# Huai-River policy and regional economic development: status and prospect

Mr. Yannan Gao

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Economics
Faculty Of Economics
Chulalongkorn University
Academic Year 2023

# นโยบายแม่น้ำฮวยและการพัฒนาเศรษฐกิจภูมิภาค: ปัจจุบันและอนาคต

นายเยี่ยนหนาน เกา

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาเศรษฐศาสตรดุษฎีบัณฑิต
สาขาวิชาเศรษฐศาสตร์
คณะเศรษฐศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2566

| | |
|---|---|
| Thesis Title | Huai-River policy and regional economic development: status and prospect |
| By | Mr. Yannan Gao |
| Field of Study | Economics |
| Thesis Advisor | Assistant Professor SAN SAMPATTAVANIJA, Ph.D. |

Accepted by the FACULTY OF ECONOMICS, Chulalongkorn University in Partial Fulfillment of the Requirement for the Doctor of Philosophy

Dean of the FACULTY OF ECONOMICS
(Associate Professor SITTIDAJ PONGKIJVORASIN, Ph.D.)

DISSERTATION COMMITTEE

Chairman
(Associate Professor Nattapong Puttanapong, Ph.D.)

Thesis Advisor
(Assistant Professor SAN SAMPATTAVANIJA, Ph.D.)

Examiner
(Assistant Professor SAWARAI BOONYAMANOND, Ph.D.)

Examiner
(Associate Professor MANACHAYA URUYOS, Ph.D.)

Examiner
(Associate Professor YONG YOON, Ph.D.)

เยี่ยนหนาน เกา : นโยบายแม่น้ำฮวยและการพัฒนาเศรษฐกิจภูมิภาค: ปัจจุบันและอนาคต . ( Huai-River policy and regional economic development: status and prospect )
อ.ที่ปรึกษาหลัก : ผศ. ดร.สันต์ สัมปัตตะวนิช

งานวิจัยนี้ ศึกษาผลกระทบของนโยบายแม่น้ำห้วย ต่อกิจกรรมของมนุษย์และเศรษฐกิจ อันดับแรก เราศึกษาผลกระทบของนโยบายการทำความร้อนจากส่วนกลาง ที่มีต่อคุณภาพอากาศ ประการที่สอง ศึกษาผลกระทบของนโยบายแม่น้ำห้วย ต่อการอพยพของมนุษย์และการลงทุนโดยตรงจากต่างประเทศ ประการที่สาม เราศึกษาถึงผลกระทบของนโยบายแม่น้ำห้วย ที่มีต่อการพัฒนาเศรษฐกิจ งานวิจัยนี้มีทั้งภาคทฤษฎีและเชิงประจักษ์ โดยเริ่มจากการศึกษาบทบาทของตัวแปรจากแบบจำลองทางทฤษฎี และใช้วิธีการทางสถิติเพื่อตรวจสอบแบบจำลอง และเสนอคำอธิบายใหม่ให้กับสภาพการณ์ที่เราพบ ข้อมูลระดับมณฑลใกล้กับเส้นเขตแดนทำความร้อน และข้อมูลระดับจังหวัดถูกนำมาใช้เพื่อตรวจสอบสมมติฐาน ด้วยวิธีการทางเศรษฐมิติ เช่น การถดถอยไม่ต่อเนื่อง การถดถอยแบบแผง การประมาณค่าตัวแปรเครื่องมือ และวิธีการควบคุมแบบสังเคราะห์

ข้อสรุปของการศึกษาครั้งนี้มีดังนี้ ประการแรก คุณภาพอากาศในพื้นที่ที่มีการทำความร้อนจากส่วนกลาง จะแย่กว่าในพื้นที่ที่ไม่มีการทำความร้อนอย่างมาก และโครงการทำความร้อนที่สะอาด มีบทบาทสำคัญในการปรับปรุงคุณภาพอากาศ ประการที่สอง การกระจายตัวของประชากรในพื้นที่ที่ไม่ทำความร้อนจากส่วนกลาง จะมีการกระจายตัวเป็นรูปตัว U เชิงบวก เมื่อมีระยะห่างจากพื้นที่ที่มีการทำความร้อนเพิ่มขึ้น สำหรับจังหวัดทั้งหมดทางใต้ของเส้นเขตแดน ประการที่สาม พื้นที่ที่มีการทำความร้อนทางตอนเหนือของเส้นเขตทำความร้อน ไม่ดึงดูดการไหลเข้าของประชากรมากขึ้น แต่พื้นที่ที่ไม่มีการทำความร้อนจากส่วนกลางทางตอนใต้ของเส้นแบ่ง สามารถดึงดูดการไหลเข้าของประชากรมากขึ้น เมื่อพิจารณาตามเทศมณฑลที่ตั้งอยู่บริเวณเส้นเขตแดนทำความร้อน ประการที่สี่ พื้นที่ทำความร้อนส่วนกลาง ไม่มีส่วนในการดึงดูดการลงทุนโดยตรงจากต่างประเทศ สุดท้ายนี้ผลิตภัณฑ์มวลรวมภายในประเทศต่อหัว และความสว่างของแสงตอนกลางคืนในพื้นที่ทำความร้อนตามแนวเขตที่มีการทำความร้อน จะสูงกว่าพื้นที่ที่ไม่มีการทำความร้อนอย่างมีนัยสำคัญ การย้ายถิ่นของประชากรไม่ใช่ปัจจัยชี้ขาด ที่ส่งผลต่อความแตกต่างในการพัฒนาเศรษฐกิจระหว่างพื้นที่ ทั้งนี้กิจกรรมทางเศรษฐกิจ และผลิตภัณฑ์ที่เกี่ยวเนื่องกับอุตสาหกรรมทำความร้อนจากส่วนกลาง เป็นปัจจัยนำมาซึ่งส่งเสริมการพัฒนาเศรษฐกิจในระดับภูมิภาค

| | | |
|---|---|---|
| สาขาวิชา | เศรษฐศาสตร์ | ลายมือชื่อนิสิต ............................................... |
| ปีการศึกษา | 2566 | ลายมือชื่อ อ.ที่ปรึกษาหลัก .............................. |

# # 6185905829 : MAJOR ECONOMICS
KEYWORD: Huai River policy, Central Heating, Air Pollution, Population Migration, Foreign Direct Investment, Night Luminosity, Economic Development

Yannan Gao : Huai-River policy and regional economic development: status and prospect . Advisor: Asst. Prof. SAN SAMPATTAVANIJA, Ph.D.

This study explores the impact of the Huai River policy on human and economic activities. First, we study the impact of central heating policy on air quality. Second, the impact of the Huai River policy on human migration and FDI is explored. Third, we further explore the impact of the Huai River policy on economic development. This study has both theoretical and empirical parts. It initially explores the role of variables by establishing a theoretical framework. Then, it applies statistical methods to verify the model and propose new explanations for the phenomenon. We use the county-level data near the heating borderline and data of prefecture-level cities to verify the hypotheses. Econometric methods such as discontinuity regression, panel regression, instrumental variable estimation, and synthetic control method are applied.

The conclusions of this study are as follows. First, the air quality in heating areas is significantly worse than in non-heating areas due to central heating. Clean heating projects have played an important role in improving air quality. Second, the population distribution of the non-heating areas has a positive U-shaped distribution with the increasing distance from the heating area. This result considers all the prefecture-level cities south of the heating borderline. Third, the heating areas north of the heating borderline do not attract more population inflows, but the non-heating areas south of the heating borderline attract more population inflows. This result is based on the counties located around the heating borderline. Fourth, central heating areas do not attract more foreign direct investment. Last, the GDP per capita and night light brightness of heating areas along the heating borderline are significantly higher than those of non-heating areas. Population migration is not the decisive factor affecting the difference in economic development between areas with and without central heating. It is the output value brought by the heating industry itself that promotes regional economic development.

| | | |
|---|---|---|
| Field of Study: | Economics | Student's Signature ............................... |
| Academic Year: | 2023 | Advisor's Signature .............................. |

# ACKNOWLEDGEMENTS

Five years of Ph.D. study flew by in an instant. The time has finally come to write the acknowledgement.

First of all, I would like to thank my supervisor, Assist. Professor San Sampattavanija,Ph.D., who gave me the greatest help in writing my thesis. He influenced me with his professional and rigorous attitude and rich knowledge. His suggestions were unique and professional. His encouragement gave me great strength when I was in difficulties. In addition to his superb professional knowledge and rigorous academic spirit, his kindness and enthusiasm also deeply influenced me and gave me great encouragement. Without his help, I would not have been able to complete my studies.

I would like to thank my doctoral committee, Associate Professor Nattapong Puttanapong, Ph.D., Associate Professor Manachaya Uruyos, Ph.D., Associate Professor Yong Yoon, Ph.D., Assistant Professor Sawarai Boonyamanond,Ph.D., Assistant Professor San Sampattavanija, Ph.D., who gave me a lot of help when drafting the topic, helped me review the topic, and gave me a lot of inspiration. During the final defense, the professors gave me a lot of suggestions for revising my thesis. In the previous progress reports, Associate Professor Dr. Manachaya Uruyos and Associate Professor Dr. Yong Yoon also game me many comments on my paper.

I would like to thank my dearest family members, my adorable children, and my parents who work hard and support me all the time. They have contributed to the achievement of this doctorate.

Thanks to all the professors, classmates and friends I have met in Chula over the past few years. Thanks to Chula for giving me the opportunity to start a wonderful life in Bangkok. No matter where I am in the world in the future, I will remember the days in Bangkok and the times of joys and tears.

Thank you all!

Yannan Gao

# TABLE OF CONTENTS

**Page**

# Chapter 1 Research question, hypothesis, and structure

## 1.1 Introduction

The Huai River policy introduced the central heating borderline between the north and south, demarcated by administrative orders across China in the 1950s [1]. The Huai River policy was first mentioned in the academic context in the study of Almond (2009) [2]. “Central heating” refers to “district heating” in the Chinese context. The central heating borderline coincides with the Qinling Mountains and Huai River line. The Qinling Mountains and Huai River line comprise the Qinling Mountains and Huai River, two well-known natural geographical landscapes in China. It is also the geographical boundary between north and south China in geography, the zero-degree isotherm of January, and the 800 mm isothermal precipitation line [3]. The government has built and implemented central heating systems in the regions north of the heating borderline since the 1950s. In contrast, central heating systems have not been introduced in the areas south of the Huai River. The Huai River policy contemplated China’s economic situation and natural climatic conditions and is still used today.

The introduction of the Huai River Policy is not only related to China's geographical and climate characteristics but also to China's economic situation at that time. This policy is based on natural factors such as geographical climate on the one hand and economic development and human activities on the other. At the same time, the environmental effects of the Huai River policy are also worthy of attention, especially in the context of China's increasing emphasis on environmental protection in recent years. Therefore, studying the Huai River policy’s economic effects is an exciting topic. At the same time, since the Huai River Policy has been issued for more than 70 years, during 70 years, China's economy and society have achieved significant development. However, the original policy may not be able to meet the actual needs of the present. More and more people have called for a revision of the Huai River policy, so assessing the impact of the Huai River policy on China's economy will provide helpful suggestions for future policy revisions [4]. The national “North and South” debate is a hot topic nationwide. In the National People’s Congress (NPC) and the Chinese People’s Political Consultative Conference (CPPCC), the two most important national conferences for policy-making and voting, many proposals for revising the Huai River policy were submitted. They evoked strong responses from citizens[1].

Considering the academic nature, interest, and importance of this topic, we choose to study the economic effects of the Huai River policy.

## 1.2 Research question

As the Huai River policy generates discrepancies between the south and the north regarding winter heating, living comfort, and family expenditures, it may further impact on other economy-related

---

[1] This news is from the website https://www.cma.gov.cn/2011xzt/2021zt/20210225/2021022503/202111/t20211119_4256117.html

terms, including population migration, GDP per capita, and FDI. This study mainly focuses on the policy's impact on economic development. What is the economic impact of the Huai River policy? This question will be fully answered in this study

This question is vital to answer. There are at least the following reasons:

First, the Huai River policy affects air quality and increases carbon emissions. It is vital in China's increasing emphasis on environmental protection and carbon neutrality. Looking at the Huai River policy from an environmental protection perspective provides an important basis for future policy revisions. At the same time, this study also evaluates the effect of clean heating policies implemented in China in recent years.

Second, this study focuses on the "population" effect of the Huai River policy, which is a brand-new angle. The labor force is one of the decisive factors for regional development. China's population is facing the risk of decreasing. The distribution of population determines the distribution of economic activity intensity. Also, it determines the policy tendency, which profoundly impacts the economic development of areas with and without central heating.

Third, the study links the Huai River policy to economic performance. The previous studies only consider the policy as an instrument to explore the policy's pollution effect. They do not analyze the economic impact of the policy. This study considers the policy's economic impact, which would be a good reference for policy revisions. The Huai River policy has been implemented for more than 70 years. In the past ten years, with the development of the economy and society, the original heating policy can no longer meet the actual needs of the existing residents. Analysis of these impacts can provide important references for future revisions of the Huai River policy.

Fourth, the designation of the Huai River policy is based on China's climate status; therefore, this study can be classified as a category of research on the impact of geography and energy policies on human economic activities, filling the research gap in this field. This is significant when the global climate is changing rapidly and energy is in short supply.

## 1.3 Hypothesis

The hypothesis of this study can be summarized as follows:

**Hypothesis 1**: Central heating areas have a higher degree of air pollution.

**Hypothesis 2**: Central heating could influence the population flow in three aspects: heating costs, air pollution, and room comfort. Households' final decisions depend on how they value these three factors. Areas near the heating borderline may be strongly motivated to flow into the heating areas.

**Hypothesis 3**: Central heating areas could have more intensified human and economic activities, thus have better economic performance.

## 1.4 Structure

The basic structure of the study refers to Figure 1.4-1. There are two types of impacts of the Huai River policy on economic development. One is what we call the aggregate effect (Impact I).

Investment in the heating industry generates GDP growth. The heating industry itself is a component of the economy. The heating industry chain has a long stretch from the upstream to the downstream industries, as described in Figure 2.5-1. Central heating industries contribute to the economic growth of heating regions. The individual impact (Impact II) refers to population and capital flow changes, corresponding to population migration and foreign direct investment. Cai and Wang (2008) explored the relationship between internal migration and economic development in China [5]. They pointed out that migration was critical for promoting urbanization and economic development.

The first half of the study focuses more on Impact II. This could generate differences in air pollution and living comfort and affect human and economic activities. We use population migration and foreign direct investment (FDI) as two proxy variables for human activities and economic activities. The second half of the study focuses on Impact I. We want to determine the policy's impact on economic growth and development. The results could help us work out more policy implications in the future.

The study overall comprises the following sub-studies:

**Section 1: Empirical studies in air pollution and the Huai River policy**

This section focuses on the relationship between central heating and air pollution. We also take the southern areas as the control group to justify the policy's impact winter heating. Meanwhile, the clean heating project evaluation will be added as a make-up for further illustration of the relationship. If the clean heating project is introduced and the air pollution could be abated, this could further explain the relationship between air pollution and the policy. In this part, we use the RD (regression discontinuity) design and synthetic control method to explore the causal relations. This study will be presented in Chapter 3.

**Section 2: The Huai River policy and population migration.**

This section links the Huai River policy with migration in three dimensions: air pollution, living comfort, and heating costs. It is a crucial part of this research. We will apply the panel data estimation method and Hausman-Taylor estimators to verify the Huai River policy's effects on migration. We first consider a Harris-Todaro-based model to include the critical variables, including air pollution, environmental tax, migration, and labor input. The model simulates the evolution of relevant variables. The model will be introduced in Section 4.1. Then we conduct two empirical studies. One is to verify the U-shaped relationship south of the heating borderline using prefecture-level data, which will be introduced in Section 4.2. The other is to explain the migration status between the heating borderline using county-level cities, which will be introduced in Section 4.3.

**Section 3: The Huai River policy and its overall economic impact**

In this section, the policy's impact on FDI is first discussed. Population migration shows households' moving and settling preferences, while FDI reveals the preferences of foreign investors. Then, we will assess policy's economic impact using econometric techniques. GDP per capita and nighttime luminosity will be taken into account. The final finding will help answer the final research question of the whole study. This section will be presented in chapter 5.

After completing all these sections, the conclusion will be drawn, and policy implications will be discussed in the final chapter 6.

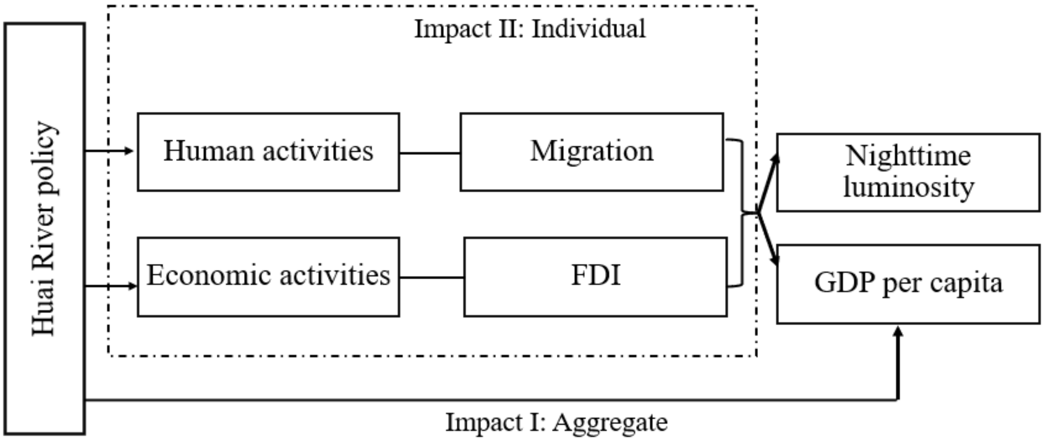

**Figure 1.4-1 Structure**

# Chapter 2 Huai River policy

## 2.1 The definition of central heating

Central heating, or district heating, is a heating system operated by a centralized heating plant with insulated pipelines buried underground extending to households. It can be serviced for both commercial and civilian users. Fossil fuels, biomass or other types of energy can be applied in central heating systems. Central heating is different from scattered heating. Scattered heating is run by residents or companies themselves. Central heating has a centralized heating plant. Users can purchase services from heating plants. The term central heating is frequently applied in China. District heating is a similar concept used mainly in Europe and North America.

Central heating has been put into use for almost 150 years. Annapolis introduced its first central heating systems in 1853 [6]. Nowadays, the EU countries, the U.S.A, and Japan have different sizes of central heating systems operated within their countries to provide heating services for their commercial and residential users [7].

## 2.2 Central heating borderline in China

China’s territory stretches from the subarctic region, reaching 55.00°N latitude and far south, reaching 4.00°S latitude. Due to its proximity to the Siberian region and its location on the biggest continent in the world, it has a rather continental climate with bitterly cold areas in the north. The coldest city in China is Genhe, Inner Mongolia, with a January average temperature of -26°C[2]. Even though this area is unbearably cold, approximately 130,000 residents live there. Thus, good heating condition is essential for people living in the north to suffer the long and cold winters. Unlike countries with relatively evenly distributed residents, China is more suitable for central heating as people are distributed mainly in the east and big cities.

For this reason, the Chinese government planned to build these heating systems a long time ago. The start of the construction of this system nationwide was in the 1950s, getting the experience of the former Soviet Union. At the time when the new country was just founded, the fiscal budgets needed to be more sufficient. Thus, the country had to choose cities that urgently needed central heating services. The criterion for setting such a borderline is the average winter temperature, especially the January temperature for Chinese cities. If the average temperature exceeds 0°C in January, the city will not have a central winter heating system. Only for the cities with January average temperature lower than 0°C can they have central heating systems. The 0-degree isotherm of January is roughly the borderline between the areas with and without central heating systems. The 0-degree isotherm line coincides with two critical geographical items, Huai-River and Qinling-Mountains. Qinling Mountains are groups of mountains extending 1600km long, with its highest point 3771.2m in Baoji, Shaanxi Province, a central province of China. Qinling Mountains

---

[2] The climate data for all cities in this section is derived from the statistical yearbooks in prefecture-level.

have played a vital role in Chinese geography as it is a geographical boundary of south and north and a cultural division of south and north. Languages, cultures, and customs are considerably distinctive through this line. Of course, we could regard these differences as significant distinctions in natural conditions, like climate and terrain. Let us look at an example: two cities in Shaanxi province, Ankang and Tongchuan. Ankang is located slightly south of the Qinling Mountains and has an annual precipitation of around 860.7mm (2018). Tongchuan is located slightly north of the Qinling Mountains and has an annual precipitation of 550.8mm (2018). The direct distance between Ankang and Tongchuan is only 240km. This also leads to a big difference in lifestyle and customs. For example, the staple food of people in Ankang is rice, while the staple food of Tongchuan is wheat because of the big gap in the precipitation. The mean temperature of Tongchuan in January is -4.2°C (2018), meaning that people can witness snowfall every winter. On the contrary, Ankang enjoys much warmer January weather, with the mean temperature reaching 2.8°C (2018), and people seldom witness heavy snow in Ankang. Meteorologists often refer to Ankang's climate as a subtropical climate.

Another critical component of this boundary is the Huai River. Huai River is located in the east with a length of ca.1000km. It was historically considered a disastrous river as it was often flooding. The Yellow River, the mother river of the Chinese people, was incorporated into the Huai River to flow into the sea, flooding the Huai River frequently. Huai River brings lots of silt from the Yellow River to form flat plains downstream. Huai River has become a natural boundary to split the south and north. These two geographical items run east-west parallel to the latitude line, forming the most important boundary of the Chinese Qinling-Mountains and Huai River boundary, which coincides with 0-degree-isotherm and 800mm isohyet.

The borderline has defined a clear boundary for winter central heating. From Figure 2.3-1, only 19 provinces (including autonomous regions) in China have central heating. There are 12 provinces which do not have this setting.

The areas south of the borderline still need heating, although their heating periods are short. Air conditioning systems or electric radiators can help. Those areas in the extreme south with milder temperatures all year round do not need heating systems, like Shenzhen, Sanya, and Haikou, whose winter average temperatures are higher than 15°C. There is an intense debate between the Southerners and Northerners in China about the necessity of central heating systems in the South. Residents of the south are also appealing to central heating because of humid and chilly weather conditions in winter, leading to uncomfortableness in south China. Guo et al. (2015) have discussed this problem [4]. This can lead to economic effects as people might take central heating as an essential factor for settling after they graduate from university.

Considering the heating costs, the unit cost of central heating is lower than that of distributed heating. However, in some areas with relatively warm winters, although the weather is sometimes very cold, the cold period is not long and the temperature is not that low. Residents can use their own facilities to heat themselves, e.g., air conditioners and electric heaters. The total costs with heating tools in some warmer places are lower. Central heating can generate more heating expenses in total, although its unit cost is lower than scattered heating. This is also an important reason why central heating has not been introduced in the southern region for a long time. Even in areas close to the heating boundary, winter temperatures are already much milder than in central heating areas. Of course, different people have different views on this issue when it comes to indoor environmental comfort.

Even in North China, central heating systems are not fully covered in rural areas. The urbanized areas are equipped with central heating systems. In many rural areas in China, households still use decentralized heating methods, like burning coal in stoves or using natural gas stoves. Central heating systems exist mainly in cities in the north. These cities include big municipalities, prefecture-level cities, and county-level cities. Whether a district has a central heating system is deemed a good criterion for judging whether a district is urbanized. The average urbanization rate for 2022 in China is around 63.9%[3]. Liaoning province has an urbanization rate of 72.14%, Inner Mongolia 69%, and Heilongjiang 66.2%. About two-thirds of the population in North China are living in urban areas and have central heating services.

The borderline between the north and the south is the Qinling Mountains and the Huai River, shown in Figure 2.2-1. This figure explains why the country set its heating borderline to the Huai River and Qinling Mountains. Although the heating borderline is roughly divided by the Qinling Mountains and Huai River, the specific heating borderline needs to be confirmed according to the actual situation of each place since the local regulations determine its real location.

---

[3] Urbanization data are retrieved from national statistical yearbooks.

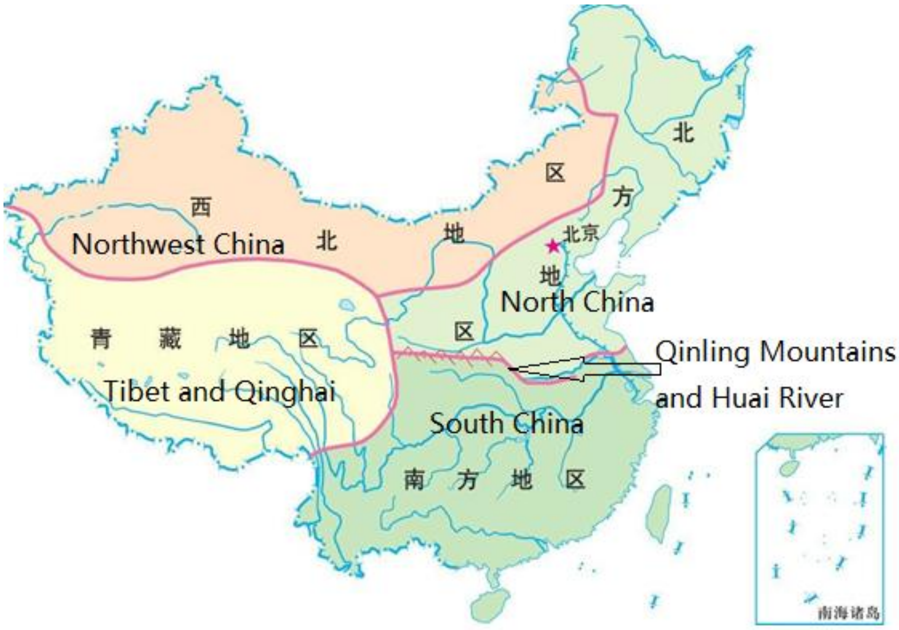

**Figure 2.2-1 China’s regional division and heating borderline.**

## 2.3 Central heating outside China

Europe has been equipped with central heating systems since the 14th century, one representative of which was the Chaudes-Aigues thermal station in France. North America’s first central heating system was in Annapolis in 1853, constructed by the US Naval Academy. The first commercial central heating system was built in New York in 1853. Canada started its first central heating system in Winnipeg in 1924 [8].

With the development of these central heating systems, many countries have been trying to build new ones. Northern European countries have taken the lead as they have long and cold winters, and their social welfare is excellent. Sweden has explored sparse district heating programs trying to cover district heating in some remote areas in the country. Denmark and Finland are considered the two top countries in the EU with advanced central heating systems. Eastern European countries once had central heating systems because of their former centralized economic systems. Examples are Bosnia and Herzegovina, Croatia, Kazakhstan, Kosovo, Macedonia, Poland, and Slovenia. South Korea developed its central heating system in 1985. The country has a constant increase in district heating. In 2013, 15% of citizens were served by district heating [7].

Figures 2.3-1 and 2.3-2 depict the world’s central heating overview. The data are retrieved from Euroheat.org. Nuorkivi (2015) [7] analyzed the current central heating situation worldwide and classified these countries into four categories: emerging, expanding, consolidating, and refurbishing. “Emerging” denotes countries with the start of central heating systems. These countries do not have a national policy or plan for developing central heating. The emerging group includes Canada, the USA, and the UK. China, South Korea, Germany and Italy belong to the "expanding" group. Countries in northern Europe like Finland, Denmark, and Sweden belong to the consolidating group, meaning that these countries' central heating systems have evolved to a mature stage. The refurbishing group denotes countries with relatively old-fashioned central heating systems that have been trying to refurbish them in recent years. These countries are mainly Eastern European countries.

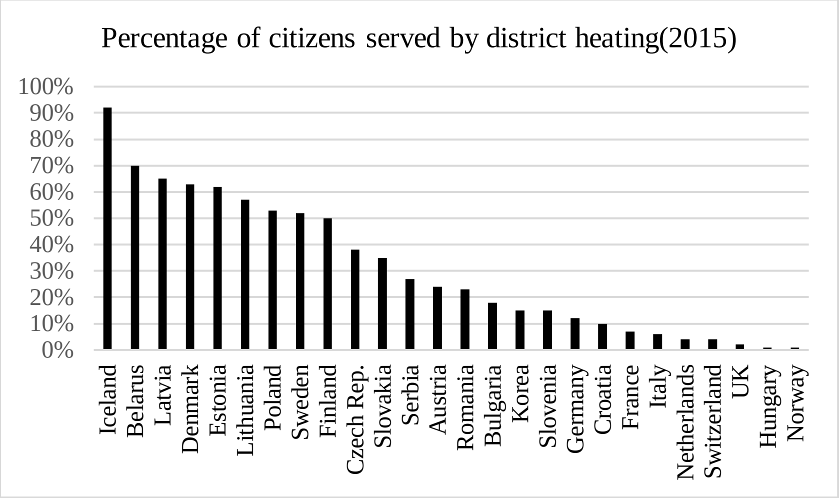

**Figure 2.3-1 Percentage of citizens served by district heating in different countries (2015)**

Source: Euroheat

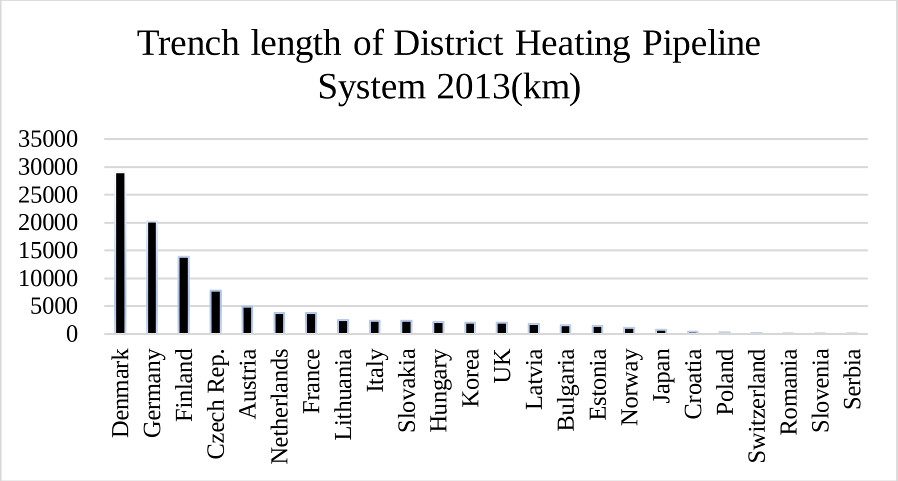

**Figure 2.3-2 Trench length of District Heating Pipeline system (km, 2013)**

Source: Euroheat

## 2.4 Huai River policy

As mentioned above, China has a clear heating and non-heating boundary. The areas slightly south of the Huai River are relatively cold, with winter temperatures often dropping below zero. In those cities, people must use heating facilities like air conditioning systems, electric heaters, and natural gas boilers. The difference is that these facilities are installed spontaneously instead of being run by the heating plants. This is scattered heating. For the cities north of Qinling Mountains and Huai

River boundary, the heating plants run the thermal companies with huge heating facilities built in a particular area. The heating pipelines are laid underground, which are large projects. That is why the southern cities cannot have this policy immediately. It takes years of construction to finish this massive project. Large heating plants run the heating industry in the north and can be partially administered and subsidized by the local governments. For the southern areas, heating is self-financed and, thus, fully market-oriented. Most cities have far shorter heating periods, and less or no expenditure has been spent because people want to save money. For instance, for areas with shorter heating periods, older people may go to the big shopping malls where extensive air conditioning heating systems are used during the coldest period of the year. So many of these families do not have heating systems in their apartments. In this paper, we call the policies centered on the central heating systems in North China the central heating policy or the Huai River policy. We conclude the Huai River policy as follows:

**(1) Borderline policy.**

Central heating in China is divided through the boundary of Qin-Mountain and Huai-River.

For areas in the north, they have this system. For areas in the south, they do not have. The discrepancy in heating use is a crucial part of the Huai River policy. Due to the heating borderline, possible differences in economic terms have been generated. This research will focus on these differences and determine their effects on China's south-north regional economic development.

**(2) City-managed heating systems.** The central heating is provided in terms of prefecture-level cities or even counties. Each city has its own and separate central heating systems. The analysis could be based either on prefecture-level cities or counties. In this study, a survey will be performed to ensure the exact location of the heating borderline.

The central government in Beijing gives legislative powers to the local cities to enact their laws and regulations for winter heating. Thus, cities will have slightly different contents for this policy and different heating periods according to the different climate conditions. Cities have different starting and ending dates for central winter heating. Cities use different types of heating facilities and fossil fuels. For example, cities nearer to Beijing use natural gas as their primary energy for heating, while cities in Shandong still use coal as their primary heating energy. The pricing of these heating services differs from city to city. However, they do not have big differences. Although the heating borderline is set according to the Qinling-Mountains and Huai River line, its concrete stretch is determined according to the regulations of cities near the borderline.

**(3) Future development.** The Chinese government is encouraging the construction of central heating systems in smaller cities and some southern cities with chilly winters. Thus, there is a development in heating areas and pipeline length. Interestingly, the southern provinces are also equipping some high-end house estates with central heating systems. However, these heating systems are very limited. There appears a clear boundary between areas with and without central heating.

**(4) Clean heating project.** The Chinese government is trying to replace the old energy with cleaner energy year by year and has put forward a series of documents and regulations.

As discussed, “Planning for Clean Heating in Winter in Northern Areas for 2017-2021”[4] was enacted in 2017. As planned, the percentage of clean heating will reach 50%, replacing the scattered coal burning by 74 million tons. This percentage will reach 70% by 2021 and replace the scattered coal burning by 150 million tons. The average comprehensive energy consumption should be reduced to 14 kg of standard coal per square meter. The schedule points out that the coal heating area has accounted for 83% of the total heating area in China. Natural gas, electricity, geothermal energy, biomass energy, and industrial waste heating account for the rest. In 2021, renewable resources should account for a large proportion of heating in China. Ground heat source pumps and air heat source pumps are encouraged. China is trying to bring down the scattered burning of coal in rural areas and introduce more central heating systems in cities or rural areas. The analysis below shows that the difference between the heating and non-heating seasons has been smaller thanks to these policies of reducing air pollution in winter heating.

**Table 2.4-1 Heating schemes since 2019**

| Date | Department | Document | Contents |
|---|---|---|---|
| **April 2019** | National Development and Reform Commission | Notice on improving the wind power trading mechanism related to wind power heating and expanding wind power heating | To carry out pilot projects for wind power clean heating |
| **April 2020** | National Development and Reform Commission | National three-year action plan for special rectification of production safety | To boost coal-to-gas campaign |
| **July 2020** | Ministry of Housing and Urban-Rural Development | Green building creation action plan | To boost clean energy central heating in North China |
| **April 2021** | National Development and Reform Commission | Energy work guidance in 2021 | To boost clean energy central heating in North China |
| **June 2021** | Ministry of Industry and Information Technology | 5g application implementation plan in the energy field | Intelligent heating systems using 5g technology |

---

[4] The document can be found on the official website of the Chinese central government.

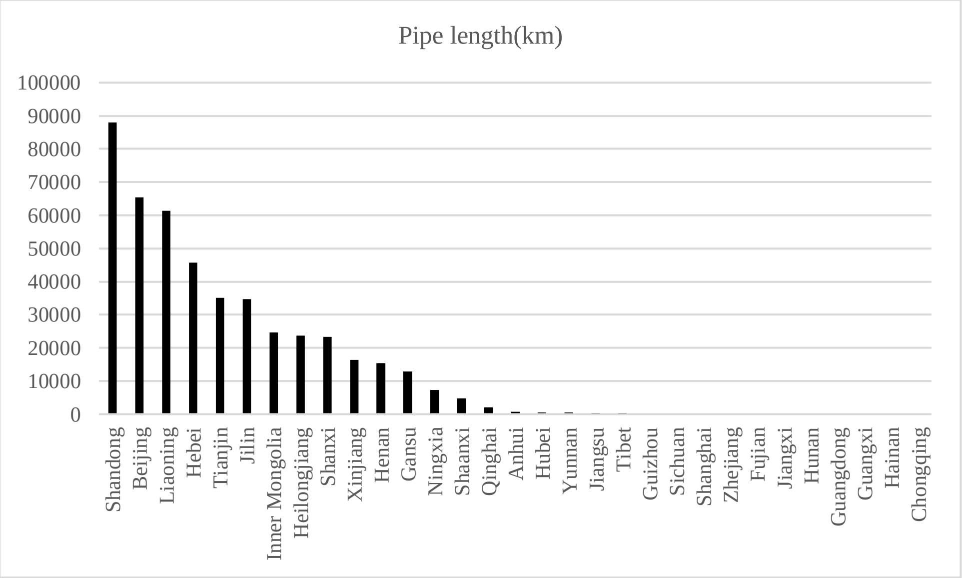

**Figure 2.4-1 Pipeline length of provinces in China (2021, km)**

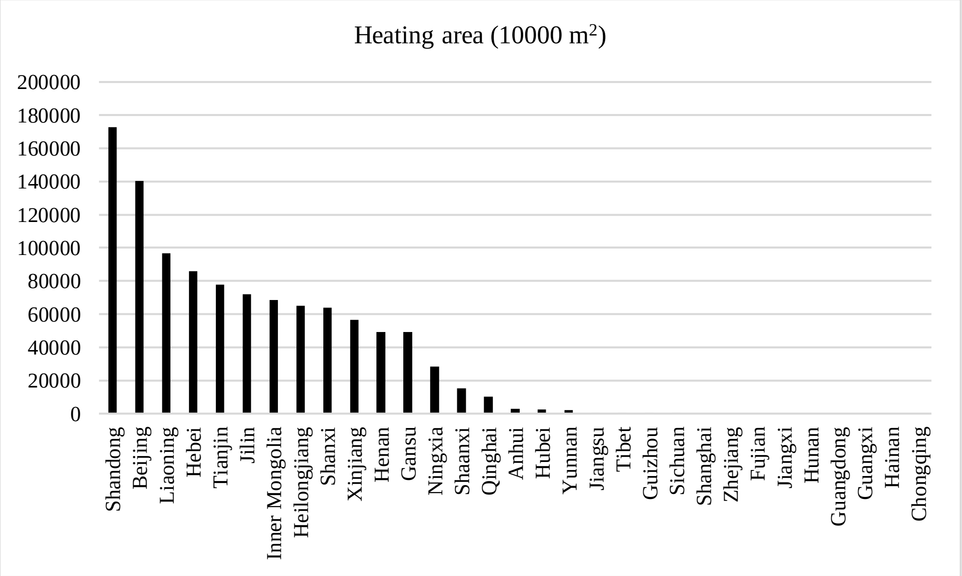

**Figure 2.4-2 Central heating area of provinces in China ($10^4$ m$^2$, 2021)**

Retrieved from National Statistical Yearbook 2022.

## 2.5 Heating industry in China

With urbanization and economic development, the central heating industry has been booming in Chinese cities. According to the report "A quick overview of China's urban heating industry in 2022[5]", in 2021, the steam-heating capacity is 118,000 t/hour, 15300 t/hour higher than that in 2020. The hot water supply capacity is 59.32 MW, 2.7MW higher than 2020. The total amount of steam heating in 2021 is 68164 GJ, 4.78% higher than last year. The heating area in 2021 is 7.3% larger than that in 2020, reaching 10.603 billion m$^2$. The scale of underground heating pipelines

---

5 The report can be viewed on the website https://www.chyxx.com/industry/1137265.html

continues to increase.

In 2021, the investment amount in central heating was 39.73 billion yuan, an increase of 0.89% over the previous year. The number of industrial enterprises above the designated size in China's electricity and heat production and supply industry was 10,372, an increase of 5.55% over the previous year.

In recent years, the country has worked out many policies to encourage the development of central heating. These policies are mainly about clean energy projects. Green, low-carbon, and intelligent central heating are the future development trends of heating.

The heating industry in China can be divided into three parts: upstream, midstream, and downstream. The upstream mainly refers to the energy market. Fossil fuels and new types of energy are introduced as main energy sources. Midstream includes the thermal power plants and pipelines. Most of the enterprises in upstream and midstream are state-owned. The downstream market is mainly privatized. They include producing heating terminal equipment using steel, plastics, and valves, family central heating control systems using electronic parts, the home furnishing products, and so on. The downstream industry is related to a series of industries, such as steel manufacturing, intelligent equipment manufacturing, and home decoration. Its development will drive the development of a series of industries.

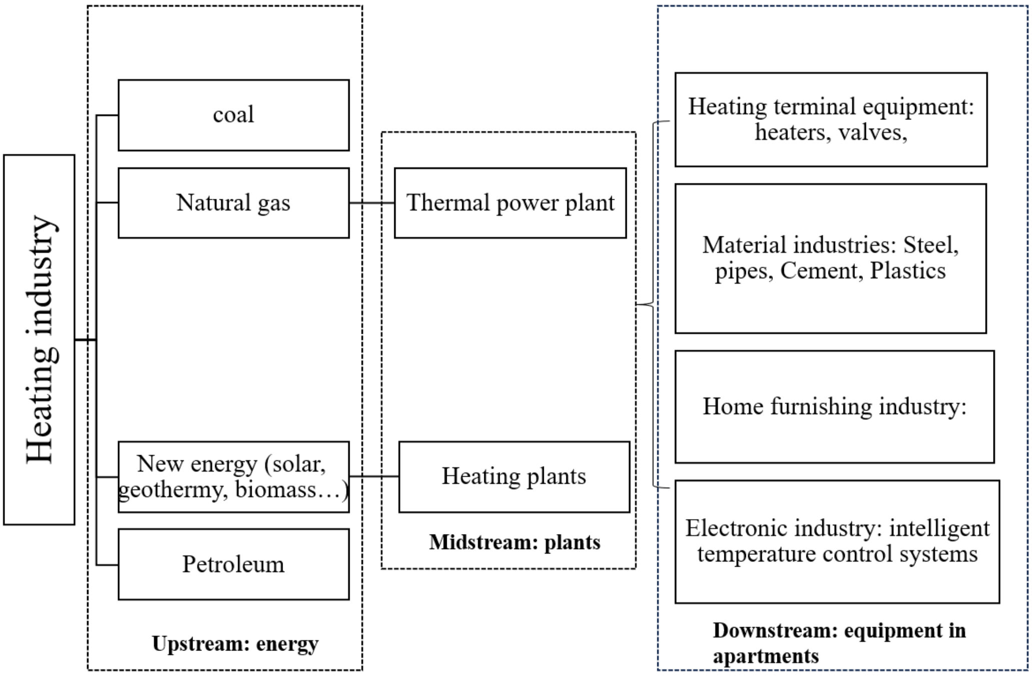

**Figure 2.5-1 Heating industry in China**

# Chapter 3 Huai River policy and air pollution

The hypothesis 1 predicts the positive relatonship between the Huai River policy and air pollution.

In this section, we conduct a series of estimations and robustness checks to further figure out the relationship between the policy and air pollution. In general, the central heating's impact on air pollution is verified, and heterogeneity among different areas are also discussed. This chapter has been published in [9].

## 3.1 Literature review

A large group of literature focuses on the linkage between winter heating and air pollution. Outside China, though there is no specific policy as Huai River policy in China, it is still common to find papers discussing winter heating and air pollution. Literature is mainly divided into two types to study the relationship between heating and air pollution using China as a case study. One is to compare the air pollution in heating areas before and after heating periods. This category usually does not emphasize the Huai River policy and the heating dividing line. The other is taking Huai River policy as the natural discontinuity and compare air pollution levels between the south and north, which requires in-depth research on the heating borderline and the Huai River policy.

We first look at the group of papers discussing pollution-related winter heating effects outside China. This group focuses mainly on case studies concerning some areas in a certain country. Most cases are from air polluting areas.

Cichowicz et al. (2017) [10] discovered that cold seasons have higher air pollution in Wielkopolska, Poland. Wang et al. (2017) [11] analyzed the levels of PM2.5 in Ulaanbaatar in the 2010 heating season. They witnessed an apparent increase in air pollutants during heating season. Guttikunda (2008) [12] also attributed heating as a main factor affecting the Ulaanbaatar's air quality. Angius et al. (1995) [13] stated that building heating emission was one of the prime sources of pollution in Milan, Italy. Ghafghazi et al. (2011) [14] took Vancouver as an example to explore emissions from wood biomass in district heating in Vancouver, Canada. They also compared the different wood fuels. They found that low-quality wood fuel produces higher PM2.5 than wood pellets or dry sawdust. Leaderer et al. (1995) [15] considered a case study in Virginia, America, to analyze the effects of Kerosene Space Heaters on indoor and outdoor air quality levels. All the studies clearly showed that winter heating could increase air pollutants in winter periods.

The following papers focus on the studies about China.

Liang et al. (2015) [16] did a case study in Beijing. They concluded that heating generally contributes to half of the increase in PM2.5 in Beijing. Apart from the winter heating, it also discussed the APEC as an event for air pollution control. Zhang et al. (2020) [17] used the example of Shijiazhuang and analyzed its air situation during heating periods. Xiao et al. (2015) [18] employed an engineering method to explore heating impacts on air pollution in North China.

They showed an apparent rise in air pollutants in winter. Winter heating increases particle pollution in North China.

Some papers do not only compare heating periods and non-heating periods. They also take the heating borderline as a quasi-natural environment and analyze the policy’s impact on air pollution. Almond et al. (2009) [2] first proposed the Huai River policy to specify the policies related to China’s central heating. This paper is considered the start of the research on China’s central heating borderline. The study used a cross-sectional regression discontinuity design method, took the latitude as the running variable, and introduced the polynomial of the variable. The result showed that the policy increased higher TSP levels in the north.

All literature shows that the Huai River policy generates significant differences in air quality levels between the heating borderline. In this section, we will examine the effects once again and consider the region heterogeneity in this problem.

## 3.2 Data

Cities located in different locations have different climatic conditions, thus they have different heating periods. We use 41 cities in central heating areas in China. We choose stations in urban areas. Appendix I lists the heating start and ending days of the cities of this dataset. The dataset includes the meteorological and air pollutants data of chosen cities.

The winter average temperatures of chosen cities are calculated according to the meteorological dataset. We group the heating cities in Table 3.2-1 and Table 3.2-2.

**Table 3.2-1 Grouping of heating cities in terms of average winter temperatures in 2016**

| Groups | Cities |
|---|---|
| Group I: <-10°C | Harbin, Karamay, Qiqihar, Mudanjiang |
| Group II:-10°C~--5°C | Baotou, Hohhot, Changchun, Chengde, Datong, Benxi, Shenyang, Urumqi |
| Group III:-5°C~0°C | Zhangjiakou, Langfang, Baoding, Yinchuan, Yingkou, Anshan, Dandong, Yan’an, Korla, Xining, Qinhuangdao, Tangshan, Taiyuan |
| Group IV: 0°C~5°C | Qingdao, Weifang, Yantai, Rizhao, Zhangqiu, Beijing, Kaifeng, Zhengzhou, Sanmenxia, Linfen, Xuzhou, Anyang, Xingtai Tianjin, Jinan, Taian |

**Table 3.2-2 Grouping of heating cities in terms of average winter temperatures in 2017**

| Groups | Cities |
|---|---|
| Group I:<10°C | Qiqihar, Mudanjiang, Urumqi, Harbin, Karamay, Changchun, |
| Group II:-10°C~--5°C | Baotou, Hohhot, Chengde, Datong, Yingkou, Shenyang, Xining, Zhangjiakou, Benxi, Korla, Dandong, |
| Group III:-5°C~0°C | Yantai, Beijing, Tangshan, Baoding, Yinchuan, Anshan, Yan’an, Tianjin, Langfang, Taiyuan, Qinhuangdao |
| Group IV: 0°C~5°C | Qingdao, Weifang, Taian, Kaifeng, Zhengzhou, Sanmenxia,Xingtai, Linfen, Xuzhou, Jinan, Rizhao, Anyang, Zhangqiu |

Dataset of air pollutants are reported by the National Urban Air Quality Real-time Publishing Platform. It is a platform from the Ministry of Ecology and Environment. This dataset contains hourly data of ground-level pollutants concentration covering approximately 2000 stations. We use the hourly data to calculate the daily data. Air quality monitoring stations are matched with meteorological stations. PM2.5, PM10, $NO_2$, $SO_2$, O3, CO, and an overall index-Air Quality Index (AQI) are included.

We introduce a group of cities in the south as control group for robustness checks. 37 cities are included and listed in Figure 3.2-2.

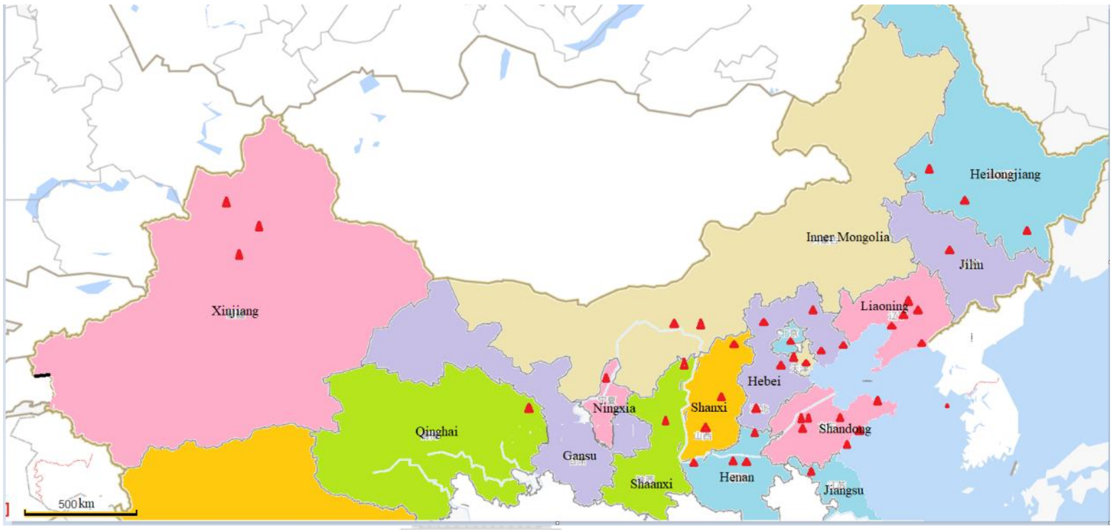

**Figure 3.2-1 Chosen cities in benchmark estimation**

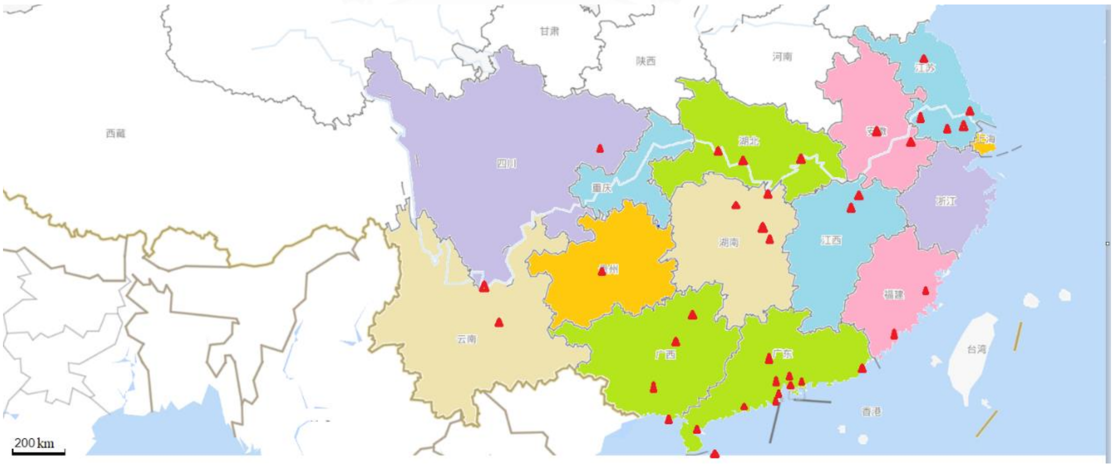

**Figure 3.2-2 Chosen cities in non-heating areas**

This section uses the meteorological data from the National Meteorological Information Center, including maximum humidity, minimum humidity, average temperature, minimum temperature, daily sunshine hours, average wind speed, wind direction at maximum wind speed, precipitation, maximum wind speed, and maximum temperature.

Both meteorological data and air pollutants data are daily panel data with 41 heating cities and 37 non-heating cities. They are divided into four years, 2015 to 2018.

### 3.3 Graphical analysis

We predict that the central heating will be a discontinuity generating differences in air pollution levels. In this section, graphs will be presented to illustrate these differences and discontinuities. This section will take Harbin as an example. The vertical line denotes the start day of central heating. Figure 3.3-1 illustrates this discontinuity intuitively.

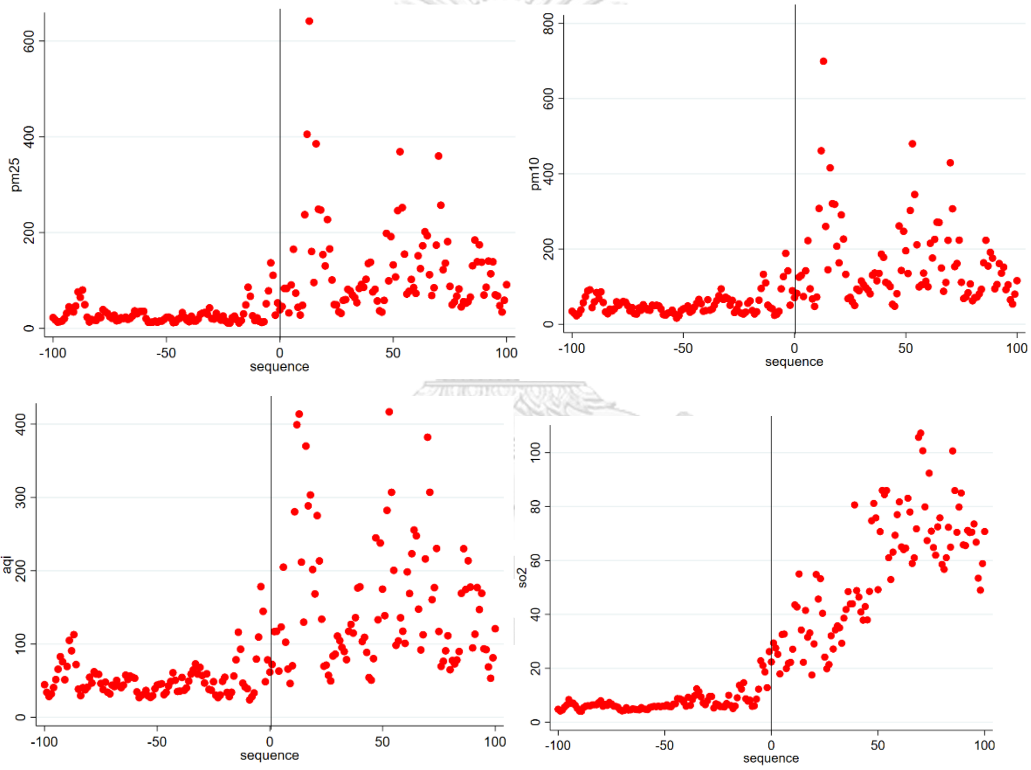

**Figure 3.3-1 Plot of average pollutants' level and days using real data**

### 3.4 Benchmark estimation.

A Regression Discontinuity (RD) design is used to verify the linkage between the policy and air pollution. We refer to a study from Chen and Whalley (2012) [19], which discusses the infrastructure impact on air pollution levels.

#### 3.4.1 Panel data estimation

The methodology is as follows:

$$\mathrm{lnAQI}_{\mathrm{it}} = \alpha_{\mathrm{i}} + \beta_1 \mathrm{Heating}_{\mathrm{it}} + \gamma \mathrm{X}_{\mathrm{it}} + \mathrm{f(t)} + \mathrm{u}_{\mathrm{it}}$$

AQI denotes all air pollutants like PM2.5, PM10, CO and so on.

Heating is an indicator variable of showing whether a place is equipped with central heating systems or not. $X_{it}$ is control variable matrix, including meteorological variables: average daily wind speed(m/s), minimum relative humidity, wind direction at the maximum wind speed, maximum daily temperature, maximum relative humidity, maximum daily wind speed(m/s), minimum daily temperature, daily precipitation (mm), average temperature (°C), daily sunshine hours (h), day of week, day of month. Considering that the control variables and pollutant level might have non-linear relations, we include the square terms of all meteorological variables to capture these characteristics.

f(t) is a third-order time polynomial. $\mathrm{lnAQI}_{it}$ represents log value of air pollutants' of PM2.5, PM10, AQI, $SO_2$, $NO_2$, CO, and $O_3$.

The estimation is conducted in terms of different years. Table 3.4.1-1 records all the estimation results of Heating's coefficients. All the coefficients are statistically significant and positive except ozone.

We can interpret the coefficients like this. The average growth of air pollutants levels for $PM_{2.5}$, and $PM_{10}$, $SO_2$, and CO are 16.68%, 8.72%, 44.96%, and 14.95% in 2015. The average growth rate is 17.62%.the growth rate for 2016 is 18.85%. 2017 and 2018 are 12.36% and 10.81%. The effects drop from 2015 to 2018, showing an improvement in air pollution levels. PM2.5, $SO_2$, and CO are the major pollutants as they have largest growth rates. Ozone is different. Central heating can reduce daily sunshine and ozone concentration.

**Table 3.4.1-1 Main estimation results-heating variable**

| **Year** | **ln(PM2.5)** | **ln(PM10)** | **ln(AQI)** | **ln($SO_2$)** | **ln($NO_2$)** | **ln(CO)** | **ln($O_3$)** |
|---|---|---|---|---|---|---|---|
| 2015 | 0.1667682 | 0.087151 | 0.1132887 | 0.4495528 | 0.090772 | 0.1495492 | -0.055211 |
| Std.Err. | .0314769 | .0276284 | .0242602 | .027895 | .0190431 | .0203425 | .0258571 |
| P-value | 0.000 | 0.002 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Overall$R^2$ | 0.3542 | 0.2778 | 0.3005 | 0.3385 | 0.2838 | 0.3736 | 0.5294 |
| Obs | 8613 | | | | | | |
| 2016 | .1843101 | .1629188 | .1693506 | .3516329 | 0804775 | .1708698 | -.0818877 |
| Std.Err. | .0278844 | .0246111 | .0219344 | .0256103 | .0164197 | .0197282 | .0291642 |
| P-value | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.005 |
| Overall$R^2$ | 0.4524 | 0.4094 | 0.4164 | 0.4076 | 0.3735 | 0.4267 | 0.2772 |
| Obs | 9579 | | | | | | |
| 2017 | .1457252 | .0647137 | .0861611 | .2662393 | .0707583 | .1073253 | -.0328374 |
| Std.Err. | .0268601 | .0227849 | .0199283 | .0225391 | .0166113 | .0178239 | .0197411 |
| P-value | 0.000 | 0.005 | 0.000 | 0.000 | 0.000 | 0.000 | 0.096 |
| OverallR2 | 0.3413 | 0.3316 | 0.3307 | 0.3524 | 0.3318 | 0.2896 | 0.5986 |
| Obs | | | 9579 | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 2018 | .079884 | 0.152648 | 0.122962 | 0.099754 | 0.064584 | 0.1268716 | -0.154291 |
| Std.Err. | .0259108 | .0224731 | .0194907 | .0212529 | .0170803 | .0173937 | .020316 |
| P-value | 0.002 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Overall r2 | 0.4613 | 0.3993 | 0.4136 | 0.3014 | 0.3912 | 0.3612 | 0.5787 |
| Obs | 9415 | | | | | | |

### 3.4.2 Pollutants and meteorological variable

This section shows the coefficients of the air pollution and other meteorological factors. Table 3.4.2-1 presents the coefficients of other meteorological variables in 2018. Wind and air pollutants have a negative relationship. Humidity and pollutants are positively correlated. Precipitation and pollution have a negative relationship. Sunshine and pollution are negatively correlated. These relationships are all reasonable.

**Table 3.4.2-1 Pollutants and meteorological variable (2018)**

| Year 2018 | Maximum humidity | Wind speed | Sunshine hours | Minimum humidity | precipitation |
|---|---|---|---|---|---|
| ln(pm25) | .0121056 | -.0048926 | -.0026536 | -.0021765 | -.0007989 |
| Std. Err. | .0007528 | .0007592 | .0001688 | .0004383 | .0000726 |
| P-value | 0.000 | 0.000 | 0.000 | .0001688 | 0.000 |
| ln(pm10) | .0054855 | -0.0036499 | -0.0023219 | -.0017969 | -.0008732 |
| Std. Err. | .0004603 | .0006582 | .0001463 | .0003816 | .0000629 |
| P-value | 0.000 | 0.000 | 0.000 | .0001463 | 0.000 |
| ln(aqi) | 0.006223 | -0.0031723 | -.0019056 | .0134135 | -.0007035 |
| Std. Err. | 0.0006339 | .0005707 | .0001268 | .0012404 | .0000546 |
| P-value | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| ln(so2) | -.007964 | -.0044712 | -.0016754 | .0190294 | -.0000782 |
| Std. Err. | .0006921 | .0006231 | .0001385 | .0013536 | .0000595 |
| P-value | 0.000 | 0.000 | 0.000 | 0.000 | 0.189 |
| ln(no2) | .0048199 | -.0095497 | -.0015365 | -.0028784 | -.000122 |
| Std. Err. | .0003501 | .0005008 | .0001113 | .0002903 | .0000478 |
| P-value | 0.000 | 0.000 | 0.000 | 0.000 | 0.011 |
| ln(co) | .0064598 | -.004311 | -.0014621 | .008706 | -.0000879 |
| Std. Err. | .0005664 | .0005099 | .0001133 | .0011077 | .0000487 |
| P-value | 0.000 | 0.000 | 0.000 | 0.000 | 0.071 |
| ln(o3) | -.0024022 | .0023781 | .0011456 | .0071929 | -.0000328 |
| Std. Err. | .0006608 | .0005948 | .0001322 | .0012929 | .0000569 |
| P-value | 0.000 | 0.000 | 0.000 | 0.000 | 0.564 |
| Obs | 9415 | | | | |

## 3.5 Robustness checks

The following robustness checks are conducted. First, we shorten the estimation window from 120 days to 70 days. Second, an imaginary date is chosen as the heating start. Third, we apply the southern cities as another dataset to conduct the benchmark estimation. Nov 15th is still selected as the heating start day. At last, we conduct a regression between the air pollutants level and daily minimum temperature. We can further verify the results in the benchmark estimation.

### 3.5.1 70-day window check

Estimation window will be adjusted from 120-day to 70-day to test the robustness. The methodology is the same as the benchmark.

In this check, most coefficients are still positive and significant. We have smaller coefficients compared with 120-day window results. $PM_{2.5}$ and $SO_2$ are the two most sensitive pollutants to heating.

**Table 3.5.1-1 70-day window estimation**

| **Year** | **2015** | **2016** | **2017** | **2018** |
|---|---|---|---|---|
| ln(pm25) | 0.1286897 | .0885972 | .2249379 | .0301077 |
| Std. Err. | .0399683 | .0337292 | .0336705 | .028117 |
| P-value | 0.001 | 0.009 | 0.000 | 0.284 |
| $R^2$ | 0.3624 | 0.4389 | 0.3606 | 0.4307 |
| ln(pm10) | .1057149 | 0.060131 | .1619521 | .0902301 |
| Std. Err. | .034782 | .0295067 | .0275496 | .0254578 |
| P-value | 0.002 | 0.042 | 0.000 | 0.000 |
| $R^2$ | 0.4184 | 0.4184 | 0.3752 | 0.3640 |
| ln(aqi) | .0879967 | 0.0591242 | .1519501 | .0523833 |
| Std. Err. | .0311665 | .0266729 | .0245397 | .0217374 |
| P-value | 0.005 | 0.027 | 0.000 | 0.016 |
| $R^2$ | 0.4229 | 0.4229 | 0.3675 | 0.3832 |
| ln($so_2$) | .3196122 | 0.282906 | .2210812 | 0.026582 |
| Std. Err. | .0338377 | .0309124 | .0263059 | .0233029 |
| P-value | 0.000 | 0.000 | 0.000 | 0.254 |
| $R^2$ | 0.3017 | 0.3481 | 0.309 | 0.2754 |
| ln($no_2$) | .0537267 | .0082169 | .0925149 | -0.0217484 |
| Std. Err. | .0235758 | .0189431 | 0.0189638 | .0162535 |
| P-value | 0.023 | 0.664 | 0.000 | 0.181 |
| $R^2$ | 0.2478 | 0.3093 | 0.3341 | 0.3193 |
| ln(co) | .0845894 | .0132024 | .0870859 | 0.0202562 |
| Std. Err. | .0252556 | .0243346 | .0223223 | .0190594 |
| P-value | 0.001 | 0.587 | 0.000 | 0.288 |
| $R^2$ | 0.4030 | 0.4486 | 0.3335 | 0.3979 |
| ln($o_3$) | .0083848 | .0166673 | -.0196448 | -.0646289 |

| | | | | |
|---|---|---|---|---|
| Std. Err. | .0300514 | .0300952 | .0253694 | .0230832 |
| P-value | 0.780 | 0.580 | 0.439 | 0.005 |
| $R^2$ | 0.5072 | 0.2698 | 0.5722 | 0.4885 |
| Obs | 5356 | 5640 | 5499 | 5639 |

### 3.5.2 Imaginary starting day

A "false" heating start day is selected as the cut-point. August 15th is the "imaginary" heating starting day. The estimation results are in Table 3.5.2-1. To make the results simpler, results for 2016 and 2017 are presented here. Results show negative coefficients, which verifies the results in the benchmark.

**Table 3.5.2-1 Another starting day**

| 100-day window August 15th | PM2.5 | PM10 | AQI | SO2 | NO2 | CO | O3 |
|---|---|---|---|---|---|---|---|
| 2016 | -0.089012 | -0.051224 | -0.025311 | 0.205359 | 0.090227 | 0.089622 | 0.129901 |
| Std. Err. | .0282223 | .0235385 | .0194928 | .0253872 | .0167546 | .0191114 | .0229918 |
| P-value | 0.002 | 0.030 | 0.194 | 0.000 | 0.000 | 0.000 | 0.000 |
| $R^2$ | 0.3030 | 0.0813 | 0.2859 | 0.1961 | 0.2766 | 0.1290 | 0.2273 |
| Obs | 8241 | | | | | | |
| 2017 | -0.078201 | -0.084632 | -0.076689 | .1088264 | -0.028814 | -0.012279 | 0.089709 |
| Std. Err. | .0274319 | .0222405 | .0183619 | .0234827 | .0164598 | .0171536 | .0204558 |
| P-value | 0.004 | 0.000 | 0.000 | 0.000 | 0.080 | 0.474 | 0.000 |
| $R^2$ | 0.3082 | 0.0817 | 0.2784 | 0.1852 | 0.3063 | 0.1109 | 0.2374 |
| Obs | 8241 | | | | | | |

### 3.5.3 Southern cities as the samples

We collect the air quality data in non-heating areas as a contrast. The chosen cities are listed in Table 3.5.3-1. A 70-day window and November 15th will be chosen as the estimation periods and start day, respectively. The estimation results are in Table 3.5.3-2.

**Table 3.5.3-1 Southern cities with average winter temperatures[6]**

| City name (average winter temperature) |
|---|

---

[6] Winter temperatures are self-calculated using the meteorological dataset mentioned before.

| | |
|---|---|
| **Cities in the south** | Beihai (20.3),Haikou (20.0), Zhanjiang (18.2) Shenzhen (17.9) Yangjiang (17.7) Zhuhai (17.7) Dongguan (17.2) Zhongshan (17.1), Shantou (17.3) Huiyang (16.6),Qingyuan (16.0) Guangzhou (15.5) Xiamen (15.3) Panzhihua (15.0) Liuzhou (14.4) Fuzhou (13.8) Nanning (13.8) Guilin (12.1) Kunming (10.5) Nanchang (9.5) Zhuzhou (9.0) Nanchong (8.8), Yueyang (8.7) Changsha (8.7) Changde (7.9) Kunshan (7.5) Jingzhou (7.4) Guiyang (7.4) Yichang (7.2) Wuxi (6.8), Wuhan (6.8), Maanshan (6.5) Nanjing (6.4) Nantong (6.2)Hefei (5.4) Huaian (4.3) Lushan (4.3) |

All the coefficients in Table 3.5.3-2 are either insignificant or negative. This shows that the southern cities do not witness a rise in pollution level when winter comes. This further illustrates the rise in air pollution is due to other common reasons that the south and north both have.

**Table 3.5.3-2 Estimation results using southern cities**

| | **2016** | **2017** |
|---|---|---|
| ln(pm25) | 0.0372473 | -0.455858*** |
| ln(pm10) | 0.057042* | -0.389328*** |
| ln(aqi) | 0.0234675 | -0.356688*** |
| ln($SO_2$) | 0.0234675 | -0.121339*** |
| ln($NO_2$) | -0.064297*** | -0.179958*** |
| ln(co) | -0.074157*** | -0.091616*** |
| ln($o_3$) | 0.0198585 | -0.290647*** |

### 3.5.4 Minimum daily temperature and air pollutant level

This section wants to explore the relationship between the minimum daily temperature and air pollution. The logic is like this: the lower the winter temperature is, the higher temperature the central heating stoves will have. Thus, it could have a higher level of air pollutant emission.

The basic methodology is

$$\ln\,(\mathrm{AQI_{it}}) = \alpha + \beta\mathrm{mintemp}_{it} + \gamma\mathbf{X}_{it} + u_{it} \tag{1}$$

mintemp is the minimum daily temperature. We exclude the variable Heating. $\mathbf{X}_{\mathrm{it}}$ is the control variable, including wind direction at maximum daily wind speed(m/s), maximum daily wind speed, maximum daily relative humidity, sunshine hours, minimum daily humidity, and average daily wind speed precipitation. Day of week and day of month are two variables controlling the weekday-effect. We also consider the square terms of these meteorological variables. The estimation results are reported in Table 3.5.4-1. We do the estimation year by year and finally, do the estimation based on the whole sample.

The estimations have negative coefficients. In the 2015-winter, when the daily minimum temperature decreases by 1°C, the concentration level of PM2.5 will increase by 1.9%. These results can further illustrate the policy’s effects on air pollution.

**Table 3.5.4-1 Minimum daily temperature and air pollutants level**

| | **lpm25** | **lpm10** | **AQI** | **$SO_2$** | **$NO_2$** | **CO** | **$O_3$** |
|---|---|---|---|---|---|---|---|
| 2015 | -.0018889 | -.0012397 | -.001838 | -.0026096 | -.0010802 | -.0008798 | .0013425 |
| St. err | .0001449 | .0001247 | | .0001585 | .0000987 | .0001007 | .0001089 |
| P-va | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Obs | 8611 | | | | | | |
| 2016 | -.001728 | -.1171531 | -.0013894 | -.0022903 | -.0008526 | -.0010885 | .0014895 |
| St. err | .0118169 | .0161691 | .0001132 | .0001667 | .0000973 | .0001095 | .0001759 |
| P-va | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Obs | | | | 9578 | | | |
| 2017 | -.00128 | -.0008743 | -.000958 | -.002392 | -.0006875 | -.0008481 | .0015922 |
| St. err | .0001335 | .0001058 | .0000946 | .0001261 | .0000825 | .0000892 | .0000892 |
| P-va | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Obs | | | | 9579 | | | |
| 2018 | -.0011143 | -.0008053 | -.000806 | -.0007318 | -.0001103 | -.0005735 | .0011956 |
| St. err | .0001509 | .0001287 | .0001118 | .0001349 | .0001014 | .0001023 | .0001005 |
| P-va | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Obs | | | | 9415 | | | |
| Overall | -.0018887 | -.0013968 | -.0014291 | -.0025908 | -.0008821 | -.0010087 | .0014418 |
| St. err | .0000722 | 0.0000605 | .0000532 | .0000774 | .0000464 | .0000503 | .0000584 |
| P-va | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| $R^2$ | 0.2443 | 0.1962 | 0.2044 | 0.2714 | 0.3136 | 0.2539 | 0.4181 |
| Obs | | | | 37183 | | | |

## 3.6 Winter heating effects of different districts

This section will consider the heterogeneity of cities in different temperature zones. We want to verify whether higher latitudes will generate higher pollution levels. Same estimation methodology of benchmark equation is used for four districts. We only report the data for 2017 for simplicity. The classification criteria are according to the discussion in Section 3.2.

The estimation results are shown in Table 3.6-1. In summary, group I, located in the far north, has always the largest effects as it has colder winter temperatures. Group IV has relatively the lowest effects.

**Table 3.6-1 Year 2017: By average winter temperature**

| **Group** | **lpm25** | **lpm10** | **AQI** | **$SO_2$** | **$NO_2$** | **CO** | **$O_3$** |
|---|---|---|---|---|---|---|---|

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| I | .5416884*** | .3354049*** | .313466*** | .6499211*** | .1764009*** | .1101177** | .059841 |
| II | .1744047*** | .1241293*** | .1068305*** | .3549782*** | .0510317 | .1075357*** | -.0678201 |
| III | .1560668*** | .1311119** | .1408444*** | .26909*** | .0892169** | .1526883*** | .0690939 |
| IV | .1232101*** | .0724796* | .2009224*** | .1622219*** | .0511519 | .0641058 | -.0787474** |

### 3.7 Results

Winter heating contributes to the air pollution growth by 17.62% from 2015-2019 on average between non- and heating seasons, excluding weather effects. There is a downward trend from 2015 to 2018 for all pollutants. PM2.5, $SO_2$, and CO are the most decisive air pollutants generated by central heating. These three pollutants are more associated with coal burning.

The coefficients are not significant when we use the imaginary heating start day. Cities in higher latitudes have more obvious air pollution due to more coal burning. 2017 sees the increase by 54.2% in PM2.5 because of central heating in the coldest areas. Moreover, another check shows a negative relationship between daily minimum temperatures and air pollutants concentration.

### 3.8 Clean heating project

Clean heating policy is also a component of the Huai River policy. Wen et al. (2021) [20] concluded that the major air pollutants dropped significantly in Beijing when clean heating projects are introduced. Zhang et al. (2020) [21] analyzed the effects of the Winter Clean Heating Pilot in China (WCHP). The main content of this campaign is to switch coal to gas or electricity. Their findings are that the major air pollutants, including AQI, PM2.5, NOx, and SO2, decreased by 20.4%, 18.59%, 34.1%, and 68.4%, respectively. Wang et al. (2022) [22] discovered that the clean heating project in North China reduced 46.6% of the total air pollutants. Tan et al. (2023) [23] found that the clean heating policy since 2017 brought down air pollution apparently. Guo et al. (2020) [24] focused their studies on Beijing-Tianjin-Hebei regions. These papers analyzed different clean heating campaigns in different parts of China, most of which used the Difference-in-difference (DID) method.

In 2017, four major Chinese ministries launched a clean-heating project, encouraging cities to use cleaner energies for winter heating. They raised 6 billion yuan as financial subsidies to those chosen cities. They encourage cities to either renovate their old winter heating technologies or work out new plans or schemes to boost the usage of clean energy. The first batch of chosen cities are Tianjin, Shijiazhuang, Tangshan, Langfang, Hengshui, Taiyuan, Jinan, Zhengzhou, Kaifeng, Hebi, and Xinxiang. The chosen cities are all distributed around the Beijing-Tianjin-Hebei area in China and belong to the cities that can transfer pollutants directly to the Beijing-Tianjin-Hebei area. Cities in this program explore different ways of introducing new types of energy into their heating packages, including coal-to-gas, coal-to-electricity, and use of biomass. Jinan introduced waste heat from the electricity plants nearby as heating energy to avoid more energy burning. Tangshan is a city with much heavy industry and steel industry is its pillar industry. It uses the waste heat from its steelmaking industries.

In order to evaluate the project's effects, it is not convincing to compare the air pollutants' level directly with the data of the previous heating season as so many variables are hard to control. It is hard to explore some policy's causal relations as countless variables are linked to the policy. In order to tackle this problem, we introduce a synthetic control method to mitigate the problem mentioned above. We use a synthetic controlled area that is synthesized by several cities that do not have this policy. We compare the policy performance of these two areas and make sure whether project-chosen cities are different from the unchosen cities. Figure 3.8-1 illustrates the difference between synthetic and real places. The project has an essential impact on air pollution for almost all four pollutants and all cities. The vertical line stands for the date when the policy is introduced. We find that the air pollution level is much stronger in the synthesized area than in the policy area.

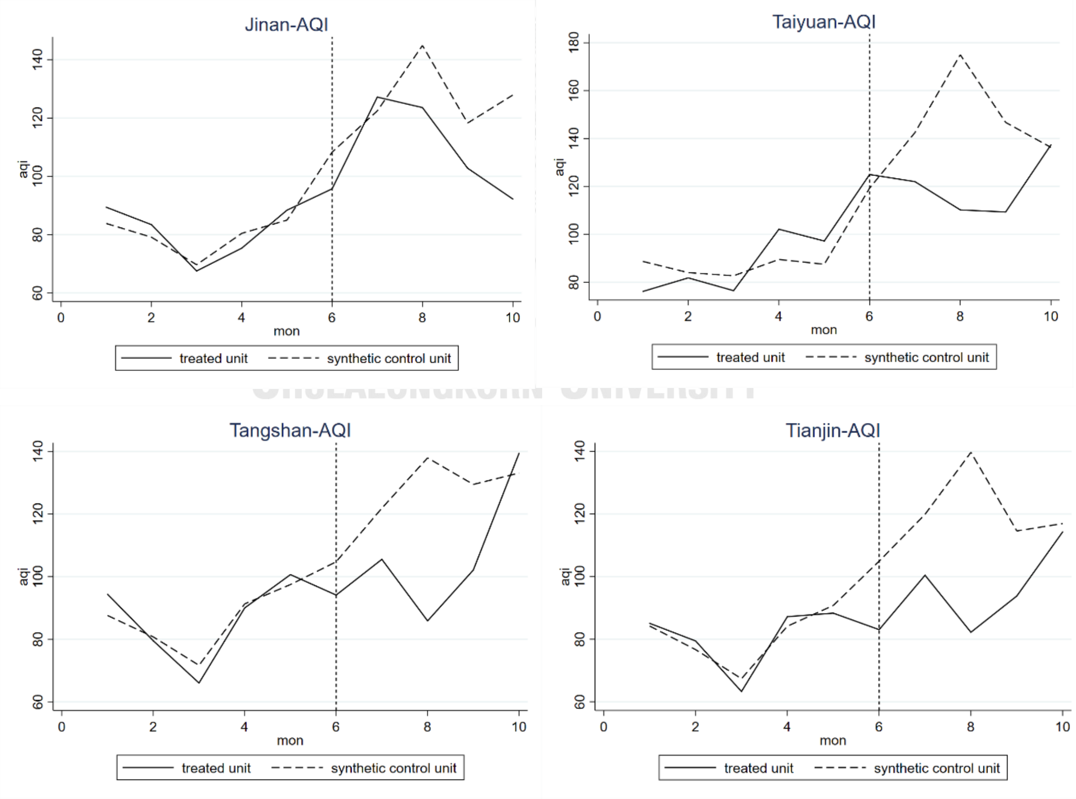

**Figure 3.8-1 AQI for Jinan, Taiyuan, Tangshan and Tianjin**

### 3.9 Summary

The above analysis shows that central heating areas have more air pollution, and clean heating projects help reduce air pollution.

# Chapter 4 Huai River policy and migration: theoretical and empirical analysis

The connection between the Huai River policy and migration is a research gap. Three factors affected by the policy can affect migration. As we discussed in the previous chapters, these three factors are heating costs, air quality, and indoor comfort. There is some research on these three factors' impact on migration.

Liu et al. (2022) [25] showed a negative relationship between air pollution and the international students' inflow in major Chinese cities. They also found that the self-financed students cared more about the air pollution levels, while students with scholarships wouldn't care that much. Liu and Yu (2020) [26] also found that air pollution could bring down the migrants' settling interest in China. Germani et al. (2021) [27] used the example of Italy and found that air pollution could influence internal migration, pushing people to move to less-polluted provinces. Wang et al. (2020) [28] found that air quality was a key factor for immigration in selected countries. Kim (2019) [29] used the example of California and found that income could affect the importance of air quality for migration.

Energy costs can be a very important factor for people to choose their settling destinations. Berman and Matthew (2017) [30] attributed the out-migration in Alaska to increased energy costs. Gale and Heath (2000) [31] found that places with more heating days could generate more costs for residents and would not be an advantage for in-migration.

Central heating brings about a higher level of room comfort and could contribute to in-migration. Fielding (2011) [32] concluded that British people prefer to concentrate in urban areas because of the benefits of district heating. Parshukovo et al. (2020) [33] used the variable of whether a house is equipped with a district heating system as an indicator to test its effects on migration.

Three factors interact to determine people's demand for heating. Although Guo et al. (2015) [4] pointed out that the unit cost of central heating is lower than that of scattered heating, the total costs of central heating could be higher than scattered heating, especially in places farther south. People sensitive to the heating costs will not take central heating as an important factor in deciding their settling destinations. Air pollution is a factor that hinders in-migration, while a more comfortable living environment is a factor that attracts the population to migrate.

Although literature discussing migration-related factors exists, no literature directly discusses the relationship between the Huai River policy and population migration theoretically and empirically. This chapter will analyze the policy and migration theoretically and empirically. Sections 4.1 and 4.2 are partly published in [34].

## 4.1 Theoretical analysis

The old Harris and Todaro (1970) [35] model in development economics is used to analyze the migration between rural and urban areas. As rural people could bear the possibility of being unemployed or having higher wages in urban areas, they have to balance these two. Many scholars are applying the Harris and Todaro framework to analyze the migration and unemployment problems. Brueckner and Zunou (1999) [36] introduced the land market area into the framework. Laing et al. (2005) [37] took the Household registration system into their framework to analyze the migration in China. Environment was introduced into this framework to analyze the impact of environmental policies on urban unemployment and migration. Copeland and Taylor (1994) [38] thought over minor changes in environmental policies and discovered that the slight rise in pollution tax may not impair welfare as spillover effects are correcting the trade distortion. Beghin et al. (1997) [39] considered pollution in consumption and incorporated it into a Harris-Todaro model. Daitoh (2003) [40] explored the welfare-improving environmental policy reform in a Harris-Todaro framework. The paper shows that a rise in pollution can generate two labor market distortions: less-than-optimal manufacturing employment and urban unemployment. It also discusses possible conditions for welfare-improving. Fukuyama and Naito (2007) [41] applied a Harris-Todaro framework to analyze the comparative statics of pollution tax rates, pollution-reduced technology, and pollution levels on migration and unemployment rates. Like Daitoh (2003) [40], they also studied the impact on social welfare.

Heating could generate more pollution, bringing down the utility of nearby households. Heating can also improve the living conditions of households and raise the utility level. Heating factories can also influence the production of other areas. Heating expenditures may also affect people’s living standards. We refer to the model of Fukuyama and Naito (2007) [40], introduce heating into this framework, and analyze the possible effects on population migration.

### 4.1.1 Model setting

Suppose there are two places, N and S. The distance between N and S is d. N has central heating systems and can produce heating (H) for itself, and S does not have any heating systems and must buy heating from N. Heating systems generate air pollution. The local government collects a tax on unit pollution of A with a tax rate of t. Residents living in N do not have to pay a transportation fee, but residents living in S must pay a rate for heating transportation. This transportation fee can also be modeled as the extra costs for residents in S to obtain heating services. The rate for one unit of distance from N is r. The production of heating in N could influence the output in S. It depends on the distance between N and S since pollutants generated in the production of N can undermine the production in S. Considering that the climate conditions can affect the demand for heating if S is located southward with higher winter temperatures, the reduction in extra heating costs can be modeled with another term c(d)-climate adjustment function. d is the distance between N and S and is modeled positive. c(d) is considered the marginal effect of the latitudes or climate on heating expenditures. With an increasing d from N, c(d) becomes larger. The overall costs [rd-c(d)d] for residents in S to buy heating decrease. We assume that the first-order derivative with respect to d is positive.

**4.1.1.1 Households**

Two plants in N and S are introduced in the model. N has a heating plant with a production function as follows:

$$\mathrm{H} = \mathrm{H}(\mathrm{L_H}, \mathrm{A})$$

$\mathrm{L_H}$ is the labor input for the production of heating, and A is the polluting goods needed to produce heating. They should be kept positive. A stands for air pollution in this scenario. We assume the basic properties of function H: H is a monotonically increasing function in both variables $\mathrm{L_H}$ and A. The second-order partial derivative with respect to $\mathrm{L_H}$ is negative. The second-order mixed partial derivative is positive. In the following context, $\mathrm{H_{L(A)}}$ stands for the partial derivative with respect to L or A. $\mathrm{H_{LA}}$ indicates the second-order mixed partial derivative.

$$\frac{\partial \mathrm{H}}{\partial \mathrm{L_H}} > 0 \quad \frac{\partial \mathrm{H}}{\partial \mathrm{A}} > 0 \quad \frac{\partial^2 \mathrm{H}}{\partial \mathrm{L_H^2}} < 0 \quad \frac{\partial^2 \mathrm{H}}{\partial \mathrm{L_H^2}} \frac{\partial^2 \mathrm{H}}{\partial \mathrm{A}^2} - \left(\frac{\partial^2 \mathrm{H}}{\partial \mathrm{L_H}\, \partial \mathrm{A}}\right)^2 > 0$$

Or they can be written as

$$\mathrm{H_L}>0 \quad \mathrm{H_A}>0 \quad \mathrm{H_{LL}}<0 \quad \mathrm{H_{LL}}\,\mathrm{H_{AA}} - \mathrm{H_{LA}}^2>0$$

We now establish the profit maximization problem of plant N. p denotes the relative price of heating. The product price in place S is normalized to 1. $\mathrm{w_H}$ represents the wage in the heating sector. The profit of plant N is equal to the return on the production of heating minus the total wage minus the air pollution tax. The profit function for plant N is

$$\pi_\mathrm{N} = \mathrm{pH}(\mathrm{L_H}, \mathrm{A}) - \mathrm{w_H L_H} - \mathrm{tA}$$

The two first-order conditions are as follows:

$$\mathrm{p}\frac{\partial \mathrm{H}}{\partial \mathrm{L_H}} = \mathrm{w_H} \qquad \mathrm{p}\frac{\partial \mathrm{H}}{\partial \mathrm{A}} = \mathrm{t} \tag{2}$$

Place S has a factory to produce product X, which can be considered other goods and services apart from heating. It is assumed that this production process generates no air pollution but is influenced by the production of heating in N. To measure the air pollution, we assume that the air pollution level is associated with the distance between S and N. d is used to represent the distance. The production function is as follows:

$$\mathrm{X} = \mathrm{Q}(\mathrm{d}, \mathrm{A})\mathrm{X}(\mathrm{L_X})$$

Q(d,A) is the air quality function correlated with the distance. Q should be always positive. It is also connected to the pollution induced by the production of heating. Q is a monotonically increasing function in d but a decreasing function in A. Q complies with the following inequalities

$$\frac{\partial \mathrm{Q}}{\partial \mathrm{d}} > 0 \qquad \frac{\partial \mathrm{Q}}{\partial \mathrm{A}} < 0 \tag{3}$$

To simplify the analyzing process, we define the functional form of Q(d,A) as follows.

$$\mathrm{Q}(\mathrm{d}, \mathrm{A}) = \left[\left(\mathrm{E} - \mathrm{A}(1-\delta)^{\mathrm{d}}\right]\right.$$

We assume the best air quality level without any pollution is E (a constant). When d rises by one unit, the air pollution level decreases by δ. δ is a percentage and falls into the interval (0,1). δ is a parameter modelling the decrease rate of air pollution. The air quality level of a place d units of kilometers from place N is $[(E - A(1-\delta)^{d}]$. We simulate the quantitative relation between Q and d in Figure 4.1.1.1-1. The function is an increasing function with a decreasing slope. It converges to E when d goes to infinity. The reason of setting a function like this is that the air quality level does not follow a linear function with distance. Areas near the heating borderline will have much worse air quality levels than areas farther away from the borderline.

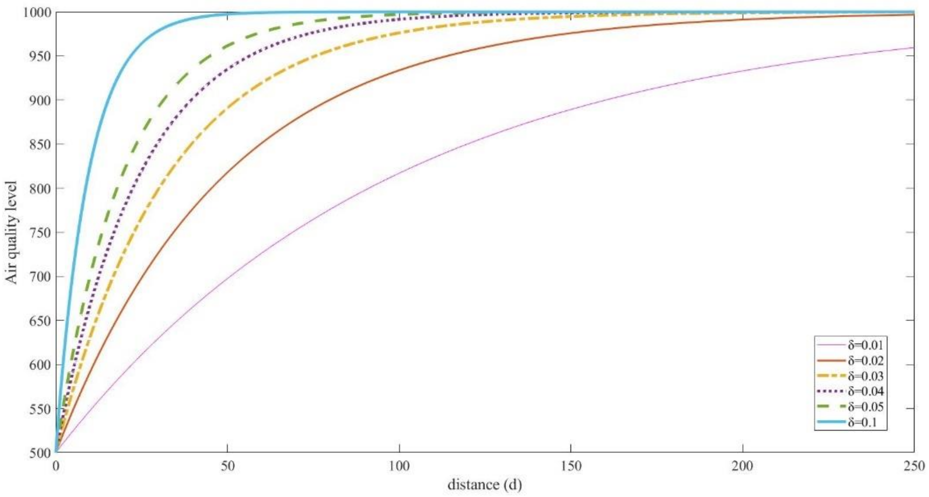

**Figure 4.1.1.1-1 Simulation of function Q and d of difference δ. (E=1000, A=500)**

It is assumed that the tax charged on heating will be a subsidy for the production of S. The total profit of plant S equals the return on selling product X minus the total wages of plant S plus the government's tax subsidies. The profit maximization for factory S is

$$\pi_S = Q(d, A)X(L_X) - w_X L_X + tA$$

The first-order condition is

$$Q(d, A)\frac{\partial X}{\partial L_X} = w_X \qquad (4)$$

**4.1.1.2 Households**

We assume factory S will not have unemployment as a sector producing the common goods X. We assume factory N has unemployment. Households can be divided into three groups: (1) households living in N and working in the heating factory; (2) households living in N and having no work; and (3) households living in S and working in the factory producing X.

The utility function for these three groups is:

$$U_j^i = u(H_j^i, X_j^i) + \gamma Q_j(d_j, A) \quad (\gamma > 0)$$

We assume a form of the utility function u(.) as homothetic, and for instance, we could further assume u(.) as $u(x, y) = x^{\alpha}y^{1-\alpha}(0 < \alpha < 1)$.

The budget constraint corresponding to the utility function i

$$pH_j^i + X_j^i + [rd_j - c(d_j)d_j] = I + \mathrm{I}(i = E)w_k$$

i represents the state of employment with the state either E (employed) or U (unemployed). j represents the location of either N or S. k represents the wages of plant N (k=H) or S (k=X). $\mathrm{I}(\cdot)$ is an indicator function: if the input is true, the function returns the value 1; if the input is false, it returns the value 0.

The following are the valuations of j, i, and k in three groups: Group (1): j=N, i=E; k=H Group (2): j=N, i=U; Group (3): j=S, i=E, k=X.

$d_j$ is the variable for measuring the distance between place j and the heating plant. If j=N, $d_N$ =0; if j=S, $d_S$=d. Thus, if j=N, $Q_N$= $E - A$; if j=S, $Q_S$= $E - A(1 - \delta)^d$. The parameter d measures the actual distance between N and S.

$rd_j$ models the extra costs the households in S should pay to buy heating, while $c(d_j)d_j$ models the cost reduction owing to the increase in winter temperatures when d becomes larger.

γ is a positive parameter that matches the utility of air quality with the utility of consumption of H and X. If γ becomes larger, households will value air quality levels more in their utility functions. We can also interpret γ as the unit monetary value of good air quality or environmental awareness in this scenario.

It is assumed that all profits are equally divided in each household. Each household is assigned income I. This assumption holds because there is no other department in this setting. Taxes are used to subsidize firms. Profits are used to subsidize households.

We perform the utility maximization problem for these three groups of households. The first-order conditions are

$$\frac{\partial U_j^i}{\partial H_j^i} \Bigg/ \frac{\partial U_j^i}{\partial X_j^i} = p$$

By solving the utility maximization problem, the equilibrium utility levels are as follows:

$$U_j^i = \left(\frac{\alpha}{P}\right)^{\alpha}(1-\alpha)^{1-\alpha}\left(I + \mathrm{I}(i = E)w_k + c(d_j)d_j - rd_j\right) + \gamma\left(E - A(1-\delta)^{d_j}\right)$$

Following the above utility function form along with the budget constraints, we can derive the equilibrium consumption level as follows:

$$H_N^E = \frac{\alpha}{P}(I + w_H) \qquad X_N^E = (1-\alpha)(I + w_H)$$

$$U_N^E = \left(\frac{\alpha}{P}\right)^{\alpha}(1-\alpha)^{1-\alpha}(I + w_H) + \gamma(E - A)$$

$$H_N^U = \frac{\alpha}{P}I \qquad X_N^U = (1-\alpha)I$$

$$U_N^U = \left(\frac{\alpha}{P}\right)^{\alpha}(1-\alpha)^{1-\alpha}I + \gamma(E - A)$$

$$H_S^E = \frac{\alpha}{P}(I + w_X - rd) \;\; X_S^E = (1-\alpha)(I + w_X - rd)$$

$$U_S^E = \left(\frac{\alpha}{P}\right)^{\alpha}(1-\alpha)^{1-\alpha}(I + w_X + c(d)d - rd) + \gamma\left(E - A(1-\delta)^d\right)$$

The migration equilibrium is attained when the utility of the household in N is equal to that of the household in S, showing that neither is motivated to migrate to the other place. $\mu = \frac{L_u}{L_u + L_H}$ can be modeled as the unemployment rate in N. The migration equilibrium is

$$(1-\mu)U_N^E + \mu U_N^U = U_s^E$$

In this theoretical analysis, we assume there is no unemployment rate. Thus, we assume $\mu=0$.

The migration equilibrium condition after simplification can be written as

$$w_H = w_X - rd + c(d)d + \gamma A\left[1 - (1-\delta)^d\right] \tag{5}$$

The left side of equation (5) is the wage of households in N, while the right side is the income in S plus some utility gain or loss (GAIN= $A\left[1 - (1-\delta)^d\right]$) due to a change in air quality level and climate conditions. When d increases, there will be a utility gain since the derivative with respect to d of GAIN is positive.

$$\mathrm{GAIN}_d = -A[\ln(1-\delta)](1-\delta)^d > 0$$

We assume that the total labor in the model is L.

Since we assume $\mu=0$, $L_H=L_N$ and $L_S=L_X$.We obtain the labor market equilibrium

$$L = L_N + L_S \tag{6}$$

**4.1.1.3 Equilibrium**

Equations (2), (4), (5), and (6) define the whole equilibrium.

By differentiating the equilibrium conditions, we obtain the following matrix:

$$\begin{bmatrix} pH_{LL} & 0 & pH_{LA} \\ pH_{AL} & 0 & pH_{AA} \\ 1 & 1-\mu & 0 \\ 0 & Q(d,A)X_{LL} & Q_A X_L + \gamma + \gamma A \end{bmatrix} \begin{bmatrix} dL_H \\ dL_S \\ dA \\ d\mu \end{bmatrix}$$

$$= \begin{bmatrix} -H_L dp \\ dt - H_A dp \\ (1-\mu)dL \\ [r - dc_d - c(d) - Q_d X_L - \gamma Q_d](dd) + d\cdot dr - [(A + Q(d,A) - E]d\gamma \end{bmatrix}$$

It is easy to prove that the determinant of the “coefficient matrix” of the left side is negative.

$$D = p^2\left(H_{LL}H_{AA} - {H_{LA}}^2\right)[(L - L_S)Q(d,A)X_{LL} - w_H] < 0$$

We apply Cramer’s rule to calculate the results of $dL_H$, $dL_S$, and d. This setting could be helpful in analyzing the effects of d on population migration.

$$\begin{aligned} dL_H = \frac{1}{D}\{&(H_A H_{LA} - H_L H_{AA})p[(L - L_S)Q(d,A)X_{LL} - (1-\mu)w_H]dp \\ &- pH_{LA}[(L - L_S)Q(d,A)X_{LL} - w_H]dt\} \end{aligned}$$

$$\begin{aligned} dL_S = \frac{1}{D}\{&p^2(L - L_S)(r - dC_d - C(d) - Q_d X_L - \gamma Q_d)\left(H_{LL}H_{AA} - {H_{AL}}^2\right)dd \\ &+ p^2(L - L_S)d\left(H_{LL}H_{AA} - {H_{AL}}^2\right)dr - p^2 w_H\left(H_{LL}H_{AA} - {H_{AL}}^2\right)dL \\ &- p(H_{LL}(L - L_S)(Q_A X_L + \gamma + \gamma A) + H_{LA}w_H)dt + p[w_H(H_A H_{LA} - H_L H_{AA}) \\ &+ (Q_A X_L + \gamma + \gamma A)(L - L_S)(H_A H_{LL} - H_L H_{AL})]dp - p^2(L \\ &- L_S)[(A + Q(d,A) - E]\left(H_{LL}H_{AA} - {H_{AL}}^2\right)d\gamma\} \end{aligned}$$

$$\begin{aligned} dA = \frac{1}{D}\{&pH_{LL}[(L - L_S)Q(d,A)X_{LL} - (1-\mu)w_H]dt \\ &- p[(L - L_S)Q(d,A)X_{LL} - w_H](H_A H_{LL} - H_L H_{AL})dp\} \end{aligned} \tag{7}$$

To analyze the effects of parameters on migration, we will try to get the derivatives of these variables and analyze their signs.

### 4.1.2 Effects of parameters on migration

#### 4.1.2.1 Effects of distance d on migration

We have the following derivatives derived from Equation (7):

$$\frac{dL_S}{dd} = \frac{1}{D}\left[p^2(L - L_S)(r - dc_d - c(d) - Q_d X_L - \gamma Q_d)\left(H_{LL}H_{AA} - {H_{AL}}^2\right)\right] \propto \frac{dMSF}{dd}$$

The simulation graph is shown in Figure 4.1.2.1-1. In the beginning, when d is very small, meaning that S is located very near N, population in S is large. When d increases, $L_S$ becomes smaller until it reaches a minimum level. In this stage, we can say that the in-migration decreases. The population starts to move to N. When d is even larger, the population in S in very large, showing an in-migration in the south.

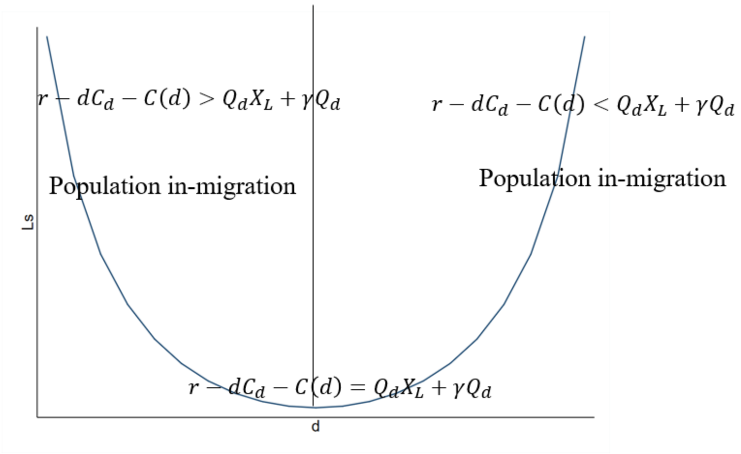

**Figure 4.1.2.1-1 A simplified graph of Ls and d**

We define the migration status of place S by taking the difference between the population in S ($L_S$) and the population in N ($L_N$). That is,

$$\text{Migration status function (MSF)} = L_S - L_N$$

As there are only two places in this model, a larger population in S would indicate a smaller size of population in N, meaning an in-migration occurs from N to S.

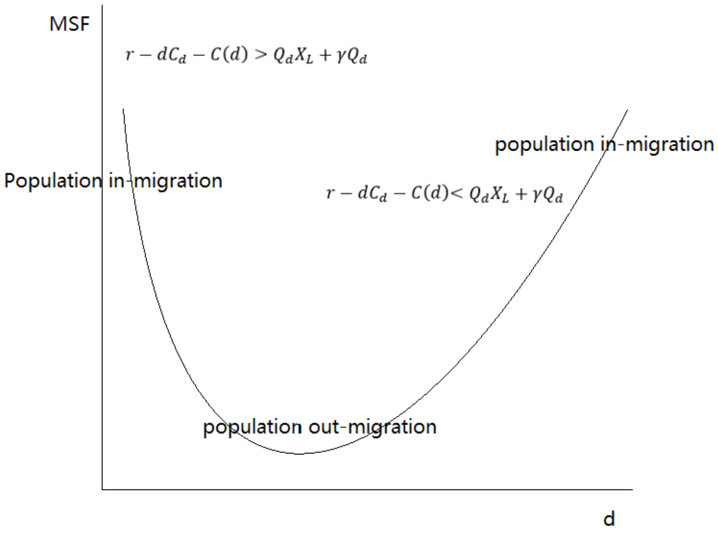

**Figure 4.1.2.1-2 A simplified graph of MSF and d**

A larger MSF (An increasing MSF) shows a larger degree of in-migration in S. The derivative of $L_S$ with respect to d is a linear function of the derivative of MSF with respect to d. To pin down the sign of $\frac{dMSF}{dd}$, we must determine the sign of the term $(r - dc_d - c(d) - Q_dX_L - \gamma Q_d)$. $[r - dc_d - c(d)]$ is the marginal cost of households in S to buy heating from the plant in N by 1-kilometer-increase. $[dc_d + c(d)]$ captures the reduction in the marginal cost of the milder winter temperature due to the one-unit increase in d. In other words, households will spend $[r - dc_d -$

c(d)] for one unit of d farther from plant N if they do not choose to immigrate. The marginal benefit $(\mathrm{Q_dX_L} + \gamma\mathrm{Q_d})$ for households in S by a 1-km increase is the sum of the gain of wages owing to more production in X and the gain of utility because of improved air quality.

Then, we compare these two parts. If the marginal benefit of moving outside S equals the marginal cost, people have no incentive to move. If the marginal cost is larger than the marginal benefit, migration of 1 unit southward will generate welfare loss. Therefore, increasing d leads to a decreasing level of in-migration. If the marginal benefit exceeds the marginal cost, migration of 1 unit southward will bring about utility gain; the increasing d will lead to an increasing level of in-migration.

In reality, N and S represent two cities from regions with and without central heating. d models the distance between them. The Huai River policy sets the borderline. We assume N is right on the heating borderline. Then, d can be modeled as the distance between S and the heating borderline. The problem has been transferred from a two-city model into the setup of the borderline and many southern cities with different values of d.

We have the following inequalities:

$$If\ r - dC_d - C(d) > Q_dX_L + \gamma Q_d,\ \ \frac{dL_S}{dd} < 0$$

$$If\ r - dC_d - C(d) < Q_dX_L + \gamma Q_d,\ \ \frac{dL_S}{dd} > 0$$

Through the above analysis, when d is very small, meaning that S is very near to the heating borderline, MSF is large, showing a large degree of in-migration to the south. Then, MSF has a first decreasing trend, then an increasing trend. MSF reaches its minimum level, showing a decreasing trend of in-migration to the south. At last, MSF becomes larger again, showing an increasing in-migration trend to the south.

In summary, the model suggests an increase in population in-migration to the south across the heating borderline. The areas south of the borderline will have a U-shaped pattern in in-migration with the increase of d.

**4.1.2.2 Tax rate t**

We get the following derivatives and analyze their signs

$$\frac{dL_N}{dt} = -\frac{1}{D}\{pH_{LA}[(L - L_S)Q(d,A)X_{LL} - w_H]\} < 0$$

$$\frac{dL_S}{dt} = -\frac{p}{D}(H_{LL}(L - L_S)(Q_AX_L + \gamma + \gamma A) + H_{LA}w_H) > 0$$

$$\frac{dA}{dt} = \frac{p}{D}H_{LL}[(L - L_S)Q(d,A)X_{LL} - w_H] < 0$$

The most noticeable effects are the relationship between A and t. $\mathrm{X_{LL}}$ is negative, $\mathrm{H_{LL}}$ is negative, and D is also negative, so $\frac{dA}{dt} < 0$, indicating that an increasing t would lead to a decreasing A. t is the tax on air-polluting goods in producing heating. t could influence the production of H by reducing the input of polluting goods.

For the relationship between $L_N$ and t, $H_{LA}$>0, $X_{LL}$<0, and finally $\frac{dL_N}{dt}$<0. This shows us that an increasing t would engender a decreasing labor input in the production of H, which is evident to see as the tax on air pollution could harm the production of H and thus influence the labor input in the heating sector in N.

In the equation of the relationship between $L_S$ and t, $Q_A$<0, $H_{LL}$<0, $X_L$>0, $H_{LA}$>0, $\frac{dL_S}{t}$>0, which is easy to interpret. As t goes up, the production of H is abated, and the emissions of air pollutants are reduced, which will boost the production of X; thus, more workers will immigrate from the N to S.

We make the summary as follows:

The rise of tax rate undermines the production in N thus impairs the heating industry in N. Households will choose to move from N to S. More broadly speaking, the environmental policy will harm the production in the north and lead to the in-migration of population to the south. The change in tax rate affects the migration decisions of households in N, while other parameters affect the migration decisions of households in S.

### 4.1.2.3 Extra cost r

Firstly, we find out the four derivatives with respect to r.

$$\frac{dL_N}{dr} = 0$$

$$\frac{dL_S}{dr} = \frac{p^2}{D}\,(L - L_S)d\left(H_{LL}H_{AA} - {H_{AL}}^2\right) < 0$$

$$\frac{dA}{dr} = 0$$

r stands for the extra costs that households have to pay for heating. In reality, place S is the place that is located south of the heating borderline, which has no access to heating services but has to buy extra heating facilities and pay for fees for heating. An increased r stands for a more expensive expenditure on heating in winter and engenders the decrease of labor input in place S, which shows that a higher r would lead to migration from place S to place N. Households will give up buying heating through paying rd but choose to move directly from S to N in order to reduce their expenditures on heating.

Similarly, if r becomes smaller, households will move from N to S.

We make a summary.

The increase in extra heating costs will lead to higher motivation of households to immigrate to areas with central heating systems to raise their overall welfare.

### 4.1.2.4 Environmental awareness γ on migration

The derivatives with respect to γ are

$$\frac{dL_N}{d\gamma} = 0$$

$$\frac{dL_S}{d\gamma} = -\frac{1}{D}[p^2(L - L_S)[(A + Q(d, A) - E](H_{LL}H_{AA} - {H_{AL}}^2)] > 0$$

$$\frac{dA}{d\gamma} = 0$$

It’s obvious to know that

$$A + Q(d, A) - E = A + E - A(1 - \delta)^d - E = A[1 - (1 - \delta)^d]$$

The signs of the above derivatives are easy to find out.

With the increase of the environmental awareness of households, more people will choose to move from the air-polluting area to the place where air quality is better. The importance of air quality level will be enhanced when $\gamma$ becomes larger. Hence more people will move from N to S. The unemployment rate in N will decline. Suppose $\gamma$ becomes very small, indicating people will not care air pollution any more, they would have large incentives to move to the area with heating facilities.

According to our model setting, the increase of environmental awareness of households will not change the air pollution level.

We make a summary as follows:

The increase in environmental awareness will enlarge the weight of good air quality in households’ minds, thus leading to possible in-migration to the south.

In the next part, we will conduct simulations to verify the results theoretically analyzed above.

#### 4.1.3 Model simulation

#### 4.1.3.1 Functional forms

As $L_H$ and A do not have specific substitution effects and the effects of air pollution on the production of H are fixed, we endow the following functional form to the production function.

$$\mathrm{H}(\mathrm{L_H}, \mathrm{A}) = \mathrm{A}^{\tau}\mathrm{L_H^{\theta}}\ (0 < \theta < 1, 0 < \tau < 1)$$

$$\mathrm{X}(\mathrm{L_X}) = \mathrm{L_X^{\beta}}\ (0 < \beta < 1)$$

Q(d,A) takes the usual form stated in 3.1.1.

We check whether those function forms satisfy the pre-set conditions of the model. Recall the conditions we assume in 3.1.

$H_L>0$ $H_A>0$ $H_{LL}<0$ $H_{LL}H_{AA}-{H_{LA}}^2>0$

The first three inequalities are as follows

$$\frac{\partial \mathrm{H}}{\partial \mathrm{L_H}} = \theta \mathrm{A}^{\tau}\mathrm{L_H^{\theta-1}} > 0 \qquad \frac{\partial \mathrm{H}}{\partial \mathrm{A}} = \tau \mathrm{A}^{\tau-1}\mathrm{L_H^{\theta}} > 0$$

$$\frac{\partial^2 \mathrm{H}}{\partial \mathrm{L_H^2}} = \theta(\theta - 1)\mathrm{A}^{\tau}\mathrm{L_H^{\theta-2}} < 0$$

It’s easy to find that the above three inequalities can be held when θ and $\tau$ are both positive.

But for the last inequality, we need to do the following calculation

$$\frac{\partial^2 \mathrm{H}}{\partial \mathrm{L}_\mathrm{H}^2}\frac{\partial^2 \mathrm{H}}{\partial \mathrm{A}^2} - \frac{\partial^2 \mathrm{H}}{\partial \mathrm{L}_\mathrm{H}\,\partial \mathrm{A}} = -\theta\tau(\theta+\tau-1)(A^{\tau-1}{\mathrm{L}_\mathrm{H}}^{\tau-1})^2$$

To keep the equation positive, the following condition should be satisfied

$$\theta + \tau < 1$$

When doing simulations, we need to take account of the above inequality to make sure the preset model conditions are fulfilled.

We assume the heating cost reduction function (climate function) represents a gradual change in the winter temperature. For simplicity, we assume a linear function of C(d)=d.

**4.1.3.2 Equilibrium conditions**

We gather all different conditions together as follows:

$$p\theta A^{\tau} L_H^{\theta-1} = w_H$$

$$p\tau A^{\tau-1} L_H^{\theta} = t$$

$$\beta\left[\left(\mathrm{E} - \mathrm{A}(1-\delta)^{\mathrm{d}}\right]\mathrm{L}_\mathrm{X}^{\beta-1} = w_X$$

$$\mathrm{w_H} = \mathrm{w_X} - \mathrm{rd} + \mathrm{c(d)d} + \gamma\mathrm{A}\left[1-(1-\delta)^{\mathrm{d}}\right]$$

$$\mathrm{L} = \mathrm{L_N} + \mathrm{L_S}$$

Then, the equation for determining the relationship between $L_S$ and d after simplification is

$$p^{-\frac{\tau}{\tau-1}}\theta\left(\frac{t}{\tau}\right)^{\frac{\tau}{\tau-1}}(L-L_S)^{\frac{\theta+\tau-1}{\tau-1}} = \beta\left[E-\left[\frac{t}{p\tau(L-L_S)^{\theta}}\right]^{\frac{1}{\tau-1}}(1-\delta)^{d}\right]L_S^{\beta-1} - rd + d^2 + \gamma\left[\frac{t}{p\tau(L-L_S)^{\theta}}\right]^{\frac{1}{\tau-1}}\left[1-(1-\delta)^{\mathrm{d}}\right] \quad (8)$$

The above equation can be numerically solved by a bisection method on computers.

We define the following function

$$F(L_S,d) = p^{-\frac{\tau}{\tau-1}}\theta\left(\frac{t}{\tau}\right)^{\frac{\tau}{\tau-1}}(L-L_S)^{\frac{\theta+\tau-1}{\tau-1}} - \beta\left[E-\left[\frac{t}{p\tau(L-L_S)^{\theta}}\right]^{\frac{1}{\tau-1}}(1-\delta)^{d}\right]L_S^{\beta-1} + rd - d^2 - \gamma\left[\frac{t}{p\tau(L-L_S)^{\theta}}\right]^{\frac{1}{\tau-1}}\left[1-(1-\delta)^{\mathrm{d}}\right]$$

We could solve the above equation $F(L_S,d)=0$ by looking for the intersection between the function graph $F(L_S,d)$ and x-axis. Figure 4.1.3.2-1 illustrates the method.

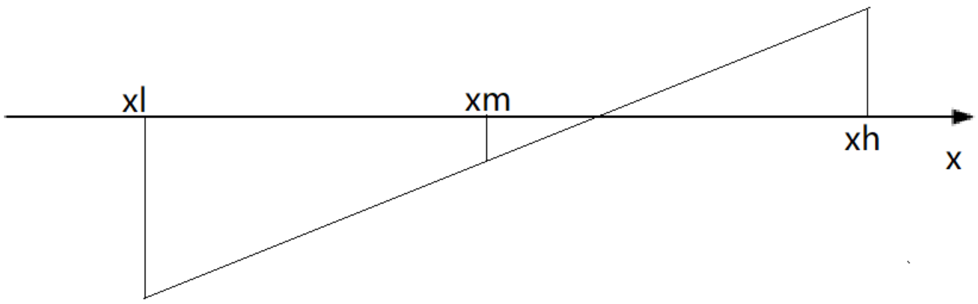

**Figure 4.1.3.2-1 Finding the numerical solutions to complicated solutions**

### 4.1.3.3 $L_S$ and d

The simulation values for all parameters are listed as follows:

t=20; p=20; θ=0.2; τ=0.1; L=1000; E=10000; δ=0.0001; β=0.2; r=10; γ=1.

For the above parameters, p=20 refers to North China's rough average unit price for heating. The total labor input is set as 1000 persons. The tax for each polluting unit is set as 20, and the extra cost r is half the price to buy heating from N (=10). The environmental status is modeled as 10000. The distance d has a unit of 100 km.

In this model, we assume that 1000 people are distributed between N and S. The MSF-d curve measures the change in population migration in the southern city S with the evolution of d. A larger MSF value stands for a higher probability of population in-migration in S.

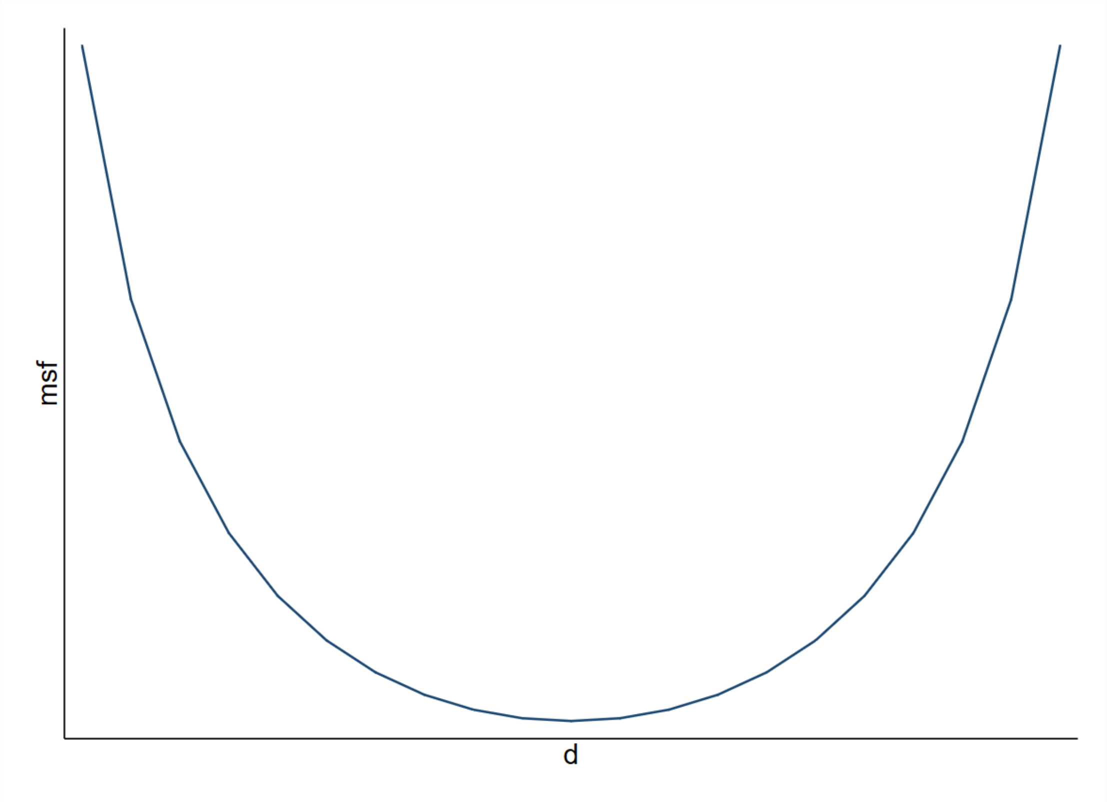

**Figure 4.1.3.3-2 Simulation between MSF and d**

The following conclusion can be drawn from the simulation with many parameters.

When d is very small, MSF is very large, showing a high degree of in-migration to the south. When d gradually grows, the population inflow of S declines; when d is larger than some point, people start to move from N to S, and the population in-migration to the south increases. Population influx begins, corresponding to the rise of $L_S$ and the drop of $L_N$. According to our calculation, MSF reaches the lowest point at around d=5(00) km under this situation.

This simulation suggests a U-shaped pattern of population migration in the south. Areas near the heating boundary will have a declining in-migration pattern and then increase gradually.

### 4.1.3.4 Parameter γ and δ for air quality improvement

In our previous analysis, we assume that the air quality level will be improved when d increases. This conforms to the reality. Southern areas are farther away from the heating areas, thus have a better air quality level. γ and δ are two parameters that measures the importance of air quality. γ is the households' overall attitudes towards the air quality. The higher γ is, the more importantly the households will think of the air quality levels. δ measures the decreasing rate of air pollutants level. A higher δ would indicate a sharper decline of air pollutants.

We will simulate the relationship between $L_S$ and $\gamma$. Theoretically, as we analyzed, an increasing $\gamma$ would suggest an increasing environmental awareness. Households would like to move to areas with cleaner air. An increasing $\gamma$ would generate an increasing in-migration of the south.

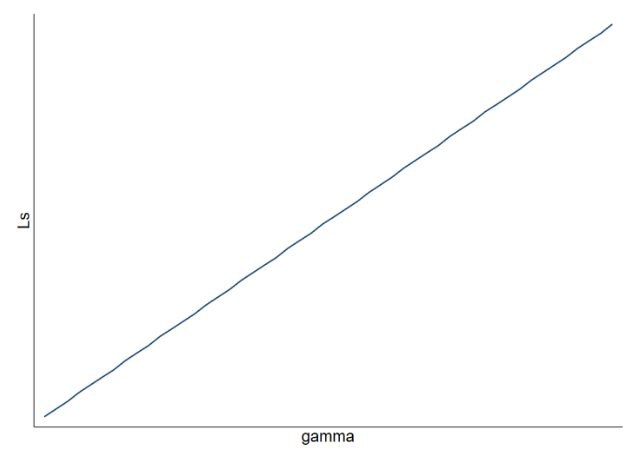

**Figure 4.1.3.4-1 Simulation between $\gamma$ and $L_S$ with parameter d=1.Other parameters same as the benchmark.**

A higher $\delta$ would suggest a higher decreasing rate of air pollutants. Theoretically, a higher would attract more population in-migration in the south. Figure 4.1.3.4-2 shows a straight increasing line.

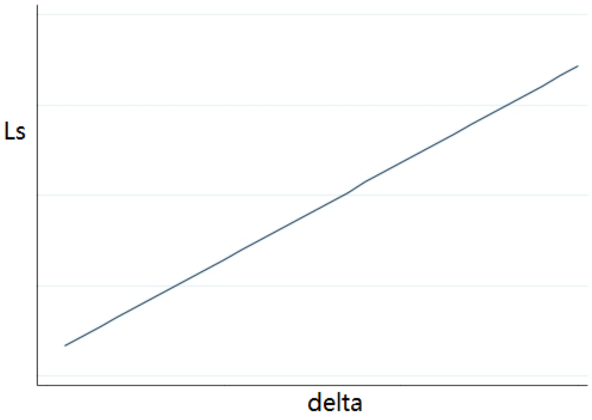

**Figure 4.1.3.4-2 Simulation between $\delta$ and $L_S$ with parameter d=1.Other parameters same as the benchmark.**

### 4.1.3.5 Parameter r

In this theoretical framework, r is modeled as the extra cost for Southerners to get central heating. In reality, the Southerners do not always need to equip central heating systems in their rooms due to milder winter temperatures. The cost of buying air conditioners or electric heaters is lower compared to introducing central heating systems. Some of the households even choose not to use any heating systems. In this setting, no matter where the households come from, the north or the south, they all should pay for a unit price p to buy heating services. In order to match the actual situation, we assume r could have negative values. The negative r suggests that people should bear lower costs in the south. This is consistent with some situations where southerners do not buy heating equipment in their room.

We simulate the relationship between $L_S$ and r. The downward trend suggests a higher r would engender a lower in-migration trend in the south.

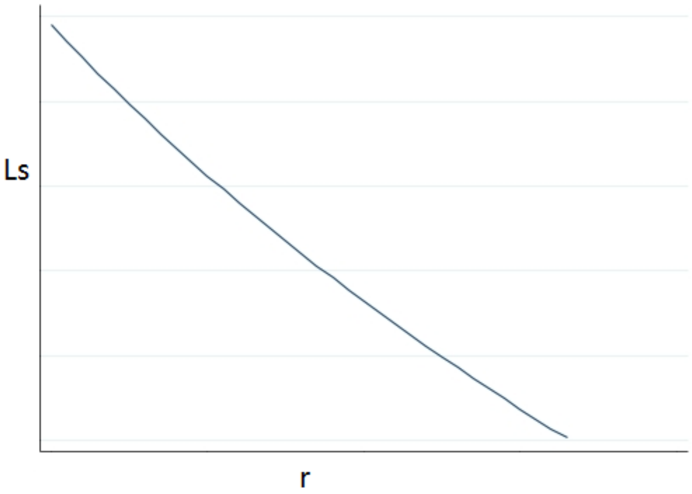

**Figure 4.1.3.5-1 Simulation between r and $L_S$ with parameter d=1.Other parameters same as the benchmark.**

### 4.1.3.6 Parameter t

According to the theoretical analysis, t is the environmental tax rate on air pollution in the north. When the tax rate increases, the development of factory N will be restrained. The southern in-migration will be enhanced. This is the same as what we analyzed in the theoretical part.

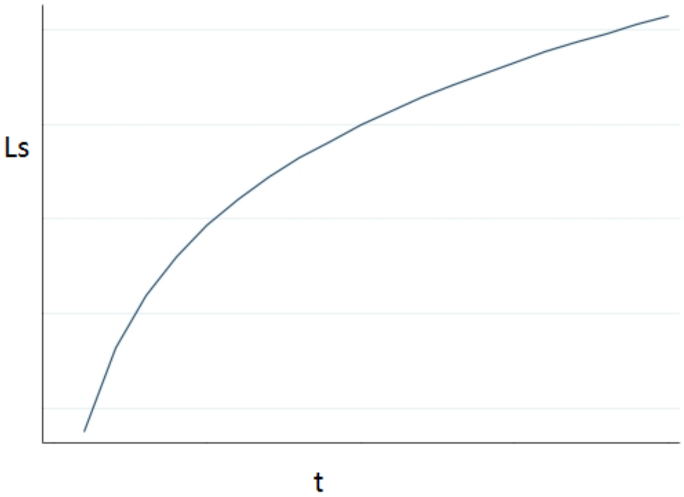

**Figure 4.1.3.6-1 Simulation between t and $L_S$ with parameter d=1.Other parameters same as the benchmark.**

### 4.1.4 Summary

The U-shaped relationship is determined by air pollution, extra costs, and room comfort. As there are only two places in the model, a higher $L_S$ will indicate a lower $L_N$. In the model, when d is very small (equal 0), $L_S$ (MSF) is very large, showing a high degree of in-migration to the south. When d is becoming larger, the effects of air pollution and living comfort change. The gain of having larger room comfort is on the rise, making population flow out from the south. When d is even larger, the temperature becomes higher. The weight of room comfort decreases, making the population flow into the south.

The model is only the simulation of the real world. It is not the real world. It cannot replicate everything in the real world. This model is merely a framework that we can put some parameters into to anzlyze the correlation between variables. There are many drawbacks of the model. It is hard to distinguish the welfare change in the real world. Some of the findings are counterintuitive. The model needs to be improved in future studies.

## 4.2 Empirical study: U-shaped relationship

### 4.2.1 Introduction

As we mentioned in Section 4.1, the relationship between the Huai River policy and migration is a research gap. Three factors affected by the Huai River policy can affect migration. These three factors are heating costs, air quality, and indoor comfort.

Section 4.1 is the theoretical model for the relationship between central heating and population migration. The model predicts an in-migration status near the borderline and a U-shaped relationship south of the heating borderline. In this chapter, we will conduct two empirical studies verifying the predictions in the model. Section 4.2 verifies the U-shaped relationship south of the heating borderline. Section 4.3 focuses on the south-north discrepancies in population in-migration, which uses the county-level data near the heating borderline.

### 4.2.2 Basic methodology

The simulation in Section 4.1 indicates that the degree of in-migration to the south is large when d is small, and the in-migration intensity will be weaker and weaker until it reaches a minimum; then the situation is gradually inversed with in-migration to the south larger and larger. As is inferred in the model, population in-migration and d have a U-shaped relationship.

Theoretically, by solving Equation (8), we obtain a solution $L_S(d, t, p, w_H, \delta, r, \gamma)$. Thus, MSF, a function of $L_S$, is also a composite function of the above parameters.

We construct the following regression equation according to the model:

$$\mathrm{DPR}_{it} = \alpha + \mu_1 d_i + \mu_2 (d_i)^2 + \tau \mathrm{Longitude}_i + \rho \mathrm{Greeningrate}_{it} + \sigma \mathrm{DI}_{it} + \omega \mathrm{Downtownarea}_{it} + \varphi \mathrm{N}_{it} + u_i + \varepsilon_{it}$$

This is a "mixed" panel with time-invariant variables of d, $d^2$, N, and Longitude and time-variant variables of Greeningrate, Downtownarea, and DI. d and the quadratic term of d ($d^2$) are introduced into the model to capture the U-shaped relationship between migration and d. $u_i$ stands for individual effects in the model. $\varepsilon_{it}$ is the random error term.

The following presents the explanations of the variables in the regression model.

**Explained variable**

In China, two methods are used to count the population: registered population (RP) and permanent population (PP). The registered population refers to the population recorded in the population registration system of the local police station. The permanent population refers to the actual population living locally as surveyed by a population census or a sampling investigation. The permanent population of some place W includes the following four groups of people:

(1) People who are registered in the population registration system of W and live in W.

(2) People who are registered in the population registration system of W but who leave W for less than half a year.

(3) People who are registered in the population registration system of W but who live abroad for further study and plan to return to W after graduation.

(4) People who live in W for more than half a year but are not registered in the population registration system of W.

Here, DPR (the difference between PP and RP) is used as the proxy variable of population in-migration to measure the in-migration status of places in the south, that is, the difference between the permanent population and the registered population.

$$\mathrm{DPR}_{\mathrm{it}} = \mathrm{PP}_{\mathrm{it}} - \mathrm{RP}_{\mathrm{it}}$$

A region with a positive DPR indicates that the permanent population is larger than the registered population, showing that a group of unregistered migrant workers flows into the region. In this sense, the region has population in-migration. Otherwise, the negative DPR shows a population out-migration. A higher DPR shows a higher degree of population in-migration.

**Explanatory variables:**

d is the key variable to explain population migration. d is modeled as the straight-line distance to the heating borderline. It is assumed that N is located right on the borderline. To capture the characteristics of the U-shaped curve, we include d and $d^2$ in the regression equation. Both d and $d^2$ are time-invariant variables.

Greeningrate is the environmental proxy variable for two parameters $\delta$ and $\gamma$. $\delta$ captures the reduction rate of the air pollution level. $\gamma$ captures the environmental awareness of people living in S. In most cases, places with a higher downtown greening rate will have good air quality levels and more heightened environmental awareness for residents. Downtownarea controls the scale of the cities.

DI means the per capita disposable income for southern cities. It is the proxy variable for $w_H$. Migration is influenced by relative wages between N and S. A rise in wages in the north essentially represents a fall in wages in the south. Therefore, we use the wages of the south instead of the north.

Longitude is a variable used to measure the location and natural conditions. Longitude does not change according to the timeline.

In the model, t, r, and p do not appear in the regression equation. These variables remain stable and can be ignored in the analysis.

### 4.2.3 Data

The dataset contains 125 cities south of the borderline, with a time span from 2011 to 2020. The dataset includes cities of various economic levels and winter weather conditions from warm to cold. It only contains cities of medium size and prefecture level. Megacities as economic hubs are excluded because they are natural centers with high population densities. The selected cities have relatively equal job opportunities, local infrastructure, and other resources.

**Table 4.2.3-2 Cities that are included in the dataset**

| Province | Cities included |
|---|---|
| Jiangsu | Suzhou, Wuxi, Nantong, Zhenjiang, Yangzhou, Taizhou, Changzhou |
| Zhejiang | Jiaxing, Shaoxing, Taizhou, Zhoushan, Huzhou |
| Fujian | Ningde, Quanzhou, Nanping, Putian, Longyan, Sanming, Zhangzhou |
| Jiangxi | Jingdezhen, Jiujiang, Xinyu, Yingtan, Fuzhou, Shangrao, Pingxiang, Yichun, Ji’an |
| Anhui | Bengbu, Huaibei, Anqing, Chuzhou, Suzhou, Lu’an, Bozhou, Chizhou, Xuancheng, Maanshan, Huangshan, Fuyang, Huainan, Wuhu |
| Guangdong | Shaoguan, Shantou, Maoming, Zhaoqing, Shanwei, Heyuan, Qingyuan, Jieyang, Yunfu, Zhuhai, Huizhou, Zhongshan, Jiangmen, Dongguan, Zhanjiang, Meizhou, Foshan, Chaozhou, Yangjiang, |
| Guangxi | Nanning, Wuzhou, Qinzhou, Guigang, Yulin, Baise, Hezhou, Hechi, Laibin, Chongzuo, Fangchenggang, Beihai, Guilin, Liuzhou |
| Sichuan | Zigong, Panzhihua, Deyang, Mianyang, Guangyuan, Neijiang, Leshan, Nanchong, Meishan Guang’an, Dazhou, Bazhong, Ziyang, Aba, Liupanshui, Luzhou, Yibin, Ya’an, Suining |
| Guizhou | Zunyi, Anshun, Qianxinan, Bijie, Tongren |
| Hubei | Huangshi, Shiyan, Yichang, Ezhou, Jingmen, Jingzhou, Huanggang, Xianning, Suizhou, Xiangyang, Xiaogan |
| Hunan | Zhuzhou, Xiangtan, Hengyang, Shaoyang, Yueyang, Changde, Zhangjiajie, Yiyang, Chenzhou, Yongzhou, Huaihua, Loudi |
| Henan | Xinyang |
| Shaanxi | Hanzhong, Ankang |

Distance (d) maps the straight-line (vertical) distance from the city to the borderline and does not change over time. We use the latitudes of cities to estimate the distance. Combined with the exact latitudes of the heating borderline, 34°N is our approximate benchmark latitude line, which coincides with our heating borderline. Each degree of latitude is approximately 111 km apart. d is then calculated as follows:

$$\mathrm{d} = (\text{actual latitude} - 34) \times 111$$

The formula is a linear combination of the actual latitude and does not affect the estimation results. The longitude and latitude data of the city can be captured from the Internet. The data on the variables Greeningrate and DI for urban residents can also be obtained from the statistical yearbooks of cities published by the provincial-level bureau of statistics.

The descriptive statistics for all variables are recorded in Table 5.1.2-3.

**Table 5.1.2-3 Descriptive statistics**

| **Variables** | **Obs** | **Mean** | **Std. Dev.** | **Min.** | **Max.** |
|---|---|---|---|---|---|
| DPR ($10^4$ persons) | 1,185 | -30.08849 | 128.9367 | -522.96 | 826.91 |
| Distance (d, km) | 1,185 | 670.485 | 396.3523 | 5.55 | 1413.03 |
| Longitude | 1,185 | 112.8205 | 5.00997 | 101.72 | 122.2 |
| Greeningrate (100%*100) | 1,185 | 40.54192 | 4.997261 | 14.76 | 58.11 |
| Downtownarea ($km^2$) | 1,185 | 100.9924 | 105.8514 | 5.06 | 1194.31 |
| Disposable income (DI) (yuan) | 1,185 | 28935.08 | 9358.247 | 13584 | 70966 |

### 4.2.4 Hausman-Taylor estimator

#### 4.2.4.1 Introduction to the method

The Hausman-Taylor estimator is one estimation method of panel data with time-invariant variables. It can also solve the endogeneity problems. The basic idea of the model is as follows.

Suppose we have two matrices of independent variables X and Z. X=$[X_1\ X_2]'$ and Z=$[Z_1\ Z_2]'$. $X_1$ and $Z_1$ are exogenous variables and $X_2$ and $Z_2$ are endogenous variables. The endogenous variables here means that they are correlated with $\mu_i$. We have the following model

$$\mathrm{y_{it}} = \boldsymbol{X}'_{it}\beta + \mathbf{Z_{1i}}\gamma + \mu_i + \varepsilon_{it}$$

Specifically,

$$\mathrm{y_{it}} = \boldsymbol{X}'_{1it}\beta_1 + \mathbf{X}'_{2it}\beta_2 + \mathbf{Z_{1i}}\gamma_1 + \boldsymbol{Z}_{2i}\gamma_2 + \mu_i + \varepsilon_{it}$$

We divide the independent variables into the following four groups: (1) $X_{1it}$ is the group of time-variant variables not correlated with $\mu_i$. This group is called time-variant exogenous variables. (2) $X_{2it}$ is the group of time-variant variables correlated with $\mu_i$. This group is called time-variant endogenous variables. (3)$Z_{1i}$ are time-invariant variable not correlated with $\mu_i$ called exogenous time-invariant variables. (4) $Z_{2i}$ are time-invariant variables correlated with $\mu_i$ called endogenous time-invariant variables.

The estimation can be conducted via the following steps.

First, perform the within regression with all time-variant variables and derive the coefficients of $\beta_1$ and $\beta_2$. Calculate the intra-temporal mean of the residuals in this regression denoted by $\tilde{e}$.

$$(y_{it} - \bar{y}_i) = \beta_1(x_{1it} - \bar{x}_{1i}) + \beta_2(x_{2it} - \bar{x}_{2i}) + (\varepsilon_{it} - \bar{\varepsilon}_i)$$

$$\tilde{e} = (y_{it} - \bar{y}) - \beta_1(x_{1it} - \bar{x}_{1i}) - \beta_2(x_{2it} - \bar{x}_{2i}) - (\varepsilon_{it} - \bar{\varepsilon_\iota})$$

We can estimate the variance of idiosyncratic error term based on the above estimation.

$$\hat{\sigma}_\varepsilon^2 = \frac{RSS}{N - n}$$

RSS is the residual sum of squares.

Second, conduct a 2-stage least square estimation (2-SLS) with the following pattern:

Take $\tilde{e}$ as the dependent variable and conduct the 2-SLS on the following variable settings. Regress $\tilde{e}$ on $Z_{1i}$ and $Z_{2i}$ taking $X_1$ and $Z_1$ as instruments and derive the following IV estimators.

$$\gamma_{\mathrm{IV}} = (\mathrm{Z'P_AZ})^{-1}\mathrm{Z'P_A}\hat{\mathrm{d}}$$

Where $P_A = A(A'A)^{-1}A'$ and $A=[X_1,Z_1]$ is a set of instruments.

Using the above results, we can derive the estimate of variance of μ, $\hat{\sigma}_\mu{}^2$.

Then we work out the weight parameter in feasible GLS regression using the following formula:

$$\theta = \sqrt{\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + T^{-1}\hat{\sigma}_\mu{}^2}}$$

The next step is to do a random transformation for all variables y, X, and Z using the parameter θ. For example, we could transform y like

$$y^* = y_{it} - \theta y_{it}$$

At last, we get the following equation

$$y_{it}^* = X_{it}^*\beta + Z_{it}^*\gamma + \varepsilon_{it}^* \qquad (9)$$

We estimate the equation (9) using IV estimation with instruments $\tilde{X}_{it}$ (within transformed $X_{it} - \bar{X}$), $Z_{1i}$, and $\bar{X}_{1i}$.

**4.2.4.2 Benchmark estimation**

As d, $d^2$, and longitude are three exogenous geographical factors that do not change according to time, it is impossible to perform a fixed-effect estimation. Considering that some endogenous variables may be correlated with the variable of fixed effects in the model, we perform a Hausman-Taylor estimator.

The regression equation is as follows:

$$\mathrm{DPR_{it}} = \mu_1 \mathrm{d_i} + \mu_2(\mathrm{d_i})^2 + \tau\mathrm{Longitude_i} + \rho\mathrm{Greeningrate_{it}} + \sigma\mathrm{DI_{it}} + \omega\mathrm{Downtownarea}_{it} + \mathrm{u_i} \\ + \varepsilon_{\mathrm{it}}$$

To perform a Hausman-Taylor estimator, we need to run a Hansen-Sargan test of overidentifying restrictions to verify the validity of the instruments. By running this test in this situation, we obtain a p value larger than 0.1, which rejects $H_0$, showing the good validity of the instruments.

$d_i$, $d_i^2$, and $\mathrm{Longtitude_i}$ are three exogenous time-invariant independent variables. $\mathrm{Greeningrate_{it}}$ is the exogenous time-variant independent variable representing residents' awareness and the air quality level in the original model. $DI_{it}$ is an endogenous time-variant independent variable. The regression results are shown in Table 4.2.4.2-1.

Regression I has a positive coefficient quadratic term and a negative linear term, but they are not statistically significant. If d is set at an interval larger than 250 (regression II), the coefficients are more significant. One explanation for this is that the central heating borderline is fuzzy. 34° N is an approximation of the heating borderline. The heating borderline zigzags and influenced by administrative divisions of local governments.

Longitude has a positive coefficient and is significant at a 1% level in the four regressions. Greeningrate has a negative coefficient. Generally, a higher greening rate usually means a better environment with more attraction for people. Here, the negative correlation may arise from the negative correlation between the greening rate and economic development. Therefore, areas with a higher greening rate are places that are more remote or less developed. In regressions I and II, DI has a positive connection with DPR, showing a positive impact.

In contrast to the results of the Hausman-Taylor estimation, a pooled OLS method is used to verify the results. $d^2$ has a positive coefficient, while d has a negative coefficient, which implies a U-shaped relationship. All coefficients of DI are positive and significant. All coefficients of Longitude are positive. The coefficients of Greeningrate in regressions of DPR are positive, showing a good connection between the natural environment and population in-migration.

The above estimation results prove the existence of a positive U-shaped relationship. The Hausman-Taylor estimation method considers the endogeneity problem and solves the shortcoming that individual effects cannot be estimated; hence, the regression results are more convincing.

**Table 4.2.4.2-1 The Hausman-Taylor estimation results**

| **Hausman-Taylor** | **Regression I: DPR (Difference between the PP and RP)** | **Regression II: DPR (distance>250)** |
|---|---|---|
| TIexogenous | | |
| d | -0.0333894 | -0.3849075* |
| (standard error) | (0.1213906) | (0.2012025) |
| $d^2$ | 0.0000557 | 0.0002549** |
| | (0.0000808) | (0.0001213) |
| Longitude | 8.756895*** | 10.61364*** |
| | (2.557981) | (2.868852) |
| Narate | ------------- | -------------- |
| TVexogenous | | |
| Greeningrate | -0.7427467*** | -0.7351811*** |
| | (0.1861152) | (0.1944103) |
| TVendogenous | | |

| | | |
|---|---|---|
| DI | 0.0004792*** | 0.0004813*** |
| | (0.0000842) | (0.0000929) |
| Constant | -1013.218*** | -1089.603*** |
| | (297.5498) | (338.1621) |
| Obs | 1185 | 984 |

**Table 4.2.4.2-2 Estimation results of the pooled OLS**

| Pooled OLS | Regression I: DPR (Difference between the PP and RP) | Regression II: DPR (distance>250) |
|---|---|---|
| d | -0.0782239** | -.3273513*** |
| | (0.7216317) | (0.0597889) |
| $d^2$ | 0.0000835*** | 0.0002241*** |
| | (0.0000236) | (0.000036) |
| Longitude | 2.356555*** | 4.14387*** |
| | (0.7216317) | (0.8138986) |
| Narate | ----------- | ----------- |
| Greeningrate | 2.604994*** | 2.649437*** |
| | (0.7049775) | (0.7428471) |
| DI | 0.0060437*** | 0.0058665*** |
| | (0.0003693) | (0.0004153) |
| Constant | -574.8324*** | -677.6136*** |
| | (76.19267) | (86.38958) |
| Obs | 1185 | 984 |

For robustness checks, we conduct an Amemiya-Macurdy (1986) estimator to verify the results in Table 4.2.4-3. As the estimation requires strongly balanced panel data, we delete the data from 2011 and 2020, which have missing data for some cities. Thus, the modified dataset is a strongly balanced panel spanning from 2012 to 2019. The new dataset deletes the data of 2020, which is the year of the COVID-19 outbreak. As large-scale lockdowns occurred in that year, its omission can help reduce the possible disturbances on population migration. The signs of the coefficients remain unchanged. As the dataset changes, the values change.

**Table 4.2.4.2-3 Estimation results: Amemiya-MaCurdy estimator**

| | **Regression I: DPR (Difference between** | **Regression II: DPR (distance>250)** |
|---|---|---|

| | the PP and RP) | |
|---|---|---|
| d | -0.066112 | -0.4223605** |
| | (0.1188512) | (0.1952225) |
| $d^2$ | 0.0000745 | 0.0002765** |
| | (0.0000792) | (0.0001179) |
| Longitude | 7.023268 *** | 8.837737*** |
| | (2.157738) | (2.420578) |
| Narate | ----------- | ----------- |
| Greeningrate | -0.5979763*** | -0.608677*** |
| | (0.1999587) | (0.2083955) |
| DI | 0.0060437*** | 0.0058665*** |
| | (0.0003693) | (0.0004153) |
| Constant | -574.8324*** | -677.6136*** |
| | (76.19267) | (86.38958) |
| Obs | 984 | 824 |

### 4.2.5 Breakpoint regressions

In this section, we further introduce another technique to verify the U-shaped relationship as robustness checks. The idea comes mainly from Simonsohn (2018) [42]. A pooled OLS is first performed between the dependent variables (DPR) with d and $d^2$ for the panel dataset. Then, the lowest point of d is calculated with the formula $d_{min} = -\frac{b}{2a}$. (It is assumed that the second-order relationship is $ad^2+bd+c=0$). The estimation results are listed in Table 4.2.5-1.

We construct the following three variables: dhigh, dlow, and low. The following are how they are generated:

$$d_{high} = d - d_{min} \text{ if } d \geq d_{min}, \text{otherwise } d_{high} = 0$$

$$d_{low} = d - d_{min} \text{ if } d \leq d_{min}, \text{otherwise } d_{low} = 0$$

$$\text{low} = 1 \text{ if } d \geq d_{min} \text{ otherwise low} = 0$$

Then, we regress DPR on dhigh, dlow, and low with the following equation (Step 2):

$$DPR = h + e \cdot d_{high} + f \cdot d_{low} + j \cdot low + \varepsilon$$

Regression results in Table 4.2.5-2 show that the signs of the two variables dhigh and dlow are different, indicating a U-shaped curve between Δ and d.

**Table 4.2.5-1 Breakpoint estimation- Step 1(Panel data)**

| Panel | I: DPR | II:DPR(d>250) |
|---|---|---|
| d | -0.1025521** | -0.4600544*** |
| | (0.041114) | (0.0692316) |
| $d^2$ | .0000911*** | 0.0002925*** |

| | | |
|---|---|---|
| | (0.0000275) | (0.0000418) |
| Constant | -15.08893 | 119.1302*** |
| | (12.47007) | (24.63183) |
| N | 1185 | 984 |

**Table 4.2.5-2 Breakpoint estimation-Step 2 (Panel data)**

| Panel | I: DPR | II:DPR(d>250) |
|---|---|---|
| $d_{high}$ | 0.1384573*** | 0.2139698*** |
| | (0.0202616) | (0.0337304) |
| $d_{low}$ | 0.0737788* | -0.220905*** |
| | (0.0389729) | (0.0381903) |
| low | -80.78148*** | 7.386236 |
| | (14.99923) | (17.91929) |
| N | 1185 | 984 |

To visualize the U-shaped curve, the data for the single year 2017 are selected. The estimation results are reported in Table 4.2.5-3 and Table 4.2.5-4. In the first-step OLS regression, the coefficients of d are negative, and those of $d^2$ are positive. In the second-step regression, two coefficients of dhigh and dlow of different signs are obtained, showing a downward trend at first and then an upward trend. The illustration in Figure 4.2.5-1 could explain the U-shaped relationship as well. Through two fitted lines, the U-shaped relationship can be viewed more intuitively. All regressions show that the population in-migration status is distorted around the 600 km line south of the central heating borderline using the formula $d_{min} = -\frac{b}{2a}$.

**Table 4.2.5-3 Breakpoint regression- Step 1 (cross-sectional data)**

| 2017 | I: DPR | II:DPR(d>250) |
|---|---|---|
| d | -0.1097829 | -0.3518182*** |
| | (0.1286503) | (0.0002318) |
| $d^2$ | 0.0000957 | 0.0002318*** |
| | (0.0000861) | (0.0000381) |
| Constant | -14.02231 | 77.66178*** |
| | (39.15039) | (21.65867) |
| Obs | 124 | 103 |

**Table 4.2.5-4 Breakpoint regression- Step 2 (cross-sectional data)**

| 2017 | I: DPR | II:DPR(d>250) |
|---|---|---|
| $d_{high}$ | 0.1398993** | 0.2139698*** |
| | (0.0633238) | (.0337304) |
| $d_{low}$ | 0.075628 | -0.220905*** |
| | (0.1216081) | (0.0381903) |

| | | |
|---|---|---|
| low | -83.02153 | 7.386236 |
| | (46.34917) | (17.91929) |
| Obs | 124 | 103 |

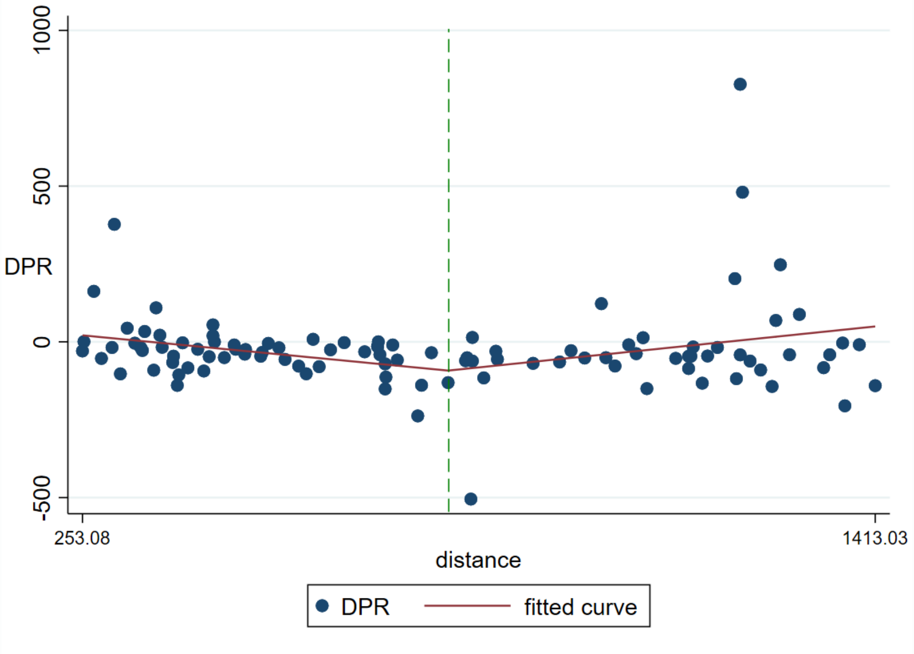

**Figure 4.2.5-1 Scatter diagrams between DPR(d>250) and d (km).**

### 4.2.6 Results

The estimation results can be concluded as follows.

In the areas south of the central heating borderline, the borderline (the Huai River policy) impacts local population migration. Specifically, with the increasing distance d from the borderline southward, the inflow of the permanent population shows a trend of decreasing at first and then increasing; that is, there is a positive U-shaped relationship between population in-migration and distance d. The Huai River policy distorts the original migration equilibrium and engenders a new population migration pattern.

### 4.2.7 Summary

In this section, we set up a “Huai River hypothesis” stating that the heating boundary policy could distort the population patterns of the country. In the model analysis, we assume that N and S are two places north and south of the borderline. The difference between the populations of N and S is modeled as the indicator of the migration status of southern cities. Through computer modeling and simulation, we obtain the following conclusion. With the increase in distance d, the inflow of population in the south will first decline and then rise, indicating a positive U-shaped relationship between the in-migration of population and distance.

The Huai River policy has changed the status of population flow in southern regions. More intuitively, the Huai River policy changes the population distribution in the country. Further analysis shows that the main reason for the positive U-shaped relationship is the distinction generated by the borderline regarding indoor living comfort (room temperature) and outdoor living environment (air quality). In the meantime, the increasing d weakens the impact of the indoor living environment (the farther south the site is, the higher the winter temperature is), making a continuous growing in-migration in the south.

## 4.3 Empirical analysis: migration across the borderline

In section 4.2, we have verified the migration pattern south of the heating borderline. The model also forecasts a high degree of in-migration near the heating borderline. This chapter will verify whether this conclusion holds in the empirical world. We first conduct a survey to determine the exact stretch of the heating borderline. Then, an empirical study is conducted to consider the in-migration status across the heating borderline. This part of the study was published in [43].

### 4.3.1 Heating borderline

We first specify the exact location of the heating borderline. Though the heating borderline is roughly set according to the 0-degree-isotherm and coincides with the Qin-Mountain and Huai River line, the specific location is determined by local governments. We refer to the official statistical yearbooks and government websites to find out the status of central heating in counties near the borderline. To set a criterion, we introduce a parameter heating density, which is defined as the ratio between the central heating area and permanent population.

$$\text{Heating dentisy} = \text{Central heating area}/\text{permanent population}$$

As the heating borderline is determined by prefecture-level government, we include all the prefecture-level cities just beside the central heating borderline. These cities are classified into three groups with different heating densities. Group I contains the cities whose heating density higher than 1. Group II contains cities without any central heating facilities (with heating densities equal to 0). Group III contains cities with heating densities between 0 and 1. Table 4.3.1-1 presents heating densities and groups of all prefecture-level cities near the borderline.

**Table 4.3.1-1 central heating status of cities near the borderline**

| Province | City (Group) | Heating area(1000 $m^2$)[7] | Heating density ($m^2$/person) |
|---|---|---|---|
| Shandong | Rizhao (I) | 24400 (2019) | 8.22 |
| | Linyi (I) | 69551 (2019) | 6.31 |
| | Zaozhuang (I) | 33560 (2019) | 8.71 |
| Jiangsu | Xuzhou (I) | 12010 (2010) | 1.33 |
| | Lianyungang(II),<br>Suqian(II), Huaian (II) | -----------[8] | 0 |
| Henan | Shangqiu (I) | 12000 (2020) | 1.54 |

---

[7] Data are found on the official websites of the local governments.

[8] ---- stands for no heating services.

| | | | |
|---|---|---|---|
| | Luohe (III) | 1500 (2020) | 0.63 |
| | Pingdingshan (I) | 18000 (2018) | 3.61 |
| | Nanyang (II) | 4900 (2016) | 0.50 |
| | Xinyang (III) | 4000 (2019) | 0.65 |
| | Zhumadian (I) | 9600 (2020) | 1.37 |
| | Zhoukou (II) | --------- | ≈0 |
| | Xuchang (I) | 15000 | 3.01 |
| | Luoyang (I) | 76370 (2021) | 11.03 |
| | Sanmenxia (I) | 13000 (2021) | 6.39 |
| | Kaifeng (I) | 36000 (2021) | 7.46 |
| Hubei | Xiangyang (II) | 1500 (2021) | 0.29 |
| | Shiyan (I) | 11050 (2019) | 5.36 |
| | Xiaogan, Suizhou (II) | ---------- | 0 |
| Anhui | Suzhou (II) | 85 (2019) | ≈0 |
| | Huaibei, Bozhou, Fuyang, Huainan, Bengbu, Chuzhou (II) | ---------- | ≈0 |
| Shaanxi | Shangluo (II) | 850 (2021) | ≈0.42 |
| | Baoji (I) | 38851 (2021) | 11.70 |
| | Xianyang (I) | 29100 (2021) | 7.35 |
| | Tongchuan (I) | 10200（2021） | 14.61 |
| | Weinan (I) | 11000 (2021) | 2.35 |
| | Hanzhong(II), Ankang(II) | ------------ | ≈0 |
| Shanxi | Yuncheng (I) | 24720 (2019) | 5.22 |
| Gansu | Longnan (II) | -------- | 0 |
| | Tianshui (I) | 24403 (2021) | 8.25 |
| Sichuan | Mianyang(II), Bazhong(II), Guangyuan(II) | ------- | 0 |

At last, a heating borderline map is drawn recording all counties near the heating borderline. Although the survey is conducted in prefecture-level cities, the map is drawn in county-level to make it clearer for subsequent estimations. Table 4.3.1-2 records the colored counties in Table 4.3.1-2. The chosen cities in the dataset refers to the Table 5.3.2-1.

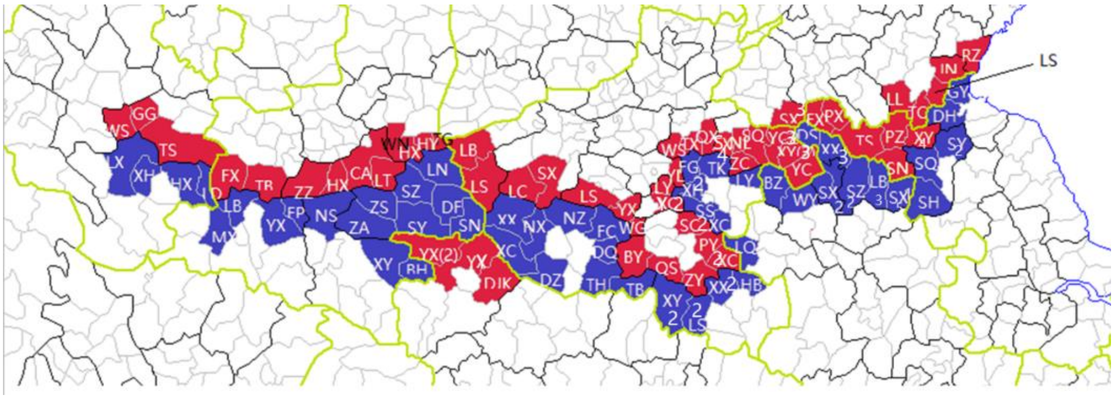

Figure 4.3.1-1 Heating borderline

Table 4.3.1-2 Counties in blue or red in Figure 4.3.1-1

| Colored counties in Figure 4.3.1-1 |
|---|
| Wushan Tongguan Gangu Lixian Tianshui Xihe Huixian Fengxian Neixiang Liangdang Liuba Mianxian Lingbao Lushi Shangnan Danjiangkou Xichuan Xixia Luanchuan Songxian Lushan Nanzhao Taibai Yangxian Zhouzhi Foping Dengzhou Huayin Danfeng Fangcheng Luonan Tanghe Changan Yexian Zhenan Wugang Zhashui Biyang Xunyang Tongbai Lantian Xinyang Luoshan Weinan Shangzhou Zhengyang Shanyang Xixian Yunxi Xincai Baihe Pingyu Huaxian Guoyang Linying Liangyuan Xihua Fugou Suixi Taikang Huaibin Linquan Shangshui Xiangcheng Shangcai Yongcheng Yancheng Yucheng Yanling Xiaoxian Weishi Xiayi Tongxu Qixian Shanxian Suixian Fengxian Ningling Peixian Shangqiu Pizhou Zhecheng Linshu Dangshan Donghai Tancheng Sixian Junan Lanling Shuyang Suiyang Sihong Suzhou Suining Xinyi Lingbi Qiaocheng |

### 4.3.2 Heating and population migration

This section covers the mechanism between heating and population migration across the heating borderline. As the model predicts, people make their moving choices according to at least three factors: heating expenditures, indoor room comfort, and outdoor air pollution. Guo et al. (2015) has pointed out that, although the unit cost for central heating is lower, the total cost in central heating areas can be larger considering the winter length and temperature levels. Heating expenditures in the north are more expensive than in the south. Central heating can help increase the indoor room comfort. Air pollution generated by central heating can reduce attraction for people. Therefore, heating expenditures and air pollution can be disadvantages for population in-migration, while indoor room comfort can be an advantage for population in-migration.

### 4.3.3 Empirical analysis

#### 4.3.3.1 Methodology

The methodology is

$$\mathrm{M_{it}} = \alpha + \beta\mathrm{heating_i} + \gamma\mathrm{longitude_i} + \delta\mathrm{GDP_{it}} + \mathrm{f(latitude_i)} + \theta\mathrm{Province}_i + \lambda\mathrm{City}_i + \varepsilon_{\mathrm{it}} \tag{10}$$

The methodology is inspired by Almond et al. (2009) and Chen and Whalley (2012). We define M as the population in-migration of the south,

$$\mathrm{M} = \text{population growth rate}(\mathrm{G}) - \text{ natural population growth rate } (\mathrm{N})$$

Natural population growth rate stands for the difference between the birth rate and death rate. The difference between the population growth rate and natural population growth rate is the population mechanical growth rate, which is equivalent to the population migration. M measures the actual state of population inflow of the chosen counties. To keep the data simple, we make M multiplied by 100.

Heating is a dummy variable. The value of cities in Group I has a value of 1 and the value of cities in Group II and III are 0.

GDP is the Gross Domestic Product for counties controlling the economic development.

Longitude is the variable representing the longitude of cities and time-invariant. These variable models the location, natural conditions.

f(latitude) is a polynomial of latitude to smooth the series. Latitude controls climate and location.

We use the regression discontinuity design (RD design) to do the estimation. Temperature is the most decisive factor affecting heating demand. Temperature is directly associated with latitude. If there is no Huai River policy, the latitude's influence on heating demand would be continuous. Because of the policy, the latitude's impact on heating demand and migration is artificially cut off. Thus, the RD design is chosen as the estimation technique. A polynomial with latitude can help smooth migration change with the variable latitude.

### 4.3.3.2 Data

China has three types of administrative divisions, including provincial-level cities, prefecture-level cities, and counties. Counties are attached to prefecture-level cities and are the smallest divisions with accurate statistics in China.

This section uses county-level data to make the estimation more convincing. In the benchmark, counties just along the borderline are selected. (See Figure 4.3.1-1 and Table 4.3.1-2 ).

The dataset is a panel with the time periods from 2007 to 2019. The data of Heating are from the survey in Table 4.3.1-1. The longitude and latitude of the counties are derived from the websites https://jingwei.supfree.net/.

G, GDP, and Natural growth rates of permanent population (N) are retrieved from the prefecture-level statistical yearbooks. M is calculated using G and N. The descriptive statistics are listed in Table 4.3.3.2-1. The dataset is unbalanced because of some missing data.

**Table 4.3.3.2-1 Descriptive statistics**

| Variables | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| M | 972 | -0.006745 | 0.0244 | -0.19538 | 0.239046 |
| Latitude | 972 | 33.7803 | 0.7207 | 32.1168 | 35.1754 |
| Longitude | 972 | 112.7609 | 3.7303 | 104.8840 | 118.8320 |
| GDP($10^8$ yuan) | 972 | 173.9910 | 147.1641 | 2.6730 | 959.7000 |
| N($10^4$ people) | 972 | 5.4644 | 2.7275 | 0.4000 | 38.7800 |

| | | | | | |
|---|---|---|---|---|---|
| Heating | 972 | 0.5442 | 0.4983 | 0 | 1 |

### 4.3.3.3 Benchmark estimation

A random effects model works out the coefficients of equations (10).

The estimation results are reported in Table 4.3.3.3-1. To avoid a collinearity problem, we use a second-order polynomial.

All these regression results show a population outflow in central heating areas. Areas without central heating have higher population inflow, while heating areas do not have higher population inflow. GDP is positively correlated to migration.

**Table 4.3.3.3-1 Benchmark estimation-$M_t$**

| | **M** |
|---|---|
| **Heating (Standard Error)** | -0.0061 ***<br>(0.0020) |
| **Latitude** | -0.1668<br>(0.1135025) |
| **Latitude$^2$** | 0.0025<br>(.0016913 |
| **Longitude** | -0.0005<br>(0.0002908) |
| **GDP** | 0.0000173**<br>(6.95e-06) |
| **Province** | .003715<br>(0.171) |
| **City** | -.0002746<br>(.0002189) |
| **Wald Chi(P-value)/F-test(p-value)** | 26.48<br>(0.000) |
| **Obs** | 972 |

### 4.3.3.4 Robustness checks

We conduct the robustness checks as follows with different specifications. In robustness check I, we only delete the variable Longitude. In check II, only the variable GDP is deleted. In robustness check III, the first-order polynomial of latitude is used. In check IV, we delete the variables Province and City. In check V, we introduce the first-order lag of GDP to remove possible endogeneity problem. The estimation results are in Table 4.3.3.4-1.

**Table 4.3.3.4-1 Robustness checks-M**

| **Robustness Check:** | **I** | **II** | **III** | **IV** | **V** |
|---|---|---|---|---|---|
| Heating | -0.0057 *** | -0.0063*** | -0.0056*** | -0.0059*** | -0.00551*** |
| Latitude | -0.1281 | -0.1443 | 0.0007 | -0.2649<br>** | -0.13588 |
| Latitude$^2$ | 0.0019 | 0.0022 | ---------- | 0.0040** | 0.0020 |
| Longitude | ---------- | -0.0001 | -0.0004 | -0.0006** | -.0005684 |
| GDP | 0.00001*** | --------- | 0.00001*** | 0.00001 ** | ----------- |
| GDP(-1) | | | | | |
| Province | 0.0032 | 0.0044* | 0.0049** | ------------ | 0.0042 |
| City | -0.0002 | -0.0003** | -0.0004 * | ------------ | -0.0003 |
| WaldChi | 24.19<br>(0.000) | 20.38<br>(0.000) | 24.17<br>(0.000) | 18.98<br>(0.001) | 26.29<br>(0.000) |
| OBS | 972 | 972 | 972 | 972 | 883 |

The estimation results are robust. All robustness checks show a negative linkage between population in-migration and Heating. The estimation results of other variables are similar.

### 4.3.3.5 Air quality and population migration

To figure out the mechanism, we conduct the following estimation to verify air pollution's impact on population migration

$$\mathrm{M_{it}} = \alpha + \beta\mathrm{Air_{it}} + \gamma\mathrm{longitude_i} + \delta\mathrm{GDP_{it}} + \theta\mathrm{Province}_i + \lambda\mathrm{City}_i + \mathrm{f(latitude_i)} + \varepsilon_{\mathrm{it}}$$

We delete the dummy variable Heating. The variable Air denotes the annual concentration of PM2.5 in these counties. All the left data remain unchanged. Air pollution data are acquired from the Dalhousie University dataset. We have two settings. The second setting includes the first-order lag of the variable Air.

**Table 4.3.3.5-1 Air-M estimation results**

| **Migration** | **I** | **II** |
|---|---|---|
| **Air** | -0.0002**<br>(0.0001) | |
| **Air (-1)** | | -0.0001<br>(0.0000879) |
| **Latitude** | -0.1244<br>(0.1168) | -0.1000<br>(0.1161899) |
| **Latitude$^2$** | 0.00185<br>(0.0017) | .0014781<br>(0.0017306) |
| **Longitude** | 0.00019 | -.0001488 |

| | | |
|---|---|---|
| | (0.0004) | (0.0004093) |
| **GDP** | .000013<br>(7.70e-06) | .0000164**<br>(7.68e-06) |
| **Province** | 0.0014<br>(0.0027) | .0021821<br>(.0026608) |
| **City** | -0.0001<br>(0.0002) | -.0001506<br>(.0002146) |
| **Obs** | 959 | 883 |

The estimation results show a negative correlation between air pollution and population in-migration (Table 4.3.3.5-1). It shows that air pollution can crowd out population in-migration. But the coefficient is very small, showing little correlation between air pollution and population migration.

### 4.3.3.6 Heterogeneity by economic development

In this section, we categorize the dataset into two groups in terms of GDP per capita and compare the results of the two groups. The estimation methodology is according to equation (10) except that the variable GDP is taken out from the regression equation.

We use the critical value c to divide the sample into two groups, one with higher GDP per capita and the other with lower GDP per capita. We compare the coefficients of the two groups to see if people's migration decisions differ. Attitudes may differ between people living in the same area but at different income levels.

GDP per capita is the ratio between GDP and permanent population. We first set the critical value c at 50,000 yuan. Counties with per capita GDP exceeding RMB 50,000 are classified into the high-income group. We conduct these regressions in two groups and compare the results with former results. Table 4.3.3.6-1 reports the estimation results.

To further verify the results, we changed the classification criteria from 20,000 to 60,000 for estimation respectively. We perform regression under the condition that per capita GDP is greater than the standard c and less than c (Table 4.3.3.6-2). We then plot the graph between the critical value and the coefficient of Heating when GDP per capita is greater than c. (Figure 4.3.3.6-1).

**Table 4.3.3.6-1 Estimation in different groups-M**

| **M** | **Group 1 (GDP per capita>50,000)** | **Group 2 (GDP per capita<50,000)** |
|---|---|---|
| **Heating** | 0.0077***<br>(0.0030) | -0.0061 ***<br>(0.0021) |
| **Obs** | 64 | 908 |

**Table 4.3.3.6-2 Regression results with changing criteria c-M.**

| **Group 1** | **Coefficient** | **Group 2** | **Coefficient** |
|---|---|---|---|

| | | | |
|---|---|---|---|
| GDP per capita >20,000 | -0.0027 | GDP per capita ≤20,000 | -0.0070 ** |
| >25,000 | -0.0047 | ≤25,000 | -0.0062** |
| >30,000 | -0.0048 | ≤30,000 | -0.0060 ** |
| >40,000 | -0.0037 | ≤40,000 | -0.0059 ** |
| >42,000 | -0.0003 | ≤42,000 | -0.0057 ** |
| >44,000 | 0.0038 | ≤44,000 | -0.0058 ** |
| >46,000 | 0.0064** | ≤46,000 | -0.0057** |
| >48,000 | 0.0069*** | ≤48,000 | -0.0058 ** |
| >50,000 | 0.0077*** | ≤50,000 | -0.0061** |
| >52,000 | 0.0089** | ≤52,000 | -0.0060 ** |
| >54,000 | 0.0091** | ≤54,000 | -0.0057 ** |
| >55,000 | 0.0093** | ≤55,000 | -0.0053 ** |
| >60,000 | -0.0021 | ≤60,000 | -0.0050** |

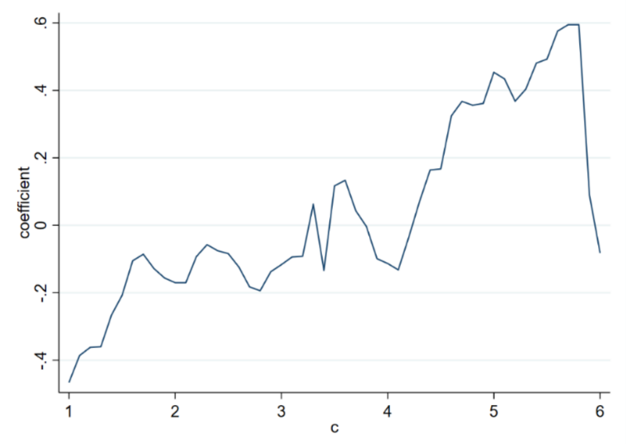

**Figure 4.3.3.6-1 Plot between criterion c ($10^4$) and coefficient of the variable Heating**

The high-income group sees a positive Heating's coefficient. It indicates a positive correlation between heating and migration. The low-income group has an opposite result that areas without central heating attract population inflow. With the increase in c, the coefficients increase according to Table 4.3.3.6-2. Moreover, the significant level also increases when c becomes larger. When c increases, the regressions with the remaining data (GDP per capita < c) have gradually decreasing coefficients. Figure 4.3.3.6-1 illustrates the finding graphically. The negative effect becomes positive when the classification criterion is larger than some point c.

Based on the above results, we find that people living in wealthier counties are more concerned about the indoor living conditions than the heating expenditures; therefore, the overall effect is positive in richer counties and negative in poorer counties.

### 4.3.4 Summary

In Section 4.3, the location of the heating boundary line is first determined. County data are then used to analyze the differences in population inflows on both sides of the heating boundary. Heating boundaries create differences in population migration. The results can be summarized in the following three points:

First, from the analysis of the total sample, population migration in heating areas is weaker than in non-heating areas. Central heating areas will not attract more population inflows.

Second, in more developed counties, the impact of district heating on population migration is positive. In contrast, the impact in less developed counties is negative. This shows that residents in relatively poor areas are more sensitive to heating expenses, while residents in wealthier areas are more concerned about living comfort.

Third, district heating affects migration partly through its effect on air quality, but this effect is small. Indoor room comfort and heating expenditure have a greater impact on population migration.

# Chapter 5：Huai River Policy and economic development

## 5.1 Introduction

According to Chapter 4, we have already derived the following results. First, the Huai River policy can affect air pollution. Second, the central heating areas do not attract more population inflow in the borderline-nearby areas. This section will link the policy to economic development.

Existing literature takes the policy as a natural instrument to justify air pollution's effects on economic terms. Most papers are developed in recent years. Some papers focus on macro-economic variables. Dong et al. (2019) [44] analyzed the relationship between the Huai River policy and inbound tourism. They found that the policy-induced air pollution could dramatically decline international inbound tourism. Li and Zhang (2019) [45] suggested that the increased air pollution could reduce foreign direct investment. Every 1% increase in particulates would engender a 0.393% decrease in FDI. Some papers are about human and economic activities. Cho et al. (2021) [46] analyzed the impact of air pollution on business ethics. It shows that firms in cities with higher air pollution exhibit higher discretionary accruals and are likelier to restate their financial statements. Lin et al. (2022) [47] discussed the air pollution's impact on human workplace choices. They find that more air pollution could make people work nearer to their homes. Qiao et al. (2023) [48] focused on the air pollution's impact on innovation. They find that air pollution could harm innovation by using the policy as a natural discontinuity. Some are more about micro-economic variables. Liu et al. (2021) [49] linked air pollution levels with China's house prices and found that the increase in air pollution would depress the increase in house prices. A similar study was constructed by Zhang et al. (2022) [50]. Huang and Du (2022) [51] found a causal effect of air pollution on cognitive bias and under-valuation in the land market.

Some papers determine the air pollution's impact on people's health and life expectancy using the policy. (Chen et al, 2013 [52]; Ebenstein et al. 2017 [53]). Some papers measure the economic costs of air pollution with the policy. Ito and Zhang (2020) [54] explored the willingness-to-pay for clean air using air purifier markets under the Huai River policy. This scenario used the policy as a tool to investigate the costs of some local amenities.

All these studies have shown that the policy might affect human and economic activities by affecting air pollution levels. These papers' limitations are that they only consider the air pollution effects induced by the Huai River policy but neglect other effects, including its comfort-improving effects and cost effects. They take the Huai River policy as an instrumental variable. They analyze the air pollution's impact on human and economic activities. However, they do not consider the policy's impact on economic variables.

In this chapter, we will examine the policy's effects on a group of economic variables that are considered relevant to the policy. First, we will consider whether foreign investment would prefer to settle in a heating city or a non-heating city. Foreign direct investment (FDI) is a variable to measure foreign companies' "moving" decisions. Second, nighttime luminosity is a term considered to be relevant to human and economic activities. At last, GDP per capita is examined as a more economy-related variable to judge the overall policy's effects. We can conclude all these variables in one graph to see their links in Figure 5.1-1. The construction of central heating systems could directly contribute to economic growth as the heating industry is directly linked to the steel industry, construction industry, power industry, transportation industry, family furnishing industry, and so on. It has a direct impact on economic growth. Migration and FDI have an indirect impact on economic terms. We have proved that the policy could affect indoor living environment (room comfort), outdoor environment (air), and heating costs and clarified the relationship between the Huai River policy and population migration. These factors combined affect people's moving decisions (migration). There are three terms left in Figure 5.1-1 to be verified. They are FDI, nighttime luminosity, and GDP per capita. This chapter will focus on these three factors and tell a complete story of the policy's economic impact.

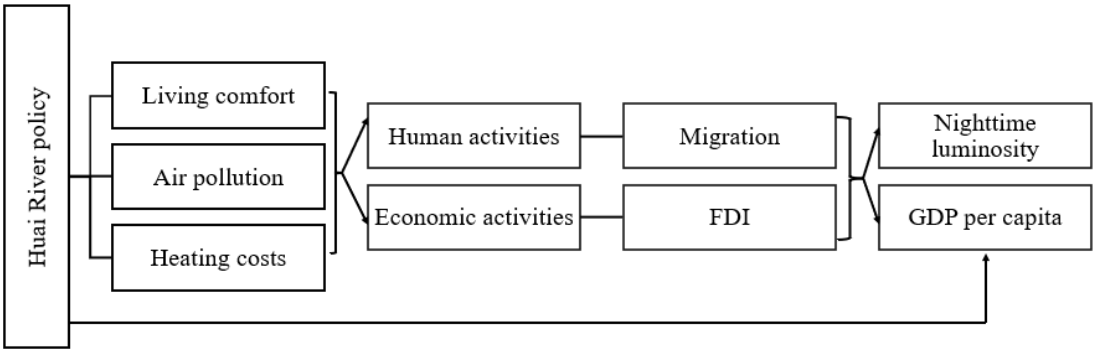

**Figure 5.1-1 Linkage of all variables in this study**

## 5.2 Foreign Direct Investment (FDI)

### 5.2.1 Introduction

In this section, we consider whether heating areas can have higher capital inflow. Population migration stands for the impact on human moving choices, which is an individual impact. In an enterprise-level, we consider the the Huai River policy's impact on foreign direct investment, which is a "settling" decision for foreign companies. The attitudes of foreign investors towards the problem are similar to those of common households. Foreign companies will consider heating costs, air quality, and living comfort. They prefer to stay in places with more pleasant office temperature, cheaper heating costs, and better air quality. Besides the heating policy, they also consider other economic terms, including the infrastructure levels, governmental policies, regional advantages, etc. There is no specific paper discussing the policy and FDI.

Section 5.2 focuses on an empirical method using prefecture-level city data near the heating borderline. We control location, infrastructure, GDP per capita, wage level, etc., and construct an econometric model to examine this question.

### 5.2.2 Data

The variables used in the model can be classified into the following categories: key indicator variable, location variables, economic variables. Key indicator variable is Heating, which indicates whether the city has central heating or not. Location variables include latitude and longitude. Economic variables include GDP per capita, Disposable income (DI), and Road Mileage, which measure economic development, wage, and infrastructure levels. Location variables and key indicator variable are time-invariant variables, while the economic variables are time-variant variables. The data source is in Table 5.2.2-2.

We apply a panel dataset with the prefecture-level cities adjacent to the heating boundary with the periods 2011 to 2020. The descriptive statistics are recorded in Table 5.2.2-1. The criteria for judging whether cities have central heating or not depend on the survey noted in Figure 4.3.1-1. The cities we apply in the dataset are kept in Table 5.2.2-3.

**Table 5.2.2-1 Descriptive statistics**

| | **Obs** | **Mean** | **Std. Dev** | **Min** | **Max** |
|---|---|---|---|---|---|
| Fdi($10^8$ yuan) | 283 | 44.29388 | 66.30882 | .0006 | 529.513 |
| heating | 283 | .5830389 | .4939298 | 0 | 1 |
| Gdppercapita (CNY,yuan) | 283 | 38590.81 | 15806.07 | 10400 | 80300 |
| Road (km) | 283 | 25878.61 | 6064.814 | 3565 | 40313 |
| Di (CNY,yuan) | 283 | 15929.47 | 7687.256 | 13051 | 43713 |
| latitude | 283 | 33.94949 | .8235773 | 32.12328 | 35.42859 |
| longitude | 283 | 113.6382 | 3.614204 | 105.725 | 119.461 |

**Table 5.2.2-2 Independent Variables**

| **Category** | **Independent variables** | | | **Source** |
|---|---|---|---|---|
| Key indicator variable | Heating | | | According to the survey in Figure 4.3.1-1. |
| Location variables (Time invariant) | Latitude | Longitude | | Latitude and longitude are retrieved from the website: http://www.jsons.cn/lngcode |
| Economic variables (Time variant) | GDP per capita (GDPpercapita) (yuan) | DI (Disposable income) (yuan) | Road mileage (Road) (km) | Retrieved from the official statistical yearbooks. |

**Table 5.2.2-3 Prefecture-level cities included in the dataset**

| **Province** | **Prefecture level city** | **Heating** |
|---|---|---|
| **Shandong** | Rizhao, Linyi, Heze | 1 |
| **Jiangsu** | Lianyungang Suqian | 0 |
| | Xuzhou | 1 |
| **Anhui** | Suzhou Bozhou Huaibei Fuyang | 0 |
| **Henan** | Shangqiu Kaifeng Xuchang Luohe<br>Luohe Zhumadian Pingdingshan | 1 |

| | | |
|---|---|---|
| | Luoyang Sanmenxia | |
| | Zhoukou Xinyang Nanyang | 0 |
| **Hubei** | Shiyan | 1 |
| **Shaanxi** | Ankang Shangluo Hanzhong | 0 |
| | Baoji Xi'an Weinan | 1 |
| **Gansu** | Longnan | 0 |
| | Tianshui | 1 |

### 5.2.3 Methodology

The methodology aims to solve two problems. One problem is that some variables have simultaneity problems. GDP per capita influences FDI, while FDI could, in turn, affect GDP per capita. So is the variable DI. This is the endogeneity problem. The second problem is that some variables are time-invariant. Using normal panel data estimation methods, we cannot estimate some variables. In order to solve these two problems, we use the Hausman-Taylor estimator. We take DI and Gdppercapita as two endogenous variables, Road as the exogenous time-variant variable. Heating, latitude, and longitude are exogenous, time-invariant variables. In addition to this, we apply several robustness checks to verify the estimation results. The estimation result of the Hausman-Taylor estimator is in Table 5.2.3-1. The benchmark estimation methodology is in Equation (11)

$$FDI_{it}=\Phi+\alpha Heating_{i}+\beta Latitude_{i}+\gamma Longitude_{i}+\delta Gdppercapita_{it}+\kappa DI_{it}+\zeta Road_{it}+u_{i}+\varepsilon_{it} \qquad (11)$$

Four robustness checks are as follows. The first one is to change the estimation technique to a normal random-effects panel estimation method. The second one is to delete the variable Gdppercapita and conduct a Hausman-Taylor estimator. The third one is to add a square term of latitude and conduct the benchmark estimation. The fourth one is to delete DI and take Road as the exogenous time- variant variable in the Hausman-Taylor estimator.

**Table 5.2.3-1 Benchmark estimation**

| Variable | Coefficient | Std. Dev | T-statistic | P-value |
|---|---|---|---|---|
| heating | 8.25345 | 27.16358 | 0.30 | 0.761 |
| Latitude | 13.61097 | 22.07243 | 0.62 | 0.537 |
| longitude | -.3970311 | 3.025256 | -0.13 | 0.896 |
| gdppercapita | 0.000794 | 0.0003452 | 2.30 | 0.021 |
| di | 0.0011986 | 0.0006343 | 1.89 | 0.059 |
| road | -0.0010317 | 0.0007056 | -1.46 | 0.144 |
| Sargan-Hansen Chi-squ(1) | | 0.898 | | |
| Wald chi2(6) | | 66.13 | | |
| Wald p value | | 0.00 | | |
| Obs | | 283 | | |

We first analyze the benchmark estimation. We find that no direct linkage between FDI and the policy exists. Although the coefficient of heating is positive, it is not significant. Areas with central heating do not always attract inflow of foreign investment. This shows foreign investors do not consider the heating a major factor in their investment decisions. Infrastructure does not play an important role. Economic terms like DI and GDP per capita are crucial for investors to make their decisions. This result is reasonable. Households always consider the policy more important than investors.

Four robustness checks draw similar conclusions, showing no strong linkage between the policy and FDI.

**Table 5.2.3-2 Robustness checks**

| | I | II | III | IV |
|---|---|---|---|---|
| heating | 16.09105 | 9.010247 | 15.4875 | 5.650705 |
| Latitude | 0.912893 | 20.49462 | -327.5103 | 12.63709 |
| Latitude^2 | ---------- | ---------- | 4.861343 | ---------- |
| longitude | -1.251202 | 0.4570523 | -1.398215 | -0.77864 |
| gdppercapita | 0.0009811*** | --------- | 0.0008515** | 0.0013379*** |
| di | 0.0010141* | 0.0024219*** | 0.0011618** | ---------- |
| road | -0.0012263** | -0.0009367 | -0.0012091 | -0.0007128 |
| Obs | 283 | 283 | 283 | 283 |

### 5.2.4 Results

The study’s conclusion shows little correlation between the inflow of foreign capital and central heating. The areas where central heating is distributed do not attract more foreign capital inflow, and the areas where central heating is not distributed do not attract more foreign capital inflow, either. Foreign capital has no particular preference for the distribution of central heating. This differs from the choice of households, for which there is an evident willingness to move to areas where central heating is not distributed.

## 5.3 Huai River policy and nighttime luminosity

It is a trend to explore economic phenomena using satellite data on lights at night as a proxy variable. One of the representative papers is from Henderson et al. (2012) [55]. It applied the satellite data on lights at night as a proxy variable for income or GDP and explored different nations’ GDP growth patterns worldwide. Using satellite data to explore facts in economics is a story that has been around for a while. Henderson et al. (2012) [55] applied the satellite data on lights at night as a proxy variable for income or GDP and explored the different nations’ GDP growth patterns worldwide. Using satellite data to explore facts in economics is not a new story. Chen and Nordhaus (2011) [56] tried to use the night luminosity as a proxy variable for output. Doll et al. (2006) [57] mapped the economic activity using nighttime light remote sensing data. Beyer et al. (2018) [58] argued that night light could serve as a good proxy variable for economic activities and applied the method in a district study in South Asia. Doll and Pachauri (2010) [59] applied the nighttime remote sensing data to estimate their rural populations. Kumar et al. (2019) [60] applied nighttime light intensity to estimate population densities of Beijing, China. The existing literature shows a good performance of nighttime light intensity data as a proxy variable to measure the intensity of human and economic activities. These data are strictly objective, hard to manipulate and reflect the real situations of economic activities. This section will explore the policy's effects on county-level nighttime luminosity.

The Huai-River policy distinguishes the areas with and without central heating systems. It can impact the economic development in two aspects. According to Section 2.5, the heating industry is a booming industry in China. The central heating industry itself can generate GDP and boost the intensity of human and economic activities, which is the Impact II. Buildings are newly constructed annually, and these buildings also provide heating services. The sales of heating pipes and equipment in the heating areas are constantly increasing. In China, there is a well-established heating industry chain. Newly built houses are equipped with floor heating systems. These systems are connected with many upstream and downstream industries, which could directly contribute to GDP growth.

The other is the Impact I (individual effect). The Huai River policy can affect air pollution, living comfort, and expenditures for households and enterprises. These factors can influence activities for both households and firms. Here, we use the nighttime luminosity as a proxy variable to measure the intensity of human and economic activities.

Areas with more frequent economic and human activities will have stronger nighttime light brightness indexes. Based on the previous findings in this study, we find that the policy's impact on FDI is not significant. Central heating zones have not attracted more foreign investment. Similar findings are also applicable to population migration. Central heating zones do not attract more population inflow because of possible increases in heating costs. We can conclude that the Impact II (individual) is difficult to capture in this study. This section will focus on the Impact I (aggregate) of the policy on economic development. That is, we will test whether the central heating industry itself, as an important component of the economy, could contribute to the intensity of human and economic activities.

### 5.3.1 Methodology

We are using the following methodology to determine whether the Huai River policy generates differences across the heating borderline in terms of night luminosity. With an endogeneity problem and exogenous variables, we will perform one type of IV estimators- Hausman-Taylor estimators with a discontinuity property. The technique has been clearly explained in Section 4.2.4.1.

$$Light_{it} = \alpha_i + \beta_1 Heating_{it} + \beta_2 Per_{it} + \beta_3 Area_i + \beta_4 Latitude_i + \beta_5 Longitude_i + \beta_6 Public_{it} + \beta_7 Enterprise_{it} + \beta_8 FDI + \varepsilon_{it}$$

**Table 5.3.1-1 Variable classification of benchmark estimation**

| **Group** | **Variables** | **Unit** | **Source** |
|---|---|---|---|
| Luminosity | Light | Index | National Centers for Environmental Information |
| Policy(1) | Heating | Either 0 or 1 | Survey |
| Population(2) | Per (permanent population) | Per ($10^4$ people) | statistical yearbooks |
|  | Area (area of the county) | Km$^2$ |  |
| Administration(3) | City |  |  |
| Location (4) | Latitude |  | http://www.jsons.cn/lngcode/ |
| Economic variables (5) | Public (general budget expenditure) | $10^4$ yuan | Statistical yearbooks |
|  | Enterprise (number of newly-established enterprises) | Number |  |
|  | FDI (Foreign direct investment) | $10^4$ yuan |  |

The independent variables can be classified into six groups. (See Table 5.3.1-1) Group 1 is the policy variable that represents the Huai River policy. Population group includes the permanent population size (Per) and area. This group measures the county scale (Per) and population concentration degree. The administrative division controls the counties' administrative affiliations. The location group includes the counties' latitude and longitude. The economic group includes variables measuring the government budget expenditures (Public), the number of newly-opened enterprises (Enterprise), and FDI (Foreign direct investment). Following are the variable explanations.

Light- annual night luminosity index is the proxy variable of the intensity of human and economic activities.

Heating- indicator variable tells whether the place is in non- or heating areas. When Heating equals 1, the place is in the north. When Heating equals 0, the place is in the south. Heating is the key explanatory variable to link the Huai River policy with the night luminosity.

Per- Permanent population. By intuition, larger permanent population size would generate higher intensities of human activities, leading to higher levels of night luminosity. This variable can control the scale of the county. The unit of the data is 10,000 people.

Area- city area. The variable records the city area retrieved from the statistical yearbooks. Its unit is $km^2$.

City: This variable can control the difference in nightlight brightness caused by policy differences between prefecture-level cities. We use the code of administrative division published by Ministry of Civil Affairs.

Latitude and Longitude report these counties' geographical locations and control these cities' natural and climatic conditions, such as humidity, precipitation, wind speed that could affect the night luminosity.

Public- this variable denotes the general budget expenditure controlling the impact of local infrastructure construction. It is usually believed that counties with higher public investment would have better infrastructure; thus, the county will have higher light intensity at night. This variable is added as an extra economic variable to control the governmental effects.

Enterprise- the variable measures the number of newly-established enterprises, which is the control variable of economic performance. FDI indicates the foreign direct investment.

In this estimation, we have several time-invariant variables, which are exogenous. Location and administrative groups are strictly exogenous geographical variables. Some time-variant variables are considered endogenous considering simultaneity problems. In the time-varying scenario, Per and public are two endogenous variables. Enterprise and FDI are two exogenous variables. In the time-invariant scenario, heating is considered as an endogenous time-invariant variable. We will conduct a Hausman-Taylor estimator to estimate the model.

### 5.3.2 Data

We use the county-level data from the National Centers for Environmental Information, which collected the night luminosity index of the global cities. The dataset records the night light intensities captured by the VIIRS sensors on the satellites of the Suomi National Polar-Orbiting Partnership (Suomi NPP). This type of data is the so-called NPP/VIIRS data. We retrieve the wanted data from the big dataset, including the county-level night luminosity data alongside the heating borderline with the time from 2013 to 2019. The original data are monthly data. We calculate the annual data by taking the average of the monthly data. 89 counties are included, which are listed in Table 5.3.2-1. 89 counties in 27 prefecture-level cities of 7 provinces are included in the dataset. Counties are one of China's basic administrative division units with accurate statistics. They are usually less developed compared to provincial-level and prefecture-level cities. They have more similar natural and political conditions than cities at higher levels. We use county-level data to remove the possible effects of administrative levels imposed on night luminosity.

Table 5.3.2-2 reports the descriptive statistics of major variables in the benchmark estimation. The source of the data is recorded in Table 5.2.2-2.

**Table 5.3.2-1 Counties that are included in the dataset**

| Province | Prefecture level city | Counties |
|---|---|---|
| **Henan** | Kaifeng | Qixian, Tongxu, Weishi |
| | Luoyang | Luoyang, Songxian, |
| | Xuchang | Yanling |
| | Luohe | Linying, Yancheng, |
| | Sanmenxia | Lushi, Lingbao, |
| | Nanyang | Nanzhao, Fangcheng, Xixia, Neixiang, Xichuan, Tanghe, Tongbai, Dengzhou |
| | Shangqiu | Suixian,Ningling,Zhecheng,Yucheng,Xiayi, Yongcheng, Liangyuan, Suiyang, |
| | Xinyang | Luoshan, Huaibin, Shihe, |
| | Zhoukou | Fugou, Xihua, Shangshui, Taikang, Xiangcheng， |
| | Zhumadian | Pingyu, Shangcai, Queshan, Biyang, Xincai |
| **Jiangsu** | Xuzhou | Fengxian, Peixian, Suining, Xinyi, Pizhou, |
| | Lianyungang | Donghai, |
| | Suqian | Shuyang, Sihong, |
| **Shaanxi** | Baoji | Fengxian, Taibai, |
| | Weinan | Linwei, Huazhou, Tongguan, Huayin, |
| | Hanzhong | Yangxian, Mianxian, Liuba, Foping, |
| | Ankang | Ningshan, Xunyang, Baihe, |
| | Shangluo | Shangzhou, Luonan, Danfeng, Shangnan, Shanyang,<br>Zhenan, Zhashui, |
| **Gansu** | Tianshui | Qingshui, Gangu, Wushan, |
| | Longnan | Xihe, Lixian, Liangdang, |
| **Anhui** | Huaibei | Suixi |
| | Fuyang | Linquan, |
| | Suzhou | Dangshan, Xiaoxian, Lingbi, Yongqiao, Sixian, |
| | Bozhou | Qiaocheng, Guoyang, |
| **Hubei** | Shiyan | Danjiangkou, Yunxi, |
| **Shandong** | Heze | Shanxian |

| | |
|---|---|
| Linyi | Tancheng, Lanling, Junan, Linshu |

**Table 5.3.2-2 Descriptive statistics of major variables**

| Variables | Obs | Mean | Std. dev. | Min | Max |
|---|---|---|---|---|---|
| Light | 520 | 11.08429 | 6.387696 | 2.38 | 41.73 |
| Heating | 520 | 0.525 | 0.4998555 | 0 | 1 |
| Per ($10^4$) | 520 | 69.18533 | 39.12143 | 3.02 | 174.9 |
| Area($km^2$) | 520 | 1942.898 | 867.7365 | 65.7 | 4300 |
| Latitude | 520 | 33.79562 | 0.7003178 | 32.1168 | 35.1754 |
| Longitude | 520 | 112.964 | 3.792301 | 104.884 | 118.832 |
| Public($10^4$ yuan) | 520 | 40.38437 | 22.89114 | 5.57 | 132.3 |
| Enterprise | 520 | 1670.337 | 1811.853 | 66 | 13235 |
| FDI($10^4$ yuan) | 520 | 519.650 | 530.245 | 15.7313 | 3465.107 |
| City | 520 | 12.91731 | 6.870877 | 1 | 26 |

### 5.3.3 Benchmark estimation

The benchmark estimation uses the Hausman-Taylor estimator. Per and Public are two time-variant endogenous variables that have simultaneity problems.

We first conduct a Sagran-Hansen test. Under 5% level, the result rejects the hypothesis, showing that the model is valid. The major estimation results are reported in Table 5.3.3-1. The positive coefficient of the variable Heating shows that the heating borderline generates distinctions on both sides. After controlling other conditions, we can say that areas in the north have 11.5 on average higher in nighttime light intensity index. Through the descriptive index, we can see that the maximum light intensity in the dataset is 41.73. The minimum of light intensity is 2.38. The estimation result witnesses a massive gap between the areas with and without heating in nighttime luminosity. In summary, areas with central heating systems have higher night luminosity after controlling other important variables that affect night light. Central heating regions have higher night light intensities.

Population affects the night luminosity positively. Higher permanent population size will lead to higher night light intensities. Population density reflects the intensity of human activities. Higher population intensity can lead to higher night light brightness.

The variable Area is negatively correlated with the variable Light, which is not hard to interpret. Larger area may generate less population density, bringing down the intensity of human and economic activities. Counties with larger areas may have a more splattered population distribution. The nighttime luminosity could be reduced.

City is positive but insignificant. We control this variable as policies issued by local governments closely influence local human and economic activities.

Control variables related to economic performance have positive and significant coefficients. The estimation shows a significant positive correlation between government expenditures and night luminosity. Enterprise contributes to the growth of nighttime light intensity.

Latitude and Longitude are positively correlated with luminosity, although latitude is not significant. These two variables are essential in controlling the location impact on the dependent variable. A larger longitude value means greater night light brightness. A larger longitude value means the county seat is closer to the eastern coastal area.

All economy group variables have positive and significant coefficients, showing a robust positive linkage between economic performance and night luminosity.

**Table 5.3.3-1 Benchmark estimation results**

| **Independent variables** | **Coefficient** | **Std. Err.** | **T-statistic** | P-Value |
|---|---|---|---|---|
| Heating | 11.50016 | 3.890074 | 2.96 | 0.003 |
| Area | -0.010045 | 0.0024207 | -4.15 | 0.000 |
| Per | 0.27743 | 0.0605801 | 4.58 | 0.000 |
| Public | 0.2229744 | 0.0166221 | 13.41 | 0.000 |
| Enterprise | 0.0004696 | 0.0001599 | 2.94 | 0.003 |
| FDI | 1.12e-06 | 5.59e-07 | 2.00 | 0.046 |
| City | 0.2273771 | 0.2136557 | 1.06 | 0.287 |
| Longitude | -4.337856 | .8465705 | -5.12 | 0.000 |
| Latitude | -18.36433 | 4.773175 | -3.85 | 0.000 |
| Constant | 1121.32 | 238.0607 | 4.71 | 0.000 |
| Wald Chi-square(9) | | 441 | | |
| Prob > chi2 | | 0.0000 | | |
| Sagran stat | | 7.241 | | |
| Obs | | 520 | | |

### 5.3.4 Robustness checks

We will re-estimate the benchmark with different specifications as follows.

In the first scenario, we consider the replacement of variables. In estimation I, we introduce the one period lagg variables of Public and Enterprise to replace the former variables considering the effects of two variables Public and Enterprise on night light are lagged.

Estimation I:

$$light_{it} = \alpha + \beta_1 Heating_{it} + \beta_2 Per_{it} + \beta_3 Area_{it} + \beta_4 Latitude_i + \beta_5 Longitude_i + \beta_6 Public_{i,t-1} + \beta_7 Enterprise_{i,t-1} + \beta_8 FDI + \varepsilon_{it}$$

We consider an imaginary heating borderline in the second scenario and conduct the benchmark methodology. We randomly pick up about half of the counties and take them as heating areas; the rest are non-heating areas. We can compare the estimation results to see whether they differ from those in the benchmark.

In the third scenario, we only leave the following variables:

Estimation III

$$light_{it} = \alpha + \beta_1 Heating_i + \beta_2 Per_{it} + \beta_3 Latitude_i + \beta_4 Enterprise_{it} + \beta_5 Longitude_i + \varepsilon_{it}$$

**Table 5.3.4-1 Robustness checks**

| **Independent variables** | **I** | **II** | **III** |
|---|---|---|---|
| Heating | 9.32065**<br>(4.061709) | 4.991757<br>(4.040197) | 12.4322***<br>(4.127828) |
| Per | 0.3004615***<br>(0.0637945) | 0.2743481***<br>(0.0596604) | 0.2713266***<br>(0.0677498) |
| Area | -0.0091353***<br>(0.0024912) | -0.0090772***<br>(0.0030707) | ------------ |
| City | 0.1464174<br>(0.2315678) | 0.0463366<br>(0.2124519) | ------------ |
| Public<br>($public_{-1}$) | 0.2637484***<br>(0.0637945) | 0.2442233***<br>(0.0175594) | ------------ |
| Enterprise | 0.0005524***<br>(0.0001568) | 0.0003103*<br>(0.0001815) | 0.0677498***<br>(0.000155) |
| FDI | 1.59e-06**<br>(5.47e-07) | 1.70e-06*<br>(6.09e-07) | -------------- |
| Longitude | -4.297375***<br>(0.8387493) | -4.04978***<br>(0.9871345) | -2.933883***<br>(0.8299123) |
| Latitude | -14.42837***<br>(4.810904) | -10.14659**<br>(4.884251) | -15.10372***<br>(4.839211) |

| | | | |
|---|---|---|---|
| Constant | 963.1093***<br>(229.6666) | 796.1128***<br>(258.5305) | 824.9904<br>(236.2363) |
| Wald Chi | 415.99 | 415.99 | 159.62 |
| P-value | 0.0000 | 0.0000 | 0.0000 |
| Obs | 520 | 520 | 520 |

The robustness checks illustrate the following important conclusion: the policy variable Heating has a significant and positive coefficient, showing a robust estimation result.

### 5.3.5 IV estimation

This section considers another estimation technique to estimate the key independent variable Heating. The model to be estimated is

$$light_{it} = \alpha + \beta_1 Heating_i + \beta_2 Per_{it} + \beta_3 Latitude_i + \beta_4 Enterprise_{it} + \beta_5 Longitude_i + \varepsilon_{it}$$

As there is endogeneity in the model, we will conduct a panel data instrument variable estimation. Enterprise and per are two endogenous variables. We introduce two instrumental variables, N and $\Delta$. N denotes the natural growth rate of the permanent population. $\Delta$ denotes the annual net increase in the permanent population. These two instrumental variables are correlated with the endogenous variables but are uncorrelated with the error terms. We will perform a 2-stage least squares estimation.

In the first stage of estimation, the coefficients of the instrumental variables are highly correlated with the endogenous variables. The Sanderson-Windmeijer multivariate F test of excluded instruments is presented in Table 5.3.5-1. The weak instrument robust inference tests are listed in Table 5.3.5-2. The two instrumental variables have passed the weak instrument tests. They are not weak instruments. The IV estimators do not have the unidentification problem in Table 5.3.5-3.

**Table 5.3.5-1 Weak instruments test: F-test**

| **Endogenous variables** | **Sanderson-Windmeijer multivariate F test** | **P value** |
|---|---|---|
| Enterprise | 14.16 | 0.000 |
| Per | 10.05 | 0.000 |

**Table 5.3.5-2 Weak instrument robust inference**

| Test name | Statistic | P value |
|---|---|---|
| Anderson-Rubin Wald test | F(3,513)= 7.96 | 0.0000 |
| Anderson-Rubin Wald test | Chi-sq(3)= 24.21 | 0.0000 |
| Stock-Wright LM S statistic | Chi-sq(3)= 23.14 | 0.0000 |

**Table 5.3.5-3 Unidentification test**

| **Unidentification test** | **Statistic** | **P value** |
|---|---|---|

| | | |
|---|---|---|
| Anderson canon. corr. LM statistic | Chi-sq(2)=14.84 | 0.0006 |

The second-stage estimation results are reported in Table 5.3.5-4. The key independent variable has a coefficient of 3.351517, which is smaller than that in benchmark estimation. As the coefficient is positive, it shows a correlation between the Huai River policy and the nighttime luminosity. The areas in heating areas have higher nighttime light intensity.

**Table 5.3.5-4 Results of IV estimation**

| | **Coefficient** | **Standard error** | **t-statistic** |
|---|---|---|---|
| Enterprise | 0.0028672*** | 0.0008265 | 3.47 |
| Per | 0.094276* | 0.0562859 | 1.67 |
| Heating | 3.351517*** | 0.9624635 | 3.48 |
| latitude | -1.508329** | 0.7059389 | 2.14 |
| Longitude | -.9523504** | 0.4907629 | -1.94 |
| Constant | 156.5693** | 69.29358 | 2.26 |
| Sagran-Hansen Test statistic | 5.002 | | |

### 5.3.6 Results

It can be concluded from the previous estimation results that after controlling for many variables that affect nighttime light brightness, the nighttime light brightness in heating areas is significantly higher than that in non-heating areas.

## 5.4 The Huai River policy and GDP per capita

### 5.4.1 Huai River policy and GDP per capita

This section will determine the relationship between the Huai River policy and GDP per capita between the south and north. We first calculate the GDP per capita of the non- and heating groups. We find that the central heating regions have essentially higher GDP per capita. This result is beyond intuition. Normally, southern areas will have better economic performance. In the database containing all the counties on both sides of the heating borderline, we found that the per capita GDP of the heating area was significantly larger than that of the non-heating area. If the statistics were made province by province, the per capita GDP of the heating area was significantly higher than that of the non-heating area, except that counties in Hubei and Shandong were located in the same area and could not be counted.

**Table 5.4.1-1 Average GDP per capita of non- and heating regions**

| | **Average GDP per capita (10,000 yuan, RMB)** |
|---|---|
| Non-heating regions | 3.037865 |
| Heating regions | 3.644007 |

| **Province** | **Heating regions** | **Non-heating regions** | **Percentage** |
|---|---|---|---|

| | | | |
|---|---|---|---|
| Henan | 3.509365 | 3.161425 | 11.00% |
| Jiangsu | 3.701365 | 3.312551 | 11.74% |
| Shaanxi | 3.606443 | 3.290689 | 9.60% |
| Gansu | 3.484387 | 2.917077 | 19.45% |
| Anhui | 3.466312 | 3.037865 | 14.10% |

**Table 5.4.1-2 Average GDP per capita by provinces (10,000 yuan, RMB)**

To further prove the result, we conduct the following estimation

$$Gdppercapita_{it} = \alpha + \beta_1 Heating_i + \beta_2 Latitude_i + \beta_3 Longitude_i + \beta_4 Enterprise_{it} + \beta_5 Public_{it} + \beta_6 City_i + \beta_7 Area_i + \beta_8 Per_{it} + u_i + \varepsilon_{it}$$

Enterprise records the newly established large enterprises in these counties. This variable is used to measure the economic activeness of these counties. This estimation controls the natural distinctions (Location group: latitude and longitude), governmental influences (Public) and economic activeness (Enterprise), political factors (administration division City), and county scale (Area and Per).

We conduct a Hausman-Taylor estimator, taking Enterprise and Public as the endogenous time-variant variables

**Table 5.4.1-3 Huai River policy and GDP per capita**

| Hausman-Taylor estimator | |
|---|---|
| Heating | 1.034397**<br>(0.5210683) |
| Per | -0.0217739**<br>(0.0107739) |
| Area | 0.0000403<br>(0.0003379) |
| Public | 0.0456623***<br>(0.0037827) |
| Enterprise | 0.0000954***<br>(0.0000346) |
| City | -0.0384978<br>(0.0255665) |
| Longitude | -0.0581336<br>(0.1383036) |
| Latitude | -0.763475<br>(0.7000194) |
| Constant | 35.11795<br>(37.03506) |
| Sargan-Hansen statistic | 0.150<br>P-value = 0.6981 |

| | |
|---|---|
| Wald chi2(8) | 267.76 |
| Prob > chi2 | 0.0000 |
| Obs | 520 |

Table 5.4.1-3 reports the estimation coefficients with a positive and significant variable of Heating, showing that the Heating area still has a larger GDP per capita. This is a fascinating finding, indicating that the Huai River policy could significantly impact the economic performance of, at least, the adjacent counties near the heating borderline.

Economic variables, including Enterprise and Public, have both positive coefficients. Permanent population and GDP per capita have a negative relationship. This is reasonable as GDP per capita is calculated by GDP divided by the permanent population.

We conduct the following robustness checks. The results are presented in Table 5.4.1-4. The coefficients remain good in robustness.

First scenario conducts an estimation in Equation (12)

Estimation I:

$$Gdppercapita_{it} = \alpha + \beta_1 Heating_{it} + \beta_2 Per_{it} + \beta_3 Area_{it} + \beta_4 Latitude_i + \beta_5 Longitude_i + \beta_6 Public_{i,t-1} + \beta_7 Enterprise_{i,t-1} + \varepsilon_{it} \quad (12)$$

We consider an imaginary heating borderline in the second scenario and conduct the benchmark methodology. We randomly pick up about half of the counties and take them as heating areas; the rest are non-heating areas. We can compare the estimation results to see whether they differ from those in the benchmark.

In the third scenario, we only leave the following variables and use a random effects estimation technique to estimate the following equation.

Estimation III

$$Gdppercapita_{it} = \alpha + \beta_1 Heating_i + \beta_2 Per_{it} + \beta_3 Latitude_i + \beta_4 Enterprise_{it} + \beta_5 Longitude_i + \varepsilon_{it}$$

**Table 5.4.1-4 Robustness checks**

| **Independent variables** | **I** | **II** | **III** |
|---|---|---|---|
| Heating | 1.230982**<br>(.632021) | 1.034397**<br>(.5210683) | 0.7180534**<br>(.3598585) |
| Per | -0.0171089<br>(.0140408) | -0.0217739**<br>(0.0107739) | -0.0228096***<br>(.0058292) |
| Area | -0.000139<br>(.0004224) | 0.0000403<br>(.0003379) | --------- |
| City | -0.0356073<br>(0.0275818) | -0.0384978<br>(0.0255665) | ---------- |
| Public<br>$(public_{-1})$ | 0.0523686<br>(0.0055616) | 0.0456623***<br>(0.0037827) | ----------- |

| | | | |
|---|---|---|---|
| Enterprise | 0.0000992** | 0.0000954*** | 0.0003167*** |
| | (0.0000398) | (0.0000346) | (0.0000333) |
| Longitude | -0.138844 | -0.0581336 | 0.127985** |
| | (0.1831247) | (0.1383036) | (0.0599158) |
| Latitude | -1.209871 | -0.763475 | -0.2205648 |
| | (0.9172723) | (0.7000194) | (0.2602569) |
| Constant | 59.16954 | 35.11795 | -2.979856 |
| | (49.53534) | (37.03506) | (10.7501) |
| Wald Chi | 163.04 | 267.76 | 99.47 |
| P-value | 0.000 | 0.000 | 0.000 |
| Obs | 520 | 520 | 520 |

### 5.4.2 IV estimator

In this section, we apply the IV estimator as a further robustness check. A simple version of the above estimation is considered:

$$Gdppercapita_{it} = \alpha + \beta_1 Heating_i + \beta_2 Latitude_i + \beta_3 Longitude_i + \beta_4 Enterprise_{it} + \beta_5 Per + \beta_6 City_i + \beta_7 Area_i + u_i + \varepsilon_{it}$$

In this model, Enterprise and Per are two endogenous variables with a simultaneity problem. We introduce the migration-related variables N (Natural growth rate), G (Growth rate), and Δ (Population increment) as three instrumental variables to eliminate the endogeneity problem. We conduct a two-stage least squares estimation. The weak instruments test shows that the instrumental variables are not weak instruments. The relevant test results are shown in Tables 5.4.2-1 to 4.

The estimation results are in Table 5.4.2-4.

**Table 5.4.2-1 Weak instruments test: F-test**

| **Estimation of Endogenous variables** | **Sanderson-Windmeijer multivariate F test** | **P value** |
|---|---|---|
| Enterprise | 10.94 | 0.000 |
| Per | 10.35 | 0.000 |

**Table 5.4.2-2 Weak instrument robust inference**

| Test name | Statistic | P value |
|---|---|---|
| Anderson-Rubin Wald test | F(3,511)= 23.14 | 0.0000 |
| Anderson-Rubin Wald test | Chi-sq(3)= 70.65 | 0.0000 |
| Stock-Wright LM S statistic | Chi-sq(3)= 62.20 | 0.0000 |

**Table 5.4.2-3 Unidentification test**

| Unidentification test | Statistic | P value |
|---|---|---|
| Anderson canon. corr. LM statistic | Chi-sq(2)= 11.920 | 0.0026 |

The key independent variable Heating has a significant and positive variable although the coefficient is smaller than that in the Hausman-Taylor estimation. Other variables have similar directions of estimation results in the benchmark.

**Table 5.4.2-4 Estimation results**

| Variables | Coefficient | Std. Err. | T-statistic | P-Value |
|---|---|---|---|---|
| Heating | .7405359 | .2401548 | 3.08 | 0.002 |
| Enterprise | .0012079 | .0002274 | 5.31 | 0.000 |
| Latitude | .1418125 | .2153436 | -0.66 | 0.510 |
| Longitude | .1536773 | .1644376 | 0.93 | 0.350 |
| Area | 0.0004313 | .0001777 | 2.43 | 0.015 |
| Per | -.0550323 | .0167061 | -3.29 | 0.001 |
| City | -.0625836 | .0155257 | -4.03 | 0.000 |

## 5.5 Light, GDP per capita and migration

We apply the following models to determine the relationship between migration and economic development in this setting. $\Delta_{it}$ is the population increment of the counties. $M_{it}$ is also used as the dependent variable to explore the relations. To simplify the estimation, we only introduce latitude and longitude to control the natural conditions, including climate, location, and administrative divisions. Enterprise is introduced to control economic development.

$$light_{it} = \alpha + \beta_1\Delta_{it} + \beta_2 Latitude_i + \beta_3 Enterprise_{it} + \beta_5 Longitude_i + \varepsilon_{it}$$

$$GDPpercapita_{it} = \alpha + \beta_1\Delta_{it} + \beta_2 Latitude_i + \beta_3 Enterprise_{it} + \beta_5 Longitude_i + \varepsilon_{it}$$

$$light_{it} = \alpha + \beta_1 M_{it} + \beta_2 Latitude_i + \beta_3 Enterprise_{it} + \beta_5 Longitude_i + \varepsilon_{it}$$

$$GDPpercapita_{it} = \alpha + \beta_1 M_{it} + \beta_2 Latitude_i + \beta_3 Enterprise_{it} + \beta_5 Longitude_i + \varepsilon_{it}$$

We conduct a random-effects panel data model.

The estimation results are reported in Table 5.5-1. The results show that migration is not correlated with economic performance. This is beyond our intuition. Normally, population inflow will generate more economic development. We can think of explanations for this. In the short term, population inflow can enlarge the population size. The per capita economic resources can be reduced. However, in the long-run, the situation might be inverted. Thus, the relationship between the two terms is complicated. To reduce the possible endogeneity problem, we introduce a lagged term of $\Delta_t$ as the robustness checks. The estimation models are as follows:

$$light_{it} = \alpha + \beta_1\Delta_{i,t-1} + \beta_2 Latitude_i + \beta_3 Enterprise_{it} + \beta_5 Longitude_i + \varepsilon_{it}$$

$$Gdppercapita_{it} = \alpha + \beta_1\Delta_{i,t-1} + \beta_2 Latitude_i + \beta_3 Enterprise_{it} + \beta_5 Longitude_i + \varepsilon_{it}$$

**Table 5.5-1 Estimation results-delta**

| | **I: Light** | **II: Gdppercapita** | **III：Light** | **IV：Gdppercapita** |
|---|---|---|---|---|
| Δ | 0.0437125<br>(0.0878776) | -0.014565<br>(0.0192119) | | |
| M | | | 4.586707<br>(7.638235) | -.5412738<br>(1.668713) |
| latitude | .1742906<br>.729739 | -0.0320921<br>(.2525) | 0.1786392*<br>(.7320297) | -0.0343867<br>(.2544385) |
| longitude | .0199132<br>.1388227 | -.0427831 (.0470287) | .0227992 (.139154) | -.0438001<br>(.0473662) |
| enterprise | .0016749 ***<br>.0001383 | .0002857 (.0000327) | .0017495 ***<br>(.0001439) | .0002868<br>(.0000327) |
| $R^2$ | 0.3198 | 0.1571 | 0.2445 | 0.1566 |
| Wald-Chi | 215.69 | 77.75 | 163.25 | 77.47 |
| Obs | 520 | 520 | 520 | 520 |

**Table 5.5-2 Robustness checks**

| | **I: Light** | **II: GDPpercapita** |
|---|---|---|
| Δ (-1) | 0.0432885<br>(0.0906496) | -.0021937<br>(0.019872) |
| $R^2$ | 0.2237 | 0.1057 |
| Wald-Chi | 94.20 | 38.89 |
| Obs | 520 | 520 |

The above analysis shows no direct relationship between the population inflow terms and economic development in this case. Combining the estimation results of these sections, we can infer that the Impact II in Figure 1.4-1 induced by FDI and migration has a minor impact on economic development.

## 5.6 Summary

Through our empirical analysis, we have the following important results.

First, areas with central heating systems have higher GDP per capita and nighttime luminosity than those without central heating. The Huai River policy generates differences in economic development on both sides. Central heating areas have better economic performance than non-heating areas after locations, population size, city scale, administrative levels, and other economy-related variables are controlled. Southern areas in China are always deemed to have better economic vitality than northern areas. This study gives us a different result: non-heating counties are situated more southward but have worse economic performance because of the Huai River policy.

Second, the policy-affected migration and FDI are not the decisive factors to affect economic performance. The results of the previous chapters tell us that heating areas just south of the heating borderline did not attract more population inflow, while non-heating areas attracted more population inflow. At the same time, heating areas have not attracted more foreign investment inflows, which shows that the relationship between heating areas, population and foreign investment inflows is unclear. The result of this chapter tells us that heating areas have better economic performance, which is reflected in economic and human activities. Heating areas have higher night light brightness and per capita GDP. This shows that it is the heating industry itself that promotes local economic development, and population migration may not directly affect the economic development and foreign investment inflow is not the key factor, either.

From the policy and practice perspective, the Huai River policy has impacted China's county economic development. Therefore, for policymakers, in this dimension of economic development, advancing the heating boundary line to the south is conducive to the economic development of the southern regions. Of course, the study also found that the Huai River policy could lead to more air pollution. Hence, policymakers need to consider the environmental costs and energy consumption when pushing the heating boundary more southward.

# Chapter 6 Conclusion and policy implications

## 6.1 Conclusion

The study takes the Huai River policy as the primary research object and explores the policy's impact on human and economic activities. This study is divided into several sub-studies as follows. Firstly, we study the impact of central heating policy on air quality. Secondly, the impact of the Huai River policy on human migration is explored. Thirdly, based on the research analyzing the Huai River policy's impact on air quality and migration activities, we further explore the impact of the policy's impact on economic development. This study combines theoretical research and empirical research. It initially explores the role of variables by establishing a theoretical framework. Then, we use data and statistical methods to verify the model and propose new explanations for the phenomenon. In this study, the county-level data near the heating borderline and the data of prefecture-level cities are used to verify the hypotheses proposed in the model. Econometric methods such as discontinuity regression and panel regression are applied. The essential conclusions of this study are summarized below.

First, it is evident that central heating increases the emission of air pollutants in the heating area, causing more significant pollution in the heating area in winter. The government's implementation of clean heating policies has apparent effects on curbing air pollution. In recent years, air pollution has been greatly abated.

Second, policy's impact on migration and FDI is complicated. People's (or enterprises') comprehensive considerations of air quality, indoor living environment, and heating costs influence their settlement decisions, affecting population migration and FDI in heating and non-heating areas. In areas around the heating borderline, counties south of the heating borderline have experienced more net population inflows. In contrast, counties north of the heating borderline have seen more net population outflows. Heating has not attracted more people to move in. One explanation is that the deterioration of air quality caused by heating may discourage people. At the same time, the high cost of central heating makes many movers not regard central heating as a primary factor in settling decisions. Empirical research shows that heating costs may be a factor that more residents consider. In areas with relatively developed economies, people pay more attention to the improvement of living conditions brought about by heating. In contrast, in areas with less developed economies, people pay more attention to the costs of heating.

Third, areas farther away from the heating borderline have a higher degree of population migration in-migration as winter temperature dramatically increases and the demand for central heating declines. Areas that are not that farther away will have a trade-off on this product, thus forming a U-shaped relationship. This result is different from the second one. The second one focuses on the areas around the heating borderline.

The last result is that the per capita GDP level and night light brightness in heating areas are significantly higher than in non-heating areas, and economic activities are significantly more active. To better explore the causal relationship between central heating and its impact on human and economic activities, we use a county-level dataset with counties like the estimation in the second conclusion. Chapter 6 and Chapter 8 use the same dataset. This result may bring some confusion. The second result indicates that heating areas do not attract more population inflow, while the non-heating areas have more population inflow. By intuition, people will think that areas with more population inflow will have higher levels of economic development. The conclusion in Chapter 8 suggests that population inflow is not positively correlated with economic development always. In the short run, the population inflow can reduce the per capita economic variables. In the long run, the impact might be positive. In this case, the influence of migration on economic development is negligible. It is the heating industry that generates GDP both in upstream and downstream industries. In the meantime, counties in heating areas will have more cogeneration plants that produce both heating and electricity and might be located nearer the county cities, thus leading to higher nighttime luminosity levels.

Figure 6.1-1 illustrates the estimation results found in this study. The Huai River policy in this graph stands for the existence of central heating systems. +/- means positive/negative correlations. × means no statistical correlation.

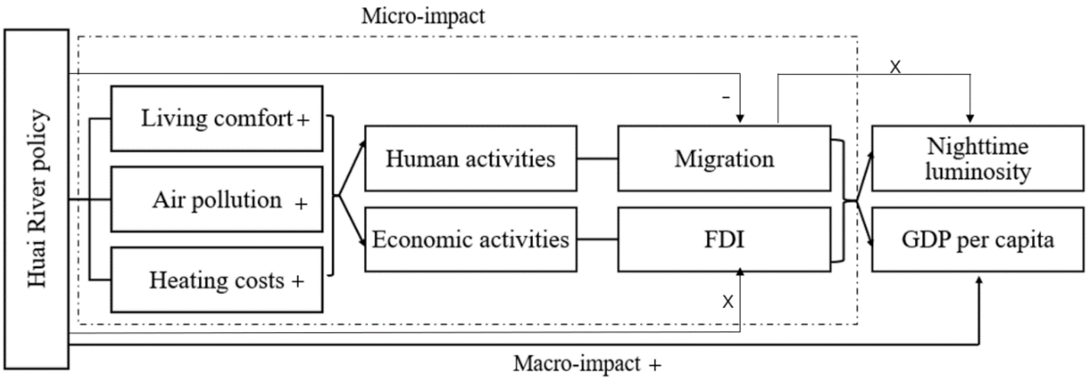

Figure 6.1-1 Policy effects

## 6.2 Policy implications and future prospect

### 6.2.1 A good industry engine in the future

The Huai River policy has dual impacts on social development. First, it undermines the air quality, increases carbon emissions, and brings social costs for environmental protection. However, it promotes nighttime luminosity and economic development. Clean heating projects can mitigate air pollution. Policymakers must balance the advantages and disadvantages of the policy when considering forwarding the systems farther south. Policymakers can take the heating industry as an excellent economic engine by promoting clean energy, the heating industry, and livable residential projects. Clean heating technology helps the energy and heating industry boost industry upgrades.

The downstream and upstream industries can bring booms to the whole supply chain. Households can be bettered off living in more comfortable apartments. After balancing the environmental and economic factors, it is a good direction for cultivating a new industry for policymakers in areas without central heating

### 6.2.2 Limited function on attraction for talents and capital

Heating, as one type of infrastructure or amenities, has limited attraction for population migration. Considering the possible higher costs and air pollution caused by central heating, heating's impact on either population migration or foreign capital investment is limited. The main factors that attract the population are economic levels, geographical locations, and the natural environment. In addition to economic factors, the main factors for attracting foreign investment include policy factors and geographical location.

### 6.2.3 Global warming and policy revisions

In the context of global warming, rising winter temperatures will also become an important consideration in future revisions of the Huai River policy. In China, there is more and more probability of experiencing warm winters. Many areas originally in the warm temperate zone are gradually transitioning to the subtropics, and the winter temperatures in some subtropical areas are gradually beginning to be consistent with those of the tropics. Promoting central heating in the southern region will bring more carbon emissions. Increased carbon emissions will further exacerbate global warming. The impact of air temperature on heating distribution is considered in our theoretical model, and our model can be used to simulate the impact of climate warming on other variables.

### 6.2.4 Policy implications

We summarize the policy implications. First, boosting the use of clean energy in central heating can further improve air quality and reduce carbon emissions in heating districts. Second, promoting the southward extension of central heating areas can improve the living conditions in non-heating areas and improve the economic performance when introducing technology innovation and upgrades. Third, people in economically developed areas are more yearning for central heating, while people in relatively less developed areas have a more ambiguous attitude towards central heating. Therefore, when the central heating dividing line advances southward, it can also be developed for different regions. Different policies formulate different price plans. Fourth, policymakers should balance the air quality, carbon emission, and economic growth when carrying on policy revisions in the future.

# REFERENCES

[1]Gong, M. and S. Werner, *An assessment of district heating research in China.* Renewable Energy, 2015. **84**: p. 97-105.

[2]Almond, D., et al., *Winter heating or clean air? Unintended impacts of China's Huai river policy.* American Economic Review, 2009. **99**(2): p. 184-190.

[3]Zheng, P., *China's Geography*. 2006: Wuzhou Publishing House.

[4]Guo, J., Y. Huang, and C. Wei, *North–South debate on district heating: Evidence from a household survey.* Energy Policy, 2015. **86**: p. 295-302.

[5]Cai, H.W., Jin *Factors influencing the migration intentions of rural workers in the Pearl River Delta.* Social Sciences in China, 2008. **29**(3): p. 157-171.

[6]Rezaie, B. and M.A. Rosen, *District heating and cooling: Review of technology and potential enhancements.* Applied energy, 2012. **93**: p. 2-10.

[7]Nuorkivi, A., *District heating and cooling policies worldwide*, in *Advanced district heating and cooling (DHC) systems*. 2016, Elsevier. p. 17-41.

[8]Marinova, M., et al., *Economic assessment of rural district heating by bio-steam supplied by a paper mill in Canada.* Bulletin of Science, Technology Society, 2008. **28**(2): p. 159-173.

[9]Gao, Y., & Sampattavanija, S. (2022). Air quality and winter heating: some evidence from China. *International Journal of Energy Economics and Policy*, *12*(4), 455-469.

[10]Cichowicz, R., G. Wielgosiński, and W. Fetter, *Dispersion of atmospheric air pollution in summer and winter season.* Environmental monitoring assessment, 2017. **189**: p. 1-10.

[11]Wang, M., et al., *Air particulate pollution in Ulaanbaatar, Mongolia: Variation in atmospheric conditions from autumn to winter.* Sola, 2017. **13**: p. 90-95.

[12]Guttikunda, S. (2008). Urban air pollution analysis for Ulaanbaatar, Mongolia.

[13]Angius, S., et al., *Evaluation of the effects of traffic and heating reduction measures on urban air quality.* Atmospheric Environment, 1995. **29**(23): p. 3477-3487.

[14]Ghafghazi, S., et al., *Particulate matter emissions from combustion of wood in district heating applications.* Renewable Sustainable Energy Reviews, 2011. **15**(6): p. 3019-3028.

[15]Leaderer, B.P., et al., *Indoor, outdoor, and regional summer and winter concentrations of PM10, PM2. 5, SO4 (2)-, H+, NH4+, NO3-, NH3, and nitrous acid in homes with and without kerosene space heaters.* Environmental Health Perspectives, 1999. **107**(3): p. 223-231.

[16]Liang, X., et al., *Assessing Beijing's PM2. 5 pollution: severity, weather impact, APEC and winter heating.* Proceedings of the Royal Society A: Mathematical, Physical Engineering Sciences, 2015. **471**(2182): p. 20150257.

[17]Zhang, W., et al., *A refined source apportionment study of atmospheric PM2. 5 during winter heating period in Shijiazhuang, China, using a receptor model coupled with a source-oriented model.* Atmospheric Environment, 2020. **222**: p. 117157.

[18]Xiao, Q., et al., *The impact of winter heating on air pollution in China.* PloS one, 2015. **10**(1): p. e0117311.

[19]Chen, Y., & Whalley, A. (2012). Green infrastructure: The effects of urban rail transit on air quality. *American Economic Journal: Economic Policy*, *4*(1), 58-97.

[20]Wen, Z., et al., *Winter air quality improvement in Beijing by clean air actions from 2014 to 2018.* Atmospheric Research, 2021. **259**: p. 105674.

[21]Zhang, Y., W. Li, and F. Wu, *Does energy transition improve air quality? Evidence derived from China's Winter Clean Heating Pilot (WCHP) project.* Energy Policy, 2020. **206**: p. 118130.

[22]Wang, S., et al., *Can clean heating in winter in northern China reduce air pollution?—empirical analysis based on the psm-did method.* Energies, 2022. **15**(5): p. 1839.

[23]Tan, X., G. Chen, and K. Chen, *Clean heating and air pollution: Evidence from Northern China.* Energy Reports, 2023. **9**: p. 303-313.

[24]Guo, X., D. Ren, and C. Li, *Study on clean heating based on air pollution and energy consumption.* Environmental Science Pollution Research, 2020. **27**: p. 6549-6559.

[25]Liu, X., et al., *Air pollution and high human capital population migration: An empirical study based on 35 major cities in China.* Sustainable Production Consumption, 2021. **27**: p. 643-652.

[26]Liu, Z. and L. Yu, *Stay or leave? The role of air pollution in urban migration choices.* Ecological Economics, 2020. **177**: p. 106780.

[27]Germani, A.R., et al., *Does air pollution influence internal migration? An empirical investigation on Italian provinces.* Environmental science policy, 2021. **120**: p. 11-20.

[28]Wang, K., Y. Wang, and C.P. Chang, *The impact of air quality on population migration.* Problemy Ekorozwoju, 2020. **15**(2): p. 23-31.

[29]Kim, B., *Do Air Quality Alerts Affect Household Migration?* Southern Economic Journal, 2019. **85**(3): p. 766-795.

[30]Berman, M., *Energy costs and rural Alaska out-migration.* 2017.

[31]Gale, L.R. and W.C. Heath, *Elderly internal migration in the United States revisited.* Public Finance Review, 2000. **28**(2): p. 153-170.

[32]Fielding, A., *The impacts of environmental change on UK internal migration.* Global Environmental Change, 2011. **21**: p. S121-S130.

[33]Parshukov, D., Z. Shaporova, and J. Koloskova, *Study Of Life Quality Indicators Influence On Rural Population Migration.* European Proceedings of Social Behavioural Sciences, 2020.

[34]Gao, Y. and S. Sampattavanija, *CHINA'S HUAI RIVER POLICY AND POPULATION MIGRATION: A U-SHAPED RELATIONSHIP.* International Journal of Energy, Environment, and Economics, 2021. **29**(4): p. 471-491.

[35]Harris, R., *J.; Todaro, P., Michael. Migration, Unemployment and Development: A Two Sector Analysis.* The American Economic Review, 1970. **60**(1): p. 126-142.

[36]Brueckner, J.K.Z., Yves *Harris-Todaro models with a land market.* Regional Science Urban Economics, 1999. **29**(3): p. 317-339.

[37]Laing, D., C. Park, and P. Wang, *A modified Harris-Todaro model of rural-urban migration for China.* Critical issues in China’s growth development, 2005: p. 245-264.

[38]Copeland, B.R. and M.S. Taylor, *North-South trade and the environment*, in *International Trade and the Environment*. 2017, Routledge. p. 205-238.

[39]Beghin, J.R.-H., David Van der Mensbrugghe, Dominique *Trade and pollution linkages: piecemeal reform and optimal intervention.* Canadian Journal of Economics, 1997: p. 442-455.

[40]Daitoh, I., *Environmental protection and urban unemployment: Environmental policy reform in a polluted dualistic economy.* Review of Development Economics, 2003. **7**(3): p. 496-509.

[41]Fukuyama, H. and T. Naito. *Unemployment, trans-boundary pollution, and environmental policy in a dualistic economy*. in *Review of Urban & Regional Development Studies: Journal of the Applied Regional Science Conference*. 2007. Wiley Online Library.

[42]Simonsohn, U., *Two lines: A valid alternative to the invalid testing of U-shaped relationships with quadratic*

*regressions.* Advances in Methods Practices in Psychological Science, 2018. **1**(4): p. 538-555.

[43]Gao, Y., & Sampattavanija, S. (2023). Central Heating Policy and Population Migration in China: An Empirical Study. *International Journal of Energy Economics and Policy*, *13*(4), 312-319.

[44]Dong, D., X. Xu, and Y.F. Wong, *Estimating the impact of air pollution on inbound tourism in China: An analysis based on regression discontinuity design.* Sustainability, 2019. **11**(6): p. 1682.

[45]Li, W. and K. Zhang, *Does air pollution crowd out foreign direct investment inflows? Evidence from a quasi-natural experiment in China.* Environmental Resource Economics, 2019. **73**: p. 1387-1414.

[46]Cho, C.H., et al., *Contaminated heart: Does air pollution harm business ethics? Evidence from earnings manipulation.* Journal of Business Ethics, 2021: p. 1-22.

[47]Lin, T., et al., *Air pollution and workplace choice: Evidence from China.* International Journal of Environmental Research Public Health, 2022. **19**(14): p. 8732.

[48]Qiao, Z., Z. Li, and Y. Wang, *Air pollution and innovation-evidence from quasi-natural experiment of China Huai River policy.* Review of Quantitative Finance Accounting, 2023. **60**(2): p. 425-443.

[49]Liu, X., Li, Q., Chand, S., & Sharpe, K. (2021). Effects of air quality on house prices: evidence from China's Huai River Policy. *New Zealand Economic Papers*, *55*(1), 52-65.

[50]Zhang, S., Zhou, Y., & Xu, P. (2022). Air quality affects house prices—Analysis based on RD of the Huai River policy. *Sustainable Cities and Society*, *85*, 104017.

[51] Huang, Z., & Du, X. (2022). Does air pollution affect investor cognition and land valuation? Evidence from the Chinese land market. *Real Estate Economics*, *50*(2), 593-613.

[52]Chen, Y., et al., *Evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River policy.* Proceedings of the National Academy of Sciences, 2013. **110**(32): p. 12936-12941.

[53]Ebenstein, A., et al., *New evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River Policy.* Proceedings of the National Academy of Sciences, 2017. **114**(39): p. 10384-10389.

[54]Ito, K. and S. Zhang, *Willingness to pay for clean air: Evidence from air purifier markets in China.* Journal of Political Economy, 2020. **128**(5): p. 1627-1672.

[55]Henderson, J.V., A. Storeygard, and D.N. Weil, *Measuring economic growth from outer space.* American economic review, 2012. **102**(2): p. 994-1028.

[56]Chen, X.N., William D *Using luminosity data as a proxy for economic statistics.* Proceedings of the National Academy of Sciences, 2011. **108**(21): p. 8589-8594.

[57]Doll, C.N., J.-P. Muller, and J. Morley, *Mapping regional economic activity from night-time light satellite imagery.* Ecological Economics, 2006. **57**(1): p. 75-92.

[58]Beyer, R., et al., *Measuring districts' monthly economic activity from outer space.* World Bank Policy Research Working Paper, 2018(8523).

[59]Doll, C.N. and S. Pachauri, *Estimating rural populations without access to electricity in developing countries through night-time light satellite imagery.* Energy policy, 2010. **38**(10): p. 5661-5670.

[60]Kumar, P., et al., *Modeling the luminous intensity of Beijing, China using DMSP-OLS night-time lights series data for estimating population density.* Physics Chemistry of the Earth, Parts A/B/C, 2019. **109**: p. 31-39.

# Appendix I Information of Heating Cities in Chapter 3

| Province name | City | Location | Start | End | Average winter temperature (2016) | Average winter temperature (2017) |
|---|---|---|---|---|---|---|
| **Shandong** | Qingdao | 36.03N 120.18E | 11.15 | 04.05 | 2.8 | 1.3 |
| | Jinan | 36.40N 117.00E | 11.15 | 03.15 | 3.1 | 1.3 |
| | Weifang | 36.43N 119.06E | 11.15 | 03.15 | 1.5 | 0.5 |
| | Yantai | 37.32N 121.24E | 11.16 | 03.30 | 0.9 | -1.1 |
| | Rizhao | 35.23N 119.32E | 11.25 | 03.25 | 2.7 | 1.4 |
| | Taian | 36.11N 117.08E | 11.15 | 03.15 | 2.0 | 0.2 |
| | Zhangqiu | 36.72N 117.53E | 11.15 | 03.15 | 1.8 | 0.5 |
| **Inner Mongolia** | Baotou | 40.58N 110.00E | 10.15 | 04.15 | -7.2 | -8.8 |
| | Hohhot | 40.82N 111.65E | 10.15 | 04.15 | -6.9 | -9.3 |
| **Heilongjiang** | Harbin | 45.80N 126.53E | 10.20 | 04.20 | -13.7 | -17.7 |
| | Mudanjiang | 44.58N 129.60E | 10.15 | 04.15 | -12.0 | -15.2 |
| | Qiqihar | 47.33N 123.95E | 10.15 | 04.15 | -14.3 | -17.7 |
| **Jilin** | Changchun | 43.90N 125.32E | 10.25 | 04.10 | -9.7 | -13.4 |
| **Beijing** | Beijing | 39.90N 116.40E | 11.15 | 03.15 | 0.2 | -1.5 |
| **Tianjin** | Tianjin | 39.12N 117.20E | 11.15 | 03.15 | 0.5 | -1.8 |
| **Henan** | Kaifeng | 34.80N 114.30E | 11.15 | 03.15 | 4.3 | 2.9 |
| | Anyang | 36.10N 114.38E | 11.15 | 03.15 | 3.1 | 1.5 |
| | Zhengzhou | 34.75N 113.62E | 11.15 | 03.15 | 4.6 | 3.2 |
| | Sanmenxia | 34.78N 111.20E | 11.15 | 03.15 | 3.4 | 1.3 |
| **Hebei** | Chengde | 40.97N 117.93E | 11.01 | 03.25 | -5.2 | -7.4 |
| | Zhangjiakou | 40.82N 114.88E | 11.01 | 03.31 | -5.0 | -7.5 |
| | Qinhuangdao | 39.93N 119.60E | 11.05 | 04.05 | -1.1 | -4.9 |
| | Tangshan | 39.63N 118.20E | 11.15 | 03.15 | -2.2 | -4.0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | Xingtai | 37.07N 114.48E | 11.15 | 03.15 | 1.6 | 0.7 |
| | Langfang | 39.52N 116.70E | 11.15 | 03.15 | -1.4 | -3.6 |
| | Baoding | 38.87N 115.47E | 10.15 | 04.15 | -0.9 | -2.2 |
| **Shanxi** | Taiyuan | 37.87N 112.55E | 11.01 | 03.31 | -1.1 | -3.3 |
| | Datong | 40.08N 113.30E | 10.25 | 04.10 | -7.2 | -7.3 |
| | Linfen | 36.08N 111.52E | 11.15 | 03.15 | 2.9 | 0.7 |
| **Ningxia** | Yinchuan | 38.47N 106.28E | 11.01 | 03.31 | -2.2 | -4.9 |
| **Liaoning** | Yingkou | 40.67N 122.23E | 11.01 | 04.01 | -3.9 | -6.7 |
| | Anshan | 41.10N 122.98E | 11.01 | 03.31 | -3.2 | -3.6 |
| | Benxi | 41.30N 123.77E | 11.01 | 03.31 | -6.7 | -9.4 |
| | Dandong | 40.13N 124.38E | 11.01 | 03.31 | -3.9 | -6.9 |
| | Shenyang | 41.80N 123.43E | 11.01 | 03.31 | -7.0 | -9.9 |
| **Shaanxi** | Yan’an | 36.60N 109.48E | 11.01 | 03.31 | -1.7 | -4.6 |
| **Jiangsu** | Xuzhou | 34.27N 117.18E | 11.21 | 03.10 | 4.2 | 2.5 |
| **Xinjiang** | Korla | 41.77N 86.15E | 11.01 | 03.31 | -3.5 | -5.8 |
| | Karamay | 45.60N 84.87E | 10.15 | 04.15 | -11.1 | -13.2 |
| | Urumqi | 43.82N 87.62E | 10.10 | 04.10 | -6.8 | -10.0 |
| **Qinghai** | Xining | 36.62N 101.78E | 10.15 | 04.15 | -4.7 | -6.7 |

Note: The average winter temperature originates from the calculation of meteorological data.

# VITA

| | |
|---|---|
| **NAME** | Yannan Gao |
| **DATE OF BIRTH** | 12 October 1990 |
| **PLACE OF BIRTH** | Weifang,China |
| **INSTITUTIONS ATTENDED** | Shandong University (Bachelor, 2008-2012)<br>University of Konstanz (Master, 2012-2014) |