumption source in China. Science of the Total Environment, 580, 641–650.

Clancy, L., Goodman, P., Sinclair, H., & Dockery, D. W. (2002). Effect of air-pollution control on death rates in Dublin, Ireland: an intervention study. The lancet, 360(9341), 1210–1214.

Davis, M. (1998). Rural household energy consumption: the effects of access to electricity—evidence from South Africa. Energy policy, 26(3), 207–217.

De Chaisemartin, C., & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. American economic review, 110(9), 2964–2996.

Démurger, S., & Fournier, M. (2011). Poverty and firewood consumption: A case study of rural households in northern China. China economic review, 22(4), 512–523.

Dispersed Coal Management Research Group. (2023). China dispersed coal management report 2023.

Dockery, D. W., Rich, D. Q., Goodman, P. G., Clancy, L., Ohman-Strickland, P., George, P., Kotlov, T., Committee, H. H. R., et al. (2013). Effect of air pollution control on mortality and hospital admissions in Ireland. Research Report (Health Effects Institute), (176), 3–109.

Edwards, J. H., & Langpap, C. (2005). Startup costs and the decision to switch from firewood to gas fuel. Land Economics, 81(4), 570–586.

Edwards, J. H., & Langpap, C. (2012). Fuel choice, indoor air pollution and children's health. Environment and Development Economics, 17(4), 379–406.

Farsi, M., Filippini, M., & Pachauri, S. (2007). Fuel choices in urban Indian households. Environment and Development Economics, 12(6), 757–774.

Gardner, J. (2022). Two-stage differences in differences. arXiv preprint arXiv:2207.05943.

Gebreegziabher, Z., & Kooten, G. C. v. (2013). Does community and household tree planting imply increased use of wood for fuel? Evidence from Ethiopia. Forest Policy and Economics, 34, 30–40.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. Journal of econometrics, 225(2), 254–277.

Gupta, G., & Köhlin, G. (2006). Preferences for domestic fuel: analysis with socio-economic factors and rankings in Kolkata, India. Ecological Economics, 57(1), 107–121.

Guta, D. D. (2014). Effect of fuelwood scarcity and socio-economic factors on household bio-based energy use and energy substitution in rural Ethiopia. Energy policy, 75, 217–227.

Heltberg, R. (2005). Factors determining household fuel choice in Guatemala. Environment and development economics, 10(3), 337–361.

Heltberg, R., Arndt, T. C., & Sekhar, N. U. (2000). Fuelwood consumption and forest degradation: a household model for domestic energy substitution in rural India. Land Economics, 213–232.

Hiemstra-Van der Horst, G., & Hovorka, A. J. (2008). Reassessing the "energy ladder": Household energy use in Maun, Botswana. Energy Policy, 36(9), 3333–3344.

Hosier, R. H., & Dowd, J. (1987). Household fuel choice in zimbabwe: An empirical test of the energy ladder hypothesis. Resources and energy, 9(4), 347–361.

Hosier, R. H., & Kipondya, W. (1993). Urban household energy use in Tanzania: prices, substitutes and poverty. Energy Policy, 21(5), 454–473.

Hughes-Cromwick, E. L. (1985). Nairobi households and their energy use: An economic analysis of consumption patterns. Energy Economics, 7(4), 265–278.

Ierza, J. V. (1985). Ordinal probit: A generalization. Communications in Statistics-Theory and Methods, 14(1), 1–11.

International Energy Agency. (2015). Buildings Energy Use in China. https : / / iea . blob . core . windows . net / assets / a63ded95 - 7da2 - 4d1d - bcf7 - 8870bb5d6aba / PARTNERCOUNTRYSERIESBuildingEnergy_WEB_FINAL.pdf.

Israel, D. (2002). Fuel choice in developing countries: Evidence from Bolivia. Economic Development and Cultural Change, 50(4), 865–890.

Jiang, L., & O'Neill, B. C. (2004). The energy transition in rural China. International Journal of Global Energy Issues, 21(1-2), 2–26.

Jingchao, Z., & Kotani, K. (2012). The determinants of household energy demand in rural Beijing: Can environmentally friendly technologies be effective? Energy Economics, 34(2), 381–388.

Johnston, F. H., Hanigan, I. C., Henderson, S. B., & Morgan, G. G. (2013). Evaluation of interventions to reduce air pollution from biomass smoke on mortality in Launceston, Australia: retrospective analysis of daily mortality, 1994-2007. Bmj, 346.

Karekezi, S., & Majoro, L. (2002). Improving modern energy services for Africa's urban poor. Energy policy, 30(11-12), 1015–1028.

Kemmler, A. (2007). Characteristics of energy access to the poor and to impoverished regions in India (tech. rep.). Swiss Federal Institute of Technology (ETH).

Lay, Y.-F., Khoo, C.-H., Treagust, D., & Chandrasegaran, A. (2013). Assessing secondary school students' understanding of the relevance of energy in their daily lives. International Journal of Environmental and Science Education, 8(1), 199–215.

Leach, G. (1992). The energy transition. Energy policy, 20(2), 116–123.

Li, J. (2018). What caused China's squeeze on natural gas. The Diplomat, Washington DC, Jan, 1.

Liao, H., Chen, T., Tang, X., & Wu, J. (2019). Fuel choices for cooking in China: Analysis based on multinomial logit model. Journal of cleaner production, 225, 104–111.

Link, C. F., Axinn, W. G., & Ghimire, D. J. (2012). Household energy consumption: Community context and the fuelwood transition. Social Science Research, 41(3), 598–611.

Louw, K., Conradie, B., Howells, M., & Dekenah, M. (2008). Determinants of electricity demand for newly electrified low-income African households. Energy policy, 36(8), 2812–2818.

Maconachie, R., Tanko, A., & Zakariya, M. (2009). Descending the energy ladder? Oil price shocks and domestic fuel choices in Kano, Nigeria. Land use policy, 26(4), 1090–1099.

Maddala, G. (1983). Limited-dependent and qualitative variables in econometrics.

Manning, D. T., & Taylor, J. E. (2014). Migration and fuel use in rural Mexico. Ecological Economics, 102, 126–136.

Masera, O. R., Saatkamp, B. D., & Kammen, D. M. (2000). From linear fuel switching to multiple cooking strategies: A critique and alternative to the energy ladder model. World development, 28(12), 2083–2103.

Mekonnen, A., & Köhlin, G. (2009). Determinants of household fuel choice in major cities in Ethiopia.

Meng, W., Zhong, Q., Chen, Y., Shen, H., Yun, X., Smith, K. R., Li, B., Liu, J., Wang, X., Ma, J., et al. (2019). Energy and air pollution benefits of household fuel policies in northern China. Proceedings of the national academy of sciences, 116(34), 16773–16780.

Mentzakis, E., & Moro, M. (2009). The poor, the rich and the happy: Exploring the link between income and subjective well-being. The Journal of Socio-Economics, 38(1), 147–158.

MEP. (2013). 2013 China Environmental Situation Bulletin [last access: 27 May 2014]. https://www.mee.gov.cn/hjzl/sthjzk/zghjzkgb/201605/P020160526564151497131.pdf.

Mirza, B., & Kemp, R. (2011). Why the rural rich remain energy poor. Consilience, (6), 133–156.

Muller, C., & Yan, H. (2018). Household fuel use in developing countries: Review of theory and evidence. Energy Economics, 70, 429–439.

Mundlak, Y. (1978). On the pooling of time series and cross section data. Econometrica: journal of the Econometric Society, 69–85.

National Energy Administration. (2013). Beijing's heating season relies heavily on clean energy, initiating a plan to replace inferior coal with superior alternatives. https://www.nea.gov.cn/2013-11/26/c_132918981.htm.

Niu, H., Wu, C., Schindler, M., Silva, L. F., Ma, B., Ma, X., Ji, X., Tian, Y., Zhu, H., Bao, X., et al. (2024). Characterization of PM2. 5 carbonaceous components in a typical industrial city in China under continuous mitigation measures. Toxics, 12(7), 461.

NRDC. (2017). Research report of scattered coal comprehensive control in China, 2017. http://coalcap.nrdc.cn/pdfviewer/web/?1536203604405867388.pdf.

Ouédraogo, I. M. (2010). Electricity consumption and economic growth in Burkina Faso: A cointegration analysis. Energy economics, 32(3), 524–531.

Özcan, K. M., Gülay, E., & Üçdoğruk, Ş. (2013). Economic and demographic determinants of household energy use in Turkey. Energy Policy, 60, 550–557.

Pachauri, S., & Jiang, L. (2008). The household energy transition in India and China. Energy policy, 36(11), 4022–4035.

Pachauri, S., Rao, N. D., & Cameron, C. (2018). Outlook for modern cooking energy access in Central America. PloS one, 13(6), e0197974.

Pandey, V. L., & Chaubal, A. (2011). Comprehending household cooking energy choice in rural India. Biomass and bioenergy, 35(11), 4724–4731.

Paudel, U., Khatri, U., & Pant, K. P. (2018). Understanding the determinants of household cooking fuel choice in Afghanistan: a multinomial logit estimation. Energy, 156, 55–62.

Peng, W., Hisham, Z., & Pan, J. (2010). Household level fuel switching in rural Hubei. Energy for sustainable development, 14(3), 238–244.

People's government of beijing municipality. (2017a). https://www.beijing.gov.cn/zhengce/gfxwj/sj/201905/t20190522_59967.html.

People's government of beijing municipality. (2017b). https://www.beijing.gov.cn/zhengce/zhengcefagui/201905/t20190522_60683.html.

People's government of beijing municipality. (2022a). https://www.beijing.gov.cn/hudong/bmwd/jsjbmyyt/jsjbncgl/jsjbncgljcsc/202207/t20220705_2763518.html.

People's government of beijing municipality. (2022b). https://www.beijing.gov.cn/zhengce/zhengcefagui/202204/t20220401_2646626.html.

People's government of beijing municipality. (2024). chrome - extension : / / efaidnbmnnnibpcajpcglclefindmkaj / https : / / www . beijing . gov . cn / zhengce / gfxwj / sj/202012/W020201201332914812668.pdf.

Pitt, M. M. (1985). Equity, externalities and energy subsidies the case of kerosine in indonesia. Journal of Development Economics, 17(3), 201–217.

Rahut, D. B., Ali, A., Mottaleb, K. A., & Aryal, J. P. (2020). Understanding households' choice of cooking fuels: evidence from urban households in Pakistan. Asian Development Review, 37(1), 185–212.

Rambachan, A., & Roth, J. (2023). A more credible approach to parallel trends. Review of Economic Studies, 90(5), 2555–2591.

Rao, M. N., & Reddy, B. S. (2007). Variations in energy use by Indian households: An analysis of micro level data. Energy, 32(2), 143–153.

Rehfuess, E. A., Puzzolo, E., Stanistreet, D., Pope, D., & Bruce, N. G. (2014). Enablers and barriers to large-scale uptake of improved solid fuel stoves: A systematic review. Environmental health perspectives, 122(2), 120–130.

Roth, J., Sant'Anna, P. H., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? a synthesis of the recent econometrics literature. Journal of Econometrics, 235(2), 2218–2244.

Scott, A. J., & Scarrott, C. (2011). Impacts of residential heating intervention measures on air quality and progress towards targets in Christchurch and Timaru, New Zealand. Atmospheric environment, 45(17), 2972–2980.

Shi, X., Nico, H., & QU, F. (2009). The role of off-farm employment in the rural energy consumption transition—A village-level analysis in Jiangxi Province, China. China Economic Review, 20(2), 350–359.

Singh, I., Squire, L., & Strauss, J. (1986). A survey of agricultural household models: Recent findings and policy implications. The World Bank Economic Review, 1(1), 149–179.

Song, C., Liu, B., Cheng, K., Cole, M. A., Dai, Q., Elliott, R. J., & Shi, Z. (2023). Attribution of air quality benefits to clean winter heating policies in China: combining machine learning with causal inference. Environmental Science & Technology, 57(46), 17707–17717.

Sternbach, T., et al. (2023). China's clean heating policy and effects on blood pressure (Working Paper). McGill University.

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. Journal of econometrics, 225(2), 175–199.

Tian, J., Ni, H., Han, Y., Shen, Z., Wang, Q., Long, X., Zhang, Y., & Cao, J. (2018). Primary PM2. 5 and trace gas emissions from residential coal combustion: assessing semi-coke briquette for emission reduction in the Beijing-Tianjin-Hebei region, China. Atmospheric Environment, 191, 378–386.

UN. (2016). Sustainable development goal 7. [last access: 10 April 2019]. https : / / sustainabledevelopment.%20un.org/sdg7.

Van der Kroon, B., Brouwer, R., & Van Beukering, P. J. (2013). The energy ladder: Theoretical myth or empirical truth? Results from a meta-analysis. Renewable and sustainable energy reviews, 20, 504–513.

Wen, H.-x., Nie, P.-y., Liu, M., Peng, R., Guo, T., Wang, C., & Xie, X.-b. (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.

WHO. (2024). Household air pollution and health [last access: 16 October 2024]. https://www.who.int/news-room/fact-sheets/detail/household-air-pollution-and-health.

Wickramasinghe, A. (2011). Energy access and transition to cleaner cooking fuels and technologies in Sri Lanka: Issues and policy limitations. Energy policy, 39(12), 7567–7574.

Wu, S., & Zheng, X. (2022). Revisit of household energy ladder –empirical evidencec from a houehold survey in rural China. Journal of Economics(Quarterly), 22, 45–66.

Wu, S., Zheng, X., Khanna, N., & Feng, W. (2020). Fighting coal—effectiveness of coal-replacement programs for residential heating in China: empirical findings from a household survey. Energy for Sustainable Development, 55, 170–180.

Xiaolin, D., Xiangzhao, F., & Mengxue, Z. (2019). Research on the optimization of capital for the governance of bulk coal in Beijing-Tianjin-Hebei region. Journal of Resources and Ecology, 10(1), 48–55.

Xu, X., Dong, J., Zhong, X., & Xie, D. (2024). Climate impact of coal-to-clean-energy shift policies in rural Northern China. Journal of Cleaner Production, 434, 139870.

Yap, P.-S., & Garcia, C. (2015). Effectiveness of residential wood-burning regulation on decreasing particulate matter levels and hospitalizations in the San Joaquin Valley Air Basin. American journal of public health, 105(4), 772–778.

Yu, C., Kang, J., Teng, J., Long, H., & Fu, Y. (2021). Does coal-to-gas policy reduce air pollution? Evidence from a quasi-natural experiment in China. Science of the Total Environment, 773, 144645.

Zhang, J., Teng, F., & Zhou, S. (2020). The structural changes and determinants of household energy choices and energy consumption in urban China: Addressing the role of building type. Energy Policy, 139, 111314.

Zhang, X., Jin, Y., Dai, H., Xie, Y., & Zhang, S. (2019). Health and economic benefits of cleaner residential heating in the Beijing–Tianjin–Hebei region in China. Energy Policy, 127, 165–178.

Zhang, X.-B., & Hassen, S. (2017). Household fuel choice in urban China: evidence from panel data. Environment and Development Economics, 22(4), 392–413.

Zheng, S., & Kahn, M. E. (2013). Understanding China's urban pollution dynamics. Journal of Economic Literature, 51(3), 731–772.

Zhu, X., Zhu, Z., Zhu, B., & Wang, P. (2022). The determinants of energy choice for household cooking in China. Energy, 260, 124987.

Appendix A: Supplemental Figures

The figures from 2.A.1 to 2.A.3 display Sankey diagrams illustrating the transitions in fuel types among households. Figure 2.A.4 depicts the number of rooms heated by these households.

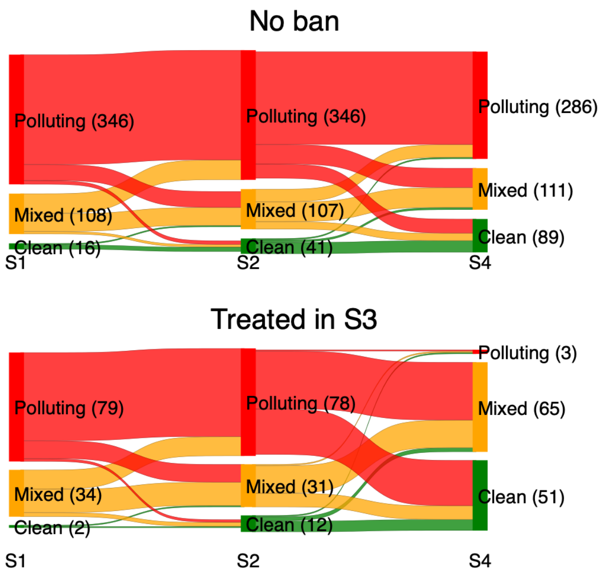


Figure 2.A.1: Heating Fuel Pattern Transition

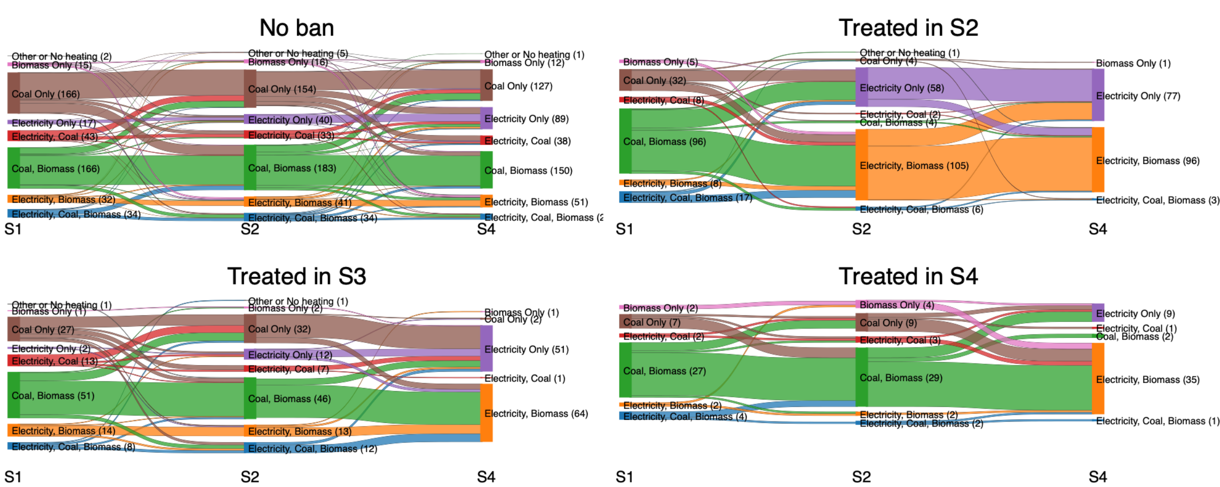


Figure 2.A.2: Heating Fuel Combination Transition

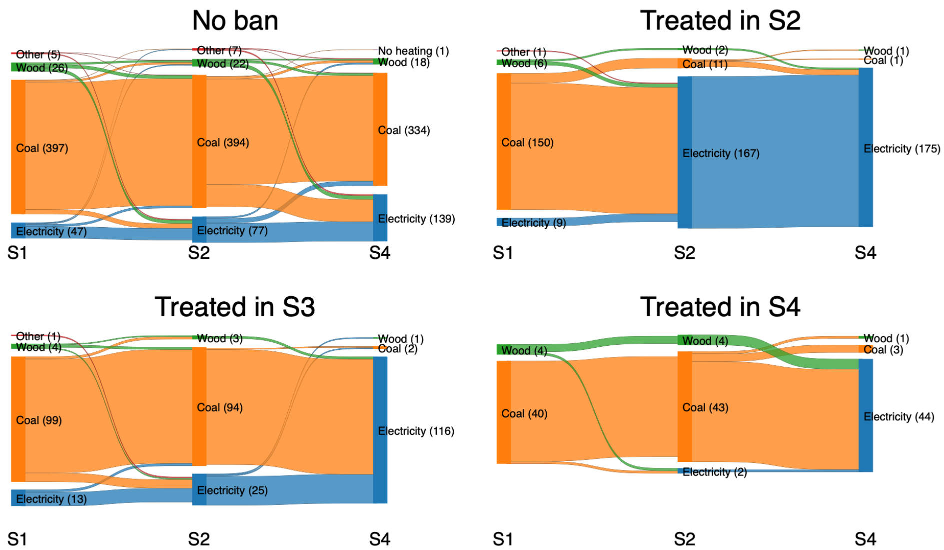


Figure 2.A.3: Primary Heating Fuel Transition

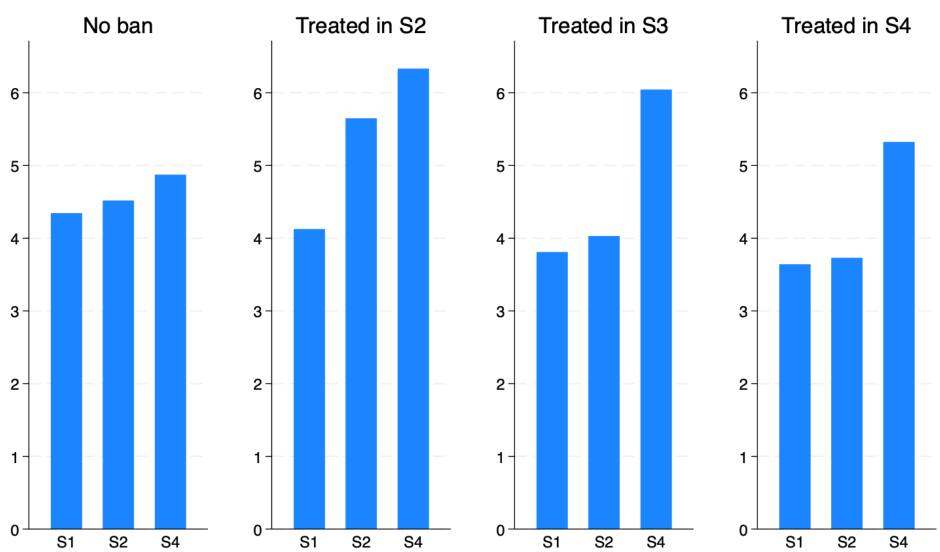


Figure 2.A.4: Number of Rooms with Daily Heating

Appendix B: Treatment Effects on Average Heating Hours per Room

B1 Determinants of Fuel Choices

One may argue that the true decision-making process is unknown to econometricians, and thus, the ranking of fuels may not be clear. As an alternative, a multinomial logit model can be used instead of ordered regression models. The standard multinomial logit model using the pooled sample assumes that households' choices are independent both within a choice (i.e., for multiple observations over time of the same choice) and across all alternative choices made by the household over time, ignoring individual heterogeneity. The random effects specification relaxes the assumption that multiple observations within a choice are independent. In this model, the choice probabilities for repeated choices made by a household share the same time-invariant unobserved heterogeneity. The household-specific effects act as a random variable, producing a correlation among the residuals for the same household within choices, but leaving the residuals independent across households.

Table 2.B.1 displays the results from the multinomial logit model, which fails to capture the effects of many variables compared to the CORE-GOP model.

Variables	Clean		Mixed		Polluting	
Log Price of Coal	0.097	**	-0.044	*	-0.053	*
	(0.032)		(0.019)		(0.026)	
Log Income per Capita	0.009		0.010		-0.019	
	(0.008)		(0.012)		(0.015)	
Age	0.000		0.001		-0.001	
	(0.001)		(0.001)		(0.002)	
Male	-0.004		-0.044	*	0.048	
	(0.014)		(0.019)		(0.025)	
Married	-0.051	*	0.008		0.044	
	(0.023)		(0.031)		(0.031)	
Secondary or Higher Education	0.024		0.023		-0.047	
	(0.017)		(0.026)		(0.027)	
Log Household Size	-0.025		0.045		-0.020	
	(0.017)		(0.026)		(0.029)	
Number of Children Under 5	0.010		0.007		-0.017	
	(0.021)		(0.025)		(0.032)	
Log Household Area	0.055	**	0.092	**	-0.147	**
	(0.020)		(0.023)		(0.029)	
Agricultural Assets						
Log Agriculture Land	0.002		0.004		-0.006	
	(0.003)		(0.004)		(0.005)	
Log Forest Land	0.000		-0.001		0.001	
	(0.002)		(0.002)		(0.003)	
Region						
Fangshan	0.118	**	0.066	*	-0.183	**
	(0.032)		(0.033)		(0.039)	
Huairou	0.086	**	0.129	**	-0.215	**
	(0.030)		(0.038)		(0.048)	
Mentougou	0.162	**	0.126	**	-0.288	**
	(0.053)		(0.034)		(0.057)	
Time						
Wave 2	0.065	**	-0.041	*	-0.025	
	(0.013)		(0.017)		(0.018)	
Wave 4	0.095	**	-0.008		-0.087	**
	(0.017)		(0.021)		(0.027)	
Number of observations	2332		2332		2332	

Table 2.B.1: Determinants of Household Heating Fuel Pattern estimated by Unordered Multi-nominal Model

[a] *** p<.01, ** p<.05, * p<.1

Linking Energy Policy and Public Health: From Fuel Choices to Sleep Patterns

The first essay of my dissertation establishes a foundational understanding of how Beijing's "Coal to Electricity" program influences household heating fuel choices and behaviors, revealing significant shifts towards cleaner energy sources. This transition, while marked by its own complexities, sets the stage for a deeper examination of the broader implications of these energy policy shifts—particularly their effects on public health, which I explore in the second essay.

Building on the findings that clean heating technologies can significantly reduce the reliance on polluting fuels, the second essay shifts focus to an often-overlooked aspect of health impacted by environmental conditions: sleep quality. While the first essay provides an analysis of the direct outcomes of policy implementation on energy consumption behaviors, the second essay delves into the indirect health outcomes resulting from these changes. Specifically, it examines how the improvements in indoor air quality and temperature stability from cleaner heating solutions influence sleep patterns among rural populations. This link between household energy use and sleep quality is important, as it highlights the multi-faceted impacts of environmental policies—not only do they contribute to reducing emissions and energy consumption, but they also potentially offer unintended public health benefits.

Together, these essays paint a comprehensive picture of how environmental interventions, while primarily aimed at sustainability and pollution reduction, also carry significant implications for individual health and well-being. This exploration is crucial for policymakers and stakeholders who must consider both the environmental and health impacts when designing and implementing energy policies. By demonstrating these interconnected effects, my research underscores the importance of integrated approaches in policy planning that consider both environmental sustainability and human health.

Chapter 3

The Impact of a Clean Heating Policy on Sleep Patterns in Rural China

Wenmei Tu

Department of Economics, McGill University, Montreal, Quebec, Canada

Abstract

Sleep is crucial for health and well-being, but it can be severely affected by environmental factors, such as air pollution, temperature and interruptions. This study analyzes the impact of Beijing's clean heating policy, which transitions households from coal to electricity, using the Difference-in-Differences method to explore how this policy affects sleep patterns. Results show that the policy predominantly impacts population subgroups with specific demographics: initial short sleepers (with an approximate increase of 1 hour in sleep duration) and long sleepers (with about a 0.5-hour decrease in sleep duration), as well as men (with around a 0.5-hour increase in sleep duration), individuals aged 60-70 (with more than a 0.5-hour increase in sleep duration), and those in regions experiencing significant shifts from polluting to clean heating fuels (with increases exceeding 0.5 hour in sleep duration). Although the policy effectively shifted bedroom heating from coal to electricity, the expected reductions in indoor air pollution and personal pollutant exposure,along with the improvement in room temperature, did not serve as the primary channels through which the policy improved sleep.

3.1 Introduction

China grapples with severe air pollution, as particulate matter concentrations frequently exceed international health standards. According to the China Environmental Quality Bulletin, in 2013, 96% of monitored cities did not meet China's National Ambient Air Quality Standard (NAAQS) of 35 $\mu\mathrm{g}/\mathrm{m}^3$ (MEP, 2013). These cities reported an average annual PM2.5 concentration of 72 $\mu g/m^3$, significantly surpassing the WHO guideline of 5 $\mu\mathrm{g}/\mathrm{m}^3$. The Beijing-Tianjin-Hebei (Jing-Jin-Ji) region, with an average PM2.5 concentration of 106 $\mu\mathrm{g}/\mathrm{m}^3$, is one of the most polluted areas in China. This level of pollution presents substantial health and environmental challenges and makes a notable contribution to global climate change.

Residential coal burning accounts for nearly 30% of outdoor PM2.5 pollution in northern China, as noted by Liu et al., 2016. This high level of pollution is largely attributed to the extensive use of coal and biomass not only for cooking but also for space heating across this region. In the rural areas of northern China, traditional heating methods such as the biomass or coal-fueled kang and stove are still widely used. The kang serves multiple purposes including space for sleeping, space heating, and occasionally cooking. Its design, shown in Figure (3.1.1), includes a fireplace, a main platform for sleeping and living activities, and a chimney. Beneath the platform, flues distribute hot air throughout the room, providing warmth (Guo, 2002). This setup, though traditionally efficient, poses health risks and may affect sleep quality due to high levels of indoor air pollution.

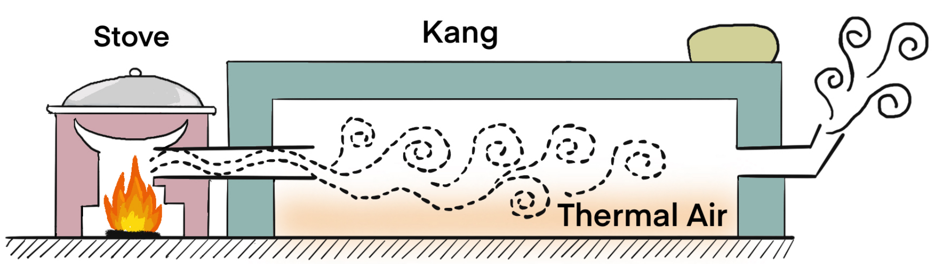
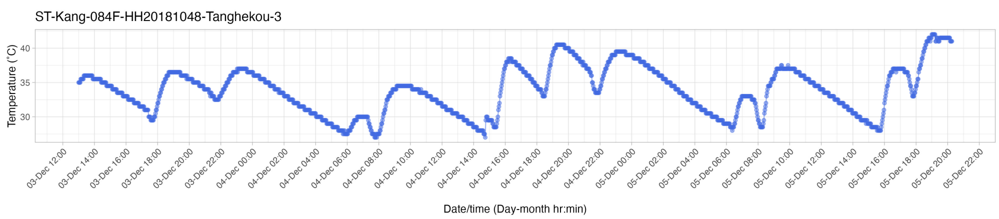


Figure 3.1.1: Illustration of Kang

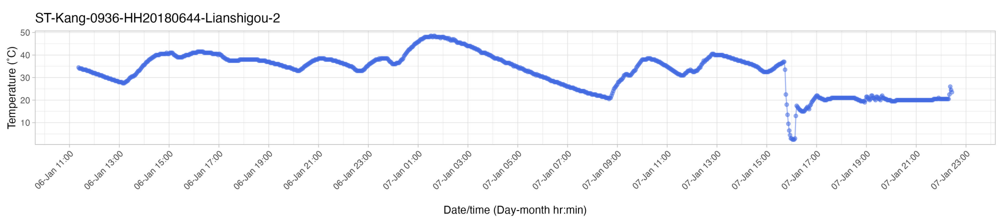
Figure 3.1.2: Coal stoves and stockpiled coal in rural Beijing.

These traditional heating methods produce significant indoor air pollution as the burning of coal and biomass releases particulate matter (PM) that can deeply penetrate the lungs, enter the bloodstream, and trigger severe cardiovascular, cerebrovascular, and respiratory issues. Prolonged personal exposure to such polluted air during the night, when individuals are typically in close proximity to these heat sources, exacerbates the risk. Additionally, the nature of these traditional heating systems, which often require periodic refueling during the night, can lead to sleep interruptions, further degrading sleep quality. Moreover, the heat produced by kangs and similar setups is not always constant; the temperature can fluctuate as the fuel burns down, leading to unstable sleeping environments. This variability in temperature, sometimes reaching up to 30°C as shown in Figure (3.1.3), not only makes it difficult to maintain optimal sleeping conditions but also may cause frequent awakenings or disrupted sleep patterns, compounding the health impacts associated with exposure to pollutants.

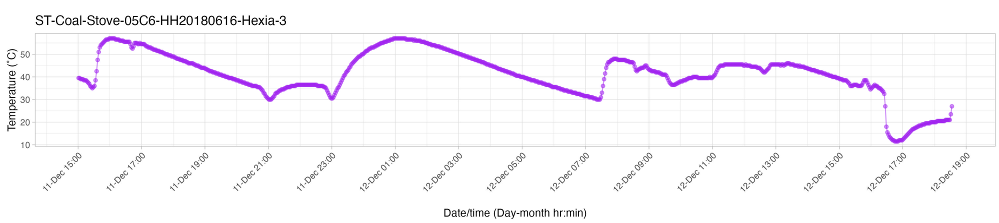
Given the crucial role of sleep in maintaining health, disruptions caused by traditional heating methods can have profound implications. Sleep serves as a fundamental pillar of health, essential for optimal cognitive, emotional, and physical functioning (Cappuccio et al., 2010; Lim and Dinges, 2010; Baglioni et al., 2011). Adequate and undisturbed sleep bolsters immune function, supports cardiac health, and plays a critical role in memory consolidation and emotional regulation. Consequently, the health risks from traditional heating methods extend beyond direct exposure to pollutants and affect overall well-being by degrading sleep quality. Research highlights that the relationship between sleep duration and health is not linear; instead, there is a recommended sleep range deemed optimal for health. The National Institutes of Health (NIH) and the Centers for Disease Control and Prevention (CDC) advise adults to aim for 7 to 9 hours of sleep per night. Deviating from this range, either through short or long sleep durations, has been linked to adverse health outcomes, including chronic diseases, mental health issues, and increased mortality rates (Itani et al., 2017; Shan et al., 2015). Further studies illustrate a U-shaped relationship between sleep duration and health risks, identifying 7-8 hours as the optimal range to minimize health risks (Wang et al., 2016; Yin et al., 2017).



Temperature Variability of Kang (1)



Temperature Variability of Kang (2)



Temperature Variability of Coal Stove

Figure 3.1.3: Temperature fluctuations of traditional heating methods
Source: Plot created by Talia Sternbach, Department of Epidemiology, McGill University.

To address the severe air pollution issues, particularly those exacerbated by residential coal burning for winter heating, the Chinese government initiated a robust series of policies to encourage

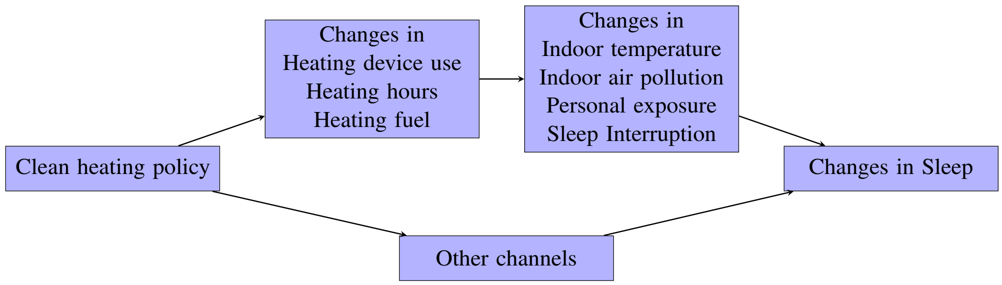


Figure 3.1.4: Impact Mechanisms of the Clean Heating Policy on Sleep Outcomes

the shift from coal to cleaner fuels such as electricity and natural gas. One pivotal initiative, the Clean Heating Policy introduced in 2017, specifically targets the Jing-Jin-Ji region and aims to facilitate a complete transition from coal to gas and electricity in designated areas. This policy mandates the cessation of coal use in these regions, supporting households with subsidies for both the purchase and installation of heat pumps, as well as for ongoing fuel costs in subsequent winters.

By July 2022, northern China reached a clean heating rate of 73.6%, surpassing the target set by the Plan to attain a 70% rate. This achievement has facilitated the replacement of use of over 150 million tons of scattered coal, according to the National Energy Administration, 2022. This substantial shift has led to a significant reduction in average PM2.5 concentrations across the region, improving the air quality index by more than one-third (National Energy Administration, 2022). The introduction of the clean heating policy results in changes to household heating fuel choices and the devices used, as well as adjustments in the allocation of heating hours among different devices. These alterations in heating practices are likely to subsequently impact indoor temperature levels, indoor air pollution, and personal exposure to pollutants. Additionally, these changes may reduce sleep disruptions previously caused by the need to add fuel during the night. The relationship between these environmental changes and individual sleep is supported by existing literature, which highlights indoor temperature, air quality, and pollutant exposure as significant factors influencing sleep quality and duration, as illustrated in Figure (3.1.4).

Given the anticipated significant benefits to sleep from the clean heating policy, our study aims to investigate its impact on the sleep durations of residents in the affected regions. This is important because sleep is crucial for individual health, influencing cognitive, emotional, and physical well-being (Cappuccio et al., 2010; Lim and Dinges, 2010; Baglioni et al., 2011). Previous research has documented varying effects of interventions aimed at improving sleep across different demographic groups, whether through pharmaceutical treatments, lifestyle modifications, or sleep hygiene improvements (Buscemi et al., 2004; Baron et al., 2021; Irish et al., 2015). This variability suggests that the impact of the clean heating intervention on sleep might also differ among populations, prompting our investigation into which demographic groups experience the most significant improvements. Our study additionally considers the mechanisms through which clean heating could influence sleep, as illustrated in Figure (3.1.4). These mechanisms include reductions in indoor air pollution and improvements in temperature stability within homes, which are believed to enhance sleep quality.

The literature's findings of a nonlinear relationship between sleep duration and health risks indicate that the clean heating intervention may have varied effects on individuals with different initial sleep durations. Relying solely on total sleep hours as an outcome may not fully capture these diverse impacts. This study aims to determine which groups are most influenced by the policy and observe how their sleep patterns change. To do this, we categorize participants into groups such as short sleepers and long sleepers, and we examine deviations from the recommended sleep durations. By analyzing a variety of sleep outcomes, we can better understand the heterogeneous effects of the clean heating policy and how it influences transitions between different types of sleep behavior.

This paper analyzes three rounds of self-reported sleep duration data collected during the clean heating program in rural Beijing. The dataset comprises 990 participants across 50 villages, none of which had received treatment at the baseline in winter 2018. During the sampling period from 2018 to 2022, 40% of these participants were scheduled for staggered treatments. The findings reveal that the policy significantly impacts specific demographics, including short sleepers, long sleepers, males, and individuals aged 60-70, especially in regions undergoing significant transitions from polluting to clean heating fuels. Although the policy effectively transitioned home heating from coal to electricity, which was expected to improve air quality and stabilize indoor temperatures, however, these changes did not significantly act as the major channel for improving the sleep quality of the individuals involved.

This paper makes a contribution by estimating the treatment effects of the clean heating policy on sleep, an area previously unexplored within the context of fuel substitution policy. Recognizing the fundamental role of sleep in individual well-being, our study extends the discourse beyond the commonly discussed environmental impacts of such policies, offering insights into their broader implications on human health. By conducting a combined analysis of various outcomes, we specifically analyze heterogeneous treatment effects across various sleeper types, recognizing that more sleep isn't always beneficial. Additionally, we investigate the potential mechanisms through which the clean heating policy could influence sleep quality, such as improvements in indoor air quality and temperature stability.

The structure of the paper is as follows: Section 2 reviews the relevant literature. Section 3 describes the data collection and the methodology employed. Section 4 outlines our empirical strategy. Section 5 discusses the results. Section 7 conducts robustness checks to validate the reliability of our findings. Section 8 discusses the broader implications of our results, integrating them into the context of existing literature and concludes the paper by suggesting avenues for future research.

3.2 Literature Review

3.2.1 Sleep Duration and Health

Sleep is indispensable for cognitive, emotional, and physical well-being, with the duration of sleep playing a critical role in maintaining health. Adequate sleep duration is linked to various benefits, while deviations from recommended sleep lengths are associated with numerous health risks. Extensive research has demonstrated that both short and long sleep durations correlate with increased risks of mortality, diabetes, cardiovascular disease, and obesity. Systematic reviews such as those by Yin et al., 2017 and Itani et al., 2017 have consistently found that deviations from the optimal 7-8 hours of sleep per night are associated with higher all-cause mortality. Similarly, Wang et al., 2016 reported that sleep durations deviating from this range correlate with a higher incidence of cardiovascular diseases.

The impact of sleep extends beyond physical health to mental well-being. The relationship between sleep duration and mental health, however, presents mixed evidence. Dong et al., 2022 found that both insufficient and excessive sleep contribute to an increased risk of depression among a large cohort of U.S. adults. In contrast, Jike et al., 2018 observed no significant association between long sleep durations and the onset of depression, while Itani et al., 2017 reported that insufficient sleep did not significantly link to depression.

Cognitive function is also affected by sleep. Research by Ma et al., 2020 tracking nearly 30,000 individuals in China and England over several years revealed that extreme sleep durations (less than 4 or more than 10 hours per night) were associated with a significantly faster decline in cognitive functions. Meanwhile, Devore et al., 2016 found mixed results among older adults, noting that longer sleep durations were more often correlated with poorer cognitive outcomes than shorter ones.

Although these studies typically do not provide evidence of a causal effect of sleep on health outcomes, it is still important to understand how variations in sleep patterns can influence overall well-being and guide effective interventions.

3.2.2 Factors Influencing Sleep

Various internal and external factors significantly influence individual sleep patterns. Internally, age, sex, health conditions, and psychological stress play critical roles in determining sleep quality and duration. Externally, environmental variables such as light exposure, noise and temperature all have substantial impacts.

In this section, we delve into how environmental and lifestyle factors, especially those related to the clean heating intervention, significantly impact sleep quality and duration. The transition prompted by clean heating policies moves households from traditional heating methods, such as kang and stoves fueled by coal, to more advanced electric heat pumps. This shift is expected to offer numerous advantages, including stable room temperatures, reduced indoor air pollution, and fewer sleep interruptions, as there is no longer a need to add fuel during the night..

Thermal comfort is a pivotal factor affecting sleep. Research consistently demonstrates that the human body is highly sensitive to air temperature during sleep, with even moderate deviations from comfort temperatures markedly impairing sleep quality. For example, Okamoto-Mizuno and Mizuno, 2012 noted that heat exposure alters core body temperature and sleep stages, significantly diminishing sleep quality. Similarly, Lan et al., 2017 observed that exposure to moderately non-neutral temperatures substantially increases sleep onset latency and reduces slow wave sleep (SWS), indicating a decline in sleep quality under such conditions. Furthermore, the optimal room temperature for sleep is generally considered to be between 15.6 to 19.4 °C (National Sleep Foundation, n.d.). Consistent with this, the World Health Organization recommends maintaining a minimum air temperature of 18°C in bedrooms to promote healthy sleep environments (Ranson, Organization, et al., 1988). These findings underscore the importance of maintaining appropriate and stable room temperatures to facilitate optimal sleep conditions, highlighting the potential benefits of the clean heating intervention in improving sleep environments.

Air pollution adversely impacts sleep quality through mechanisms such as airway irritation, inflammation, or stress response activation, and there is strong evidence linking ambient air pollution to poor sleep quality globally. Studies like Zanobetti et al., 2010 found that increases in PM10 levels in the U.S. were associated with poorer sleep quality. Similarly, Abou-Khadra, 2013 linked higher PM10 exposure to disorders of sleep among children in Egypt. This pattern repeats in findings from Turkey (Lawrence et al., 2018) and China, where studies by Yu et al., 2019 and Chen et al., 2019 observed reductions in sleep duration and deteriorations in sleep quality correlated with increased pollution levels. These studies collectively imply the potential benefits of transitioning from coal to electricity for heating, highlighting the likely improvements in both ambient air pollution and personal exposure, thereby enhancing individual sleep quality.

Interruption during sleep can impact health, and this is particularly relevant in the context of clean heating policies. Traditional heating methods like the kang or coal stove require nocturnal refueling, which disrupts sleep. According to Medic et al., 2017, sleep disruptions can lead to serious health issues. In the short term, they increase stress, emotional distress, and impair cognitive functions. Over the long term, disrupted sleep can contribute to hypertension, cardiovascular disease, and type 2 diabetes, and can exacerbate symptoms of various medical conditions. Thus, policies that promote continuous heating could improve sleep continuity and overall well-being.

Again, the aforementioned studies identified associations between sleep and various factors, rather than establishing causal relationships.

3.2.3 Optimal Sleep Duration

Building on the discussion in the previous sections, the relationship between sleep duration and health outcomes is complex and warrants detailed examination. Extremes in sleep duration—both short and long—have been associated with a variety of health risks, underscoring the importance of an ”optimal” sleep duration. Major health organizations have stepped forward with recommendations to guide the public. The National Institutes of Health (NIH) advises that adults aim for 7-9 hours of sleep daily. Similarly, the Centers for Disease Control and Prevention (CDC) provides tiered recommendations based on age: individuals between 18 and 60 years should get at least 7 hours, those aged 61 to 64 years should sleep 7-9 hours, and those over 65 years should target 7-8 hours nightly to maintain optimal health.

In the literature, short sleep duration is typically defined as less than 6 or 7 hours per day (Uehli et al., 2014). This condition has been linked to numerous adverse health outcomes, including an increased risk of chronic diseases such as obesity, diabetes, and cardiovascular diseases, as well as detrimental effects on mental health, including depression and anxiety (Itani et al., 2017).

Conversely, long sleep, generally classified as exceeding 8 or 9 hours per day (Jike et al., 2018), is also associated with negative health impacts. These include an elevated risk of cardiovascular diseases, cognitive decline, depression, and even increased mortality rates (Shan et al., 2015).

Furthermore, research on dose-response relationships reveals an inverted U-shaped curve between sleep duration and health risks, suggesting a nonlinear association. Optimal health outcomes, including lower risks of mortality, cardiovascular disease, mental health issues, obesity, and bone health issues, are most frequently observed at sleep durations of 7-8 hours per night (Wang et al., 2016; Yin et al., 2017). These findings emphasize the importance of balanced sleep duration to maximize health benefits and minimize risks.

3.2.4 Heterogeneous Treatment Effect for Sleep Intervention

In recent years, there has been a growing body of evidence highlighting the prevalence of sleep disorders. This increasing awareness has spurred the development of a variety of strategies and treatments aimed at enhancing sleep quality and addressing sleep disorders. These interventions encompass behavioral, pharmacological, and environmental approaches. Among the behavioral strategies, cognitive-behavioral therapy for insomnia, sleep hygiene education, and relaxation training are particularly noteworthy. The efficacy of these sleep interventions has been the subject of extensive research, revealing that their impacts on sleep quality, duration, and broader health outcomes vary significantly across different demographic groups. For instance, Buscemi et al., 2004 observed that the effectiveness of melatonin varies widely, particularly in relation to the severity and type of sleep disorders faced by individuals. Specifically, melatonin was observed to reduce sleep onset latency by 10.7 minutes in individuals with primary sleep disorders, but it did not have a statistically significant effect on normal sleepers or those with secondary sleep disorders. Baron et al., 2021 reported that the effectiveness of behavioral interventions in extending sleep duration varies across different demographic groups, with larger effects observed in adults and youth compared to college students. Research by Irish et al., 2015 suggests that individual differences play a critical role in determining the effectiveness and applicability of specific sleep hygiene practices. These factors include age, gender, genetic polymorphisms, education level, comorbid health conditions, and social or occupational demands.

The effects of sleep interventions are not only diverse across demographic groups but are also influenced by the initial sleep quality and patterns of the individuals involved.

Clean heating interventions, primarily aimed at reducing air pollution, may also inadvertently improve sleep quality, thereby functioning as an indirect form of ”sleep intervention.” Although these measures are not designed with sleep improvements as their main objective, the potential spillover effects on sleep cannot be overlooked. Given the varied nature of responses to direct sleep interventions, it is reasonable to anticipate that clean heating interventions would similarly exhibit heterogeneous treatment effects. These effects likely vary across different demographic groups and depend on the initial sleep statuses of individuals.

3.3 Data

3.3.1 Beijing Household Energy Transition Project

The Beijing Household Energy Transition Project[1] was initiated to comprehensively assess the impacts of the clean heating policy on a variety of dimensions — including air pollution, indoor temperature, public health, household energy usage, and overall well-being. This project employed a multi-level approach to data collection: at the household level, data on indoor air temperature, PM2.5 levels, and energy use and expenditures were gathered through sensors and questionnaires. At the individual level, questionnaires collected data on health status, behaviors, medication use, as well as personal exposure to PM2.5 and black carbon. Notably, this study also included the key outcome of this paper–a behavioral measure of self-reported sleep duration, obtained via the questionnaire query: "How many hours do you typically sleep per day (including naps)?"

A unique strength of this dataset lies in its comprehensive inclusion of information on policy treatment, sleep outcomes, and additional variables that facilitate an in-depth investigation of heterogeneous treatment effects and potential mediators. This richness allows for a multifaceted analysis within a single study, enabling us to understand not only the direct impacts of the policy but also the varied responses among different demographic groups and the underlying mechanisms at play.

The study was conducted in four districts of Beijing, encompassing 50 villages that were randomly selected to participate in the project. In each village, approximately 20 households were chosen with the assistance of the village head, and within each household, one individual was randomly selected for the individual-level measurements. Should a participant drop out, another

[1]The project is primarily led by Prof. Jill Baumgartner and Sam Harper at McGill University, with funding from the Health Effects Institute (HEI), the Canadian Institutes of Health Research (CIHR), the Social Sciences and Humanities Research Council (SSHRC), and the Government of Canada.

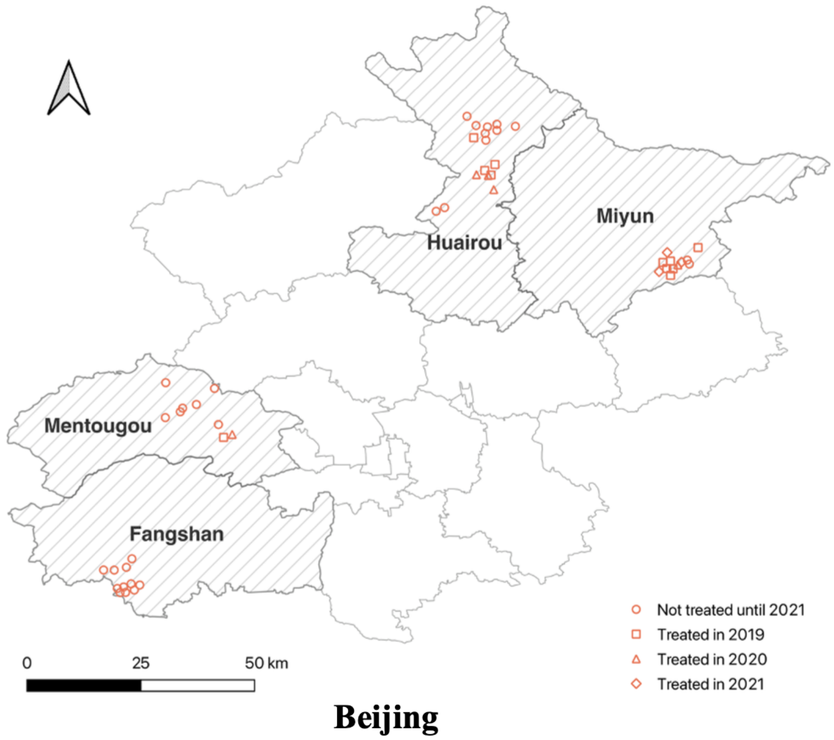


Figure 3.3.1: Treatment Map

individual from the same household, or from a newly enrolled household, will be selected to continue the study.

Data collection occurred over four waves, spanning from winter 2018 to winter 2021, with the clean heating policy being implemented at village level and in a staggered manner throughout this period. The timing and distribution of policy implementation are illustrated in Figure (3.3.2). At the outset (baseline), no villages had received the intervention, but by winter 2021, 20 out of the 50 villages had been treated. The third wave of data collection coincided with the peak of the COVID-19 pandemic, which restricted the collection to only air quality and temperature measurements. Therefore, for the purpose of analyzing sleep data, only the information from waves 1, 2, and 4 will be utilized, as these waves provide the most complete datasets relevant to this analysis.

3.3.2 Variables Description

3.3.2.1 Definitions

In this study, we analyze multiple sleep outcomes to assess the effects of clean heating policy on sleep patterns. The primary outcome we examine is Sleep Duration, calculated based on self-reported typical daily sleep duration. To categorize individuals into different sleeper types, we establish three binary indicators aligned with sleep duration recommendations from health institutions and existing literature. Specifically, we define Short Sleep as having a sleep duration of less than 7 hours per day, and Long Sleep as a duration exceeding 9 hours per day. Consequently, Normal Sleep is defined as a sleep duration ranging from 7 to 9 hours per day.

To further analyze deviations from the normal sleep range, we introduce two additional variables. The first, Deviation from Normal Sleep, is calculated by subtracting 7 hours from the sleep duration for short sleepers, or 9 hours for long sleepers, depending on the category. The second variable, Absolute Deviation from Normal Sleep, measures the shortest distance to the normal sleep range, calculated as the absolute value of the Deviation from Normal Sleep. These metrics allow us to quantify the different ways in which individual sleep patterns deviate from recommended norms. The definitions of all outcome variables are presented in Panel A of Table (3.3.1).

To explore potential mechanisms by which the clean heating policy might influence individual sleep patterns, as illustrated in Figure (3.1.4), we have identified several sets of mediating variables as detailed in Panel B of Table (3.3.1). The first set includes heating fuel choices and heating hours by fuel type. This encompasses patterns of bedroom heating fuels categorized as completely clean, mixed, or entirely polluting. Clean fuels include electricity, natural gas, liquefied petroleum gas (LPG), and solar energy; polluting fuels comprise coal, firewood, charcoal, and crop residues; and mixed fuels represent a combination of both clean and polluting sources. We further analyze the primary heating fuel used in bedrooms, determined by the fuel utilized by the device accounting for the most heating hours across all bedrooms. Additionally, we assess both the total and average bedroom heating hours by fuel type (clean or polluting) on a typical day.

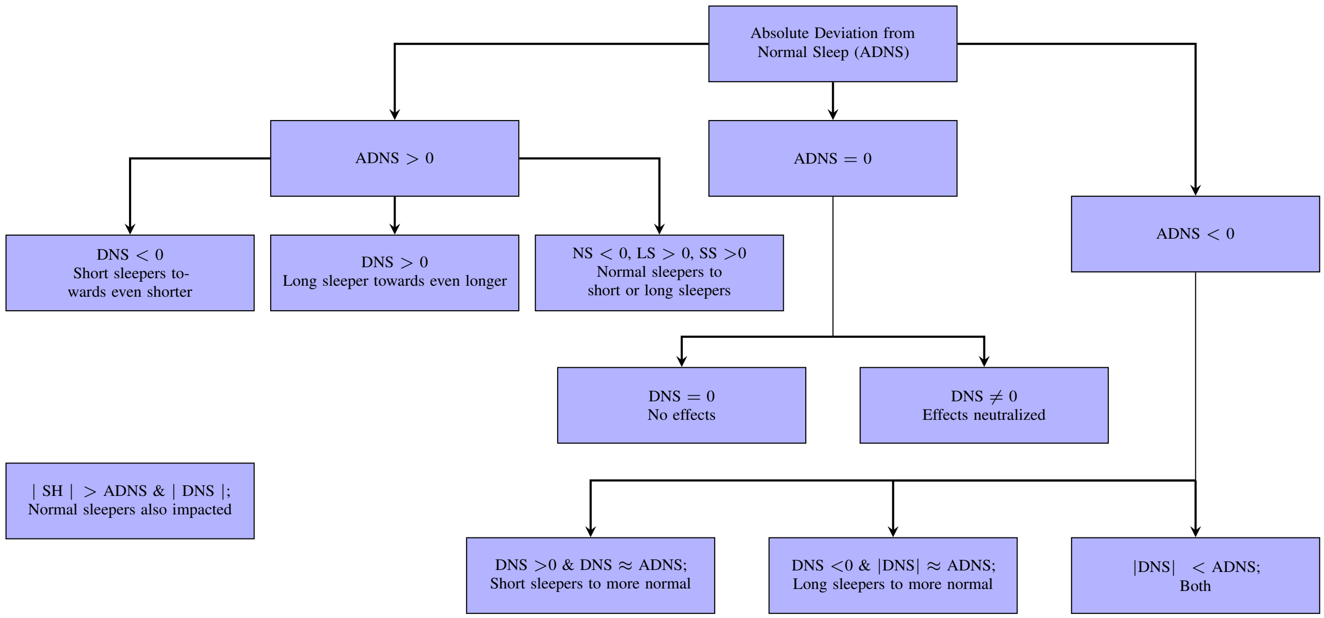


Figure 3.3.2: Diagram illustrating combined analysis

The second set of mediating variables comprises average indoor air pollution during the heating season, personal exposure to PM2.5, average indoor temperature during typical sleep hours, and variability of indoor temperature during sleep hours. Detailed definitions for these variables are provided in Panel B-Set B of Table (3.3.1).

3.3.2.2 Combined Analysis

The rationale for utilizing multiple sleep outcomes in this study stems from the understanding that sleep duration is not a simple matter of 'more is always better.' Instead, there exists an optimal or recommended range of sleep that is considered healthy. Given that the intervention may have varied effects on different types of sleepers—short, normal, and long—relying solely on total sleep hours as an outcome may not fully capture these nuanced impacts. By analyzing a spectrum of sleep outcomes, we can more effectively identify which categories of sleepers are most influenced by the intervention. This comprehensive approach allows us to observe patterns of how individuals may transition between different sleep types as a result of the intervention.

Variables	Definition
Outcome Variables	
Sleep Duration	= Self-reported typical sleep duration per day
Short Sleep	= 1 if sleep duration $<$ 7 hours; 0 otherwise
Normal Sleep	= 1 if sleep duration between 7 $\sim$ 9 hours; 0 otherwise
Long Sleep	= 1 if sleep duration $>$ 9 hours; 0 otherwise
Deviation from Normal Sleep	= Sleep Duration - 7 if sleep duration $<$ 7 hours; = Sleep Duration - 9 if sleep duration $>$ 9; 0 otherwise
Absolute Deviation from Normal Sleep	= $\min$ { |Sleep Duration - 7| , |Sleep Duration - 9|} if sleep duration $<$ 7 or $>$ 9; 0 otherwise
Mediation Variables	
Set A: Heating Fuel Choices and Behaviors	
Bedroom Heating Fuel Pattern	
Clean	= 1 if all bedroom heating fuels chosen by the household are clean; 0 otherwise
Polluting	= 1 if all bedroom heating fuels chosen by the household are polluting; 0 otherwise
Mixed	= 1 if household uses a combination of both clean and polluting fuels for bedroom heating; 0 otherwise
Bedroom Primary Heating Fuel	
Electricity	= 1 if electricity is the primary fuel utilized by the household, accounting for the most heating hours across all bedrooms; 0 otherwise.
Coal	= 1 if coal is the primary fuel utilized by the household, accounting for the most heating hours across all bedrooms; 0 otherwise.
Wood	= 1 if wood is the primary fuel utilized by the household, accounting for the most heating hours across all bedrooms; 0 otherwise.
Total Bedroom Heating Hours by Fuel Type	
Clean	= total heating hours by clean fuels across all bedrooms in the household
Polluting	= total heating hours by polluting fuels across all bedrooms in the household
Average Bedroom Heating Hours by Fuel Type	
Clean	= average heating hours by clean fuels across bedrooms with daily heating in the household
Polluting	= average heating hours by polluting fuels across bedrooms with daily heating in the household
Set B: Environmental Variables	
Average Indoor Air Pollution Level During Heating Season	= the average concentration of indoor air pollutants in micrograms per cubic meter ($\mu g/m^3$) during the heating season
Personal Exposure to PM2.5	= individual exposure to particulate matter of 2.5 micrometers or smaller, measured in micrograms per cubic meter ($\mu g/m^3$)
Average Indoor Temperature during Typical Sleep Hours	= the average temperature inside a dwelling during typical sleep hours, from 10:00 PM to 8:00 AM, measured in degrees Celsius (°C)
Variability of Indoor Temperature during Typical Sleep Hours	= standard deviation of indoor temperature measured during typical sleep hours, from 10:00 PM to 8:00 AM

Table 3.3.1: Definition of Variables

Note: Clean fuels include electricity, natural gas, liquefied petroleum gas (LPG), and solar energy; polluting fuels comprise coal, firewood, charcoal, and crop residues.

Figure 3.3.2 illustrates this approach. To analyze the impact of sleep-related variables, the coefficient of the Absolute Deviation from Normal Sleep (ADNS) can be examined in three distinct scenarios: ADNS equals zero, greater than zero, and less than zero.

For the case where ADNS $= 0$, further examination involves analyzing the Deviation from Normal Sleep (DNS). If DNS is also zero, it implies there are no effects across all categories of sleepers. However, if DNS is not zero, this suggests that the effects on sleep are counterbalanced by opposing yet equivalent effects from both short and long sleepers.

In the scenario where ADNS > 0, if DNS is less than zero, it indicates that short sleepers are experiencing shorter sleep durations than usual. Conversely, if DNS is greater than zero, this suggests that long sleepers are tending towards longer sleep durations. Further insights can be derived by examining the signs of binary variables such as Normal Sleep (NS), Short Sleep (SS), and Long Sleep (LS). For example, a negative NS combined with a positive LS suggests that normal sleepers are transitioning towards longer sleep durations; similarly, a negative NS with a positive SS indicates a shift towards shorter sleep durations.

When ADNS is less than zero, several results are possible depending on the values of DNS. If DNS is greater than zero and closely matches ADNS, this suggests that short sleepers are moving towards normal sleep durations by increasing their hours of sleep. Conversely, if DNS is less than zero and its absolute value is similar to ADNS, it indicates that long sleepers are reducing their sleep times towards normal durations. If DNS is approximately zero and its absolute value matches ADNS, then both short and long sleepers are adjusting their sleep towards normal durations. This analysis helps identify how deviations from normal sleep are being corrected among different groups, providing a clear picture of the directional changes in sleep behavior relative to normal sleep standards.

Additionally, examining the signs of NS, LS, and SS will clarify whether the impacted individuals have achieved or are approaching normal sleep durations. By analyzing the magnitude of Sleep Hours (SH) and deviation variables (ADNS, DNS), we can ascertain the broader impact on sleep patterns. Particularly, if the absolute value of SH is greater than ADNS and greater than the absolute value of DNS, it can be inferred that normal sleepers are affected. This structured approach offers a detailed understanding of the myriad ways different factors influence sleep patterns across various categories of sleepers.

3.4 Empirical Strategy

Given the quasi-random design of the study, we utilize the Difference-in-Differences (DID) method (Lechner et al., 2011) to estimate the causal effect of the clean heating policy on sleep. This econometric technique is particularly suited to this context as it compares the before-and-after differences in outcomes between treated and untreated groups, effectively controlling for time-invariant unobserved confounders that could bias the results, as well as for changes that are common to both groups over time.

The model employed in this analysis is the Two-way Fixed-effect (TWFE) Model, which is specified as follows:

$$Y_{ijt} = \alpha_j + \gamma_t + \beta \times D_{jt} + X_{ijt} + \epsilon_{ijt} \tag{3.1}$$

In this equation, Y_{ijt} represents the sleep outcome for individual i in village j in year t, D_{jt} is the treatment status indicator, where $D_{jt} = 1$ if village j has received the clean heating policy treatment in year t, —signifying a ban on coal use and the installation of heat pumps with subsidies for purchase, installation, and subsequent electricity bills. $D_{it} = 0$ indicates the absence of treatment. X_{ijt} includes individual time-variant control variables that might affect sleep, such as socioeconomic status, γ_t represents fixed effects for each time period to control for external factors affecting all units, such as seasonal variations or economic cycles. α_j is the village fixed effect, capturing unobserved characteristics that are constant over time but vary across villages, such as village location and environmental conditions. ϵ_{ijt} is the error term. β represents the estimated effect of the treatment on sleep outcomes.

A crucial assumption underpinning the Difference-in-Differences (DID) method is the parallel trends assumption (Imbens and Wooldridge, 2009). This posits that, in the absence of the intervention, the outcome trends for both the treatment and control groups would have progressed similarly over time. Establishing this assumption is essential, as it ensures that any observed post-intervention differences can be causally attributed to the intervention itself, rather than to pre-existing trends. To validate this assumption, we will conduct pre-trend tests. Additionally, it is important to ensure that there are no other time-varying confounders that could affect the results.

This empirical strategy, employing the Two-Way Fixed Effects (TWFE) Model, is applied to all treatment effect estimations in this study. This includes analyzing the average treatment effects on sleep, heterogeneous treatment effects on sleep, and treatment effects on bedroom heating fuel choices and heating hours.

3.5 Empirical Results

3.5.1 Summary Statistics

Table (3.5.1) presents the summary statistics for sleep outcomes. Across three waves, the dataset comprises 3,110 observations. The average sleep duration for participants throughout the study period is 7.77 hours, ranging from a minimum of 1 hour to a maximum of 20 hours. Of these observations, 56% fall within the normal sleep duration, 25% are categorized as short sleep, and 19% as long sleep. The average deviation from normal sleep is -0.18 hours, and the average absolute deviation is 0.74 hours. Among the observations classified as short sleep, the average duration is 5.17 hours, with more than half of these observations ranging between 4.5 and 6 hours. For long sleep observations, the average duration is 10.51 hours, with most falling between 9.5 and 11 hours.

Figure (3.5.1) illustrates the baseline distribution of all sleep outcomes. The initial distribution of sleep outcomes shows a marked similarity between the control and treatment groups, indicating that both groups were comparable prior to the intervention. Table (3.5.2) provides the summary statistics for the control and treatment groups at baseline. The data reveal that individual characteristics, such as age, gender, socioeconomic status, and health conditions, are closely aligned between the groups. This similarity supports the effectiveness of the quasi-random assignment in creating comparable groups, which is crucial for the validity of the Difference-in-Differences analysis. Such alignment at baseline add support to the assumption that any post-intervention

	mean	sd	min	p25	p50	p75	max	N
Full Sample								
Sleep Hours	7.77	1.98	1.00	6.50	8.00	9.00	20.00	3110
Normal Sleep	0.56	0.50	0.00	0.00	1.00	1.00	1.00	3110
Short Sleep	0.25	0.43	0.00	0.00	0.00	1.00	1.00	3110
Long Sleep	0.19	0.39	0.00	0.00	0.00	0.00	1.00	3110
Deviation from Normal Sleep	-0.18	1.32	-6.00	-0.50	0.00	0.00	11.00	3110
Absolute Deviation from Normal Sleep	0.74	1.10	0.00	0.00	0.00	1.00	11.00	3110
Short Sleep								
Sleep Hours	5.17	1.12	1.00	4.50	5.50	6.00	6.50	781
Long Sleep								
Sleep Hours	10.51	1.00	9.50	10.00	10.00	11.00	20.00	579

Table 3.5.1: Summary Statistics of Sleep Outcomes

differences in sleep outcomes can be attributed to the effects of the clean heating policy rather than pre-existing differences between the groups.

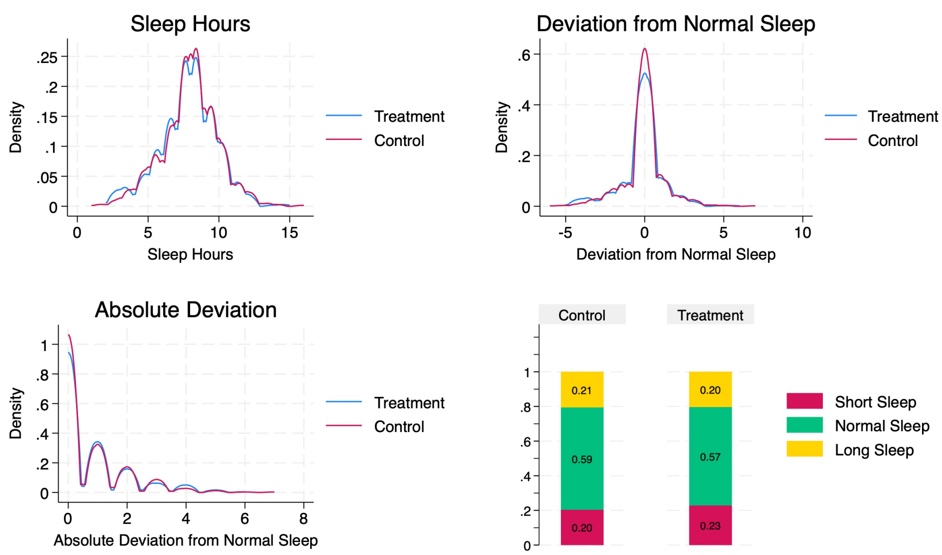


Figure 3.5.1: Sleep Outcomes Distribution at Baseline

	Control			Ever Treated		
	Mean	SD	N	Mean	SD	N
Age	60.36	9.30	592	60.79	9.15	398
Male	0.40	0.49	594	0.40	0.49	398
Married	0.90	0.30	582	0.87	0.34	389
Self-reported Health Status						
Excellent	0.04	0.18	595	0.03	0.18	398
Good	0.15	0.36	595	0.22	0.42	398
Fair	0.45	0.50	595	0.33	0.47	398
Poor	0.36	0.48	595	0.41	0.49	398
Respirotory Health						
Good	0.74	0.44	594	0.74	0.44	396
Fair	0.22	0.41	594	0.21	0.41	396
Poor	0.05	0.21	594	0.05	0.22	396
Weekly Drinking	0.25	0.43	595	0.27	0.44	398
Regular Exercising	0.76	0.43	595	0.74	0.44	398
Regular Farming	0.49	0.50	595	0.55	0.50	398
Satisfaction						
Life	7.63	2.16	593	7.22	2.38	394
Living Condition	7.47	4.43	593	7.41	5.24	394
Income	5.74	6.03	593	5.19	5.38	393
Log Family Income	9.88	0.69	595	9.82	0.73	398
Wealth Index	0.01	4.01	574	0.07	4.49	388

Table 3.5.2: Summary Statistics at Baseline

3.5.2 Treatment Effects of Clean Heating Policy on Sleep

3.5.2.1 Parallel Trends

To verify the parallel trends assumption, we employ a pre-trends test. This test involves estimating dynamic treatment effects separately for each pre-intervention and post-intervention period within the same regression model. By doing so, we can assess whether any significant treatment effects exist before the actual treatment implementation. If these pre-treatment effects are statistically insignificant, it supports the contention that the parallel trends assumption has not been violated.

Figure (3.5.2) presents the results from these pre-trend tests for various sleep outcomes. The emergence of significant pre-treatment effects for ”Short Sleep” and ”Sleep Hours” indicate pre-existing differences in sleep patterns. However, it is important to note that the direction of these pre-treatment effects is opposite to that observed in the post-treatment period. This contrasting directional trend between the pre- and post-treatment periods implies that the actual impact of the treatment might be underestimated.

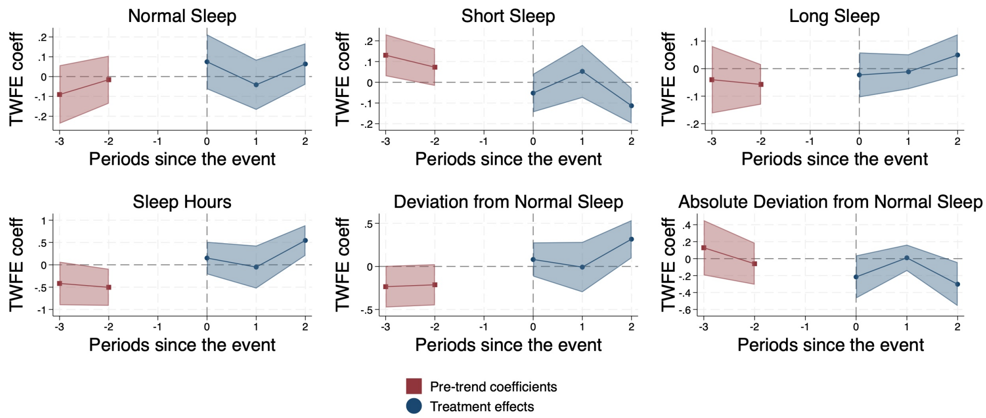


Figure 3.5.2: Pre-trends test for sleep outcomes

Note: Dynamic treatment effects estimated by the TWFE Model are displayed alongside their 95% confidence intervals. The -1 period is omitted as the reference period. Units for Sleep Hours and Deviation Variables are presented in hours, and the units for sleep categories (normal, short, and long sleep) are shown as probabilities.

3.5.2.2 Average Treatment Effects

Figure (3.5.3) presents the estimated average treatment effects of the clean heating policy on various sleep outcomes. The analysis indicates only minor treatment effects: on average, the policy results in less than a 15-minute change in sleep duration for all participants, and the probability of transitioning to a normal sleep duration has increased by less than 5% from either short or long sleep durations. Moreover, none of the estimated effects are significant at the 5% significance level. This suggests that the intervention may not have had a measurable impact on the overall sleep patterns of the study participants. However, it remains valuable to employ a combined analysis approach to determine if certain types of sleepers benefit from the intervention. This will provide valuable insights for the analysis of heterogeneous treatment effects in subsequent sections.

Using the approach illustrated in Figure (3.3.2), we find that the coefficients for the Absolute Deviation from Normal Sleep are negative, while those for Deviation from Normal Sleep are positive, both of similar magnitude. This pattern suggests that the primary impact of the intervention was on short sleepers, encouraging them towards longer sleep durations, rather than reducing sleep among long sleepers. This interpretation is further supported by the coefficients on binary outcomes: ”Short Sleep” shows a negative coefficient, ”Long Sleep” is almost zero, and ”Normal

Sleep" is positive and roughly the same magnitude as "Short Sleep." These findings imply that the intervention has helped transition some of the short sleepers to the normal sleep category.

Moreover, the larger magnitude of the coefficient for "Sleep Hours" compared to "Deviation" variables indicates that normal sleepers may also be experiencing benefits from increased sleep duration, albeit modestly.

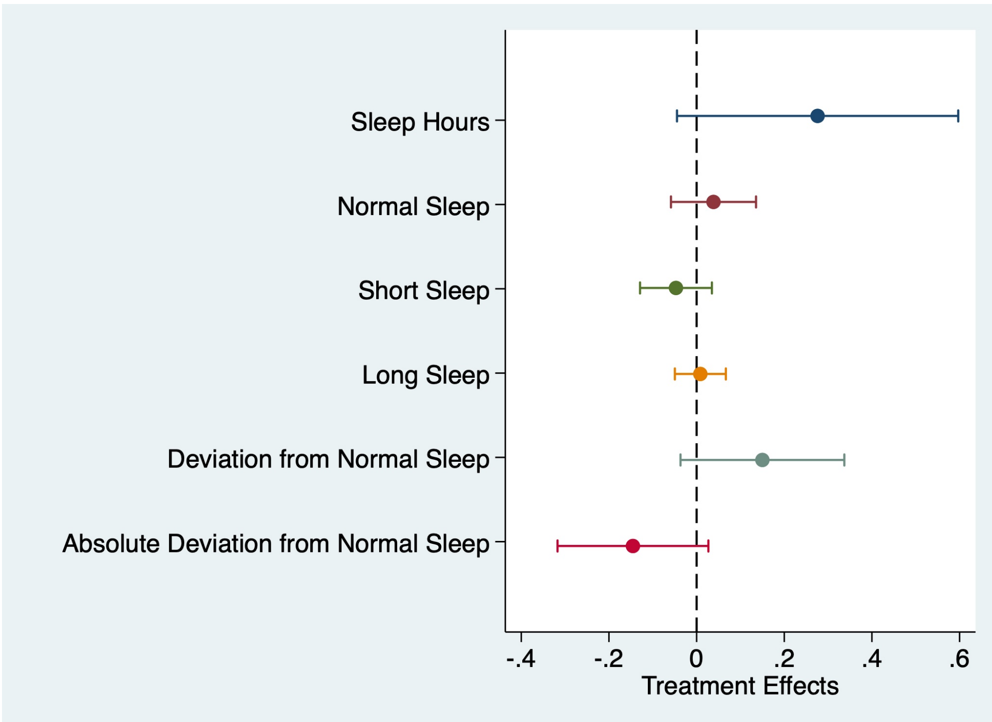


Figure 3.5.3: Average Treatment Effects

Note: Treatment effects estimated by the TWFE Model along with the 95% confidence intervals are plotted. The units for Sleep Hours and Deviation Variables are in hours, while the units for sleep categories (normal, short, and long sleep) are expressed as probabilities.

3.5.2.3 Heterogeneous Treatment Effects

The average treatment effects of the clean heating policy on sleep patterns have proven to be statistically insignificant. However, it is plausible that the policy may have impacted specific subgroups differently. To gain a clearer understanding, further analysis is necessary to identify which subgroups have potentially benefited most from the coal ban in relation to sleep.

3.5.2.3.1 By Type of Sleepers Figure (3.5.4) displays the estimated treatment effects of the clean heating policy on different types of sleepers at baseline, revealing significant heterogeneous effects across sleeper types. The differences in treatment effects suggest that both short and long sleepers have adjusted towards more normative sleep durations post-intervention. Specifically, individuals identified as short sleepers prior to the implementation of the policy are now experiencing an average increase of approximately one hour in sleep duration, whereas those categorized as long sleepers are reducing their sleep by about half an hour. This divergence in effects likely explains the small and insignificant overall average treatment effect shown in Figure (3.5.3). For individuals with normal sleep patterns at baseline, the intervention's effects on sleep hours are minor and statistically insignificant, indicating that the clean heating policy primarily influences those with initially extreme sleep durations.

Further analysis using sleep type and deviation variables indicates that the beneficial impacts of the intervention were limited to a fraction of participants, as evidenced by the smaller magnitude of deviation variables' coefficients compared to the "Sleep Hours." Additionally, the coefficients for "Normal Sleep" for initial short and long sleepers are considerably less than 1, suggesting the fractional change within those groups. For certain participants, however, the impacts of the intervention might not be beneficial. For instance, the coefficient for 'Long Sleep' among initially normal sleepers is significant at approximately 0.1, suggesting an increase of about 10% in the probability that these individuals extend their sleep duration excessively, thereby falling into an unhealthy long sleep range.

These findings highlight the complexity of the intervention's impact on sleep patterns, with benefits for some but potentially adverse effects for others.

3.5.2.3.2 By Gender Figure (3.5.5) illustrates the gender-specific heterogeneous treatment effects of the clean heating policy on sleep duration. The graph shows that, post-policy, male participants experienced an average increase of half an hour in sleep duration. Further combined analysis reveals that this impact was primarily observed in males who were categorized as short or normal sleepers at baseline. In contrast, the intervention appears to have minor ad insignificant

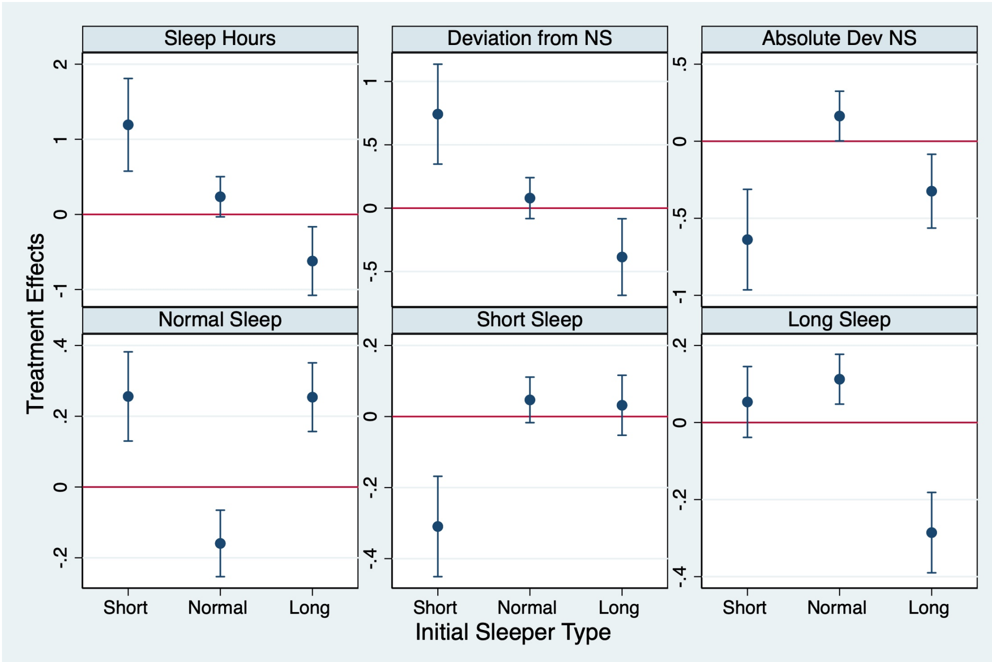


Figure 3.5.4: Heterogeneous Treatment Effects by Initial Sleeper Type

Note: Heterogeneous treatment effects estimated by the TWFE Model are displayed alongside their 95% confidence intervals. Units for Sleep Hours and Deviation Variables are presented in hours, and the units for sleep categories (normal, short, and long sleep) are shown as probabilities.

impact on the sleep duration of female participants. These findings suggest that gender differences may exist in the health benefits derived from such interventions.

3.5.2.3.3 By Age Figure (3.5.6) illustrates the heterogeneous treatment effects of the clean heating policy on sleep across various age groups. Among individuals aged 60 to 70 years, an average increase of half an hour in sleep duration is observed, accompanied by a 0.2 hour change in deviation variables. These results primarily benefit short sleepers, as indicated by a significant coefficient of -0.1 for Short Sleep. For individuals younger than 50, there is approximately a 15% decrease in the probability of being categorized as a long sleeper, with a similar magnitude increase in the likelihood of being a normal sleeper. Additionally, a clear age-related trend is evident: younger individuals tend to experience reduced sleep durations, whereas older individuals generally report increased sleep durations following the intervention. This trend is suggestive that age-specific factors might influence how sleep patterns are affected by policy interventions.

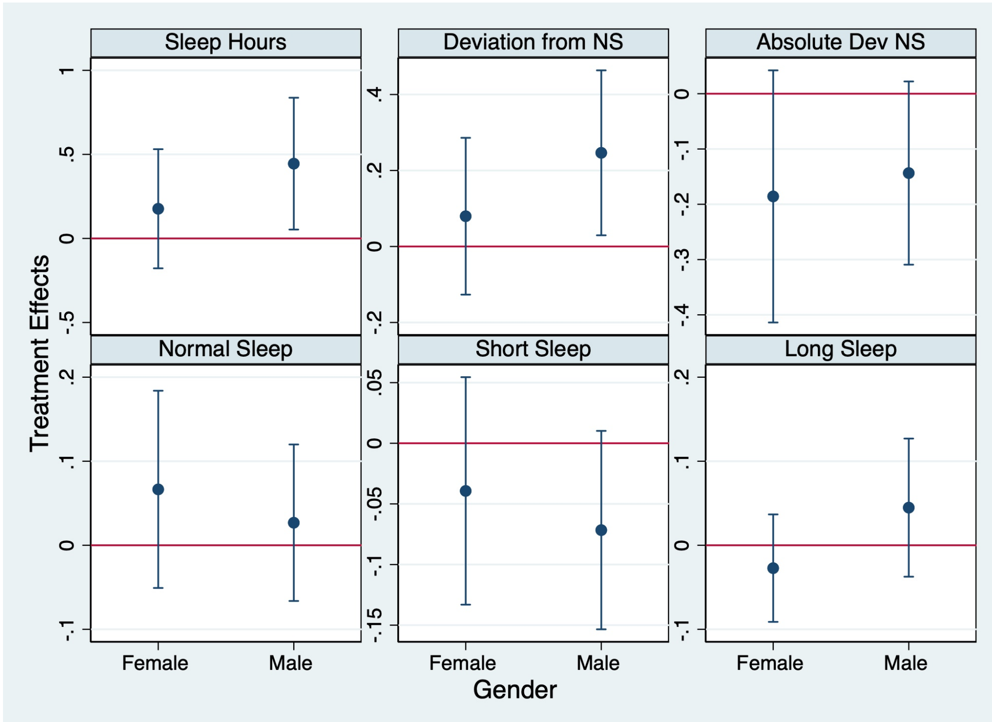


Figure 3.5.5: Heterogeneous Treatment Effects by Gender

Note: Heterogeneous treatment effects estimated by the TWFE Model are displayed alongside their 95% confidence intervals. Units for Sleep Hours and Deviation Variables are presented in hours, and the units for sleep categories (normal, short, and long sleep) are shown as probabilities.

3.5.2.3.4 By Region Figure (3.5.7) displays the region-specific treatment effects of the clean heating policy on sleep outcomes. Notably, the Miyun district shows improvements in sleep, with an increase of more than half an hour in sleep duration, over a 15% increase in the probability of being classified as a normal sleeper, and approximately a 15% decrease in the likelihood of being categorized as a short sleeper. These results suggest that the benefits are primarily due to short sleepers transitioning towards normal sleep durations, a finding further supported by the opposite but similar magnitudes of coefficients of two deviation variables. Conversely, the other two districts do not show significant or major effects. Miyun district had substantially higher reliance on polluting fuels, such as coal and wood, for bedroom heating prior to the implementation of the coal ban. At baseline, 91.6% of households exclusively used these polluting fuels for bedroom heating, while the remainder utilized a combination of clean and polluting fuels. This district not only had the most significant dependence on polluting fuels but also the highest average heating hours, leading

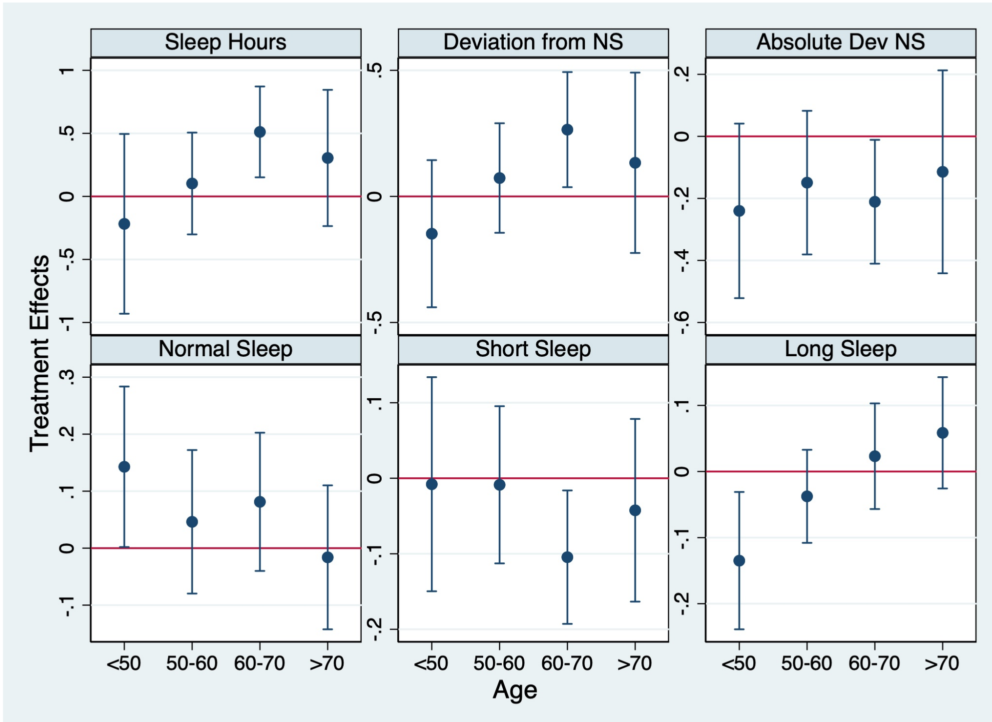


Figure 3.5.6: Heterogeneous Treatment Effects by Age

Note: Heterogeneous treatment effects estimated by the TWFE Model are displayed alongside their 95% confidence intervals. Units for Sleep Hours and Deviation Variables are presented in hours, and the units for sleep categories (normal, short, and long sleep) are shown as probabilities.

to the greatest exposure to pollution from heating sources. The introduction of heat pumps likely significantly reduced the use of polluting fuels, which could be part of the mechanism behind the more pronounced impact on improving sleep outcomes in Miyun. We will further explore this hypothesis in a subsequent section.

3.5.3 Mediation Analysis Exploration

In this section, we explore the potential mechanisms through which the clean heating policy could influence individual sleep patterns, as illustrated in Figure (3.1.4). Understanding these pathways is crucial for assessing the comprehensive impact of the clean heating policy on sleep and can provide valuable insights for shaping future environmental and public health policies.

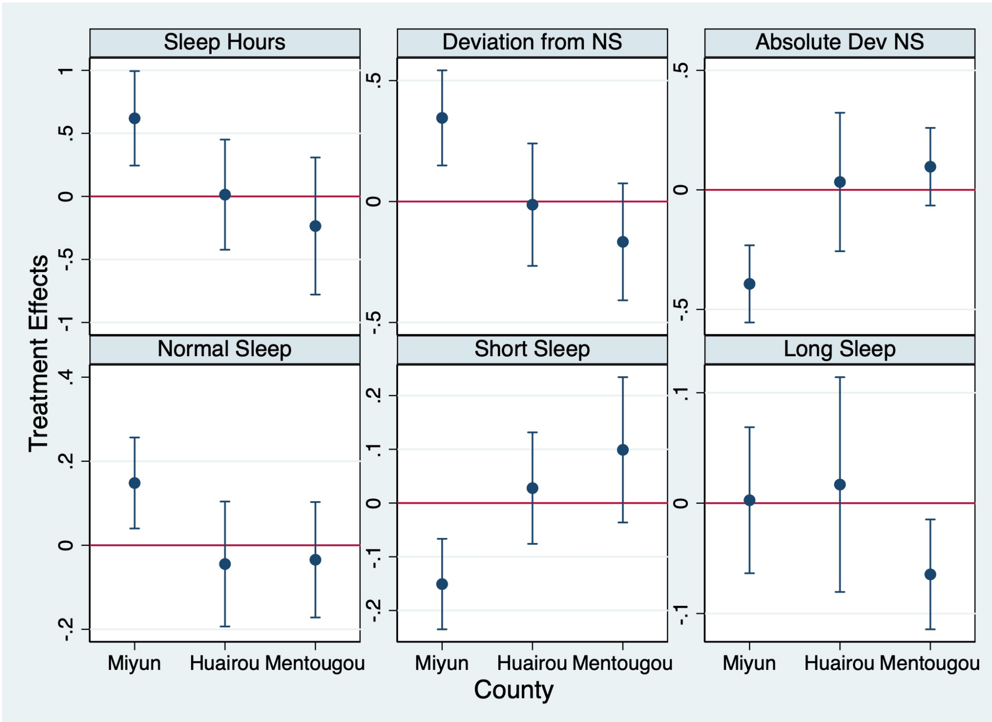


Figure 3.5.7: Heterogeneous Treatment Effects by Region

Note: Heterogeneous treatment effects estimated by the TWFE Model are displayed alongside their 95% confidence intervals. Units for Sleep Hours and Deviation Variables are presented in hours, and the units for sleep categories (normal, short, and long sleep) are shown as probabilities.

3.5.3.1 Treatment Effects of Clean Heating Policy on Bedroom Heating

This subsection initiates the first step of our mechanism analysis by evaluating whether the clean heating policy has influenced household choices regarding heating devices and fuels, as well as heating behaviors. Specifically, we test the pathway from the clean heating policy to changes in heating device usage, heating hours, and heating fuels, as outlined in the upper path of Figure (3.1.4). Once we establish that the clean heating policy has indeed impacted households' choices and behaviors regarding heating, we will then explore how environmental changes induced by the policy, such as alterations in air pollution and indoor temperature, may act as mechanisms for improving sleep.

We assess several outcomes related to bedroom heating fuel choices and behaviors, as presented in Panel B-Set A in Table (3.3.1). Figure (3.5.8) to (3.5.11) display the pre-trends tests for bedroom heating outcomes. Some differences between the control and treatment groups are observed, as evidenced by significant pre-treatment effects for certain outcomes. Notably, these pre-treatment effects are in the opposite direction to the post-treatment effects again, suggesting a potential underestimation of the actual impact.

The detailed results for bedroom heating outcomes are presented in Tables (3.5.3) and (3.5.4). Following the intervention, significant shifts in heating practices were observed: there was a 68% reduction in the probability of using entirely polluting heating fuels in bedrooms, a 45% increase in the probability of households adopting a mixed heating fuel approach, and a 23% increase in the likelihood of choosing exclusively clean heating fuels. Additionally, for primary heating fuels in bedrooms, there was a 68% decrease in the likelihood of households choosing coal and a 5% decrease in the selection of wood. These reductions contributed to a 73% increase in the probability of choosing electricity as the primary heating fuel.

For total heating hours across all bedrooms, the estimated treatment effect is an increase of 42 hours in clean fuel usage coupled with a decrease of 32 hours in polluting fuel usage as shown in Table (3.5.4), resulting in a net gain of 10 additional total heating hours. Regarding average heating hours in bedrooms that are heated daily, there is a decrease of 14 hours in the use of polluting fuels, offset by a nearly equivalent increase in clean fuel usage. This pattern suggests that households are opting to heat more bedrooms rather than extending the heating duration within each room.

Therefore, the clean heating intervention has effectively transitioned households away from polluting fuels, such as coal, towards cleaner alternatives for heating bedrooms. As a result, households are not only utilizing more environmentally friendly heating options but are also benefiting from increased heating hours.

3.5.3.2 Heterogeneous Treatment Effects by Bedroom Heating Fuel Pattern

Having confirmed that the clean heating policy significantly influenced households' choices and behaviors towards cleaner heating solutions, we will next explore how environmental changes resulting from these heating adjustments—such as reductions in air pollution and modifications in indoor temperature—may serve as mechanisms for improving sleep.According to our hypothesis, illustrated in the upper path of Figure (3.1.4), more pronounced treatment effects should be observed

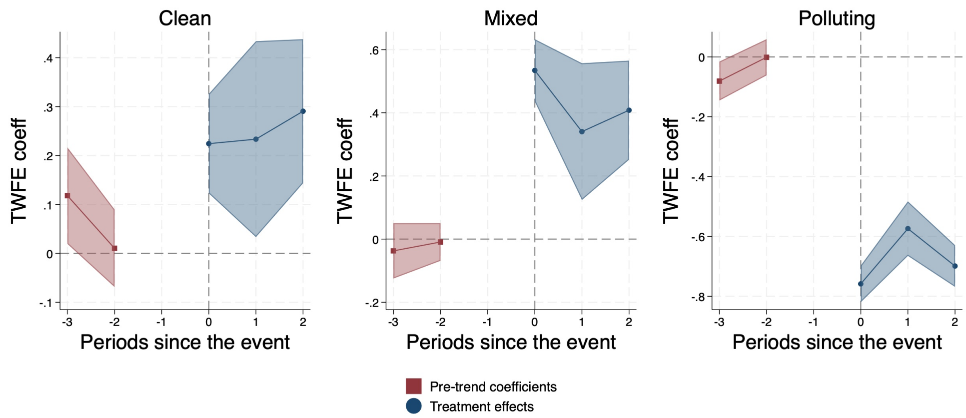


Figure 3.5.8: Pre-trends test for bedroom heating fuel pattern

Note: Dynamic treatment effects estimated by the TWFE Model are displayed alongside their 95% confidence intervals. The -1 period is omitted as the reference period. Units for all variables are shown as probabilities.

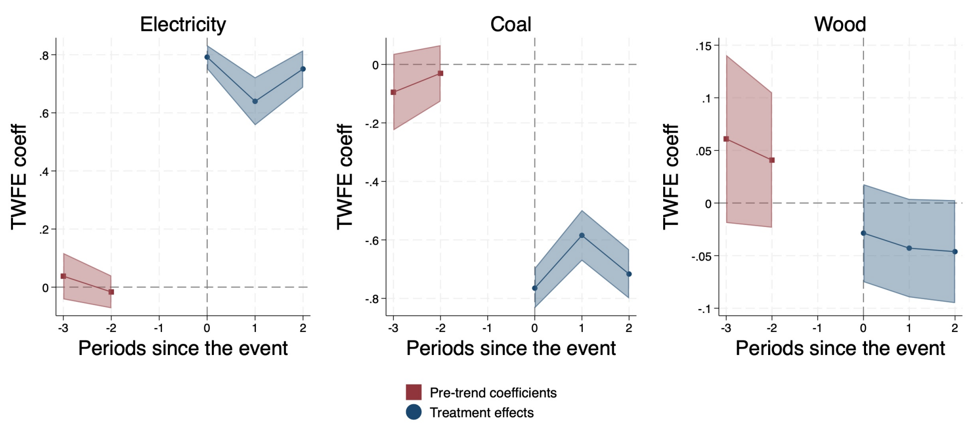


Figure 3.5.9: Pre-trends test for bedroom primary heating fuel

Note: Dynamic treatment effects estimated by the TWFE Model are displayed alongside their 95% confidence intervals. The -1 period is omitted as the reference period. Units for all variables are shown as probabilities.

among households that initially used more polluting heating portfolios in their bedrooms, as these households are likely to experience greater environmental benefits from the policy.

3.5.3.2.1 By Heating Fuel Pattern Figure (3.5.13) depicts the heterogeneous treatment effects on sleep based on individuals' initial bedroom heating portfolios. It is not surprising that the treatment effects for individuals initially using clean heating fuels are close to zero and insignificant, given the minimal environmental improvement they would experience post-intervention. Conversely,

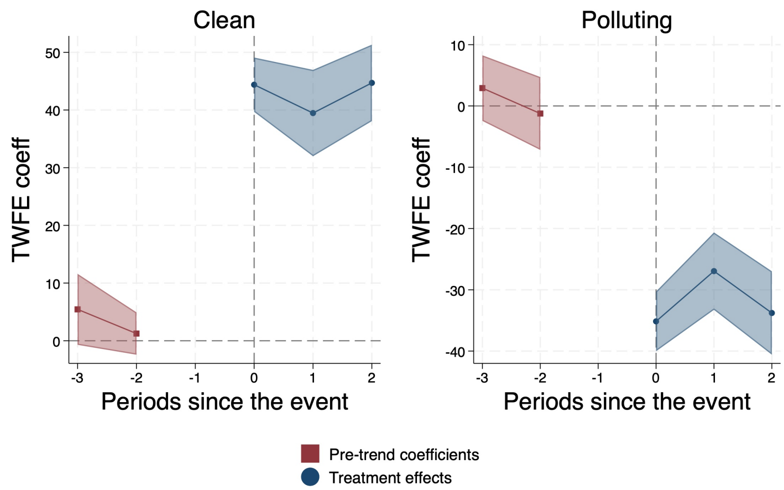


Figure 3.5.10: Pre-trends test for total bedroom heating hours

Note: Dynamic treatment effects estimated by the TWFE Model are displayed alongside their 95% confidence intervals. The -1 period is omitted as the reference period. Units for all variables are in hours.

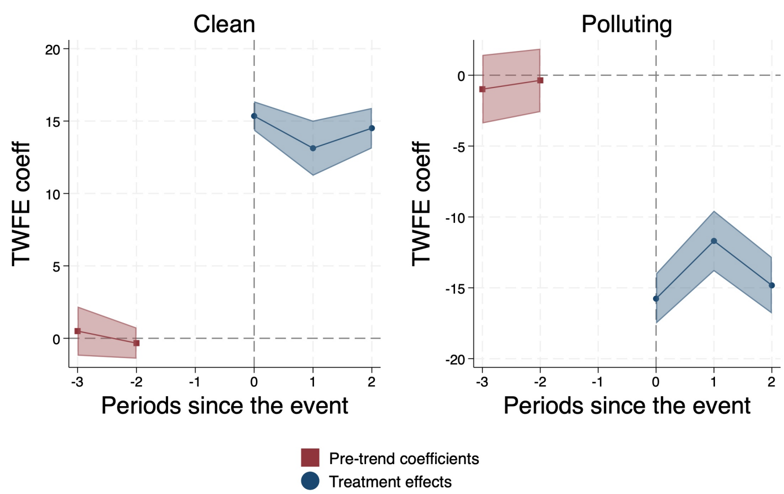


Figure 3.5.11: Pre-trends test for mean bedroom heating hours

Note: Dynamic treatment effects estimated by the TWFE Model are displayed alongside their 95% confidence intervals. The -1 period is omitted as the reference period. Units for all variables are in hours.

we observe an average half-hour increase in sleep duration among individuals initially employing mixed heating fuels, attributed to both normal and short sleepers excessively extending their sleep beyond the normal range.Surprisingly, those initially using polluting heating fuels display only minor and insignificant treatment effects, which deviates from our expectations. However, this lack

	Panel A: Bedroom Heating Fuel Pattern		
	(1)	(2)	(3)
	Clean	Mixed	Polluting
Treatment	0.227 ***	0.451 ***	-0.678 ***
	(0.052)	(0.056)	(0.032)
	Panel B: Bedroom Primary Heating Fuel		
	(1)	(2)	(3)
	Electricity	Coal	Wood
Treatment	0.733 ***	-0.684 ***	-0.047 **
	(0.027)	(0.035)	(0.019)

Table 3.5.3: Treatment Effects on Binary Heating Fuel Pattern

Note: [a] Treatment effects estimated by the TWFE Model are displayed with their standard errors in brackets. The units for all variables are expressed as probabilities. [b] *** p<.01, ** p<.05, * p<.1

	Panel A:Total Heating Hours by All Bedrooms	
	(1)	(2)
	Clean	Polluting
Treatment	42.122 ***	-31.878 ***
	(2.296)	(2.349)
	Panel B:Mean Heating Hours by Bedrooms with Daily Heating	
	(1)	(2)
	Clean	Polluting
Treatment	14.480 ***	-14.057 ***
	(0.557)	(0.806)

Table 3.5.4: Treatment Effects on Bedroom Heating Hours by Sources

Note: [a] Treatment effects estimated by the TWFE Model are displayed with their standard errors in brackets. The units for all variables are expressed as probabilities. [b] *** p<.01, ** p<.05, * p<.1

of significant findings does not negate the possibility of differential impacts; it may instead indicate that additional statistical power or precision might be required to detect these nuanced effects. The variation in responses is suggestive that the impacts of the clean heating policy may differ based on individuals' pre-existing heating choices.

3.5.3.2.2 By Bedroom Primary Heating Fuel Figure (3.5.13) displays the heterogeneous treatment effects on sleep based on individuals' initial primary bedroom heating fuel. Interestingly, the increase in sleep duration is more pronounced for individuals with clean primary heating fuels at baseline compared to those with polluting fuels. However, this does not necessarily imply a

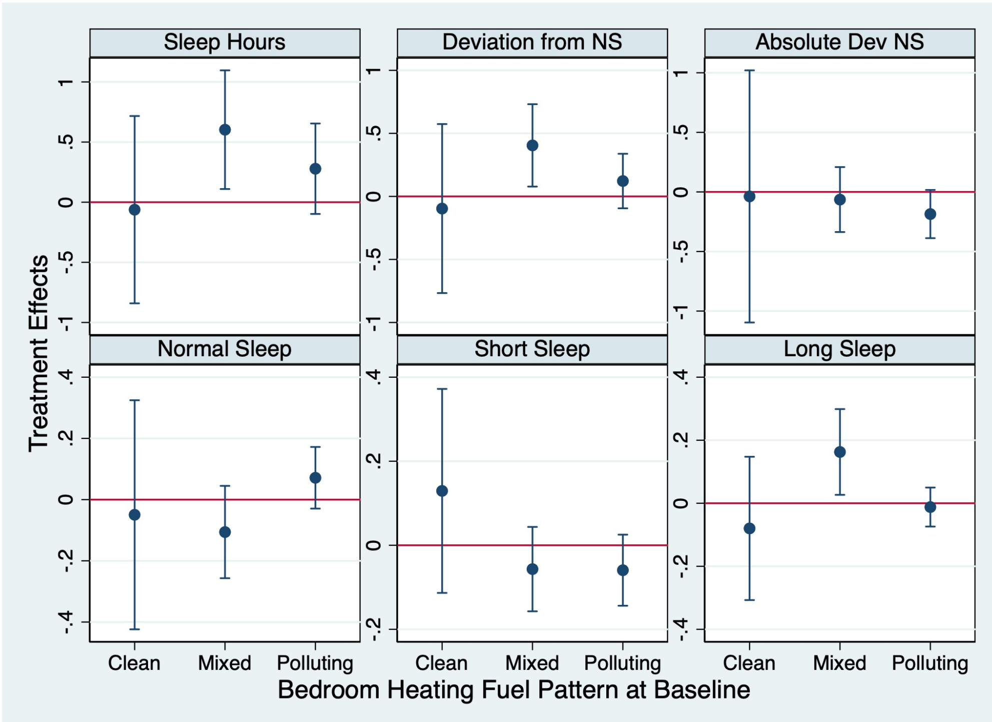


Figure 3.5.12: Heterogeneous Treatment Effects by Baseline Heating Fuel Pattern

Note: Heterogeneous treatment effects estimated by the TWFE Model are displayed alongside their 95% confidence intervals. Units for Sleep Hours and Deviation Variables are presented in hours, and the units for sleep categories (normal, short, and long sleep) are shown as probabilities.

greater benefit from the intervention. Further analysis using a combined approach reveals that for the group with clean heating fuels at baseline, the increase in sleep duration is predominantly among normal sleepers extending into the long sleep range—a shift that is not considered beneficial to health. In contrast, the changes observed in the group with polluting heating fuels are more subtle; yet, the direction of these changes suggests benefits primarily accrue to short sleepers transitioning towards normal sleep durations. While these results do not achieve statistical significance, the trends indicate a potential improvement in sleep quality for those initially using more polluting heating fuels. These individuals tend to experience fewer instances of short sleep, an increase in the normal sleep category, and a reduced deviation from the normal sleep range, suggesting that the impact of clean heating policies may be more beneficial as the initial heating portfolio is more polluting.

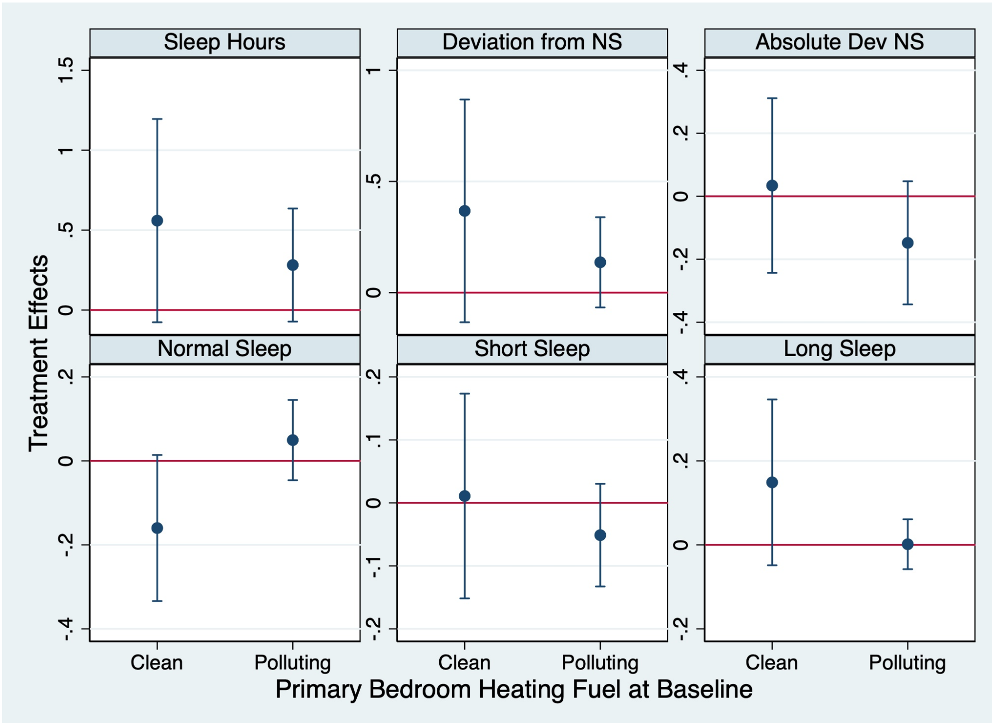


Figure 3.5.13: Heterogeneous Treatment Effects by Baseline Primary Bedroom Heating Fuel

Note: Heterogeneous treatment effects estimated by the TWFE Model are displayed alongside their 95% confidence intervals. Units for Sleep Hours and Deviation Variables are presented in hours, and the units for sleep categories (normal, short, and long sleep) are shown as probabilities.

3.5.3.3 Mediation Analysis Exploration

In this section, we explore potential mediators that could explain the average impact of the coal ban on sleep outcomes. While data on these mediators were not available for all respondents, this analysis serves as a preliminary exploration of the potential channels through which the intervention may influence sleep. Possible mediators include the average indoor air pollution level during the heating season, personal exposure to PM2.5, average indoor temperature during typical sleep hours, and the variability of indoor temperature during these hours, as indicated by the standard deviation. Details on these mediators can be found in Panel B-Set B of Table (3.3.1).

The mediation analysis for examining the potential channels through which the coal ban affects sleep can be structured using the following equations:

$$\begin{tikzcd} D \arrow[r, "\alpha_1"] & Y \end{tikzcd} \qquad \begin{tikzcd} D \arrow[rr, "\beta_1"] \arrow[dr, "\gamma_1"'] & & Y \\ & M \arrow[ur, "\beta_2"'] & \end{tikzcd}$$

Figure 3.5.14: Diagram of Mediation Effect

$$Y = \alpha_0 + \alpha_1 D + \varepsilon_{Y_1} \tag{3.1}$$

$$Y = \beta_0 + \beta_1 D + \beta_2 M + \varepsilon_{Y_2} \tag{3.2}$$

$$M = \gamma_0 + \gamma_1 D + \varepsilon_M \tag{3.3}$$

where D represents the treatment (clean heating policy), M is the mediator (e.g., indoor air pollution, personal PM2.5 exposure, or indoor temperature), and Y is the outcome variable (sleep). Equation (2.1) models the direct effect of the intervention on sleep without considering the mediator, while Equation (2.2) models the outcome considering both the direct effect of the ban and its effect through the mediator. Equation (2.3) represents how the treatment affects the mediator.

The diagram illustrating this mediation effect is shown in Figure (3.5.14). The direct effect of the clean heating policy on sleep is represented by α_1, and the indirect effect through the mediator is calculated as $\beta_2\gamma_1$. The relationship between these effects is captured by the equation: $\alpha_1 = \beta_1 + \beta_2\gamma_1$, which helps in understanding the proportion of the treatment effect that is mediated. This analytical framework allows for the decomposition of the overall effect of the intervention on sleep into components that are directly due to the intervention and those that operate through changes in environmental and personal conditions.

The analysis presented in Figure (3.5.15) did not identify any significant mediators, with indirect effects approaching zero and contributing to less than 10% of the total effect. This lack of significant mediation may be attributed to excessive noise in the data and the minimal initial gains in sleep, which are too small to detect mediation factors reliably. Consequently, we extended the mediation analysis to subgroups that demonstrated larger and more significant direct effects. However, this pattern persists even among these subgroups. For instance, the mediation effects for specific subgroups, such as initial short sleepers and long sleepers, detailed in the appendix Figures

(3.8.1) and (3.8.2), remain insignificant despite significant direct effects. This outcome suggests the possibility of other underlying reasons, such as the presence of another unobserved factor impacting sleep that was not included in our analysis. We will further explore these possibilities in the discussion section, following the robustness checks.

3.6 Robustness Check

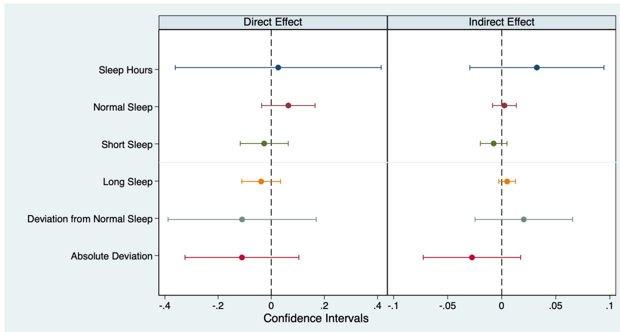
3.6.1 DID Estimators for Staggered Treatment

In this paper, we estimated the treatment effects of clean heating policy on sleep across various groups using the Two-Way Fixed Effects (TWFE) Model. However, recent research has highlighted certain limitations of this model, especially concerning the analysis of staggered treatment implementations (Goodman-Bacon, 2021; Callaway and Sant'Anna, 2021; Gardner, 2022; Borusyak et al., 2024; Sun and Abraham, 2021; De Chaisemartin and d'Haultfoeuille, 2020). This model generally presumes that the treatment effects are consistent across all units, suggesting that each unit is affected equally by the treatment regardless of when it starts or how long it has been in place. However, this assumption can lead to inaccuracies, especially since units that received treatment earlier are used as controls for those treated later, potentially skewing the results if the effects of the treatment are not actually uniform.

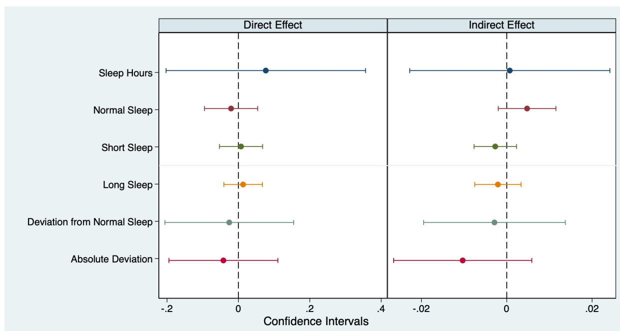
In response to these challenges, a variety of enhanced Difference-in-Differences (DID) estimators have been developed. These estimators discard the assumption of uniform treatment effects, allowing for variations in how the treatment impacts different units and at different times. This approach refrains from using units that have already received treatment as controls, facilitating the estimation of treatment effects that are specific to each unit and time period. These effects can then be compiled to reflect an overall effect or detailed by treatment cohort, specific times, or the length of time since the treatment began.

Among these innovative methods, the approach suggested by Callaway and Sant'Anna, 2021 stands out. We employed this methodology as a robustness check in our study. It offers a way to

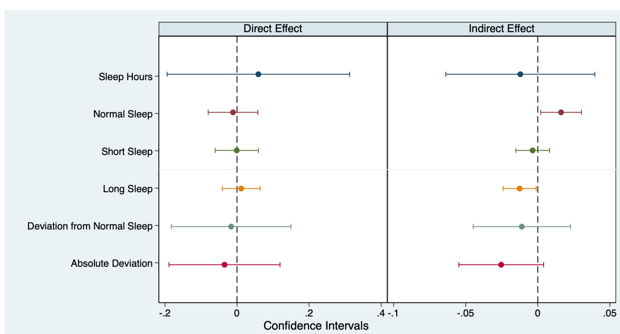
Mediator: Indoor air pollution during heating season



Mediator: personal exposure to PM2.5



Mediator: average indoor temperature during sleep time(22:00-8:00)



Mediator: standard deviation of indoor temperature during sleep time(22:00-8:00)

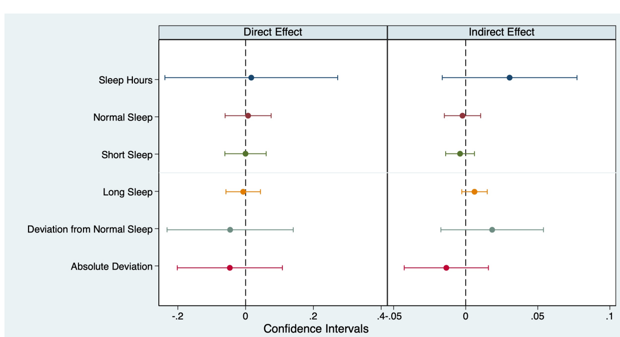


Figure 3.5.15: Mediation Analysis for ATT

Note: Direct effects $= \alpha_1$ estimated by Eq (3.1) and Indirect effects $= \beta_2\gamma_1$ estimated by Eq (3.2) and (3.3) using full sample.

mitigate the potential biases associated with traditional TWFE DID estimators and more accurately reflect the diverse effects of the treatment across various settings and times.

Figure (3.6.1) displays the results obtained using the Callaway and Sant'Anna, 2021 method alongside those derived from traditional Two-Way Fixed Effects (TWFE) estimators. The similarity between these two sets of results underscores a robust consistency in the measured treatment effects. This consistency extends to analyses of heterogeneous treatment effects; however, for brevity, we only present the comparison for average treatment effects here.

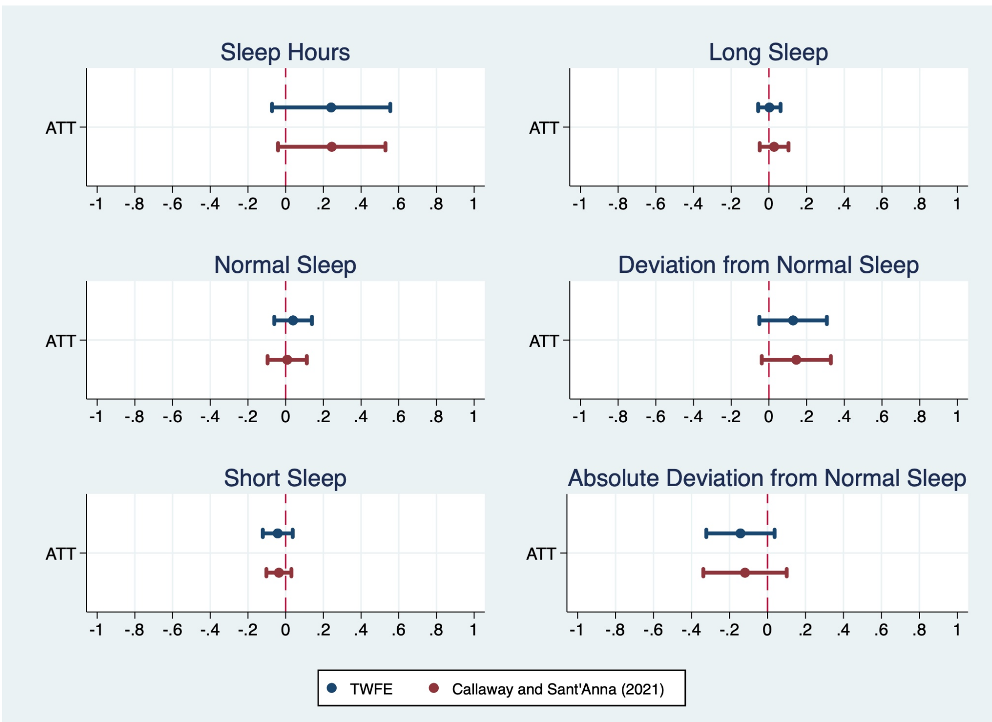


Figure 3.6.1: Comparison Between TWFE and Callway Sant'Anna Estimators

3.6.2 Definition for Short and Long Sleep

We initially defined Short Sleep as less than 7 hours per day and Long Sleep as more than 9 hours per day. To conduct a sensitivity analysis, we adjusted the criteria for Short Sleep to less than 6 hours and for Long Sleep to more than 10 hours, and then re-estimated the average treatment effects. Figure (3.6.2) compares the results using these two sets of cutoffs. It shows that when we broaden the range of what is considered normal sleep duration, the estimated treatment effects are slightly reduced in magnitude and feature narrower confidence intervals. Importantly, the direction of all coefficients remains consistent. This consistency is expected because the deviation from healthy sleep duration is now smaller due to the expanded range of normal sleep. The analysis suggests that the clean heating policy's effects extend beyond those with extreme sleep durations to include individuals closer to the recommended sleep range. Further, we applied these definitions to analyze heterogeneous treatment effects by initial sleep type. Figure (3.6.3) displays these results, with the original cutoffs 7/9 in Panel A and the new cutoffs 6/10 in Panel B. The direction of effects remains consistent across all metrics, yet the magnitude of effects is larger with the new cutoffs for initial short and long sleepers, indicating that the policy's benefits are not uniformly distributed across all individuals; rather, larger benefits are observed in those with more extreme initial sleep hours.

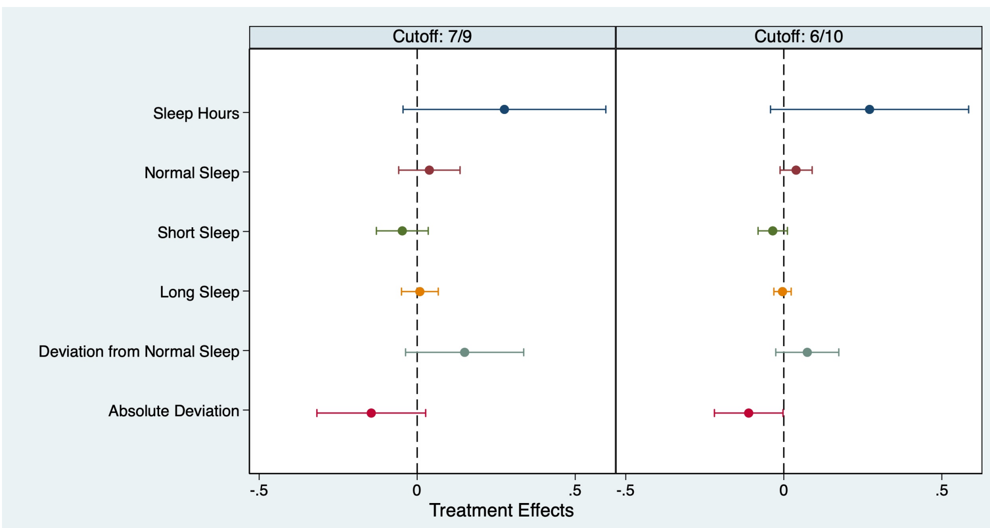


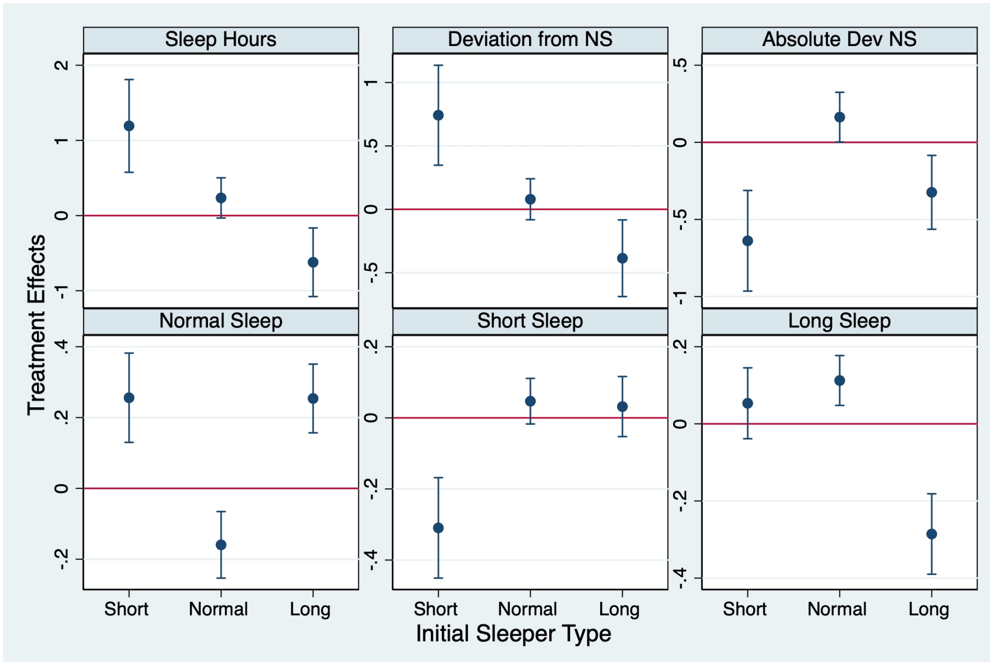
Figure 3.6.2: ATT Comparison Between Different Definition of Healthy Sleep

Note: Average treatment effects estimated by the TWFE Model are displayed alongside their 95% confidence intervals. Units for Sleep Hours and Deviation Variables are presented in hours, and the units for sleep categories (normal, short, and long sleep) are shown as probabilities.

3.7 Discussion and Conclusion

The Beijing clean heating policy, which includes a coal ban and financial subsidies for the purchase and installation of heat pumps as well as subsequent electricity bills, aims to provide residents with

Panel A: Cutoff: 7/9



Panel B: Cutoff: 6/10

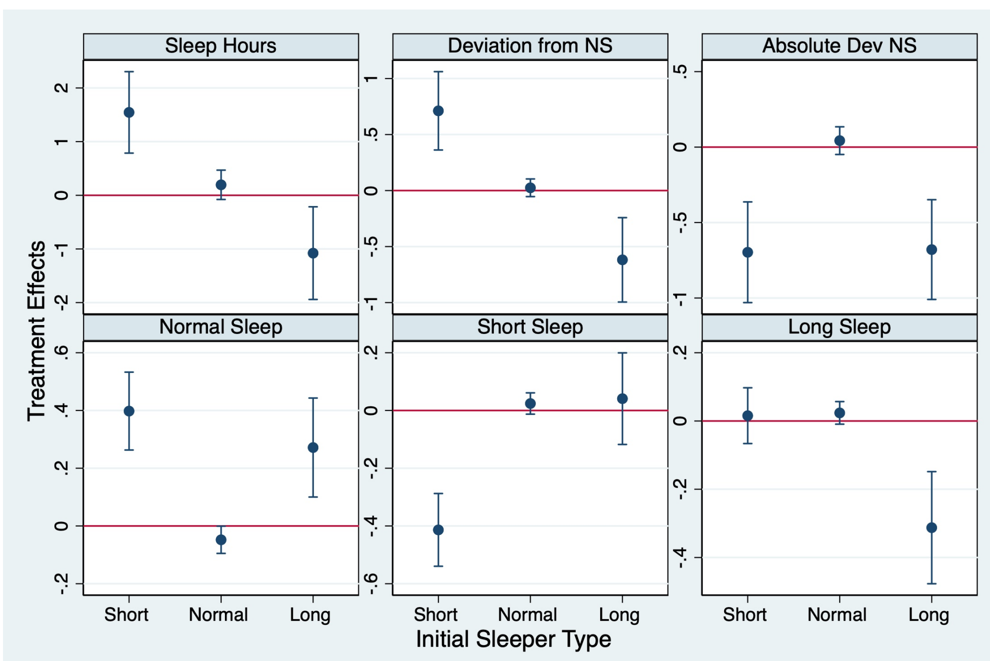


Figure 3.6.3: HE Comparison Between Different Definition of Healthy Sleep

Note:Heterogeneous treatment effects estimated by the TWFE Model are displayed alongside their 95% confidence intervals. Units for Sleep Hours and Deviation Variables are presented in hours, and the units for sleep categories (normal, short, and long sleep) are shown as probabilities.

a cleaner and warmer winter. According to the literature, these benefits are expected to improve individuals' sleep. In this paper, we estimate the treatment effects of the clean heating policy on various sleep outcomes, including sleep hours, sleep types categorized as short, normal, and long sleep, and deviations from normal sleep. We found no statistically significant average treatment effects. However, when applying a combined analysis in Figure (3.3.2) to various outcomes, we identified that certain subgroups, particularly initial short sleepers, benefited from the policy. Our heterogeneous treatment analysis further substantiates these findings, revealing that the demographic groups benefiting most from the policy are: initial short sleepers, who experienced an approximate increase in sleep duration of 1 hour; long sleepers, with an decrease of about 0.5 hours; men, who gained around 0.5 hours; individuals aged 60-70, with increases exceeding 0.5 hours; and those in regions undergoing significant shifts from polluting to clean heating fuels, also with increases exceeding 0.5 hour.

We also conducted a mediation analysis to explore potential mechanisms through which the intervention could influence individual sleep. We confirmed that the clean heating policy has successfully transitioned households from polluting heating fuels to cleaner heating options, which should improve indoor air quality, provide stable room temperatures, and minimize sleep disruptions caused by the need to add fuel. However, these environmental changes, such as improvements in indoor air pollution, room temperature stability, and reduced personal exposure to pollutants, could not account for even 10% of the changes in sleep patterns . This leaves uncertainty about the exact mechanisms through which the clean heating policy affects sleep, suggesting further research is needed to explore other potential mediators.

Several factors may have contributed to the insignificant mediation effects observed in our study. First and foremost, there were limitations related to the data on potential mediators. The temperature and air pollution data were not collected for each participant, limiting our analysis to a relatively small sample size. This may have restricted our ability to detect any existing mediation effects, even if they were present, due to insufficient statistical power.

Furthermore, the impact of the intervention on some mediators might have been minimal, resulting in inadequate variation among the mediators to effectively assess their role in transmitting the policy's effects. For instance, our preliminary estimates suggest that the personal exposure to PM2.5 among sampled individuals showed no significant change, with an estimated effect of -1 μg/m³ against a pre-intervention average concentration of 105 μg/m³ (Brehmer and Li, 2023). This indicates that while the clean heating intervention may have altered heating behaviors, the resulting environmental changes were not substantial enough to significantly influence sleep outcomes through the pathway of personal exposure to PM2.5. Future research should investigate why the shift to clean bedroom heating is not associated with a meaningful reduction in PM2.5 exposure as hypothesized.

Additionally, the mediators we considered may be counteracted by other factors we were not able to consider. Although the intervention decreased indoor PM2.5 concentrations by approximately 30.9 μg/m³ (Brehmer and Li, 2023; Baumgartner et al., 2025)—more than one-third of the average indoor concentration before the intervention (86.8 μg/m³)—and increased room temperatures by roughly 2°C from an average of 14°C (Baumgartner et al., 2025), with a decrease in temperature variability by 0.27 (approximately one-third of the mean standard errors), these notable improvements might still be countered by other factors. For instance, the study by Okamoto et al. (2012) suggests that the impact of room temperature on sleep can vary depending on the use of bedding and clothing, which are crucial for thermoregulation and maintaining sleep during cold exposure. Similarly, severe outdoor air pollution could mitigate the benefits gained from indoor air quality improvements and losses due to personal preferences for traditional heating methods might outweigh the gains from more stable and increased room temperatures.

Lastly, there might be additional pathways, such as minimized sleep disruptions caused by the need to add fuel, which we were unable to measure effectively. These considerations suggest the complexity of assessing mediation in the interventions and suggest the need for further research to explore other potential pathways and contextualize our findings.

We acknowledge several limitations that merit consideration for future research. First, the measurement of sleep employed in our study—primarily focusing on sleep duration—captures only one dimension of sleep quality. Sleep quality is a multifaceted construct that includes other important aspects such as sleep latency, sleep depth, number of awakenings, and sleep efficiency etc(Morin et al., 2009; Ohayon et al., 1997; Ohayon, 2002; Yi et al., 2006). Relying solely on sleep duration may not fully capture the potential benefits of the clean heating policy on overall sleep quality. Moreover, the use of self-reported sleep duration may introduce inaccuracies due to recall or rounding bias. However, the rounding errors are expected to be random, suggesting that any resulting biases in the estimated effects on sleep would likely be attenuated toward zero. Enhancements to our methodology could include more comprehensive sleep assessments using validated questionnaire tools like the Pittsburgh Sleep Quality Index (PSQI) (Buysse, 2003) and the Epworth Sleepiness Scale(Scale, 1997), or objective measures such as Polysomnography (PSG) and actigraphy, which would provide a more robust evaluation of sleep quality.

Second, while our combined analysis provides insights about which types of sleepers benefit most from the intervention, the approach of testing multiple hypotheses independently increases the risk of committing Type I errors—falsely rejecting true null hypotheses. This issue of multiple hypothesis testing requires careful management, possibly through statistical adjustments such as Bonferroni correction or controlling the false discovery rate, to ensure the reliability of our conclusions.

Third, the mediation analysis conducted in this study is preliminary and serves primarily as an exploratory tool. The potential pathways through which the clean heating policy affects sleep were not exhaustively considered, and the analysis did not adequately control for potential confounders that could influence both the implementation of the clean heating policy and sleep outcomes. For example, geographic location could correlate with policy implementation due to local regulations and also influence sleep patterns through regional climate variations. Future work should aim to identify and adjust for these confounders to better isolate the effects of the clean heating policy.

Further research should aim to address these limitations by incorporating more detailed sleep assessments, applying rigorous methods to control for multiple testing, and enhancing the robustness of mediation analysis by considering a wider range of mediators and controlling for potential confounding variables. This will strengthen the validity of findings and provide clearer insights into the impacts of environmental policies on sleep health.

Our findings suggest that promoting clean heating solutions is advantageous not only for their environmental benefits but also for their potential health benefits, including improved sleep. In conducting cost-benefit analyses to justify the upfront costs of clean heating technologies, it's important to consider the long-term benefits derived from improved population health. These analyses should explicitly account for enhanced sleep quality and other health benefits associated with better sleep, such as reduced risk of chronic diseases and improved mental health. Such a comprehensive evaluation could provide important evidence regarding investments in clean heating technologies.

Our research identifies specific subgroups that derive the most benefit from clean heating policies, such as the elderly and those with pre-existing sleeping problems. Future health-related policies could be specifically tailored to prioritize these groups for subsidies or support programs. By focusing on these vulnerable populations, policies can maximize health outcomes and ensure that the benefits of interventions are equitably distributed. This targeted approach not only enhances the effectiveness of public health interventions but also ensures that the limited resources are utilized where they can have the greatest impact. Tailoring policies in this manner is crucial for addressing the specific needs of different demographic groups and for achieving significant improvements in public health at the community level.

References

Abou-Khadra, M. K. (2013). Association between PM 10 exposure and sleep of Egyptian school children. Sleep and Breathing, 17, 653–657.

Baglioni, C., Spiegelhalder, K., Nissen, C., & Riemann, D. (2011). Clinical implications of the causal relationship between insomnia and depression: How individually tailored treatment of sleeping difficulties could prevent the onset of depression. Epma Journal, 2, 287–293.

Baron, K. G., Duffecy, J., Reutrakul, S., Levenson, J. C., McFarland, M. M., Lee, S., & Qeadan, F. (2021). Behavioral interventions to extend sleep duration: A systematic review and meta-analysis. Sleep medicine reviews, 60, 101532.

Baumgartner, J., Harper, S., Barrington-Leigh, C., Brehmer, C., Carter, E. M., Li, X., Robinson, B. E., Shen, G., Sternbach, T. J., Tao, S., Xue, K., Yuan, W., Zhang, X., & Zhang, Y. (2025, January 27). How do household energy transitions work? (Research Report). Health Effects Institute. https://github.com/sbh4th/bhet-report/blob/main/hei-report-final-submission.pdf.

Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event-study designs: Robust and efficient estimation. Review of Economic Studies, rdae007.

Brehmer, C., & Li, X. (2023). Multi-year, field-based assessment on the impacts of a Coal-to-Clean Energy policy on indoor, outdoor, and personal exposure to PM2.5 in Beijing, China (Working Paper).

Buscemi, N., Vandermeer, B., Pandya, R., Hooton, N., Tjosvold, L., Hartling, L., Baker, G., Vohra, S., & Klassen, T. (2004). Melatonin for treatment of sleep disorders: Summary. AHRQ evidence report summaries.

Buysse, D. (2003). The pittsburgh sleep quality index: A new instrument for psychiatric practice and research. Psychiatry, 18, 498.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. Journal of econometrics, 225(2), 200–230.

Cappuccio, F. P., D'Elia, L., Strazzullo, P., & Miller, M. A. (2010). Sleep duration and all-cause mortality: A systematic review and meta-analysis of prospective studies. Sleep, 33(5), 585–592.

Chen, G., Xiang, H., Mao, Z., Huo, W., Guo, Y., Wang, C., & Li, S. (2019). Is long-term exposure to air pollution associated with poor sleep quality in rural China? Environment international, 133, 105205.

De Chaisemartin, C., & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. American economic review, 110(9), 2964–2996.

Devore, E. E., Grodstein, F., & Schernhammer, E. S. (2016). Sleep duration in relation to cognitive function among older adults: A systematic review of observational studies. Neuroepidemiology, 46(1), 57–78.

Dong, L., Xie, Y., & Zou, X. (2022). Association between sleep duration and depression in US adults: a cross-sectional study. Journal of Affective Disorders, 296, 183–188.

Gardner, J. (2022). Two-stage differences in differences. arXiv preprint arXiv:2207.05943.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. Journal of econometrics, 225(2), 254–277.

Guo, Q. (2002). The Chinese domestic architectural heating system [Kang]: origins, applications and techniques. Architectural History, 45, 32–48.

Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. Journal of economic literature, 47(1), 5–86.

Irish, L. A., Kline, C. E., Gunn, H. E., Buysse, D. J., & Hall, M. H. (2015). The role of sleep hygiene in promoting public health: A review of empirical evidence. Sleep medicine reviews, 22, 23–36.

Itani, O., Jike, M., Watanabe, N., & Kaneita, Y. (2017). Short sleep duration and health outcomes: A systematic review, meta-analysis, and meta-regression. Sleep medicine, 32, 246–256.

Jike, M., Itani, O., Watanabe, N., Buysse, D. J., & Kaneita, Y. (2018). Long sleep duration and health outcomes: A systematic review, meta-analysis and meta-regression. Sleep medicine reviews, 39, 25–36.

Lan, L., Tsuzuki, K., Liu, Y., & Lian, Z. (2017). Thermal environment and sleep quality: A review. Energy and Buildings, 149, 101–113.

Lawrence, W. R., Yang, M., Zhang, C., Liu, R.-Q., Lin, S., Wang, S.-Q., Liu, Y., Ma, H., Chen, D.-H., Zeng, X.-W., et al. (2018). Association between long-term exposure to air pollution and sleep disorder in Chinese children: the Seven Northeastern Cities study. Sleep, 41(9), zsy122.

Lechner, M., et al. (2011). The estimation of causal effects by difference-in-difference methods. Foundations and Trends® in Econometrics, 4(3), 165–224.

Lim, J., & Dinges, D. F. (2010). A meta-analysis of the impact of short-term sleep deprivation on cognitive variables. Psychological bulletin, 136(3), 375.

Liu, J., Mauzerall, D. L., Chen, Q., Zhang, Q., Song, Y., Peng, W., Klimont, Z., Qiu, X., Zhang, S., Hu, M., et al. (2016). Air pollutant emissions from Chinese households: A major and under-appreciated ambient pollution source. Proceedings of the National Academy of Sciences, 113(28), 7756–7761.

Ma, Y., Liang, L., Zheng, F., Shi, L., Zhong, B., & Xie, W. (2020). Association between sleep duration and cognitive decline. JAMA network open, 3(9), e2013573–e2013573.

Medic, G., Wille, M., & Hemels, M. E. (2017). Short-and long-term health consequences of sleep disruption. Nature and science of sleep, 151–161.

MEP. (2013). 2013 China Environmental Situation Bulletin [last access: 27 May 2014]. https://www.mee.gov.cn/hjzl/sthjzk/zghjzkgb/201605/P020160526564151497131.pdf.

Morin, C. M., Bélanger, L., LeBlanc, M., Ivers, H., Savard, J., Espie, C. A., Mérette, C., Baillargeon, L., & Grégoire, J.-P. (2009). The natural history of insomnia: A population-based 3-year longitudinal study. Archives of internal medicine, 169(5), 447–453.

National Energy Administration. (2022). Press Conference on ”Accelerating the construction of a strong energy nation and ensuring energy security” held by the State Council Information Office of the People’s Republic of China. http://www.nea.gov.cn/2022-07/27/c_1310646889.html.

National Sleep Foundation. (n.d.). Best temperature for sleep. https://www.sleepfoundation.org/bedroom-environment/best-temperature-for-sleep.

Ohayon, M. M. (2002). Epidemiology of insomnia: What we know and what we still need to learn. Sleep medicine reviews, 6(2), 97–111.

Ohayon, M. M., Caulet, M., Philip, P., Guilleminault, C., & Priest, R. G. (1997). How sleep and mental disorders are related to complaints of daytime sleepiness. Archives of internal medicine, 157(22), 2645–2652.

Okamoto-Mizuno, K., & Mizuno, K. (2012). Effects of thermal environment on sleep and circadian rhythm. Journal of physiological anthropology, 31, 1–9.

Ranson, R. P., Organization, W. H., et al. (1988). Guidelines for healthy housing.

Scale, S. (1997). Epworth sleepiness scale. About the ESS.

Shan, Z., Ma, H., Xie, M., Yan, P., Guo, Y., Bao, W., Rong, Y., Jackson, C. L., Hu, F. B., & Liu, L. (2015). Sleep duration and risk of type 2 diabetes: A meta-analysis of prospective studies. Diabetes care, 38(3), 529–537.

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. Journal of econometrics, 225(2), 175–199.

Uehli, K., Mehta, A. J., Miedinger, D., Hug, K., Schindler, C., Holsboer-Trachsler, E., Leuppi, J. D., & Künzli, N. (2014). Sleep problems and work injuries: A systematic review and meta-analysis. Sleep medicine reviews, 18(1), 61–73.

Wang, D., Li, W., Cui, X., Meng, Y., Zhou, M., Xiao, L., Ma, J., Yi, G., & Chen, W. (2016). Sleep duration and risk of coronary heart disease: A systematic review and meta-analysis of prospective cohort studies. International Journal of Cardiology, 219, 231–239.

Yi, H., Shin, K., & Shin, C. (2006). Development of the sleep quality scale. Journal of sleep research, 15(3), 309–316.

Yin, J., Jin, X., Shan, Z., Li, S., Huang, H., Li, P., Peng, X., Peng, Z., Yu, K., Bao, W., et al. (2017). Relationship of sleep duration with all-cause mortality and cardiovascular events: A systematic review and dose-response meta-analysis of prospective cohort studies. Journal of the American Heart Association, 6(9), e005947.

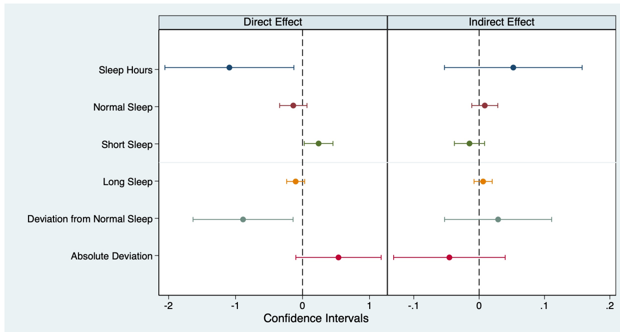
Yu, H., Chen, P., Paige Gordon, S., Yu, M., & Wang, Y. (2019). The association between air pollution and sleep duration: a cohort study of freshmen at a university in Beijing, China. International journal of environmental research and public health, 16(18), 3362.

Zanobetti, A., Redline, S., Schwartz, J., Rosen, D., Patel, S., O'Connor, G. T., Lebowitz, M., Coull, B. A., & Gold, D. R. (2010). Associations of PM10 with sleep and sleep-disordered breathing in adults from seven US urban areas. American journal of respiratory and critical care medicine, 182(6), 819–825.

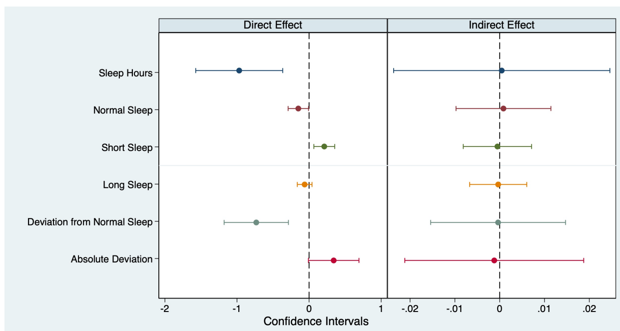
3.8 Appendix

3.8.1 Mediation Analysis for subgroups

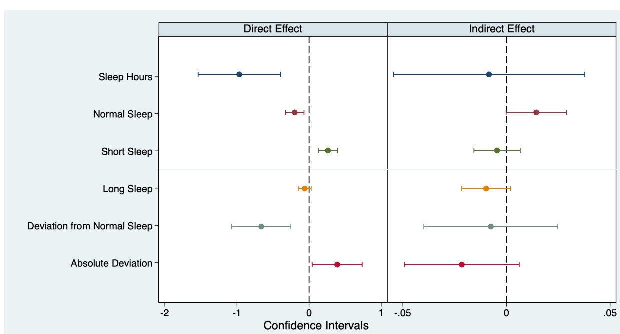
Mediator: Indoor air pollution during heating season



Mediator: personal exposure to PM2.5



Mediator: average indoor temperature during sleep time(22:00-8:00)



Mediator: standard deviation of indoor temperature during sleep time(22:00-8:00)

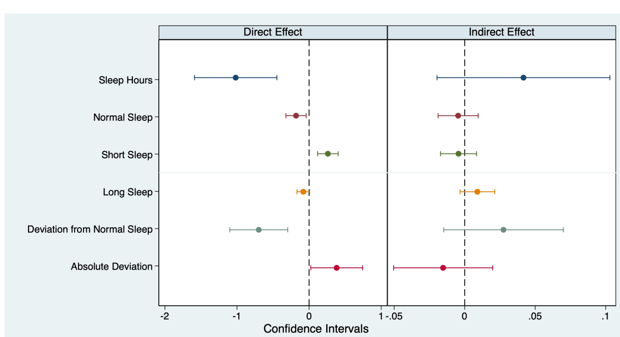
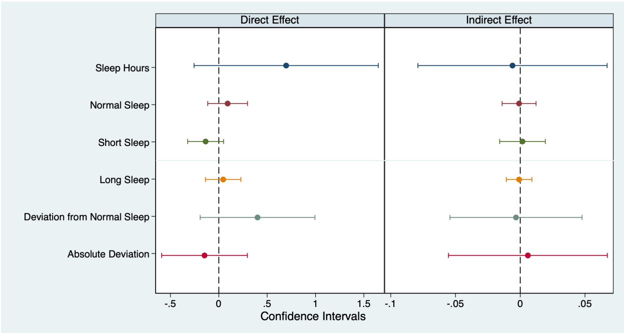


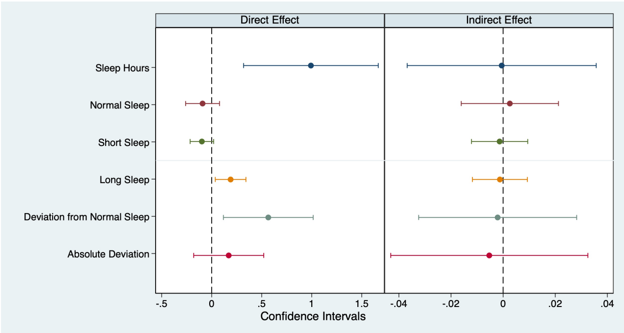
Figure 3.8.1: Mediation Analysis for Subgroup: Initial Short Sleepers

Note: Direct effects $= \alpha_1$ estimated by Eq (3.1) and Indirect effects $= \beta_2\gamma_1$ estimated by Eq (3.2) and (3.3) for individuals with a short sleep duration at baseline.

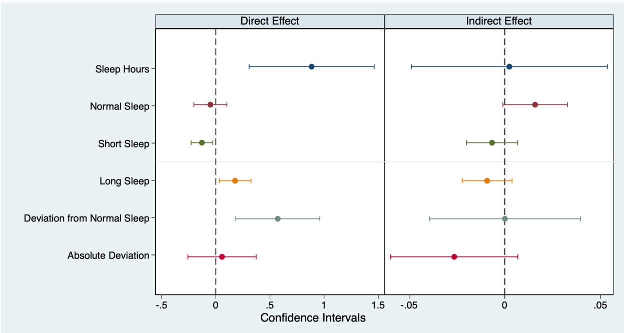
Mediator: Indoor air pollution during heating season



Mediator: personal exposure to PM2.5



Mediator: average indoor temperature during sleep time(22:00-8:00)



Mediator: standard deviation of indoor temperature during sleep time(22:00-8:00)

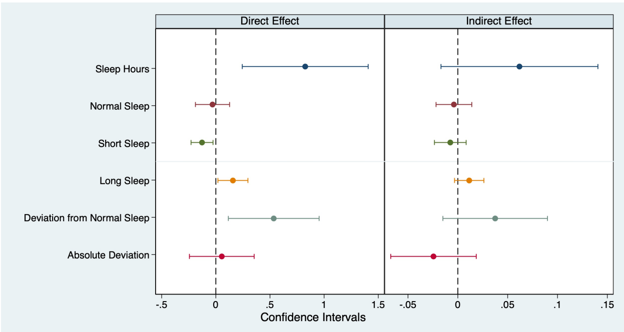


Figure 3.8.2: Mediation Analysis for Subgroup: Initial Long Sleepers

Note: Direct effects = α_1 estimated by Eq (3.1) and Indirect effects = $\beta_2\gamma_1$ estimated by Eq (3.2) and (3.3) for individuals with a long sleep duration at baseline.

From Environmental Policy Impacts to Environmental Shocks: Bridging Public Health and Development with Policy Interventions

Building on the exploration of indirect health effects from environmental policies examined in the second essay, where we observed how cleaner heating systems impacted sleep patterns in rural China, the dissertation transitions to the third essay which broadens the scope of environmental impact. This essay explores how natural environmental shocks, specifically rainfall, affect early-life conditions and subsequent human capital development in rural Indonesia. Furthermore, it investigates the interaction between these environmental shocks and the Early Childhood Education and Development (ECED) program, assessing how policy interventions can potentially mitigate or amplify the long-term effects of such shocks on child development.

While the second essay focuses on the implementation of a specific policy and its unintended health benefits, the third essay shifts to the natural environment's role in shaping life outcomes through early developmental stages. Both essays emphasize the profound and often unexpected ways in which environmental factors influence personal health and well-being. In the second essay, the reduction in indoor air pollution from cleaner heating technologies showed complex effects on sleep quality, an essential component of health. Similarly, the third essay examines how environmental shocks, like rainfall, serve as external stressors that have long-term effects on human capital development, highlighting the vulnerability of rural populations to their environmental conditions. This essay further delves into how strategic policy interventions, such as ECED programs, interact with these environmental shocks to either buffer or exacerbate their impacts, offering insights into the nuanced dynamics between environmental conditions and policy efficacy.

This progression underscores the importance of understanding both policy-induced changes and natural environmental conditions to fully grasp their implications on public health and development. Such insights are crucial for designing interventions that not only address the immediate environmental issues but also mitigate the long-term developmental impacts on future generations.

Chapter 4

How do Early-life Shocks Interact with Subsequent Human Capital Investment? Evidence from Indonesia

Wenmei Tu

Department of Economics, McGill University, Montreal, Quebec, Canada

Abstract

This study investigates the dynamic interplay between early-life shocks and subsequent investments in human capital, focusing on children in Indonesia. It examines how early-life shocks and later educational investments interact to influence long-term human capital development. Understanding this interaction is crucial for both parents and policymakers aiming to optimize investment strategies that maximize lifelong human capital development. This paper leverages the triple difference method to examine the combined effects of two sources of exogenous variation on children's human capital in rural Indonesia : early-life rainfall shocks, which significantly influence rural household income and consequently initial investments in children's human capital, and access to free Early Childhood Education and Development (ECED) services that aim to prepare children for school. The analysis reveals that positive early-life rainfall shocks significantly enhance children's mental and behavioral development, resulting in at least a 15% decrease in the likelihood of poor development in areas such as emotional maturity, conduct problems, and pro-social behavior issues. Conversely, the evaluation of the ECED program indicates that its effects on early childhood development are mild or insignificant. This minor effect does not necessarily imply that ECED services are ineffective but may reflect challenges in program delivery or alignment with children's needs. Furthermore, the interaction between rainfall shocks and the ECED program indicates that while they act as substitutes in improving certain anthropometric and emotional outcomes with minor or modest effects, their impact on behavioral and cognitive outcomes is mixed and remains unclear. The ECED program demonstrates limited capacity to counteract the adverse effects of early-life shocks across nearly all developmental dimensions.

4.1 Introduction

The extensive body of literature underscores the long-term negative impacts of early-childhood shocks on subsequent life outcomes, such as health, education, income, and other socioeconomic variables (Almond et al., 2018; Agüero, 2014; Lavy et al., 2016; Lee, 2014; Miller and Urdinola, 2010). However, the question remains: Can early life shocks be effectively mitigated by later investments or interventions? Understanding the dynamics between shocks and investments at different stages is vital for both parents and policymakers, as determining whether investments at different times act as substitutes or complements can influence the effectiveness of parental efforts and government interventions. Addressing this question could also guide strategies for optimal investment allocation to maximize lifetime human capital or to assist disadvantaged groups in catching up.

Addressing the interplay between shocks and investments poses significant challenges. Investments in children may correlate with both the children's inherent endowments and unobserved parental factors (Almond and Mazumder, 2013), complicating the attribution of direct causal effects. Therefore, identifying causal impacts requires leveraging at least two sources of exogenous variation that affect children's developmental stages distinctly.

This paper investigates two such exogenous variations in rural Indonesia. The first is early-life rainfall shocks, which are closely linked to agricultural output. Increased rainfall during early childhood typically boosts family income, enabling greater investment in children's initial human capital (Maccini and Yang, 2009). The second source is the implementation of the Early Childhood Education and Development (ECED) program in 2008, which expanded access to educational and developmental services for children aged 0-6 in economically disadvantaged districts. Enrollment in this program is quasi-randomly selected by the government, making it an ideal exogenous variable for study. By exploring these two factors, this study aims to illuminate how early interventions and environmental factors collectively contribute to human capital formation.

Despite the significance of understanding how shocks and investments interact to shape children's human capital development, few studies have successfully estimated these interactions causally, primarily due to empirical challenges. The existing evidence on the subject is mixed. For instance, Aguilar and Vicarelli, 2022 and Malamud et al., 2016 found no significant interactions affecting later life outcomes, suggesting that shocks and investments may operate independently of each other. On the other hand, research by Gilraine, 2016, Johnson and Jackson, 2019, and Duque et al., 2018 suggests that investments at different stages can act as complements, enhancing the overall impact on human capital. Conversely, Rossin-Slater and Wüst, 2015 and studies by Adhvaryu and Nyshadham, 2015, Gunnsteinsson et al., 2014, and Sviatschi, 2022 indicate a substitution effect, where later investments can mitigate the adverse impacts of early shocks. This paper contributes further evidence to this debate, enhancing our understanding of the nuanced processes involved in human capital formation.

Moreover, this study is pertinent to the evaluation of policies or programs aimed at assisting the disadvantaged, not merely by assessing the direct effects of such interventions but by examining whether there are heterogeneous impacts across different levels of human capital. Specifically, it investigates whether programs like Early Childhood Education and Development (ECED) can help bridge the development gap caused by early-life adversities. While few evaluations have approached ECED programs from this angle, findings by Jung and Hasan, 2016 suggest that the introduction of the ECED program in Indonesia led to a narrowing of the achievement gap between children from different socioeconomic backgrounds in project villages. The results of this paper could thus provide valuable insights for future policies or programs targeting disadvantaged groups, particularly if effects are found to be more pronounced among children with lower initial endowments, thereby underscoring the cost-effectiveness and high returns of focusing on these vulnerable populations.

This paper also contributes evidence on the mechanisms by which early-life shocks influence adult outcomes. While extensive research has documented the long-term consequences of early-life adversities on adulthood (Almond, 2006; Almond and Mazumder, 2011; Maccini and Yang, 2009; Scholte et al., 2015), fewer studies have explored their medium-term effects (Leight et al., 2015; Shah and Steinberg, 2017; Aguilar and Vicarelli, 2022). This study fills this gap by providing comprehensive indicators of early childhood development for children aged 2 to 6 years, encompassing physical, cognitive, language, social, and behavioral domains. The breadth and variety of these indicators allow us to determine which aspects of early childhood development are most susceptible to shocks and which can be effectively mitigated by subsequent investments in Early Childhood Education and Development (ECED).

In this research, we analyze the impact of early-life rainfall shocks in Indonesia on child development for ages 4-6 and evaluate the randomized ECED program initiated in 2008 using a Difference-in-Difference model. Crucially, we also explore the interaction between early-life rainfall shocks and later ECED investments through a triple difference approach. Our primary assumption is that both the occurrence of rainfall shocks and the exposure to the ECED program are exogenous, thus influencing changes in children's human capital. We linked rainfall data to individual children in the ECED dataset based on their birth dates and locations. Our findings indicate that positive early-life rainfall shocks correlate with better developmental outcomes across multiple dimensions. However, the ECED program evaluation reveals that it has mild or insignificant effects on early childhood development and does not significantly counteract the effects of early adversities in most developmental dimensions. Regarding the interaction between early-life rainfall shocks and the ECED program, we found substitutable effects on anthropometric and health outcomes, though the results are mixed and ambiguous for behavioral and cognitive outcomes.

The paper is organized as follows: Section 2 introduces the conceptual framework. Section 3 describes the empirical context, including the nature of the weather shocks and the ECED program. Section 4 outlines the empirical methods employed, Section 5 presents the main findings, Section 6 discusses robustness checks, and Section 7 concludes the paper.

4.2 Conceptual Framework

Heckman, 2007 developed a comprehensive model of human capital investment that integrates diverse strands of empirical evidence related to human capital development. The core of Heckman's framework is the human capital production function, which is formulated as follows:

$$h_{t+1} = f_t(p, h_t, I_t)$$

where h_t represents the stock of human capital at time t, p denotes parental capabilities, and I_t is the parental investment at that time. The function f_t is assumed to be strictly increasing and concave with respect to I_t, highlighting diminishing returns on investments. It is also twice differentiable in all its arguments, with $\partial f_t / \partial I_t > 0$. Through recursive substitution, the human capital stock can be represented as:

$$h_{t+1} = m_t(p, h_1, I_1, \ldots, I_t)$$

indicating that the stock of human capital depends on parental capabilities, initial human capital, and all past investments.

To delve deeper into the interaction between investments at different stages, Heckman, 2007 delineates two crucial properties within the human capital development process: self-productivity and dynamic complementarity. Self-productivity is described by:

$$\frac{\partial f_t(p, h_t, I_t)}{\partial h_t} > 0$$

implying that higher levels of human capital in one period enhance capabilities in subsequent periods. Dynamic complementarity, expressed as:

$$\frac{\partial^2 f_t(p, h_t, I_t)}{\partial h_t \partial I_t} > 0$$

suggests that children with greater early capabilities benefit more efficiently from later investments, embodying the notion that "skill begets skill."

Consider a simplified model where childhood has only two stages, leading to adult human capital expressed as:

$$h_3 = m(p, h_1, I_1, I_2)$$

Assuming a constant elasticity of substitution production function for investments, adult human capital can be modeled as:

$$h_3 = m\left(p, h_1, [\gamma(I_1)^{\phi} + (1-\gamma)(I_2)^{\phi}]^{\frac{1}{\phi}}\right)$$

where $\phi \leq 1$ and $0 \leq \gamma \leq 1$. Here, ϕ measures the substitutability of late investments for early ones, with $\phi = 1$, indicating perfect substitutes and $\phi \to -\infty$denoting perfect complements. The parameter γ acts as a capability multiplier reflecting both self-productivity and dynamic complementarity effects.

The optimal investment ratio can be derived based on these parameters. For $\phi = 1$, i.e., when early and late investments are perfect substitutes, the optimal strategy is influenced by discounting effects and self-productivity; it's advantageous to invest early if $\gamma > (1-\gamma)(1+r)$. For $\phi = -\infty$, i.e., when they are perfect complements, equal investments ($I_1 = I_2$) are optimal, emphasizing the necessity of balanced investment across stages. In more general scenarios where $-\infty < \phi < 1$, the optimal ratio of early to late investment is given by:

$$\frac{I_1}{I_2} = \left[\frac{\gamma}{(1-\gamma)(1+r))}\right]^{\frac{1}{1-\phi}}$$

As the degree of complementarity varies, so does the influence of the capability multiplier on the optimal investment strategy. A higher degree of complementarity dictates substantial investment at both early and late stages, whereas decreasing complementarity shifts focus towards early investment.

This two-stage model highlights the importance of whether investments in different periods act as substitutes or complements, shaping the efficacy of parental and governmental interventions. When dynamic complementarity is present, it suggests that for parents and policy-makers alike, early intervention is crucial and should be complemented by subsequent investments to maintain and enhance early gains.

4.3 Background

4.3.1 ECED Program

4.3.1.1 The Different Types of ECED Services Available in Indonesia

According to the World Bank report by Hasan et al., 2013, Indonesia offers a wide range of Early Childhood Education and Development (ECED) services that are overseen by various ministries. These services are comprehensively listed in Table (4.3.1). These services are designed to cater to specific age groups, as illustrated in Figure (4.3.1). The predominant forms of pre-primary education are playgroups and kindergartens, each targeting distinct age brackets. Playgroups primarily serve children aged 3 and 4, while kindergartens are designed for children aged 5 and 6. However, the report also noted that adherence to these age-specific guidelines can be flexible in practice. It is not uncommon for some children to remain in playgroups beyond the intended age limit of 4 years. Similarly, early enrollment in primary school is a frequent occurrence, with children sometimes starting as early as age 5 or 6, driven by the fact that unlike kindergartens, which often require tuition fees, primary education is compulsory and provided at no cost. By the age of 7, children are expected to transition to primary school, with almost universal enrollment achieved by this milestone.

In Indonesia, the intensity of Early Childhood Education and Development (ECED) services varies significantly between different types (Hasan et al., 2013; Brinkman et al., 2017). Both playgroups and kindergartens typically operate from 8 a.m. to 11 a.m., but they differ in frequency; playgroups usually meet only three days a week, while kindergartens operate five to six days a week. Consequently, children in kindergartens receive more weekly hours of intervention compared to those attending playgroups.

Brinkman et al., 2017 also pointed out that despite the range of early childhood services available in Indonesia, their implementation faces several persistent challenges. First, coverage remains low across the country, limiting access to many children who could benefit from early education. Second, the majority of ECED services are privately provided, which is problematic given the low levels of public investment in this area. This often results in high costs that can exclude

	Ministry of Education and Culture (MoEC)	Ministry of Relligious Affairs (MoRA)	Ministry of Home Affairs with Ministry of Health Staff	National Familly Planning Board
Formal	Kindergartens (Taman Kanak-kanak, TK)	Islamic Kindergartens (Raudhotul Atfal, RA)		
	Playgroups (Kelompok Bermain, KB)	Islamic Kindergartens Gaman Pendidikan Quran, TPQ)	Integrated Health Service Units (Posyandu)	Toddler Family Groups (Bina Keluarga Balita, BKB)
Non-formal	ECED Posts (Pos PAUD)			
	Childcare centers (Taman Penitipan Anak, TPA)			
	Other early childhood units (Satuan PAUD Sejenis, 5PS)			

Table 4.3.1: ECED services are provided in different formats by different ministries

[a] Source: From Atfal, 2012, p.1

economically disadvantaged families. Lastly, the quality of education is frequently compromised by the use of volunteer teachers who typically have minimal or no formal training in early childhood education, as training opportunities for teachers in this sector are scarce. These issues collectively hinder the effectiveness of early childhood interventions in Indonesia.

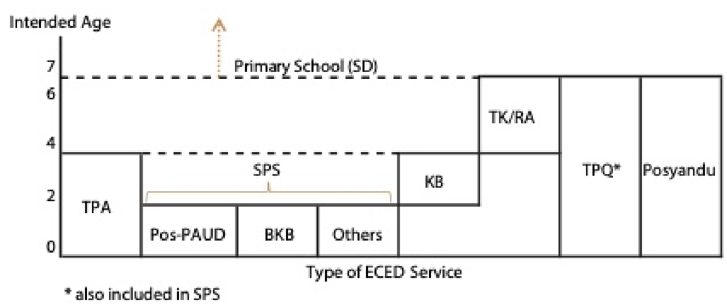


Figure 4.3.1: The intended age for various ECED services in Indonesia

Source: From Atfal, 2012, p.1; for the full name of ECED service, please refer to Table (4.3.1)

4.3.1.2 The Early Childhood Development and Education Project

In Indonesia, access to Early Childhood Education and Development (ECED) services and subsequent child developmental outcomes are significantly influenced by socioeconomic disparities (Jung and Hasan, 2016). To address these disparities, the Indonesian government, collaborating with the World Bank and the Dutch government, initiated an ECED project in 2008 targeting 3,000 underprivileged villages across 50 districts. These districts were selected based on a composite score primarily focused on poverty rates and gross enrollment rates, aiming to improve ECED access for children aged 0–6 years in areas with low participation rates.

As part of the initiative, each district was required to establish a district-wide office for early childhood services to ensure a structured support system. The program provided a comprehensive support package, including the assignment of community facilitators, block grants, and teacher training, detailed further in Hasan et al., 2013, Jung and Hasan, 2016, and Hasan et al., 2021. This intervention was crafted to foster sustainable, community-based services tailored to meet the specific needs of each village, thus promoting the holistic development of young children.

A central component of this initiative was the formation of playgroups, which typically operated several days a week. These groups utilized a play-based curriculum that emphasized learning through play, the use of tangible, hands-on materials, teacher scaffolding, and a balanced mix of structured and child-initiated activities. This approach aimed to create an engaging and supportive learning environment that accommodates the varied developmental needs of young children.

4.3.1.3 Evaluation Design

This study utilizes data from 310 villages across nine districts in Indonesia, all participating in the Early Childhood Development and Education Project. These districts were selected based on their willingness to cooperate with a randomized rollout of the program and their geographical diversity, ensuring a representation of the regional variety within the project locales. Within each district, villages were grouped into three categories: 10 villages were randomly chosen to receive the project in the first round, another 10 in a subsequent round, and a third group of 10 villages served as a matched comparison group. These comparison villages were identified by local administrators as being similar to those in the treatment groups but were not scheduled to receive the project.

As a result, the study comprised 218 treatment villages and 92 comparison villages. Of the treatment villages, 105 received the project in the first batch in 2009, and 113 received it 11 months later in 2010. The comparison villages were selected by district governments based on similar poverty levels to the treatment villages but were never included in the project. This non-random selection of the 10 comparison villages represents a limitation of the study's design.

Data collection occurred over four rounds: a baseline survey conducted from March to June 2009 for children aged 2-6 years, and three follow-up surveys in 2010, 2013, and 2016. This dataset includes vital information such as children's birth dates and locations, enabling linkage with rainfall data, along with various developmental outcomes. The timeline of project implementation and the survey schedule are detailed in Table (4.3.2). Due to procurement delays with the survey firm, the baseline survey was conducted six months after the initial batch of villages received their block grants. Most villages in the late-treatment group received their grants about five months post-baseline survey.

Event	Date	Description
Project Start	Jan 2008	Initiation of the project with early village treatments starting.
Baseline Survey	Mar 2009	Initial survey conducted to gather baseline data.
2nd Survey	Jan 2010	Follow-up survey to assess mid-term project impact.
Later Village Treatments Begin	April 2010	Start of treatments in later villages.
3rd Survey	Mar 2013	Third survey conducted to mid-term project impact.
4thd Survey	2016	Final survey conducted to assess overall project impact.

Table 4.3.2: Timeline of Project Implementation and Survey Schedule

For the purposes of this analysis, the paper focuses on comparing the late-treatment villages with the comparison group, utilizing only the pre-treatment data from the late-treatment group. This approach facilitates a difference-in-differences analysis, allowing for robust estimation of the project's impacts by contrasting changes over time between the two groups under different treatment conditions.

By the second follow-up, the late-treatment villages had experienced 9 months of exposure to the program, and by the third follow-up, this exposure extended to 40 months. Meanwhile, the comparison villages had not been exposed to the project at all.

4.3.1.4 Data

This paper concentrates on data from the later-treated villages and the matched control group, specifically during the first two survey waves conducted in 2009 and 2010. The developmental measures used in these surveys are consistent due to their temporal proximity. In contrast, the measures used in the third and fourth surveys were modified to account for the increasing age of the children, resulting in changes to the developmental assessments. The exposure to the Early Childhood Education and Development (ECED) program is represented by a binary variable, assigned a value of 1 for children residing in later treated villages, indicating their exposure to the program.

The primary group of interest comprises children who were aged 4-6 at the time of the baseline survey. This age range is particularly relevant as the most commonly established services in the treatment villages were playgroups, which predominantly cater to children ages 4 and 5, according to Brinkman et al., 2017. This age specification helps to ensure that the analysis focuses on the subset of the population most likely to be directly influenced by the playgroup interventions, thereby providing a clearer assessment of the ECED program's impact on early childhood development in these communities.

4.3.1.5 Outcome Variables

Child development was assessed using a comprehensive suite of internationally validated and locally adapted measures within the ECED dataset. The descriptive statistics for these outcomes are summarized in Table (4.3.4) and include the following components:

(1) **Anthropometric and Health Indicators**: These measures are critical for assessing balanced nutrition and overall physical health. During each survey wave, height and weight data are collected for children aged 2-6 years. Stunting is defined as weight falling more than two

		All		Treatment=1		Treatment=0	
		Mean	SE	Mean	SE	Mean	SE
Panel A: Anthropometrics and health							
(1)	Weight	14.29	2.17	14.27	2.12	14.31	2.22
(2)	Height	99.12	5.72	99.09	5.71	99.16	5.73
(3)	Height-for-age Z-score	-1.63	1.24	-1.64	1.22	-1.61	1.26
(4)	Weight-for-age Z-score	-1.49	1.05	-1.50	1.02	-1.48	1.09
(5)	BMI-for-age Z-score	-0.67	1.43	-0.67	1.46	-0.66	1.39
(6)	Stunting	0.29	0.46	0.29	0.46	0.29	0.46
(7)	Short	0.36	0.48	0.37	0.48	0.36	0.48
(8)	Healthy	0.89	0.31	0.89	0.31	0.89	0.32
Panel B: Poor Early Development reported by caregivers							
(1)	Any Poor Development	0.44	0.50	0.43	0.50	0.46	0.50
(2)	Poor Developments > 1	0.93	0.25	0.94	0.25	0.93	0.26
(3)	Physical health	0.17	0.38	0.15	0.36	0.20	0.40
(4)	Social Competence	0.11	0.31	0.12	0.32	0.09	0.29
(5)	Emotional Maturity	0.37	0.48	0.36	0.48	0.37	0.48
(6)	Language and Cognitive S1	0.89	0.32	0.89	0.32	0.88	0.32
(7)	Communication and Gener	0.06	0.23	0.07	0.25	0.05	0.21
Panel C: Strengths and Difficulties Questionnaire(SDQ) rep. by caregivers							
(1)	Total difficulties	0.38	0.48	0.38	0.49	0.37	0.48
(2)	Emotional Symptoms	0.34	0.47	0.33	0.47	0.34	0.47
(3)	Conduct Problems	0.48	0.50	0.49	0.50	0.48	0.50
(4)	Hyperactivity/Inattention	0.03	0.18	0.03	0.18	0.03	0.18
(5)	Peer Problems	0.27	0.44	0.27	0.44	0.27	0.44
(6)	Pro-social behavior	0.33	0.47	0.34	0.47	0.31	0.46
Panel D: Controls							
(1)	Wealth Index	0.06	1.66	-0.05	1.58	0.20	1.75
(2)	Poor	0.48	0.50	0.48	0.50	0.47	0.50
(3)	Age in month	54.03	3.40	54.09	3.46	53.95	3.32
(4)	Sex	1.50	0.50	1.50	0.50	1.49	0.59

Table 4.3.3: Descriptive Statistics at Baseline

[a] Variables in Panels B and C are binary, where a value of 1 indicates poor development or difficulties in a specific dimension for the child.

standard deviations below the age-specific mean, often indicative of insufficient nutrition during early development and potentially leading to long-lasting adverse effects, including premature mortality (UNICEF., 2009). Being below the height norm (referred to as ”short”) by more than two standard deviations is considered less severe than stunting but may still suggest nutritional or health issues. Additionally, caregiver-reported health status and instances of illness in the past month are also recorded.

(2) **Strength and Difficulties Questionnaire (SDQ)**: This tool is used to screen for potential mental health problems in children and adolescents aged 3-16 years. Developed by Goodman in 1997, the SDQ comprises five scales—emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems, and prosocial behaviors—each scored from 1 to 10 and consisting of five items. Studies have confirmed the SDQ's good concurrent and discriminant validity and moderate test-retest reliability (Muris et al., 2003; Lundh et al., 2008; Yao et al., 2009).

(3) **Early Development Instrument (EDI) (short version)**: Created by Dr. Janus and Dr. Offord at the Offord Centre for Child Studies in 2007, the EDI evaluates whether children meet age-appropriate developmental expectations across five domains: physical health and well-being, social competence, emotional maturity, language and cognitive development, communication skills, and general knowledge. A 47-item short version of the standard 104-item EDI is used in Indonesia, recognized for its predictive validity concerning later educational outcomes (Pradhan et al., 2013).

(4) **Dimensional Change Card Sort (DCCS) Game**: The DCCS task is a standard test for assessing executive function in children aged 3-7 years. It involves asking children to switch sorting strategies for cards, for example, from color to shape, measuring cognitive flexibility (Zelazo, 2006).

(5) **Drawing Tasks**: Children are tasked with drawing a human figure and a house, with the complexity and detail of these drawings used as indicators of cognitive development. The number of identifiable parts included in each drawing is scored, with more detailed representations receiving higher scores.

(6) **Expressive and Receptive Language Tasks**: These tasks assess expressive language by asking children to name everyday items and receptive language by requiring them to identify body parts as named by the interviewer.

These diverse measures provide a robust framework for evaluating the multifaceted aspects of child development, capturing both physical and cognitive dimensions critical to early childhood growth and education.

4.3.1.6 Previous literature

Numerous studies have evaluated Early Childhood Education and Development (ECED) services and programs across various countries, both developing and developed, with mixed findings regarding their effectiveness. In developed nations, research has yielded conflicting results. For instance, Spiess et al., 2003 found that kindergarten attendance among children of German citizens did not significantly affect later school placement. Similarly, Driessen, 2017 observed that ECED services in the Netherlands did not significantly enhance children's cognitive and non-cognitive skills after accounting for parental characteristics. Conversely, Reynolds et al., 2004 reported that preschool attendance markedly improved educational achievements and social outcomes for low-income children in a 15-year study in Chicago.

In developing countries, the evidence is equally varied. Mwaura et al., 2008 found that African children who attended preschool exhibited significant improvements on ability and intelligence tests. Raine et al., 2010 observed better behavioral outcomes among children who participated in an enriched preschool program in Mauritius. However, Bouguen et al., 2013 studied three scaled-up ECED programs in Cambodia and noted little difference in development outcomes over three years.

Research consistently shows steep socio-economic gradients in early childhood development—children from poorer or less-educated families generally exhibit poorer developmental outcomes (Schady et al., 2015; Driessen, 2017; Berlinski et al., 2008; Behrman et al., 2006; Grantham-McGregor et al., 2007; Naudeau et al., 2011; Fernald et al., 2012). Studies also document heterogeneous treatment effects of ECED services across different socio-economic backgrounds, often finding larger impacts for disadvantaged groups (Barnett, 1997; Dhuey, 2007; Spiess et al., 2003; Peisner-Feinberg et al., 2001). Yet, few studies have systematically investigated whether ECED services have narrowed capability gaps across socio-economic groups. Burger, 2010 conducted a systematic review on the effects of early childhood education on cognitive development and its potential to compensate for social inequalities due to socio-economic status differences. He found that children from lower socio-economic backgrounds made more substantial improvements than their more advantaged peers, though the development gap could not be completely bridged by ECED interventions. Similarly, Jung and Hasan, 2016 assessed the ECED program in Indonesia using a fixed-effects model with a DIDID estimator and observed decreased achievement gaps in early childhood development between poorer and richer children compared with control villages.

However, endogeneity issues often complicate these studies, as children's enrollment in ECED services tends to correlate with their socio-economic status, which is closely linked to their pre-intervention human capital stock. Thus, attributing causal effects of ECED programs on reducing or exacerbating capability gaps remains challenging. This paper aims to causally estimate the impact of a community-based program that increases access to ECED services in rural areas of Indonesia on reducing the capability gaps generated by exogenous early-life shocks, providing clearer insights into the effectiveness of such interventions in leveling the playing field for children from varying socio-economic backgrounds.

4.3.2 Rainfall shocks in Indonesia

4.3.2.1 Description

Weather in Indonesia is characterized primarily by variations in precipitation rather than temperature, due to its equatorial location, which ensures minimal temperature fluctuations throughout the year. However, the amount of rainfall can vary significantly both within and across years across the country. Indonesia's climate typically alternates between a wet season, with monthly rainfall exceeding 200 mm, and a dry season each year. The duration and intensity of these seasons can vary greatly among different regions and are subject to annual changes. This variability is largely due to the monsoons, whose paths can shift each year. Early or delayed monsoons can significantly affect the timing and amount of rainfall during the wet season, which in turn impacts agricultural conditions.

Several studies highlight the relationship between higher rainfall and increased agricultural output and productivity in Indonesia (Levine and Yang, 2006; Kishore et al., 2000). In rural areas, where agriculture remains a primary source of income, greater precipitation can lead to higher household incomes. This economic boost provides families with more resources to invest in nutrition and healthcare for children. This paper focuses on this pathway, examining how variations in rainfall influence child human capital through changes in agricultural productivity and subsequent shifts in household income.

4.3.2.2 Previous Literature

The literature on the consequences of early-life rainfall presents mixed evidence. Maccini and Yang, 2009 find that higher rainfall in early life positively affects long-term socioeconomic outcomes in adulthood, but this effect is observed predominantly among women. In contrast, other studies highlight the drawbacks of irregular rainfall patterns, such as delays in monsoons. For instance, Korkeala, 2012 notes that delayed monsoons are associated with an increase in child labor, while Thiede and Gray, 2020 report that such delays lead to poorer health outcomes in children. Collectively, these findings suggest that lower rainfall during early childhood in Indonesia can have detrimental effects on children's human capital, impacting both health and educational outcomes. Thus, a deficit in early-life rainfall can be considered as a form of negative investment in human capital.

4.3.2.3 Rainfall Data

The rainfall data utilized in this study are sourced from the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS). CHIRPS is a quasi-global rainfall dataset that covers a wide geographical range from 50°S to 50°N across all longitudes. It spans from 1981 to near-present, providing high-resolution estimates of rainfall in a time series format. The data are segmented into 0.05-by-0.05-degree cells, which offers detailed precipitation records. For the purposes of this study, monthly rainfall data for each village included in the ECED dataset have been extracted from

January 1981 through March 2015, allowing for comprehensive analysis of precipitation patterns and their potential impacts on the ECED outcomes.

4.3.2.4 Exposure to Early-life Shocks

A rainfall shock is defined as a significant shortfall in rainfall during one's birth year, specifically falling below the 20th percentile of the historical norm for the individual's birth village. This threshold was determined after assessing the nonlinear impacts of rainfall on various outcomes both in simple OLS and non-parametric settings, where rainfall below and above the 20th percentile demonstrated notably different effects. Additionally, while rainfall during other critical periods such as in-utero or the second year of life can also influence outcomes, the effects of shocks during these periods tend to be smaller in magnitude or statistically insignificant compared to those during the birth year. Consequently, the analysis primarily focuses on shocks experienced in the birth year.

For analytical clarity, a ”positive rainfall” dummy variable is introduced to indicate the absence of a negative shock. This facilitates the interpretation of interaction coefficients within the model. When calculating ”rainfall in one's birth year,” the analysis considers the complete wet and dry seasons relevant to the agricultural cycles rather than the calendar year. Following the methodology of Maccini and Yang, 2009, the ”birth season” for each individual is determined based on their birth month and province. ”Rainfall in one's birth year” is then calculated as the total precipitation during the birth season plus the subsequent season, effectively covering the 12 consecutive months that span a child's first wet and dry seasons. The ”positive rainfall” dummy is assigned a value of 1 if the total rainfall during this period exceeds the 20th percentile for the individual's birth village, based on historical data from 1981 to 2015.

4.4 Empirical Strategy

4.4.1 Effects of Early-life Shocks on Human Capital

This paper follows the approach of Maccini and Yang, 2009 to estimate the impact of birth year rainfall on child developmental outcomes, specifying the following reduced-form linear relationship

		ALL (N = 1608)		Positive Rainfall=1 (N = 1046)		Positive Rainfall=0 (N = 562)	
		Mean	SE	Mean	SE	Mean	SE
Panel A: Anthropometrics and health							
(1)	Weight	14.29	2.17	14.26	2.13	14.35	2.24
(2)	Height	99.12	5.72	99.06	5.59	99.23	5.95
(3)	Height-for-age Z-score	-1.63	1.24	-1.55	1.20	-1.78	1.29
(4)	Weight-for-age Z-score	-1.49	1.05	-1.45	1.05	-1.57	1.06
(5)	BMI-for-age Z-score	-0.67	1.43	-0.68	1.39	-0.64	1.51
(6)	Stunting	0.29	0.46	0.28	0.45	0.31	0.46
(7)	Short	0.36	0.48	0.34	0.47	0.41	0.49
(8)	Healthy	0.89	0.31	0.88	0.33	0.91	0.28
Panel B:Poor Early Development reported by caregivers							
(1)	Any Poor Development	0.44	0.50	0.45	0.50	0.43	0.49
(2)	Poor Developments > 1	0.93	0.25	0.94	0.24	0.92	0.27
(3)	Physical health	0.17	0.38	0.18	0.38	0.16	0.37
(4)	Social Competence	0.11	0.31	0.09	0.29	0.13	0.34
(5)	Emotional Maturity	0.37	0.48	0.37	0.48	0.36	0.48
(6)	Language and Cognitive Skills	0.89	0.32	0.91	0.29	0.84	0.37
(7)	Communication and General Knowledge	0.06	0.23	0.06	0.23	0.06	0.24
Panel C: Strengths and Difficulties Questionnaire(SDQ) rep, by caregivers							
(1)	Total difficulties	0.38	0.48	0.39	0.49	0.35	0.48
(2)	Emotional Symptoms	0.34	0.47	0.35	0.48	0.31	0.46
(3)	Conduct Problems	0.48	0.50	0.48	0.50	0.49	0.50
(4)	Hyperactivity/Inattention	0.03	0.18	0.04	0.19	0.02	0.16
(5)	Peer Problems	0.27	0.44	0.27	0.44	0.27	0.45
(6)	Pro-social behavior	0.33	0.47	0.29	0.45	0.40	0.49
Panel D: Controls							
(1)	Wealth Index	0.06	1.66	0.10	1.59	-0.02	1.77
(2)	Poor	0.48	0.50	0.48	0.50	0.48	0.50
(3)	Age in month	54.03	3.40	53.18	3.31	55.61	2.99
(4)	Sex	1.50	0.50	1.49	0.50	1.50	0.50

Table 4.3.4: Descriptive Statistics at Baseline

[a] Variables in Panels B and C are binary, where a value of 1 indicates poor development or difficulties in a specific dimension for the child.

for the developmental outcome Y_{ijst} of child i born in village k,district j, season s and year t:

$$Y_{ijst} = \beta R_{ikt} + \mu_{js} + \gamma_{js} TREND + \delta_{st} + \theta' X_{ijkt} + \varepsilon_{ijkst} \tag{4.1}$$

Here, the coefficient of interest is β, which measures the impact of rainfall in the birth year R_{ikt} on the developmental outcome. The model controls for potential timing of births by parents, who may choose specific seasons for childbirth based on varying factors. These choices might correlate with parental characteristics that could influence child development, necessitating robust controls in the analysis: μ_{js} represents fixed effects for each district-season combination, capturing variations across different geographic and temporal contexts within the same district. $\gamma_{js}TREND$ captures a linear time trend specific to each district-season, allowing the model to adjust for long-term changes over time that might affect developmental outcomes. δ_{st} is a fixed effect for the specific birth year-season combination, accounting for cohort effects that vary across different seasons. X_{ijkt} includes additional covariates that may influence the developmental outcome, with θ' representing the vector of coefficients for these variables. ε_{ijkst} is a mean-zero error term, which is assumed to have a potential arbitrary variance-covariance structure due to spatial and serial correlation within each birth district. Thus, standard errors are clustered by birth district to account for this correlation.

This specification provides a comprehensive approach to analyze how early-life environmental factors, like rainfall, affect child development, while adequately controlling for a range of fixed effects and trends that might confound the estimated relationships.

4.4.2 Effects of ECED Program on Human Capital

Before exploring the interaction between positive rainfall shocks and the ECED project, it is essential to assess the direct impact of the ECED project on child development outcomes. We utilize children from later treated villages as the treatment group and those from non-project villages as the comparison group, estimating the project's impact using the difference-in-differences (DID) method. This approach involves tracking two groups—similar at baseline—and comparing the difference in their outcomes at follow-up after one group has received an intervention and the other has not. For this study, the treatment group had no intervention at baseline but received the intervention for nine months by the time of the follow-up survey, whereas the comparison group had no intervention at either point.

The regression model is structured as follows:

$$Y_{ikt} = \alpha_i + \delta_0 Treat_{ik} + \delta_1 Post_t + \delta_2 Treat_{ik} \times Post_t + \epsilon_{ikt} \tag{4.2}$$

Where α_i controls for all time-invariant observed and unobserved characteristics of child i in village k. $Treat_{ik}$ is a binary indicator, with 1 representing children in the treatment group and 0 for those in the comparison group. $Post_t$ is a binary time indicator, set to 0 for the baseline period and 1 for the follow-up survey, conducted approximately 14 months later. δ_1 captures the general time effect on child development outcomes, reflecting natural progression in children's development and ECED enrollment as they age, independent of the project. $Treat_{ik} \times Post_t$ is the interaction term between the treatment group indicator and the time indicator. The coefficient δ_2, the DID estimator, measures the impact of the ECED project that had been implemented for about nine months at the time of the follow-up.

This model captures the effects not on children who enrolled in the ECED services but on those who were merely offered the opportunity to enroll. This perspective is particularly relevant for policymakers since it aligns with the operational reality of many social programs, which are based on voluntary participation by eligible individuals. A critical assumption underlying this model is that any differences in outcomes between the treatment and comparison villages at follow-up would mirror those at baseline if there were no ECED intervention or if the intervention had no effect.

In summary, using the DID methodology, we analyze how developmental outcomes of children living in treatment group villages differ from those in comparison group villages as they age from four to five years old, focusing on the implications of being offered access to ECED services rather than on direct participation alone.

4.4.3 Interaction between Early-Life Shocks and ECED Program

To examine potential interactive effects between rainfall shocks and ECED investments, we employ a difference-in-difference-in-difference (DIDID) approach, structured by the following model:

$$Y_{ijst} = \gamma_1 R_{ikt} + \gamma_2 Treat_{ik} + \gamma_3 Post_t + \gamma_4 R_{ikt} \times Treat_{ik} + \gamma_5 Treat_{ik} \times Post_t + \gamma_6 R_{ikt} \times Post_t \tag{4.3}$$

$$+ \gamma_7 R_{ikt} \times Treat_{ik} \times Post_t + \theta' X_{ijkt} + \mu_{js} + \gamma_{js} TREND + \delta_{st} + \varepsilon_{ijkst} \tag{4.4}$$

In this model, $Y_{ijst\tau}$ represents the developmental outcomes for child i, born in season s, year t, residing in village k and district j, measured at time τ; R_{ikt} s a dummy variable indicating a positive rainfall shock, set to 1 if the rainfall during the birth year of child i in village k was above the 20th percentile; $Treat_{ik}$ indicates whether child i resides in a village where the ECED program was implemented (1 for treatment villages, 0 for control); $Post_t$ is a dummy variable indicating the timing of the outcome measurement (1 for follow-up, 0 for baseline). Interaction terms $R_{ikt} \times Treat_{ik}$, $Treat_{ik} \times Post_t$ and $R_{ikt} \times Post_t$ explore various combined effects of rainfall, ECED treatment, and survey timing. $R_{ikt} \times Treat_{ik} \times Post_t$ is the triple interaction term of primary interest, examining the differential impact of the ECED program on children who experienced a positive rainfall shock.

The model controls for fixed effects related to district (μ_{js}), season (δ_{st}), and linear time trends $(\gamma_{js}TREND)$, as well as other covariates ($\theta' X_{ijkt}$), aiming to isolate the unique contributions o each variable to child development outcomes.

The coefficient γ_7 is pivotal as it captures the differential effect of the ECED program specifically for childre who experienced early positive rainfall conditions. Assuming both psositive rainfall shock and ECED program are beneficial investments, a positive γ_7 would indicate that the ECED program had a more substantial impact on children with a higher initial endowment (those who experienced a positive rainfall shock), suggesting a potential complementarity between early advantageous conditions and later educational investments. Conversely, a negative γ_7 suggests that the ECED program may act as a substitute for early advantages or might help mitigate the negative impacts of early-life adversities, narrowing capability gaps between children from different early-life conditions. This analysis helps in understanding whether interventions like ECED can effectively equalize or exacerbate initial disparities in child development outcomes.

4.5 Empirical Results

4.5.1 Effects of Early-life Rainfall Shocks on Human Capital

Before we delve into the effects of rainfall on developmental outcomes, let's review the descriptive statistics in Table (4.3.4), comparing children who experienced a positive rainfall shock with those who did not, based on data from the baseline survey. According to Panel A, children who experienced a positive rainfall shock show slightly worse outcomes in all anthropometric and health indicators than those without such a shock. Panel B and Panel C, which report early development measures assessed by caregivers, indicate that the group without the positive rainfall shock generally exhibits fewer developmental difficulties across most dimensions, except for Social Competence and Conduct Problems. These findings suggest that despite the assumption that higher rainfall in the birth year might boost child development through increased family income, it could also have adverse effects on development through other channels. For example, higher rainfall may increase the prevalence of waterborne diseases, which could adversely affect child health and development.

Additionally, it's noteworthy that the group without a positive rainfall shock is on average 2.5 months older than their counterparts, which might account for some of the observed differences; therefore, it is essential to control for age in the regressions. The higher wealth index in the group with a positive rainfall shock provides some support for the hypothesis that rainfall influences child development through income effects.

The detailed results from Panel A of Table (4.5.1) show that a positive rainfall shock negatively impacts weight, leading to lower weight-for-age Z-score, BMI-for-age Z-score, and Weight-for-height Z-score. This finding contradicts the hypothesis that increased rainfall, by boosting agricultural income, should result in higher investments in children and thus a higher initial human capital stock. Potential measurement errors in weight could be a factor behind these results. Additionally, this negative effect may indicate channels other than income through which rainfall could impact children's initial human capital. For instance, excessive rainfall may change the disease environment, impacting nutrient absorption critical in early growth. It can also affect environmental conditions that influence economic activity and public health, including forest fires,

		Positive Rainfall Shock Treatment Effects (γ_{js})	SE
Panel A: Anthropometric and health			
(1)	Weight	-0.441^{**}	(0.14)
(2)	Height	0.918^{***}	(0.14)
(3)	Height-for-age Z-score	0.195^{***}	(0.03)
(4)	Weight-for-age Z-score	-0.130^{*}	(0.07)
(5)	BMI-for-age Z-score	-0.357^{***}	(0.10)
(6)	Weight-for-height Z-score	-0.372^{***}	(0.10)
(7)	Stunting	0.016	(0.03)
(8)	Short	-0.043	(0.03)
(9)	Healthy	-0.025	(0.02)
(10)	Sick Last Month	0.034	(0.02)
Panel B: Poor Early Development reported by caregivers			
(1)	Any Poor Development	-0.170^{***}	(0.03)
(2)	Poor Developments > 1	-0.040^{*}	(0.02)
(3)	Physical health	0.022	(0.01)
(4)	Social Competence	0.015	(0.04)
(5)	Emotional Maturity	-0.152^{***}	(0.03)
(6)	Language and Cognitive Skill	-0.047^{*}	(0.02)
(7)	Communication and General Knowledge	-0.003	(0.01)
Panel C: Strengths und Difficulties Questionnaire(SDQ) rep, by caregivers			
(1)	Total difficulties	-0.111^{***}	(0.03)
(2)	Emotional Symptoms	0.051^{*}	(0.03)
(3)	Conduct Problems	-0.209^{***}	(0.02)
(4)	Hyperactivity/Inattention	-0.002	(0.00)
(5)	Peer Problems	-0.200^{***}	(0.01)
(6)	Pro-social behavior Difficulties	-0.194^{***}	(0.03)
Panel D: Early Development Measured by Interviewers			
(1)	Gross Motor	-0.015^{*}	(0.01)
(2)	Soft Motor(Score)	1.119^{***}	(0.20)
(3)	Language Skills	-0.001	(0.00)
(4)	Card sorting	1.443^{***}	(0.34)

Table 4.5.1: Effects of Early-life Positive Rainfall Shock on Child Development

[a] *** p<.01, ** p<.05, * p<.1
[b] Variables in Panels B and C are binary, where a value of 1 indicates poor development or difficulties in a specific dimension for the child.
[c] Treatment effects are γ_{js} estimated by Eq (4.4.1)

floods, landslides, water availability, and pest control (Maccini and Yang, 2009). These factors might negate the positive effects of improved agricultural output. However, positive rainfall shocks do positively impact height, albeit less than 1% of the average height, and also improve the Height-for-age Z-score by 0.15 standard deviations. There are no significant effects on stunting, short, or caregiver-reported health status. Overall, positive rainfall shocks have relatively small but adverse effects on children's anthropometric and health outcomes.

Panel B of the results, as reported by caregivers, addresses early developmental concerns including physical health, social competence, emotional maturity, language and cognitive skills, communication, general knowledge, any poor development across these aspects, and multiple poor developments. The data indicates that positive rainfall shocks have minimal and statistically insignificant effects on physical health, social competence, and communication and general knowledge. This aligns with the findings from Panel A, suggesting that rainfall shocks do not significantly influence anthropometric and health outcomes.

However, the influence of rainfall shocks is more pronounced in other developmental areas. Notably, there is a significant reduction in the likelihood of poor development in emotional maturity, decreasing by 15%, although the effects on language and cognitive skills are modest, improving by only 5%. Collectively, these results suggest that a positive rainfall shock reduces the likelihood of any poor development by 17% and decreases the occurrence of multiple poor developments by 4%.

Panel C details the outcomes from the Strength and Difficulties Questionnaire (SDQ) as reported by children's caregivers. This assessment divides into five dimensions of mental health in children, encompassing emotional symptoms, conduct problems, hyperactivity/inattention, peer problems, and pro-social behavior. For analytical clarity, binary indicators were constructed according to recommended scoring methods, which aggregate responses into these categorical outcomes.

The results show that positive rainfall shocks substantially reduce the likelihood of difficulties across almost all mental health dimensions. Notable reductions include a 11% decrease in total difficulties, a 15% reduction in emotional symptoms, a 21% decrease in conduct problems, a 5% reduction in peer problems, and a 20% decrease in pro-social behavior problems. The only dimension where the rainfall shocks did not significantly affect outcomes was hyperactivity/inattention. This exception can likely be attributed to the inherently low incidence rate of these difficulties among the sample—only 3% of children exhibited issues in this area, as opposed to the 30% to 50% facing challenges in the other dimensions.

Panel D of the report offers insights into early developmental assessments conducted by interviewers, covering areas such as gross motor skills, fine motor skills, language abilities, and executive functions evaluated through card sorting tests. The findings reveal that positive rainfall shocks have a statistically significant, albeit modest, beneficial impact on fine motor skills, language abilities, and executive functions—improving fine motor skills by an average of 3% and executive functions by 10%. In contrast, a negative effect is observed on gross motor skills.

In summary, the analysis reveals that positive rainfall shocks significantly benefit behavioral outcomes and modestly improve language and cognitive skills. However, these benefits are not uniformly observed across all developmental domains. For instance, anthropometric and health indicators, as well as gross motor skills, experience negative impacts from these shocks, albeit these effects are minor in magnitude. These findings highlight the variable impact of positive rainfall shocks on different aspects of early childhood development. The complexity of these environmental influences, such as rainfall, is evident as they can enhance certain developmental areas while having neutral or adverse effects on others. This nuanced response emphasizes the importance of a detailed assessment of developmental outcomes when evaluating the impact of environmental conditions on early childhood development.

4.5.2 Effects of ECED Program on Human Capital

Considering potential non-randomness in the selection of control groups, it's important first to verify baseline comparability between treatment and control groups. As presented in Table (4.3.4), initial developmental outcomes between the two groups are comparable, suggesting similar starting points. However, there is a slight difference in socioeconomic status, with the treatment group showing a somewhat lower wealth index compared to the control group.

Table (4.5.2) presents the findings from the DID analysis, examining the impact of the Early Childhood Education and Development (ECED) program across three distinct groups: (a) all children, (b) children who experienced a positive rainfall shock in the birth year, and (c) children who did not experience a positive rainfall shock in early life. The results in column (a) indicate that the ECED program does not have statistically significant impacts on most developmental measures, with the exception of social competence as detailed in Panel B. However, looking at the 95% confidence intervals, positive and minor impacts of the ECED program are found on height,

		(a) All		(b) Positive Rainfall=1		(c) Positive Rainfall=0		Better with rainfall?
		δ_2	SE	δ_2	SE	δ_2	SE	
Panel A: Anthropometrics and health								
(1)	Weight	-0.209	(0.21)	-0.080	(0.23)	-0.587**	(0.20)	YES
(2)	Height	0.333	(0.30)	0.420	(0.36)	-0.171	(0.53)	YES
(3)	Height-for-age Z-score	0.072	(0.06)	0.097	(0.07)	-0.051	(0.12)	YES
(4)	Weight-for-age Z-score	-0.092	(0.10)	-0.035	(0.09)	-0.246*	(0.11)	YES
(5)	BMI-for-age Z-score	-0.224	(0.17)	-0.167	(0.15)	-0.319	(0.22)	YES
(6)	Weight-for-age Z-score	-0.193	(0.19)	-0.126	(0.17)	-0.317	(0.21)	YES
(7)	Stunting	0.023	(0.03)	0.000	(0.03)	0.071**	(0.03)	YES
(8)	Short	-0.023	(0.03)	-0.033	(0.04)	0.018	(0.04)	YES
(9)	Healthy	0.002	(0.02)	0.010	(0.03)	-0.016	(0.02)	YES
Panel B: Poor Early Development reported by caregivers								
(1)	Any Poor Development	0.013	(0.03)	0.030	(0.04)	-0.019	(0.05)	NO
(2)	Poor Developments > 1	-0.044	(0.03)	-0.076*	(0.04)	0.029	(0.03)	YES
(3)	Physical health	0.012	(0.02)	0.009	(0.03)	0.014	(0.03)	YES
(4)	Social Competence	-0.048*	(0.02)	-0.042	(0.03)	-0.047*	(0.02)	NO
(5)	Emotional Maturity	0.001	(0.02)	0.029	(0.02)	-0.050	(0.06)	NO
(6)	Language and Cognitive Skills	-0.010	(0.04)	-0.037	(0.04)	0.047	(0.07)	YES
(7)	Communication & General Knowledge	-0.026	(0.01)	-0.029	(0.02)	-0.021	(0.03)	YES
Panel C: Strengths and Difficulties Questionnaire(SDQ) rep. by caregivers								
(1)	Total difficulties	-0.010	(0.03)	-0.002	(0.04)	-0.025	(0.05)	NO
(2)	Emotional Symptoms	0.025	(0.02)	0.037	(0.02)	-0.008	(0.03)	NO
(3)	Conduct Problems	-0.044	(0.03)	-0.046	(0.04)	-0.029	(0.05)	YES
(4)	Hyperactivity/Inattention	0.010	(0.01)	0.024	(0.02)	-0.021	(0.03)	NO
(5)	Peer Problems	0.011	(0.03)	0.005	(0.04)	0.029	(0.05)	YES
(6)	Pro-social behavior Difficulty	-0.082	(0.06)	-0.059	(0.05)	-0.094	(0.08)	NO

Table 4.5.2: Effects of ECED Program on Child Development

[a] *** p<.01, ** p<.05, * p<.1
[b] Variables in Panels B and C are binary, where a value of 1 indicates poor development or difficulties in a specific dimension for the child.
[c] Treatment effects are δ_2 estimated by Eq (4.4.2)

social competence, communication and general knowledge, conduct problems, and pro-social behaviors, though statistically insignificant. Conversely, some developmental outcomes, such as weight, emotional symptoms, and hyperactivity/inattention, show slight negative impacts. Other outcomes present ambiguous effects.

These results align partially with those reported in Jung and Hasan, 2016, who evaluated the same program in Indonesia using a similar set of child developmental indicators. When assessing the overall effects of the program across all socioeconomic backgrounds, the impacts are small and generally insignificant for most developmental outcomes.

Several factors could explain the minor and non-significant impacts of the Early Childhood Education and Development (ECED) program observed in this study. The short duration of exposure to the program by the time of the second-round survey in 2010 was about 9 months. This relatively brief period may be insufficient for the full benefits of educational interventions, which often take longer to manifest, to become evident.

Additionally, the ECED project did succeed in increasing enrollment rates from 20% to approximately 28% according to Brinkman et al., 2017, but such an increase might still be too small to significantly alter average treatment effects. The voluntary nature of participation means that not all eligible children were reached, potentially diluting the observable impacts.

The project also allowed considerable flexibility in how villages implemented the ECED services, which could lead to heterogeneous treatment effects across villages and a corresponding close-to-null average effect. Variations in implementation quality, focus, and resource allocation across different villages might result in varied outcomes, complicating the evaluation of the program's overall effectiveness.

Another complicating factor was the substitution between different types of educational services. As reported by Brinkman et al., 2017 and Hasan et al., 2021, the expansion of government-sponsored playgroups resulted in decreased enrollment in kindergartens, which typically provide more frequent educational engagement—5-6 days per week—compared to playgroups that meet only three times a week. Since most kindergartens charge fees while playgroups are often free or cheaper, parents may opt for the less expensive option even if it offers less frequent educational opportunities. This shift might partly explain why some developmental outcomes did not improve as expected.

Lastly, it's possible that the developmental measures used in this analysis did not adequately capture the dimensions most sensitive to the program's impacts. If the chosen indicators do not align closely with the program's influence, the analysis might underestimate its effectiveness.

Overall, these elements suggest that while the ECED program aimed to enhance human capital development among children in rural areas, its effectiveness as measured in this study was limited. The design and implementation of the program, along with the metrics used to evaluate its success, may not have adequately captured its potential benefits.

Next, we conducted separate DID regressions for groups with and without an early life rainfall shock, revealing heterogeneous treatment effects between the two groups. The magnitude and significance of DID estimators varied, indicating differential responses to the ECED program based on early environmental conditions.

In examining the results, children who experienced a positive rainfall shock consistently showed better outcomes across all developmental measures in Panel A and about half of the indicators in Panels B and C. This trend suggests that children with a higher initial endowment due to early-life rainfall shocks also have larger beneficial impacts from later ECED investments. Such findings lend support to the hypothesis of dynamic complementarity between human capital investments at different stages.

This dynamic complementarity implies that early advantages can enhance the effectiveness of subsequent educational interventions, suggesting that initial conditions play a critical role in shaping the impact of later investments. This pattern underscores the importance of considering both the timing and nature of interventions when designing programs aimed at enhancing human capital development.

4.5.3 Interaction between Early-Life Shocks and ECED Program

Having established the relevance of rainfall shocks to early life endowments and identified heterogeneous treatment effects between children who experienced a positive rainfall shock and those who did not, we now explore the interaction effects between these early life rainfall shocks and the ECED program using the DDD method. This approach also allows us to assess whether the introduction of the ECED program has helped to narrow the human capital gap generated by early-life rainfall shocks.

The varying effects of rainfall and the ECED program across different developmental outcomes complicate the interpretation of the results. To accurately interpret these results, it is necessary to first ascertain whether the effects of the rainfall shock and the ECED program on each developmental outcome are beneficial or detrimental. Subsequently, one should evaluate whether these differential effects (represented by γ_7) either exacerbate or mitigate disparities among different groups. Given that the ECED program shows minor and statistically insignificant effects on most outcomes, for simplicity, we assume a beneficial impact of the ECED program on all developmental outcomes. We can then analyze the direction of the differential effects (γ_7) alongside the effects of rainfall. If they are in the same direction, the factors are complementary; if in opposite directions, they act as substitutes.

The DDD estimates reported in Table (4) show only a few significant interaction effects, which is consistent with the generally insignificant effects of the program itself. However, where interaction effects are significant, the results indicate a complex relationship between early-life positive rainfall shocks and subsequent ECED investments. Specifically, for weight, weight-for-age Z score, and stunting, the results suggest that early-life rainfall shocks and later ECED investments act as substitutes, but he magnitude of these effects is relatively minor. In contrast, for the Multiple Poor Developments dummy, they appear to complement each other. This mixed evidence highlights the nuanced interplay between early and later investments in child development.

While most of the DDD estimates are statistically insignificant, examining the confidence intervals reveals some trends in the effects. The results are varied, indicating different relationships between early adversities and subsequent interventions. For most anthropometric and health indicators, except for those related to height in Panel A, the interactions suggest substitution effects between early and later investments, implying that early-life adversities might be mitigated by later educational interventions.

The outcomes for behavioral and cognitive indicators in Panels B and C also present a complex picture. Here, there appears to be a substitution effect for emotional maturity, emotional symptoms, hyperactivity/inattention, and pro-social behaviors. This suggests that these areas might not see as much benefit from the ECED program among children who already experienced favorable early conditions. In contrast, language and cognitive skills exhibit complementary effects, indicating that positive early conditions, when paired with subsequent educational interventions, can enhance these specific areas of development.

The interaction effects for other dimensions of development remain ambiguous, underscoring the complexity of how early environmental factors and educational interventions interplay to influence various aspects of child development. This complexity highlights the need for a nuanced approach to designing and implementing early childhood education programs, taking into account the diverse initial conditions and potential for different types of investment interactions across different areas of development.

The analysis shows that the ECED program has enabled relatively disadvantaged children, as defined by exposure to early-life rainfall shocks, to catch up in certain dimensions such as anthropometric and emotional development, while showing complementary or ambiguous effects for other outcomes. This observation is inconsistent with the findings from Jung and Hasan, 2016, who explored the compensatory effects of the ECED program in Indonesia with a focus on socioeconomic disparities. Their research concluded that the ECED program was beneficial for children from poorer families, improving outcomes across several developmental dimensions such as social competence (which was ambiguous in our study), language and cognitive development (found to be complementary in our study), and executive function. This inconsistency suggests the impact of the ECED program can vary depending on the specific disadvantages or adversities faced by the children it aims to support.

The inconsistencies between these findings and those reported in Jung and Hasan, 2016 can be attributed to the different definitions of disadvantage being considered. While Jung and Hasan focused on children from economically disadvantaged backgrounds, our study considers children affected by early-life environmental shocks. These two types of disadvantage differ fundamentally: one is based on socio-economic status, which is well-documented to correlate with various child development outcomes, and the other is an exogenous environmental factor, which might have less straightforward or consistent impacts on development.

An important takeaway from this analysis is the ECED program's limited ability to mitigate the negative impacts of early-life rainfall shocks. This suggests that while ECED interventions can be effective in addressing some forms of disadvantage, such as those associated with socio-economic status, they may not be as effective in overcoming setbacks attributed to adverse environmental conditions experienced early in life.

		DIDID estimator (γ_7)		Rainfall Effect	Interpretation
Panel A: Anthropometrics and health		Coeff	SE		
(1)	Weight	0.474*	(0.23)	-	substitute
(2)	Height	0.455	(0.57)	+	largely complement
(3)	Height-for-age Z-score	0.119	(0.12)	+	complement
(4)	Weight-for-age Z-score	0.196*	(0.10)	-	substitute
(5)	BMI-for-age Z-score	0.162	(0.18)	-	largely substitute
(6)	Weight-for-height Z-score	0.195	(0.16)	-	substitute
(7)	Stunt	-0.066**	(0.03)	+	substitute
(8)	Short	-0.044	(0.05)	-	largely complement
(9)	Healthy	0.026	(0.04)	-	largely substitute
Panel B: Poor Early Development reported by caregivers					
(1)	Any Poor Development	0.051	(0.06)	-	largely substitute
(2)	Poor Developments > 1	-0.107**	(0.05)	-	complement
(3)	Physical health	-0.006	(0.04)	+	ambiguous
(4)	Social Competence	0.005	(0.02)	+	ambiguous
(5)	Emotional Maturity	0.077	(0.06)	-	substitute
(6)	Language and Cognitive Skills	-0.085	(0.08)	-	complement
(7)	Communication and General Knowledge	-0.008	(0.04)	-	ambiguous
Panel C: Strengths and Difficulties Questionnaire(SDQ) rep. by caregivers					
(1)	Total difficulties	0.017	(0.06)	-	ambiguous
(2)	Emotional Symptoms	0.040	(0.04)	+	substitute
(3)	Conduct Problems	-0.021	(0.07)	-	ambiguous
(4)	Hyperactivity/Inattention	0.043	(0.03)	-	substitute
(5)	Peer Problems	-0.026	(0.06)	-	ambiguous
(6)	Pro-social behavior	0.039	(0.06)	-	largely substitute

Table 4.5.3: Interaction between Early-Life Shocks and ECED Program

[a] *** p<.01, ** p<.05, * p<.1
[b] Variables in Panels B and C are binary, where a value of 1 indicates poor development or difficulties in a specific dimension for the child.
[c] Treatment effects are γ_7 estimated by Eq (4.4.3)
[d] Rainfall effects present the sign of γ_{js} estimated by Eq (4.4.1)

4.6 Conclusion

This paper investigates the effects of early-life rainfall shocks and the Early Childhood Education and Development (ECED) program on children in rural Indonesia, along with the interaction between these two factors. The analysis of rainfall shocks reveals that they impact various dimensions of childhood development, although the effects differ across these dimensions. Anthropometric and health indicators, as well as gross motor skills, generally experience minor negative impacts from these shocks. Conversely, positive rainfall shocks are shown to significantly benefit behavioral outcomes and modestly enhance language and cognitive skills.

The evaluation of the ECED program generally indicates statistically insignificant and minor effects across most developmental dimensions. Nonetheless, there are notable heterogeneous treatment effects between children who experienced early-life rainfall shocks and those who did not, highlighting the interaction between these two factors.

Finally, the study explores the interaction between rainfall shocks and the ECED program intervention, finding mixed or ambiguous evidence regarding their substitutability or complementarity.

The generally insignificant effects of the ECED program across a broad array of developmental outcomes suggest that the program's design and implementation may not adequately address the specific needs of children. The short exposure period of approximately 9 months, noted before the second survey round, is likely insufficient for realizing the full potential benefits of the program. This is a important consideration in the context of early childhood development programs, where the duration and continuity of intervention can significantly influence effectiveness.

The voluntary nature of participation and the modest increase in enrollment rates (from 20% to about 28%) also suggest that the program may not be reaching or engaging the most vulnerable children to a sufficient extent. Furthermore, the variability in how services are delivered across different villages introduces heterogeneity in treatment effects, complicating the assessment of the program's overall impact. For future policy implementation, it is crucial to focus on enhancing both the participation rate and the quality of the services provided to ensure more consistent and impactful outcomes.

The interaction analysis also offers valuable insights for both parents and policymakers regarding optimal strategies for human capital development. It suggests that while early-life adversities in physical and emotional dimensions can potentially be mitigated by subsequent investments, poor development in language and communication skills during early stages may have long-lasting effects that are more challenging to address later. Therefore, it is important for parents to prioritize and monitor the development of these skills during their children's early years. For governments, these findings underscore the need for targeted interventions that address specific developmental delays at the earliest stages, ensuring that children have the foundational skills needed for later success.

We must acknowledge the limitations of our analysis. The interpretation of interaction effects requires caution due to the variable impacts of early-life rainfall shocks and the minor and statistically insignificant effects of the ECED program. To more accurately understand the interaction between investments at different stages, it may be necessary to study not just two exogenous factors, but two investments that clearly benefit children's development. Therefore, future research should focus on interventions with clear and significant benefits to more effectively investigate these interaction effects.

The research enhances our understanding of the medium-term impacts of early-life rainfall shocks, the effects of ECED programs on child development, and the interaction between these factors. The findings offer valuable insights into the optimal investment strategies for maximizing human capital development within certain constraints.

References

Adhvaryu, A., & Nyshadham, A. (2015). Return to treatment in the formal health care sector: Evidence from Tanzania. American Economic Journal: Economic Policy, 7(3), 29–57.

Agüero, J. M. (2014). Long-term effect of climate change on health: Evidence from heat waves in Mexico.

Aguilar, A., & Vicarelli, M. (2022). El niño and children: Medium-term effects of early-life weather shocks on cognitive and health outcomes. World Development, 150, 105690.

Almond, D. (2006). Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 US population. Journal of political Economy, 114(4), 672–712.

Almond, D., Currie, J., & Duque, V. (2018). Childhood circumstances and adult outcomes: Act ii. Journal of Economic Literature, 56(4), 1360–1446.

Almond, D., & Mazumder, B. (2011). Health capital and the prenatal environment: The effect of ramadan observance during pregnancy. American Economic Journal: Applied Economics, 3(4), 56–85.

Almond, D., & Mazumder, B. (2013). Fetal origins and parental responses. Annu. Rev. Econ., 5(1), 37–56.

Atfal, R. (2012). The Indonesia early childhood education and development (ECED) project: Findings and policy recommendations.

Barnett, W. S. (1997). Costs and financing of early child development programs. Early child development. Investing in our children's future, 287–295.

Behrman, J. R., King, E. M., Armecin, G., Duazo, P., Ghuman, S., Gultiano, S., & Lee, N. (2006). Early Childhood Development through an Integrated Program: Evidence from the Philippines (Policy Research Working Paper No. 3922). World Bank. Washington, DC. https://doi.org/10.1596/1813-9450-3922

Berlinski, S., Galiani, S., & Manacorda, M. (2008). Giving children a better start: Preschool attendance and school-age profiles. Journal of public Economics, 92(5-6), 1416–1440.

Bouguen, A., Filmer, D., Macours, K., & Naudeau, S. (2013). Impact evaluation of three types of early childhood development interventions in Cambodia. World Bank Policy Research Working Paper, (6540).

Brinkman, S. A., Hasan, A., Jung, H., Kinnell, A., & Pradhan, M. (2017). The impact of expanding access to early childhood education services in rural Indonesia. Journal of Labor Economics, 35(S1), S305–S335.

Burger, K. (2010). How does early childhood care and education affect cognitive development? an international review of the effects of early interventions for children from different social backgrounds. Early childhood research quarterly, 25(2), 140–165.

Dhuey, E. (2007). Essays on the economics of education. University of California, Santa Barbara.

Driessen, G. (2017). Early childhood education intervention programs in the Netherlands: Still searching for empirical evidence. Education Sciences, 8(1), 3.

Duque, V., Rosales-Rueda, M., Sanchez, F., et al. (2018). How do early-life shocks interact with subsequent human-capital investments? evidence from administrative data. IZA world of labor conference.

Fernald, L. C., Kariger, P., Hidrobo, M., & Gertler, P. J. (2012). Socioeconomic gradients in child development in very young children: Evidence from India, Indonesia, Peru, and Senegal. Proceedings of the National Academy of Sciences, 109(supplement_2), 17273–17280.

Gilraine, M. (2016). School accountability and the dynamics of human capital formation. URL: http://tinyurl. com/Gilraine-JMP.

Grantham-McGregor, S., Cheung, Y. B., Cueto, S., Glewwe, P., Richter, L., & Strupp, B. (2007). Developmental potential in the first 5 years for children in developing countries. The lancet, 369(9555), 60–70.

Gunnsteinsson, S., Adhvaryu, A., Christian, P., Labrique, A., Sugimoto, J., Shamim, A. A., & West Jr, K. P. (2014). Resilience to early life shocks (tech. rep.). Technical report.

Hasan, A., Hyson, M., & Chang, M. C. (2013). Early childhood education and development in poor villages of Indonesia: Strong foundations, later success. World Bank Publications.

Hasan, A., Jung, H., Kinnell, A., Maika, A., Nakajima, N., & Pradhan, M. (2021). Contrasting experiences: Understanding the longer-term impact of improving access to pre-primary education in rural Indonesia. Journal of Research on Educational Effectiveness, 14(1), 28–56.

Heckman, J. J. (2007). The economics, technology, and neuroscience of human capability formation. Proceedings of the national Academy of Sciences, 104(33), 13250–13255.

Johnson, R. C., & Jackson, C. K. (2019). Reducing inequality through dynamic complementarity: Evidence from head start and public school spending. American Economic Journal: Economic Policy, 11(4), 310–349.

Jung, H., & Hasan, A. (2016). The impact of early childhood education on early achievement gaps in Indonesia. Journal of Development Effectiveness, 8(2), 216–233.

Kishore, K., Subbiah, A. R., Sribimawati, T., Diharto, S., Alimoeso, S., Rogers, P., & Setiana, A. (2000). Indonesia country study (tech. rep.). Asian Disaster Preparedness Center.

Korkeala, O. (2012). The effect of weather shocks and risk on schooling and child labour in rural Indonesia (tech. rep.).

Lavy, V., Schlosser, A., & Shany, A. (2016). Out of Africa: Human capital consequences of in utero conditions (tech. rep.). National Bureau of Economic Research.

Lee, C. (2014). Intergenerational health consequences of in utero exposure to maternal stress: Evidence from the 1980 Kwangju uprising. Social Science & Medicine, 119, 284–291.

Leight, J., Glewwe, P., & Park, A. (2015). The impact of early childhood rainfall shocks on the evolution of cognitive and non-cognitive skills. Repository. upenn. edu. https://scholar. google. com/scholar_lookup.

Levine, D. I., & Yang, D. (2006). A note on the impact of local rainfall on rice output in Indonesian districts. Unpublished manuscript.

Lundh, L.-G., Wångby-Lundh, M., & Bjärehed, J. (2008). Self-reported emotional and behavioral problems in swedish 14 to 15-year-old adolescents: A study with the self-report version of the strengths and difficulties questionnaire. Scandinavian journal of psychology, 49(6), 523–532.

Maccini, S., & Yang, D. (2009). Under the weather: Health, schooling, and economic consequences of early-life rainfall. American economic review, 99(3), 1006–1026.

Malamud, O., Pop-Eleches, C., & Urquiola, M. (2016). Interactions between family and school environments: Evidence on dynamic complementarities? (Tech. rep.). National Bureau of Economic Research.

Miller, G., & Urdinola, B. P. (2010). Cyclicality, mortality, and the value of time: The case of coffee price fluctuations and child survival in Colombia. Journal of Political Economy, 118(1), 113–155.

Muris, P., Meesters, C., & Van den Berg, F. (2003). The strengths and difficulties questionnaire (sdq) further evidence for its reliability and validity in a community sample of dutch children and adolescents. European child & adolescent psychiatry, 12, 1–8.

Mwaura, P. A., Sylva, K., & Malmberg, L.-E. (2008). Evaluating the Madrasa preschool programme in East Africa: a quasi-experimental study. International Journal of Early Years Education, 16(3), 237–255.

Naudeau, S., Martinez, S., Premand, P., & Filmer, D. (2011, February). Cognitive development among young children in low-income countries. The World Bank.

Peisner-Feinberg, E. S., Burchinal, M. R., Clifford, R. M., Culkin, M. L., Howes, C., Kagan, S. L., & Yazejian, N. (2001). The relation of preschool child-care quality to children's cognitive and social developmental trajectories through second grade. Child development, 72(5), 1534–1553.

Pradhan, M., Brinkman, S. A., Beatty, A., Maika, A., Satriawan, E., de Ree, J., & Hasan, A. (2013). Evaluating a community-based early childhood education and development program in Indonesia: study protocol for a pragmatic cluster randomized controlled trial with supplementary matched control group. Trials, 14, 1–16.

Raine, A., Liu, J., Venables, P. H., Mednick, S. A., & Dalais, C. (2010). Cohort profile: The mauritius child health project. International Journal of Epidemiology, 39(6), 1441–1451.

Reynolds, A. J., Ou, S.-R., & Topitzes, J. W. (2004). Paths of effects of early childhood intervention on educational attainment and delinquency: A confirmatory analysis of the Chicago Child-Parent Centers. Child development, 75(5), 1299–1328.

Rossin-Slater, M., & Wüst, M. (2015). Are different early investments complements or substitutes? Long-run and intergenerational evidence from Denmark. In Technical report. Mimeo, University of California Santa Barbara.

Schady, N., Behrman, J., Araujo, M. C., Azuero, R., Bernal, R., Bravo, D., Lopez-Boo, F., Macours, K., Marshall, D., Paxson, C., et al. (2015). Wealth gradients in early childhood cognitive development in five Latin American countries. Journal of Human Resources, 50(2), 446–463.

Scholte, R. S., Van Den Berg, G. J., & Lindeboom, M. (2015). Long-run effects of gestation during the Dutch Hunger Winter famine on labor market and hospitalization outcomes. Journal of health economics, 39, 17–30.

Shah, M., & Steinberg, B. M. (2017). Drought of opportunities: Contemporaneous and long-term impacts of rainfall shocks on human capital. Journal of Political Economy, 125(2), 527–561.

Spiess, C. K., Büchel, F., & Wagner, G. G. (2003). Children's school placement in germany: Does kindergarten attendance matter? Early childhood research quarterly, 18(2), 255–270.

Sviatschi, M. M. (2022). Making a narco: Childhood exposure to illegal labor markets and criminal life paths. Econometrica, 90(4), 1835–1878.

Thiede, B. C., & Gray, C. (2020). Climate exposures and child undernutrition: Evidence from Indonesia. Social Science & Medicine, 265, 113298.

UNICEF. (2009). State of the world's children: Celebrating 20 years of the convention on the rights of the child. Unicef.

Yao, S., Zhang, C., Zhu, X., Jing, X., McWhinnie, C. M., & Abela, J. R. (2009). Measuring adolescent psychopathology: psychometric properties of the self-report strengths and difficulties questionnaire in a sample of Chinese adolescents. Journal of Adolescent Health, 45(1), 55–62.

Zelazo, P. D. (2006). The dimensional change card sort (dccs): A method of assessing executive function in children. Nature protocols, 1(1), 297–301.

Chapter 5

Conclusion

In this section, I will consider three studies collectively, drawing lessons from the presented analysis, assessing the strengths and weaknesses of the work and draw conclusions as a whole.

The first paper provides an analysis of how the Clean Heating Policy in Beijing has reshaped household energy behaviors. While the policy successfully reduced coal use and increased adoption of electricity and heat pumps, it did not completely eliminate coal or other polluting fuels like firewood. Further investigation on the determinants of fuel choices and heating behaviors underscores the complex dynamics of fuel transitions, where economic, cultural, and logistical factors play significant roles. The findings suggest that policies aimed at energy transition should consider these multifaceted influences and support households in moving entirely up the energy ladder, beyond mere initial adoption of cleaner technologies.

The second paper shifts focus to the public health implications of the Clean Heating Policy, particularly its impact on sleep in rural Beijing. Although improvements in heating efficiency and reductions in air pollution were hypothesized to enhance sleep quality, the findings reveal that the policy did not significantly alter overall sleep patterns, except in specific demographic subgroups, such as short sleepers and the elderly. Moreover, the environmental changes induced by the policy were not the primary drivers of the observed improvements in sleep. This suggests that while clean energy policies can yield health benefits, their effectiveness is likely influenced by individual baseline characteristics and potentially other unmeasured environmental or behavioral factors. The complexity uncovered in these results underscores the need for further research to elucidate the mechanisms by which such policies influence health outcomes.

The third paper examines how early-life environmental shocks, such as rainfall variability, interact with human capital interventions like the ECED program in Indonesia. It demonstrates that early-life rainfall shocks significantly impact childhood development across various dimensions. However, the ECED program displays limited effectiveness in enhancing childhood development and it does not effectively mitigate the developmental challenges posed by early-life environmental disadvantages. These findings provide crucial insights into formulating optimal investment strategies for parents and governments, and they offer guidance for refining future educational interventions to achieve improved outcomes.

This thesis has several limitations that need to be acknowledged. A primary limitation is the reliance on available data. For instance, heating hours were used as a proxy for fuel consumption, which, although informative, does not perfectly measure fuel use. Additionally, the measurement of sleep employed in the second study focuses mainly on sleep duration, capturing only one dimension of sleep quality. However, sleep quality is a multifaceted construct that encompasses various other important aspects. The time frames of these studies also limit the ability to observe long-term impacts and trends, with a maximum observation period of three years following the implementation of the clean heating policy and only nine months of exposure to the ECED program. Furthermore, the studies are primarily focused on specific regions—rural Beijing and rural Indonesia—potentially limiting the generalizability of the findings to other areas with different socio-economic and environmental characteristics. These limitations underscore the need for cautious interpretation when applying the conclusions of this research to broader settings or different policy scenarios.

The strengths of this thesis are rooted in its methodological rigor, particularly through the use of quasi-experimental designs across the studies. This approach enhances the reliability of the conclusions by minimizing potential biases and enabling more accurate estimations of causal effects. Additionally, the unique data sets utilized, which are not commonly found in other studies, strengthen the analysis. For instance, the detailed recording of energy choices for specific purposes (such as cooking, heating, and boiling water) and the comprehensive tracking of heating hours for various devices in different rooms provide a granular view of household choices and behaviors. Furthermore, the comprehensive inclusion of information on policy treatments, sleep outcomes, and additional variables allows for an in-depth investigation of heterogeneous treatment effects and potential mediators on sleep. These methodological advantages, combined with distinctive data, offer valuable insights into the dynamics of household decision-making and the efficacy of policies. Together, they offer a foundation for future research and policy development in similar developmental and environmental contexts, enhancing our understanding of how targeted interventions can influence household behaviors and outcomes.

Overall, this dissertation highlights the intricate relationship between policy design and human behavior. It demonstrates to policymakers and stakeholders that effective interventions extend beyond mere financial incentives and technological solutions. Instead, they necessitate a deep understanding of local cultural norms, economic circumstances, and individual needs. Policies must be flexible and carefully tailored, aiming to address specific vulnerabilities while capitalizing on local knowledge about policy-related practices.

References

Almond, D. (2006). Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 US population. Journal of political Economy, 114(4), 672–712.

Baglioni, C., Spiegelhalder, K., Nissen, C., & Riemann, D. (2011). Clinical implications of the causal relationship between insomnia and depression: How individually tailored treatment of sleeping difficulties could prevent the onset of depression. Epma Journal, 2, 287–293.

Cappuccio, F. P., D'Elia, L., Strazzullo, P., & Miller, M. A. (2010). Sleep duration and all-cause mortality: A systematic review and meta-analysis of prospective studies. Sleep, 33(5), 585–592.

Cheng, M., Zhi, G., Tang, W., Liu, S., Dang, H., Guo, Z., Du, J., Du, X., Zhang, W., Zhang, Y., et al. (2017). Air pollutant emission from the underestimated households' coal consumption source in China. Science of the Total Environment, 580, 641–650.

Guo, Q. (2002). The Chinese domestic architectural heating system [Kang]: origins, applications and techniques. Architectural History, 45, 32–48.

Heckman, J. J. (2007). The economics, technology, and neuroscience of human capability formation. Proceedings of the national Academy of Sciences, 104(33), 13250–13255.

Lan, L., Tsuzuki, K., Liu, Y., & Lian, Z. (2017). Thermal environment and sleep quality: A review. Energy and Buildings, 149, 101–113.

Lavy, V., Schlosser, A., & Shany, A. (2016). Out of Africa: Human capital consequences of in utero conditions (tech. rep.). National Bureau of Economic Research.

Lawrence, W. R., Yang, M., Zhang, C., Liu, R.-Q., Lin, S., Wang, S.-Q., Liu, Y., Ma, H., Chen, D.-H., Zeng, X.-W., et al. (2018). Association between long-term exposure to air pollution and sleep disorder in Chinese children: the Seven Northeastern Cities study. Sleep, 41(9), zsy122.

Lee, C. (2014). Intergenerational health consequences of in utero exposure to maternal stress: Evidence from the 1980 Kwangju uprising. Social Science & Medicine, 119, 284–291.

Lim, J., & Dinges, D. F. (2010). A meta-analysis of the impact of short-term sleep deprivation on cognitive variables. Psychological bulletin, 136(3), 375.

Liu, J., Mauzerall, D. L., Chen, Q., Zhang, Q., Song, Y., Peng, W., Klimont, Z., Qiu, X., Zhang, S., Hu, M., et al. (2016). Air pollutant emissions from Chinese households: A major and under-appreciated ambient pollution source. Proceedings of the National Academy of Sciences, 113(28), 7756–7761.

Medic, G., Wille, M., & Hemels, M. E. (2017). Short-and long-term health consequences of sleep disruption. Nature and science of sleep, 151–161.

UN. (2016). Sustainable development goal 7. [last access: 10 April 2019]. https : / / sustainabledevelopment.%20un.org/sdg7.

Wang, D., Li, W., Cui, X., Meng, Y., Zhou, M., Xiao, L., Ma, J., Yi, G., & Chen, W. (2016). Sleep duration and risk of coronary heart disease: A systematic review and meta-analysis of prospective cohort studies. International Journal of Cardiology, 219, 231–239.

WHO. (2024). Household air pollution and health [last access: 16 October 2024]. https://www.who.int/news-room/fact-sheets/detail/household-air-pollution-and-health.

Yin, J., Jin, X., Shan, Z., Li, S., Huang, H., Li, P., Peng, X., Peng, Z., Yu, K., Bao, W., et al. (2017). Relationship of sleep duration with all-cause mortality and cardiovascular events: A systematic review and dose-response meta-analysis of prospective cohort studies. Journal of the American Heart Association, 6(9), e005947.