University of Southampton

School of Economic, Social and Political Sciences
Department of Economics

Environmental and Socioeconomic Dynamics: How Air Pollution and Migration Shape Behavior and Aging in China

by

Yu Qin

Mater of Economics

ORCiD: 0009-0000-9780-5523

A thesis for the degree of
Doctor of Philosophy

December 2025

University of Southampton

Abstract

School of Economic, Social and Political Sciences
Department of Economics

Doctor of Philosophy

Environmental and Socioeconomic Dynamics: How Air Pollution and Migration Shape Behavior and Aging in China

by Yu Qin

This Thesis explores how environmental challenges, particularly air pollution, shape migration, retirement, and behaviour in China. Using robust econometric methods and nationally representative data, it examines three interrelated aspects.

First, it investigates how rural-to-urban migration affects environmental behaviours among left-behind families. Current migration reduces recycling willingness by 71.9% and fixed garbage placement by 21.8%, while return migration increases these behaviours by 79.0% and 46.3%, respectively. Mechanism analysis highlights the roles of green infrastructure and social remittance in driving these effects.

Second, it examines air pollution's effect on urban-to-urban migration flows. A doubling of the destination-to-origin relative PM2.5 concentration between destination and origin cities reduces migration inflows by 42%. The relationship is influenced by migration distance, infrastructure, and settlement costs, with older, middle-educated, married male migrants more affected by pollution disparities.

Third, it explores how air pollution impacts retirement expectations. A 1% increase in PM2.5 concentration reduces the expected retirement age by 5.11 months, with rural residents facing larger declines (9.03 months) than urban residents (4.45 months). Mechanisms such as financial support, green infrastructure, and welfare systems mitigate pollution's adverse effects, while dynamic pollution shocks amplify early retirement adjustments.

This research contributes to understanding the socioeconomic impacts of environmental degradation. By integrating migration, environmental, and labour economics, it offers insights into policies promoting sustainable urbanization, green infrastructure, and welfare systems, particularly for vulnerable groups.

Keywords: Migration Dynamics, Air Pollution, Green Behaviours, Sustainability and Development, Retirement Expectations

Contents

List of Figures

List of Tables

Declaration of Authorship

Print Name: Yu Qin

Title of thesis: Environmental and Socioeconomic Dynamics: How Air Pollution and Migration Shape Behavior and Aging in China

I declare that this thesis and the work presented in it is my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. None of this work has been published before submission

Signed:.. Date: 12/10/2025

Acknowledgements

This thesis represents the culmination of years of research, learning, and personal growth, made possible through the guidance, support, and encouragement of many individuals. I would like to take this opportunity to express my deepest gratitude to those who have contributed to this journey.

First and foremost, I extend my sincere appreciation to my supervisors, **Professor Corrado Giulietti** and **Professor Jackie Wahba**, for their invaluable guidance, intellectual insights, and continuous encouragement throughout my doctoral studies. Their constructive feedback and unwavering support have been instrumental in shaping this research and refining my academic skills.

I am also deeply grateful to my thesis committee members for their valuable advice, critical discussions, and insightful suggestions, which have significantly strengthened this work.

My heartfelt thanks go to my colleagues and friends, **Dr.Dingzhi Chen, Dr. Fengtian Yang, Dr. Yijie Chen**, and **Dr. Rick Boidurjo**, whose stimulating discussions and shared experiences have made this journey more enriching and enjoyable. The academic environment and collaboration have provided both inspiration and motivation during challenging moments.

Special thanks to my father, **Mr. Zhanjun Qin**, my mother, **Mrs. Huaijun Li**, my dearly loved late grandfather, **Mr. Liang Huai**, and other family members, for their unconditional love, patience, and encouragement. Their financial support made this research possible, allowing me to focus on my studies and conduct the necessary analyses without financial constraints. Above all, their belief in my abilities and unwavering support have been my greatest source of strength throughout this journey.

Finally, I would like to extend my appreciation to all those who have, in one way or another, contributed to this thesis. Their support — whether academic, emotional, or logistical — has been invaluable, and I am truly grateful for their presence on this journey.

Symbols and Abbreviations

Air Pollution The presence of harmful substances in the atmosphere, including particulate matter (PM2.5), nitrogen dioxide (NO), and other pollutants, which adversely affect human health, economic productivity, and environmental quality.

CHARLS China Health and Retirement Longitudinal Study – A nationally representative longitudinal survey focusing on middle-aged and older Chinese adults.

CLDS China Labour Dynamics Survey – A comprehensive dataset covering employment, income, and social behaviours, including environmental practices.

CMDS China Migration Data Survey – A dataset providing information on migration flows and characteristics across Chinese cities.

Differential Gravity Model An econometric framework that examines migration flows based on relative conditions, such as the pollution gap between origin and destination cities.

Dynamic Pollution Shock Sudden increases in pollution levels over a short period, often used to study immediate behavioural or economic responses.

GDP Gross Domestic Product – The total economic output of a region, used to measure economic activity and development.

Green Behaviours Environmentally sustainable practices, such as recycling, proper waste disposal, and reduced resource consumption, that contribute to environmental protection.

Green Infrastructure Infrastructure designed to promote sustainability and environmental health, such as recycling facilities, clean energy systems, and waste management programs.

Hukou System China's household registration system, which ties access to public services and benefits to an individual's place of official residence, creating urban-rural disparities.

Instrumental Variable (IV) A statistical technique used to address endogeneity by identifying exogenous variation that affects the dependent variable through the endogenous regressor.

NASA National Aeronautics and Space Administration – Provides satellite-based environmental data, including air pollution metrics.

NBS National Bureau of Statistics of China – The central agency for statistical data collection and dissemination in China.

NCCA National Civilised City Award – A program in China aimed at improving urban sustainability through environmental and social initiatives.

PM2.5 Fine particulate matter with a diameter of 2.5 microns or less. These particles are small enough to penetrate the respiratory system, causing significant health risks such as respiratory diseases, cardiovascular issues, and premature death.

Retirement Expectations Anticipated age at which individuals plan to exit the labour force, influenced by economic, social, and environmental factors.

Return Migration The process by which individuals return to their place of origin after living in another region, often bringing back skills, behaviours, and practices acquired during their time away.

Social Capital The networks, norms, and trust that facilitate coordination and cooperation within a community, often influencing migration and green behaviours.

Social Remittances The transmission of cultural, social, and behavioural norms from host to origin regions through migration. This includes practices related to environmental sustainability, health behaviours, and social capital.

SWB Subjective Well-Being – An individual's self-reported evaluation of their happiness and life satisfaction, often used as an outcome variable in socioeconomic studies.

Thermal Inversion Days Meteorological phenomena where cooler air is trapped beneath warmer air, leading to the accumulation of pollutants near the surface. Used as an instrumental variable for air pollution in econometric analysis.

Urban-Rural Disparities Systematic differences in infrastructure, healthcare, and economic opportunities between urban and rural areas, often exacerbated by migration and environmental stressors.

WHO World Health Organization – An international body overseeing global health and providing air quality standards.

WU Washington University in St. Louis – A source of city-level PM2.5 data for environmental research.

Chapter 1

Introduction

The interplay between environmental challenges and socioeconomic behaviours has become an increasingly important area of research in light of rapid urbanisation, demographic transitions, and rising environmental stressors. Air pollution, in particular, poses significant threats to public health, economic productivity, and social stability, necessitating a deeper understanding of its broader implications. Migration and retirement decisions, shaped by environmental and economic factors, reflect complex adaptation mechanisms that influence both individual well-being and societal development. This thesis investigates these dynamics, focusing on how air pollution and migration interact to shape behavioural and economic outcomes in China, providing insights into sustainable development and policy responses.

1.1 Background and Motivation

Environmental challenges have emerged as defining issues of the 21st century, influencing economic stability, societal well-being, and public health on a global scale. Among these challenges, air pollution stands out as a critical concern due to its far-reaching impacts. As economies grow and urbanise, the pressures on environmental systems intensify, creating complex feedback loops that affect migration patterns, labour markets, and demographic structures. These dynamics are particularly pronounced in developing economies, where rapid industrialisation often comes at the expense of environmental and social sustainability.

China exemplifies the dual pressures of economic transformation and environmental degradation. As the world's largest developing economy, its remarkable growth has been accompanied by substantial increases in air pollution, contributing to public health crises and socioeconomic disparities. Meanwhile, the country's internal migration system—one of the largest in the world—plays a pivotal role in shaping its labour markets and social structures. Combined with a rapidly aging population, these factors create a unique context for studying the interplay between environmental stressors, migration, and socioeconomic outcomes.

This thesis explores these interconnections, focusing on the ways air pollution influences migration flows, reshapes environmental behaviours, and impacts retirement planning. By integrating perspectives from environmental economics, migration studies, and labour economics, it aims to address critical gaps in understanding how environmental stressors drive socioeconomic behaviours and outcomes. The following section provides the necessary background and motivation, outlining the broader global context before narrowing the focus to China's unique challenges and contributions to these fields of study.

1.1.1 Global Environmental Challenges

Environmental degradation has emerged as one of the most pressing global challenges, with air pollution posing a critical threat to human health, economic stability, and social well-being. According to the World Health Organization's 2021 report, ambient air pollution causes approximately seven million premature deaths each year, with the vast majority occurring in low- and middle-income countries (World Health Organization, 2021). Beyond its devastating health toll, air pollution imposes substantial economic costs. It reduces labour productivity, increases healthcare expenditures, and erodes human capital, particularly in developing economies where institutional constraints limit the capacity for effective environmental governance (World Bank and Institute for Health Metrics and Evaluation, 2016).

From an economic perspective, air pollution affects individuals and societies both directly and indirectly. Acute exposure to fine particulate matter ($PM_{2.5}$) significantly impairs worker performance. Graff Zivin and Neidell (2013) find that exposure to $PM_{2.5}$ reduces cognitive and physical function across various occupations, especially in labour-intensive sectors. For instance, Chang et al. (2019) show that a 10 $\mu g/m^3$ increase in $PM_{2.5}$ concentration leads to a 6% decline in daily worker output in China, highlighting the disproportionate burden on outdoor workers. At the macroeconomic level, these productivity losses aggregate into substantial GDP reductions and widen income inequality, especially in regions with limited environmental enforcement.

Migration has become an important adaptation strategy to environmental pressures. While traditionally seen as a consequence of economic or social factors, environmental stressors—such as air pollution and climate variability—are increasingly recognised as determinants of migration decisions. Studies by Cai et al. (2016) and Deschênes and Greenstone (2011); Fan and Li (2019) provide evidence that environmental shocks, including pollution-induced declines in agricultural productivity and extreme weather events, drive population displacement. Nevertheless, migration is not universally accessible.

Economic constraints and social barriers limit mobility for many low-income households, making them more vulnerable to environmental risks (Boyce, 2013). At the same time, the presence of historical migration networks plays a critical role in facilitating relocation by lowering informational and financial barriers to movement.

Although migration can offer improved living standards for movers, its impact on left-behind households is complex, particularly in the environmental domain. The departure of family members may disrupt daily routines, including waste management, recycling, and land use practices, which can lead to a deterioration in environmental behaviours (Fan, 2007). In contrast, return migration often introduces new dynamics. Migrants who return from urban areas may bring with them pro-environmental norms and practices learned through exposure to urban infrastructure and regulatory systems, thus fostering behavioural change in their home communities (Williams and Paudel, 2020). This dual mechanism underscores the need to examine both the disruptive and transformative effects of migration on environmental practices in origin areas.

This thesis investigates how these challenges intersect in the Chinese context by analysing the dynamic relationships between air pollution exposure, migration behaviour, and retirement expectations. Through a series of empirical studies, it examines how pollution affects both mobility decisions and intergenerational behavioural change, offering new insights into the socio-environmental impacts of migration under conditions of ecological stress.

1.1.2 Environmental and Socioeconomic Transitions in China

As the world's largest developing economy, it faces acute environmental challenges following decades of rapid industrialisation.In 2015, 96% of Chinese cities reported $PM_{2.5}$ levels exceeding WHO guidelines, with some cities experiencing concentrations more than four times higher than recommended thresholds (Deschenes et al., 2020), reflecting the severity of urban pollution. The consequences of air pollution extend beyond public health, influencing labour supply decisions, shaping migration flows, and accelerating retirement among vulnerable populations. The country's unique hukou (household registration) system compounds these effects, generating structural disparities in mobility, employment, and access to urban services between rural and urban populations. These institutional and environmental complexities make China an ideal setting for exploring the intersection of air pollution, migration, and individual behavioural responses.

Air pollution remains one of China's most urgent environmental issues. Prolonged exposure to high levels of air pollution has resulted in substantial public health burdens, including increased incidences of respiratory and cardiovascular diseases. These health challenges, in turn, affect labour productivity and force individuals, particularly those nearing retirement, to reconsider their labour market participation (Graff Zivin and

Neidell, 2013). Recent research also highlights air pollution's role in reducing cognitive performance, further exacerbating economic losses (Zhang et al., 2018).

Migration plays a pivotal role in this context. For urban migrants, air pollution in destination cities has become a critical determinant of migration flows, as individuals weigh the economic opportunities of urban areas against the health risks posed by poor air quality. Studies have shown that relative $PM_{2.5}$ concentration between destination and origin cities significantly affects migration decisions, with higher pollution levels at destinations deterring inflows (Cai et al., 2016; Chen et al., 2022) Meanwhile, rural communities face challenges in maintaining sustainable practices as migration removes key members of households, disrupting traditional environmental behaviours. However, return migration can mitigate some of these disruptions by introducing urban-acquired knowledge and norms back to rural areas, fostering improvements in urbanisation (Zhu et al., 2021).

At the same time, China is undergoing one of the largest internal migration movements in history. With over 292 million rural-to-urban migrants as of 2021, migration has become a defining feature of China's socioeconomic landscape (National Bureau of Statistics, 2022). Migrants not only contribute to urban economies but also alter social structures in their home communities. However, the benefits of migration are often unevenly distributed. Migrants face significant hurdles in accessing social services due to the dualistic hukou system, which restricts entitlements to education, healthcare, and housing based on one's household registration status. This system creates disparities in how environmental and socioeconomic factors affect urban and rural populations (Fan, 2007).

Also, China's ageing population presents significant socioeconomic challenges, with individuals aged 60 and above accounting for 18.9% of the population in 2021, a figure expected to rise to nearly 30% by 2050 (National Bureau of Statistics, 2022). This demographic shift, coupled with chronic exposure to air pollution, is placing increasing strain on the labour force and public health systems. Studies show that pollutants like $PM_{2.5}$ and NO exacerbate respiratory and cardiovascular conditions, disproportionately affecting older workers and accelerating decisions to retire earlier than planned (Li and Li, 2022; Zhang et al., 2018). Such health shocks not only reduce labour market participation but also impose rising healthcare costs, intensifying the economic dependency ratio and challenging the sustainability of pension systems. For example, a study on China's urban workers found that worsening pollution led to earlier retirement, particularly in rural areas where access to healthcare is limited (Chen et al., 2023). Addressing these interlinked pressures requires integrated policies that improve air quality while supporting ageing-related welfare systems to mitigate deepening inequalities and ensure economic stability.

Together, these transitions reflect China's struggle to balance economic growth, environmental protection, and social equity. Air pollution, migration, and demographic change are not isolated phenomena but interconnected challenges that require integrated policy responses. The following sections will explore these dynamics in greater depth, focusing on how air pollution affects migration flows, reshapes environmental behaviours in rural areas, and alters retirement decisions in an ageing society.

1.1.3 Unique Institutional Context: Why China Matters

China offers a distinctive and analytically rich context to examine the intersection between environmental stressors, migration dynamics, and ageing-related behaviours. Its institutional architecture—particularly the household registration (*hukou*) system, a fragmented pension regime, and the world's largest internal migration—distinguishes it sharply from many developed and developing countries.

The *hukou* system classifies individuals as rural or urban residents based on their place of birth registration rather than actual residence, fundamentally shaping access to education, healthcare, housing, and pension entitlements. As of 2022, approximately 65% of the population retained rural hukou status (National Bureau of Statistics of China, 2023). Migrants from rural areas typically remain ineligible for full urban welfare benefits, even after years of urban employment. Consequently, migration decisions are driven not only by economic incentives, but also by institutional exclusion and uneven access to social protection.

China's pension system reinforces this duality. The Urban Employee Basic Pension Scheme (UEBPS) covered around 480 million contributors in 2022, offering a replacement rate exceeding 40% for urban formal-sector workers (MOHRSS, 2023). In contrast, rural and informal workers are enrolled in a flat-rate pension programme with an average monthly benefit of approximately ¥180 (about US$25) (Li and Gustafsson, 2021). Similar structural divides exist in other developing countries. In India, fewer than 15% of workers participate in any formal pension scheme (International Labour Organization, 2021), while in Indonesia and Bangladesh, coverage remains below 10% and is largely restricted to public employees (World Bank, 2020).

China's internal migration scale is likewise unparalleled. As of 2022, an estimated 295 million people—more than 20% of the population—were living outside their hukou registration area (National Bureau of Statistics of China, 2023). This dwarfs India's estimated 140 million internal migrants in 2020, whose movements are often limited by linguistic, caste, and legal constraints (IIPS, 2020). Indonesia, with around 30 million annual internal migrants, lacks an equivalent institutional barrier such as the hukou (UNESCAP, 2019), though informal barriers persist.

Unique China's internal migration is based on rural-to-urban and urban-to-urban migration two distinct and equally significant implications for China's development trajectory and behavioural transformations. Rural-to-urban flows comprise the bulk of China's internal migration and represent the foundational labour force driving industrialisation and urban expansion. These migrants often engage in physically demanding, environmentally exposed occupations while remaining excluded from urban welfare provisions. By contrast, urban-to-urban migration, though smaller in volume, increasingly reflects the movement of high-skilled workers and professionals across metropolitan regions. This group is more likely to be formally employed, integrated into social insurance systems, and responsive to institutional and environmental incentives. As such, they play a disproportionately large role in shaping urban sustainability practices, knowledge diffusion, and policy uptake. In one word, the former speaks to structural vulnerability and welfare exclusion, while the latter relates to the mobilisation of human capital and the behavioural adaptation capacity of cities.

Environmental behaviours in China also reflect strong spatial and institutional disparities. Urban areas have pioneered the rollout of green infrastructure and policy mandates, including waste sorting, air quality monitoring, and electric vehicle subsidies. A 2021 national survey found that 75% of urban residents reported regular waste separation, compared to just 34% in rural areas (China Environmental Awareness Index, 2021). Comparable gaps are evident in other developing countries: in India, only 28% of urban residents engage in waste sorting, and fewer than 10% do so in rural areas (Pew Research Center, 2018); in Indonesia, less than 20% of households consistently separate waste (World Bank, 2019).

One distinguishing feature of China's environmental governance is its centralised, top-down approach. National initiatives—such as the plastic ban, emissions trading pilots, and nationwide waste-sorting mandates—have been implemented with relatively swift enforcement. In contrast, India and Indonesia exhibit more decentralised governance, resulting in fragmented implementation and regionally heterogeneous outcomes. While all face rising environmental stress, China's institutional capacity to steer collective behavioural change is arguably stronger.

Taken together, these demographic, institutional, and environmental conditions make China a compelling empirical setting to study how migration intersects with welfare systems and environmental behaviour. At the same time, comparison with other large developing countries highlights both the uniqueness and transferability of China's experience. This thesis thus contributes not only to the understanding of individual responses under environmental and institutional constraints in China, but also offers broader lessons for countries undergoing similar transitions in population mobility, welfare provision, and sustainability governance.

1.1.4 From the Chinese Context to Global Relevance

While the institutional and demographic context of China is unique, the underlying behavioural mechanisms identified in this thesis have broad relevance. The central pathways explored—how environmental stressors shape migration decisions, influence household behaviour, and alter retirement expectations—are applicable across settings, even if their magnitude or mediating factors differ.

First, the health consequences of air pollution are universal. Numerous studies across diverse countries have linked pollution exposure to declines in cognitive performance, physical health, and life expectancy. The behavioural responses—whether to migrate, alter consumption patterns, or exit the labour market early — are therefore likely to be observed wherever pollution levels exceed safe thresholds.

Second, China's institutional configuration allows for the examination of these mechanisms under extreme variation. The country's large rural–urban divide, segmented welfare system, and substantial pollution disparities create a "stress test" environment in which responses to environmental and social pressures are more visible. Insights drawn from such a setting can inform policy design in countries experiencing growing urbanisation, rising environmental concerns, or institutional inequality—even if at a smaller scale.

Third, the empirical strategies and conceptual frameworks adopted in this study—such as differential migration models, retirement response to health shocks, and the role of social remittances—can be replicated or adapted in other contexts. The mechanisms are sufficiently general to allow for testing in countries with different welfare states, environmental conditions, or demographic pressures. Doing so would enrich our understanding of how global challenges like pollution and ageing intersect in diverse policy environments.

In this way, the Chinese case serves not only as a country-specific investigation but also as a lens through which broader socio-environmental dynamics can be understood and studied globally.

1.2 Literature Review and Research Gap

The interplay between migration, air pollution, and socioeconomic behaviours has gained increasing attention in recent years, yet significant gaps remain in understanding their nuanced relationships. This thesis addresses these gaps by exploring three interconnected themes: the impact of migration on green behaviours, the role of air pollution in internal migration flows, and the influence of pollution on retirement planning. In doing so, it contributes to the broader literature on environmental and socioeconomic dynamics.

1.2.1 Migration and Green Behaviours

The relationship between migration and environmental behaviours is a relatively under-explored field. Traditional migration research has largely focused on the economic and social effects of migration, particularly financial remittances, which have been shown to enhance education, healthcare, and living standards in origin communities (Abramitzky et al., 2012; Rapoport and Docquier, 2006). More recently, attention has shifted to social remittances, a term coined by Stigler (1961) to describe the transfer of norms, values, and practices from host to origin regions. Social remittances influence various domains, such as political preferences, entrepreneurship, and gender roles are transmitted through communication, financial support, and return migration (Giulietti et al., 2013; Levitt and Lamba-Nieves, 2011; Tuccio and Wahba, 2018).

However, environmental behaviours have received relatively little attention in the social remittance literature. While studies in green economics often focus on narrow topics, such as energy consumption (Alem et al., 2016) or pollution mitigation policies (Wu et al., 2021a), the mechanisms through which migration shapes environmental awareness and practices remain poorly understood. The socio-economic divide between urban and rural areas in China provides a unique context for exploring this dynamic. Urban environments, characterised by superior infrastructure, such as public waste management systems and sanitation services, create opportunities for migrants to adopt green behaviours that may be transmitted back to rural households.

Previous research has highlighted how migration can foster entrepreneurship (Giulietti et al., 2013), improve health knowledge (Hildebrandt et al., 2005), and influence political norms (Tuccio et al., 2019). Yet, its role in promoting sustainability remains largely unexplored. In Chapter 2, this thesis addresses this gap by examining how rural-to-urban migration in China impacts green behaviours among left-behind households. It explores the dual effect of migration—both disrupting traditional routines and introducing urban-acquired practices—offering new insights into how social remittances can drive environmental behavioural change in developing contexts.

1.2.2 Air Pollution and Migration Flows

Air pollution is a growing concern in urbanising economies, influencing migration decisions through its impact on physical and mental well-being. Existing research highlights how environmental quality is a critical factor in migration intentions, particularly for urban residents exposed to high levels of pollution. The Rosen-Roback hedonic framework (Rosen, 1979; Roback, 1982) offers a theoretical foundation for understanding how amenities, including air quality, influence migration by balancing economic benefits and environmental costs. However, this framework often assumes perfect information and rational decision-making, which may not hold in developing economies where pollution data is incomplete or inaccessible (Gao et al., 2023).

Empirical studies have shown that air pollution negatively affects migrants' well-being and drives decisions to relocate (Cai and Wang, 2007; Chen et al., 2022). For instance, Lai et al. (2021) find that high-skilled workers are particularly sensitive to pollution and prefer cleaner environments when choosing migration destinations. Similarly, Sun et al. (2019) document how urban pollution encourages low-income migrants to leave their current cities. However, most research focuses on net migration flows or single-city factors, neglecting the relative pollution levels between origin and destination cities. Furthermore, structural barriers such as China's hukou system restrict mobility, particularly for rural migrants, compounding the effects of environmental disparities.

In Chapter 3, this thesis addresses these gaps by employing a differential-based gravity model to analyse bidirectional migration flows. By using the ratio of destination-origin $PM_{2.5}$ concentrations as a proxy for pollution gaps, it disentangles the push and pull dynamics of pollution in migration decisions. The focus on urban-to-urban migration extends the literature, which has historically prioritised rural-to-urban flows, providing new insights into how pollution influences mobility among high-skilled workers and urban populations.

1.2.3 Air Pollution and Retirement Planning

The economic and health impacts of air pollution are well-documented, with studies linking exposure to pollutants such as $PM_{2.5}$ and NO2 to respiratory diseases, cardiovascular problems, and cognitive decline (Burnett et al., 2014; Chen et al., 2020). Vulnerable populations, such as older adults and low-income groups, are disproportionately affected due to declining health and limited access to healthcare services (Boyce, 2013; Shi et al., 2016). These health shocks often lead to productivity losses, absenteeism, and increased healthcare costs, as demonstrated by Chang et al. (2019) and Neidell (2023).

Despite this robust body of research, the implications of air pollution for retirement planning remain underexplored. Health deterioration caused by pollution may compel individuals to exit the labour market early, as highlighted by Maestas and Zissimopoulos (2010). However, financial insecurity often delays retirement, creating a paradox where individuals are simultaneously driven to leave and forced to stay in the workforce (Shen et al., 2021). This tension is particularly acute in China, where a rapidly ageing population and strained pension systems exacerbate the challenges of balancing health and financial considerations in retirement decisions.

In Chapter 4, this thesis examines the dynamic effects of pollution shocks on retirement age expectations, providing empirical evidence on how pollution influences labour market exits. It explores how sudden increases in pollution accelerate adjustments to retirement planning, highlighting the dual pressures of health deterioration and financial insecurity. By linking environmental health to labour economics, this research contributes to a deeper understanding of how pollution shapes life-course decisions in ageing societies.

1.2.4 Significance and Contribution

This thesis bridges multiple disciplines—including migration studies, environmental economics, and labour economics—to provide a holistic perspective on how air pollution interacts with migration and socioeconomic behaviours. Through its three key themes, this research addresses critical gaps in the literature and offers significant theoretical and practical contributions.

First, the study expands the concept of social remittances by incorporating environmental behaviours, an area often overlooked in migration literature. While previous research has extensively examined financial remittances and the socio-political effects of migration, little attention has been paid to how migration influences environmental practices. This thesis demonstrates how rural-to-urban migration in China facilitates the transfer of green norms — such as recycling and waste management, from urban to rural households. Urban environments, with their superior infrastructure and stricter regulations, expose migrants to sustainability practices that are later communicated or adopted by left-behind families. By quantitatively analysing this process, the study not only broadens the scope of social remittances but also highlights migration's potential role in bridging the urban-rural sustainability divide. This finding offers a valuable contribution to the fields of migration studies and green economics, particularly in the context of developing economies where rural areas often lack environmental awareness and infrastructure.

Second, the thesis provides new insights into the role of air pollution in shaping internal migration flows, focusing specifically on urban-to-urban migration. Unlike most existing studies, which emphasise rural-to-urban migration or net migration flows, this research captures the bidirectional nature of urban migration by employing a differential-based gravity model. Using relative $PM_{2.5}$ levels between origin and destination cities as a proxy for pollution gaps, it reveals how environmental quality acts as both a push and pull factor in mobility decisions. High-skilled workers, in particular, are shown to prioritise health and environmental conditions when choosing where to live and work, underscoring the importance of pollution reduction in retaining talent. By integrating the often-neglected role of environmental disparities into migration decision-making, this study extends traditional migration theories and provides actionable insights for urban planners and policymakers seeking to improve city competitiveness and quality of life.

Finally, the thesis explores the complex interplay between pollution, health, and retirement planning, addressing a significant gap in the literature on long-term socioeconomic impacts of air pollution. While the health consequences of pollution are well-documented, their implications for retirement decisions have received limited attention. This research shows how pollution-induced health shocks can accelerate early retirement, particularly among vulnerable populations such as older adults and low-income groups. At the same time, financial insecurity often forces individuals to delay retirement, creating a paradox with profound implications for labour market dynamics and social welfare systems. By examining these dual pressures in the context of an ageing society like China, the study contributes to labour economics and offers valuable perspectives on the sustainability of pension systems in regions facing severe environmental and demographic challenges.

The significance of this thesis lies not only in its theoretical advancements but also in its practical implications. Academically, it bridges gaps between migration studies, environmental and labour economics, and the emerging discourse on sustainability, providing a comprehensive framework for understanding how environmental stressors shape human behaviours and societal outcomes. Its contributions extend existing theories by incorporating environmental dimensions into the study of migration and retirement, thereby enriching the interdisciplinary understanding of these complex phenomena.

Practically, the findings offer evidence-based strategies for addressing some of the most pressing challenges in environmental governance, urban planning, and social welfare. By demonstrating how migration can act as a channel for environmental awareness, the research highlights the potential of leveraging urbanisation to promote sustainability in rural areas. The analysis of air pollution's impact on migration flows emphasises the need for policies aimed at reducing pollution to attract and retain talent in urban centres, enhancing their economic and social competitiveness. Additionally, the study's insights into pollution-driven retirement adjustments underscore the importance of aligning labour and pension policies with environmental realities, ensuring the resilience of social welfare systems in the face of ageing populations and environmental degradation.

Overall, this thesis advances the discourse on sustainable development by linking environmental stressors to migration and socioeconomic behaviours in novel ways. Its integrated approach not only deepens theoretical understanding but also provides a practical foundation for policies aimed at fostering sustainability, reducing inequalities, and promoting resilience in rapidly transforming societies.

1.3 Research Objectives and Questions

Building on the pressing challenges highlighted in the background and motivation, this thesis aims to deepen our understanding of how air pollution interacts with migration and socioeconomic behaviours, particularly within the context of China. The research focuses on three interconnected areas: the environmental impacts of migration, air pollution's influence on migration flows, and pollution's effects on retirement decisions. These objectives are motivated by gaps in the existing literature and the need to address pressing policy challenges related to sustainability, urbanisation, and aging populations.

1.3.1 Research Objectives

This thesis seeks to explore the intricate interactions between air pollution, migration, and socioeconomic behaviours, focusing on their dynamic relationships within the Chinese context. These themes are deeply interconnected, as environmental stressors influence individual decisions, such as migration and retirement, while migration itself drives changes in social and environmental behaviours. By addressing these interdependencies, this research provides a holistic perspective on how air pollution shapes human behaviour and social outcomes, with implications for sustainability, urbanisation, and social welfare. The three objectives of the thesis are distinct but closely linked, reflecting these overlapping dynamics and structuring the empirical analyses presented in subsequent chapters.

(1) To explore the impact of rural-to-urban migration on green behaviours in left-behind and return-migrant households

The first objective of Chapter 2 is to analyse how rural-to-urban migration shapes the environmental behaviours of rural households, with explicit attention to two distinct groups: (i) households with current migrants (left-behind households) and (ii) households with return migrants. Migration is not merely an economic decision; it is also a social process that can disrupt or transform household practices through changes in labour availability, household composition, and exposure to external norms.

For left-behind households, the analysis assesses whether the absence of working-age members due to out-migration reduces engagement in pro-environmental behaviours such as waste sorting, recycling, and land management. The expected mechanism is that labour loss and altered intra-household dynamics may weaken the capacity or incentives for green practices.

For households with return migrants, the focus shifts to whether returnees bring back urban-acquired environmental norms—shaped by greater exposure to green infrastructure, environmental regulations, and pro-sustainability social norms—and diffuse these behaviours within their origin communities. Return migration may therefore serve as a channel for the transmission of "environmental social remittances," enhancing local sustainability practices.

By applying instrumental variable strategies to address potential selection bias in migration, the chapter provides causal evidence on the contrasting effects of outward and return migration. This dual perspective underscores the nuanced role of migration in rural sustainability: while out-migration can erode environmental engagement, return migration has the potential to strengthen it.

(2) To examine how air pollution shapes internal urban-to-urban migration flows in China

The second objective examines the role of air pollution in shaping internal urban-to-urban migration flows. Air pollution acts as both a push factor, driving individuals away from highly polluted cities, and a pull factor, attracting migrants to cleaner environments. This dual dynamic creates a feedback loop, where migration patterns influenced by environmental disparities can either exacerbate or mitigate urban challenges. For instance, cities with high pollution levels may experience a loss of skilled workers, undermining their economic competitiveness. In contrast, cleaner cities may attract large influxes of migrants, leading to population pressures that strain infrastructure and threaten environmental sustainability.

This research focuses on the relative pollution levels between origin and destination cities, emphasising the importance of environmental quality comparisons in shaping urban mobility. Chapter 3 explores how this "relative pollution strength" influences migrants' destination choices, moving beyond a narrow one-sided focusing on conditions at either the origin or the destination. By analysing these dynamics, the study reveals how pollution influences destination preferences, underscores the role of green information in migration decisions, and connects this objective to the first.

This approach provides valuable insights into how migrants respond to differing norms and informational shocks compared to their places of origin. Furthermore, it clarifies the mechanisms by which green social remittances are generated, linking urban migration behaviour to broader environmental transformations in origin regions.

(3) To investigate the long-term effects of air pollution on retirement planning.

The third objective examines how long-term pollution exposure and short-term pollution shocks influence retirement planning, particularly among older individuals in an ageing society. Chronic exposure to pollutants such as $PM_{2.5}$ accelerates health deterioration, often leading to earlier-than-expected exits from the labour market. However, financial constraints, exacerbated by inadequate pension systems, can delay retirement decisions, creating a paradoxical interaction between health and economic pressures.

Chapter 4 investigates individuals' responses to pollution in terms of retirement expectations. By analysing the reactions of those nearing retirement age and retirees to varying air pollution levels, this study constructs a comprehensive picture of how ageing populations make labour supply decisions under environmental stress. Given the heterogeneity between rural and urban regions, both groups are examined, with a focus on the inequalities faced by rural labourers in dealing with pollution. Furthermore, the role of macro-level factors, such as social capital and infrastructure development, as well as micro-level channels like family ties, is explored in shaping both reactions to and defences against pollution-induced health deterioration.

This research generalises the effects of pollution on migrants to encompass all residents, revealing how air pollution influences labour supply and urbanisation while highlighting the inequality and vulnerability of rural areas. These concerns extend beyond domestic rural-urban or urban-urban migration patterns, emphasising the broader implications for sustainability in developing countries.

1.3.2 Summary of Research Questions

To achieve the objectives outlined above, this thesis seeks to answer the following research questions. These questions are designed to capture the complex and dynamic interplay between air pollution, migration, and socioeconomic behaviours, while also addressing the gaps identified in the existing literature:

Q1. How does rural-to-urban migration influence the environmental behaviours of left-behind and return-migrant households, and through what mechanisms are these behaviours transmitted?

This question investigates the environmental effects of rural-to-urban migration across two household types: left-behind households with current migrants, and households with return migrants. The aim is to identify whether out-migration diminishes engagement in green behaviours due to labour loss and household restructuring, and whether return migration promotes such behaviours by transferring environmental norms acquired in urban settings.

The analysis explores several mechanisms of behavioural transmission, including the accumulation of social capital, exposure to advanced urban green infrastructure, and the use of communication technologies to maintain information flows. Comparing the impacts of current versus return migration allows for a clear understanding of migration's dual role in rural environmental sustainability—acting as a disruptor when human and social capital are depleted, and as a catalyst when they are replenished through return.

Q2. How do relative pollution levels between origin and destination cities shape internal urban-to-urban migration flows in China?

This question examines the role of "relative pollution strength" in migration decisions, focusing on how migrants weigh environmental quality when choosing destinations. It investigates whether pollution acts as a push factor in polluted cities or as a pull factor in cleaner cities, and how these dynamics influence skilled and unskilled labour differently. The study also considers how information availability about environmental conditions mediates these choices, linking the migration decision-making process to broader transformations in environmental awareness and urban development. By exploring these pathways, the research clarifies how urban-to-urban migration reinforces or disrupts environmental and economic disparities between regions.

Q3. How do long-term pollution exposure and short-term pollution shocks influence retirement planning, particularly in ageing societies?

This question addresses the paradox of pollution's dual effect on retirement: the health risks that accelerate early labour market exits and the financial pressures that delay retirement. It examines how individuals nearing retirement respond to varying levels of pollution, with particular attention to differences between rural and urban residents. The research also investigates the macro-level factors (such as pension systems and urban infrastructure) and micro-level channels (such as family networks and access to healthcare) that mediate these decisions. The analysis sheds light on the broader implications of pollution-induced health shocks for labour supply, urbanisation, and social equity.

These research questions not only align with the specific objectives of this thesis but also highlight the interconnections between chapters. The first question addresses the behavioural shifts initiated by migration, particularly the diffusion of environmental practices from urban to rural areas. The second builds on this foundation, exploring how environmental disparities influence mobility decisions, which in turn shape the socioeconomic and environmental landscapes of both origin and destination cities. The third question extends the analysis to life-course decisions, examining how air pollution impacts retirement planning and, by extension, labour market dynamics and social security systems. Overall, these questions provide a comprehensive framework for understanding the interactions between environmental stressors, migration, and socioeconomic behaviours in a rapidly transforming society like China. Beyond the Chinese context, the findings have global relevance, offering insights into how environmental and social policies can address shared challenges of sustainability, urbanisation, and demographic change in other developing and industrialised economies.

1.3.3 Integrated View and Feedback Loops

Regarding all empirical research questions in this thesis. Figure 1.1 illustrates the core feedback loops that connect pollution exposure to key socioeconomic behaviours explored in this thesis. At the centre, pollution acts as a foundational environmental stressor, exerting influence through three primary pathways: norms, health, and income.

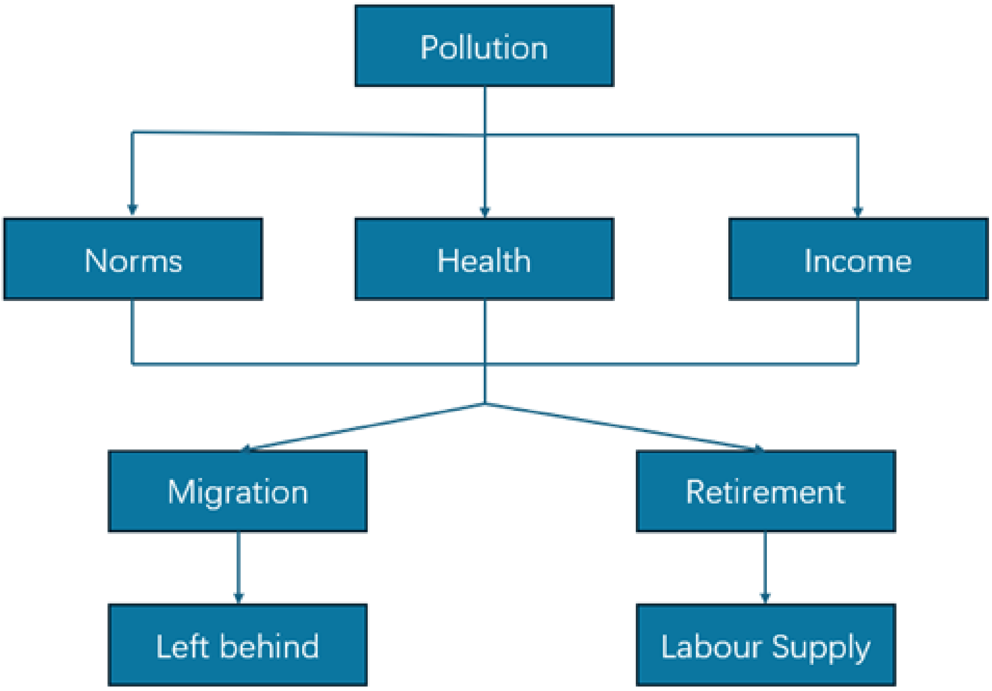


FIGURE 1.1: **Interconnections Between Pollution, Migration, and Socioeconomic Dynamics**

Pollution affects environmental norms by shaping public awareness and behavioural expectations, particularly through exposure to cleaner infrastructure or stricter environmental regulation in urban areas. These altered norms are often transmitted back to rural households through migration, especially via returnees, facilitating the adoption of greener practices in left-behind communities.

Simultaneously, pollution degrades health and reduces earning capacity. These effects influence migration decisions, as households seek to escape unhealthy environments or access better healthcare and jobs elsewhere. For those left behind, the departure of working-age members can disrupt established environmental routines and resource management. Yet, the same migration also generates opportunities for norm diffusion, creating a dynamic tension between disruption and transformation.

The health and income consequences of pollution also influence retirement planning, especially among older individuals. Poor health can accelerate early exit from the labour force, while insufficient income or weak pension access may delay retirement. This creates a paradox: pollution simultaneously pushes people out of work for health reasons while pulling them back in due to financial necessity. The result is a strain on both individual well-being and broader labour supply dynamics.

These interconnected processes form feedback loops. For example, migration reshapes household structure and labour allocation, which in turn affects green behaviour and retirement decisions. Likewise, changes in labour supply or retirement timing can influence family migration strategies. These multi-directional relationships reinforce the central thesis of this study: that pollution is not merely an environmental issue but a systemic force that interacts with mobility, behaviour, and demographic change.

1.4 Data and Methodology

This section outlines the methodological framework and data sources employed in this thesis, which comprises three distinct yet interconnected empirical studies. Given the interconnected nature of these topics, the research integrates micro-level panel data with macro-level environmental and socioeconomic indicators to capture both individual and aggregate-level dynamics. Each study addresses specific research questions related to the impact of pollution on migration, green behaviours, and retirement planning. The methodologies adopted reflect the nature of the research questions, combining econometric modelling, causal inference techniques, and robustness checks to ensure empirical rigour.

1.4.1 Research Design

The research is structured into three empirical studies, each addressing a distinct but interconnected dimension of the interactions between pollution, migration, and socioeconomic behaviours. The design reflects the diverse pathways through which environmental stressors affect human behaviour and incorporates methods capable of capturing these complexities.

(1) Migration and Green Behaviours (Chapter 2):

This study focuses on how rural-to-urban migration reshapes environmental behaviours among left-behind households through the diffusion of social remittances—non-material transfers of norms, values, and practices. It applies linear probability models (LPM) and fixed-effects and migration network as instrumental variables (IV) to identify whether migrants' exposure to better environmental infrastructure in urban areas influences the green behaviours of origin households. By distinguishing between current migration and return migration, the analysis uncovers differences in the transmission of sustainable practices.

(2) Pollution and Migration Flows (Chapter 3):

The second study investigates how differences in pollution levels between origin and destination cities shape migration flows. It adopts a differential gravity model with instrumental variable - thermal inversion, which captures the relative pollution gaps driving migration decisions. This model accounts for push-and-pull factors, allowing the analysis to disentangle the role of environmental quality from economic motivations. In particular, it highlights how highly polluted cities lose skilled workers, exacerbating urban inequality.

(3) Pollution and Retirement Planning (Chapter 4):

The third study evaluates the impact of long-term pollution exposure and sudden pollution shocks on retirement planning. It employs panel fixed-effects models and instrumental variable (IV) approaches to address endogeneity, using thermal inversion days as an exogenous instrument for pollution. The analysis reveals how pollution-induced health deterioration influences retirement expectations, highlighting the tension between health-driven exits and economic constraints that delay retirement. Together, these three studies provide a holistic view of how environmental stressors shape migration patterns, green behaviours, and labour market decisions. They also reinforce the broader argument that pollution-induced socioeconomic effects are interconnected and cumulative, demanding integrated policy responses.

1.4.2 Data Sources

Empirical Study 1: Migration and Green Behaviours

- **Micro Survey Dataset:** China Labour-force Dynamics Survey (CLDS)

  The CLDS is a biennial panel dataset designed to capture labour market activities, household dynamics, and social behaviours. It provides extensive information on economic conditions, migration patterns, and environmental practices.

  - **Scope and Coverage**
    - **Waves:** 2014 and 2016.
    - **Sample Size:** Over 15,000 households and 23,000 individuals covering both urban and rural regions.
  - **Key Variables**
    - **Environmental Practices:** Waste sorting and recycling willingness.
    - **Rural-to-urban Migration Status:** Includes current and return migration.
    - **Economic and Social Indicators:** Household and village characteristics.
  - **Relevance**
    - Investigates the impact of migration on green behaviours among left-behind households.
    - Focuses on the diffusion of social remittances, highlighting how urban-acquired environmental norms are transmitted to rural areas.

- **City-Level Data**

  Environmental infrastructure indicators, such as public sanitation services and waste management facilities, were sourced from the National Bureau of Statistics and China Statistical Yearbooks.

Empirical Study 2: Pollution and Migration Flows

- **Primary Dataset:** China Migrants Dynamic Survey (CMDS)

  The CMDS is a nationally representative cross-sectional survey focused on migration patterns, labour mobility, and settlement decisions.

  - **Scope and Coverage**
    - **Year:** 2017 wave.
    - **Sample Size:** Over 170,000 migrants, capturing urban-to-urban and rural-to-urban migration.

  - **Key Variables**

    Migration Patterns: Aggregated migration flows based on origin-destination relationships.

  - **Relevance**

    * Analyses how pollution gaps between origin and destination cities influence migration decisions.
    * Examines push-and-pull dynamics driven by environmental disparities.

- **Pollution Data**

  $PM_{2.5}$ concentrations sourced from NASA's satellite data, providing high-resolution pollution measures across Chinese cities.

  Thermal inversion days as an instrumental variable (IV) to address endogeneity and ensure causal identification.

- **City-Level Controls**

  - **Climate Variables:** Temperature and rainfall from the China Meteorological Administration.
  - **Economic Indicators:** GDP per capita, industrial composition, unemployment rates, population density, average wage, and housing prices, sourced from the National Bureau of Statistics.

Empirical Study 3: Pollution and Retirement Planning

- **Primary Dataset:** China Health and Retirement Longitudinal Study (CHARLS)

  The CHARLS is a nationally representative panel dataset designed to study the economic and health conditions of middle-aged and older adults (aged 45 and above).

  - **Scope and Coverage**

    * **Waves:** 2011, 2013, 2015, 2018, and 2020.
    * **Sample Size:** Over 17,000 respondents across 450 communities.

  - **Key Variables**

    * **Retirement Planning:** Expected retirement age.
    * **Health Indicators:** Chronic illnesses, disabilities, and self-assessed health status.
    * **Demographics:** Education, income, marital status, and family composition.

  - **Relevance**

    * Evaluates the effect of pollution exposure on retirement expectations and labour market behaviour.
    * Tracks long-term trends and short-term shocks using panel data.
  - **Pollution Data**
    * $PM_{2.5}$ concentrations sourced from Washington University in St. Louis, providing historical pollution trends.
    * Thermal inversion days as an instrumental variable (IV) to address endogeneity and ensure causal identification.
  - **Macro Data**
    * GDP, unemployment rates, and health services, sourced from the National Bureau of Statistics.

Integration of Macro-Level Environmental and Economic Data

To complement the household surveys, this thesis integrates macro-level indicators from multiple sources to control for structural differences across cities and regions.

Pollution Indicators

- NASA Satellite Data (Chapter 3): City-level $PM_{2.5}$ measurements to analyse pollution-driven migration flows.
- Washington University Data (Chapter 4): Historical $PM_{2.5}$ data to track long-term exposure for retirement planning.

Instrumental Variables

- Thermal Inversion Days (Chapters 3 and 4) to address adverse causality between air pollution and economic development.
- Migration Network (Chapter 2) to address the selection bias of current migration and return migration.

Climate and Geography

- Rainfall, temperature, and altitude data from the China Meteorological Administration.

Economic Indicators

- GDP per capita, unemployment rates, and industrial composition from the National Bureau of Statistics.
- Infrastructure data covering hospitals, sanitation systems, and green facilities.

Data Integration and Preparation

The datasets are merged using geographical identifiers (e.g., province and city codes) to establish a multi-level panel structure that links individual behaviours with regional trends.

Panel Structure

- Tracks longitudinal changes for retirement planning and green behaviours while examining cross-sectional migration patterns.

Scaling and Standardisation

- Variables such as $PM_{2.5}$ are log-transformed to account for non-linear effects, while migration and retirement outcomes are standardised for comparability.

Causal Inference Tools

- Fixed effects models, instrumental variables, and robustness checks (e.g., lagged variables) are applied to ensure reliability.

Overall Contribution

By integrating micro-level panel data with macro-level environmental and economic indicators, this thesis constructs a multi-dimensional empirical framework to explore the socioeconomic impacts of pollution. Table 1.1 summarises the data source and structure. Each chapter leverages a carefully selected dataset—CLDS for green behaviours, CMDS for migration decisions, and CHARLS for retirement planning—ensuring that the analysis is both comprehensive and policy-relevant.

Table 1.1: **Summary of Data Sources**

Study	Dataset	Key Features
Migration and Green Behaviours **Focus**: Impact of migration on green behaviours, diffusion of urban norms to rural areas	• National-representative China Labour-force Dynamics Survey (CLDS) • China City Statistical Yearbook	• Repeated cross-section based on 2014 and 2016 • 23,000 individuals, urban and rural • Variables: Waste sorting, recycling willingness, rural-to-urban migration, household and village characteristics
Pollution and Migration Flows **Focus**: Pollution gaps between cities driving migration; push-and-pull dynamics	• National-representative China Migrants Dynamic Survey (CMDS) • NASA's MODIS, MISR, and SeaWiFS • China City Statistical Yearbook	• Cross-Section Data of 2017 • 170,000 migrants • Variables: Aggregated migration flows, $PM_{2.5}$ pollution, climate variables, economic indicators
Pollution and Retirement Planning **Focus**: Impact of long-term pollution exposure on retirement planning	• National-representative China Health and Retirement Longitudinal Study (CHARLS) • Satellite-derived $PM_{2.5}$ data from WashU • China City Statistical Yearbook • Climate data from the China Meteorological Administration	• Panel Data based on 2011, 2013, 2015, 2018 and 2020 waves • 17,000 respondents, aged 45+ • Variables: Expected retirement age, $PM_{2.5}$ pollution, individual, household, and city-level indicators

Note: This table summarises the data source and structure, providing sample size, key variables and scopes.

1.4.3 Empirical Strategy

(1) Analytical Framework

The thesis applies a multi-level analytical framework that links individual behaviours and macroeconomic trends to environmental stressors. It utilises panel data econometrics to capture both cross-sectional variations and within-unit changes over time, offering insights into behavioural responses and policy implications.

Each chapter uses tailored models:

1. **Chapter 2:** Investigates how rural-to-urban migration influences the adoption of green behaviours among left-behind households. It applies Linear Probability Models (LPM) and Fixed Effects Regressions to estimate the effects of migration status on environmental practices.
2. **Chapter 3:** Examines how pollution gaps drive urban-to-urban migration decisions. It employs a differential-based gravity model to analyse migration flows based on relative environmental quality between origin and destination cities.
3. **Chapter 4:** Explores how pollution shocks affect retirement expectations. Using Panel Fixed Effects Models and Instrumental Variable (IV) approaches, it isolates the causal effects of long-term pollution exposure on retirement planning.

This framework provides flexibility to address endogeneity, unobserved heterogeneity, and dynamic adjustments, ensuring robust and credible results.

(2) Empirical Specification

Migration and Green Behaviours (Chapter 2)

To examine whether migration transmits environmental practices to left-behind households, this chapter uses the following fixed-effects regression model:

$$Y_{it} = \alpha + \beta Migration_{it} + \delta X_{it} + \mu_i + \epsilon_{it} \tag{1.1}$$

where:

- Y_{it} represents the green behaviours (e.g., waste sorting and recycling) of household i at time t.
- $Migration_{it}$ is a dummy variable for whether the household has a migrant member.

- X_{it} includes household controls (e.g., income, education, family size).
- μ_i accounts for individual fixed effects, and ϵ_{it} is the error term.

Key Hypothesis: Exposure to urban infrastructure and practices facilitates the transmission of social remittances, leading to the adoption of sustainable behaviours in rural areas.

Pollution and Migration Flows (Chapter 3)

To evaluate how pollution disparities influence migration flows, this chapter uses a differential-based gravity model:

$$M_{ij} = \alpha + \beta \frac{PM_{2.5j}}{PM_{2.5i}} + \gamma Z_{ij} + \epsilon_{ij} \tag{1.2}$$

where:

- M_{ij} represents migration flows from origin city i to destination city j.
- $\frac{PM_{2.5j}}{PM_{2.5i}}$ is the pollution gap ratio between destination and origin.
- Z_{ij} includes economic controls (e.g., GDP, wages, infrastructure).

Key Hypothesis: Cities with higher relative pollution act as push factors, while cleaner cities attract migrants, creating environmental inequality between regions.

Pollution and Retirement Planning (Chapter 4)

To examine the effects of pollution exposure on retirement expectations, this chapter applies the following panel fixed-effects model:

$$R_{it} = \alpha + \beta PM_{2.5it} + \gamma Z_{it} + \mu_i + \theta_t + \epsilon_{it} \tag{1.3}$$

where:

- R_{it} denotes the expected retirement age of individual i at time t.
- $PM_{2.5it}$ represents pollution levels at the city level.
- Z_{it} includes health, income, and demographic controls.
- μ_i and θ_t control for individual and time fixed effects, respectively.

Key Hypothesis: Higher pollution levels accelerate retirement decisions due to deteriorating health, while financial insecurity may delay retirement.

(3) Identification Strategies

To ensure causal inference, this thesis employs several identification strategies across the three empirical chapters:

- **Endogeneity Problems and Solutions**

  The empirical analyses in this thesis face potential endogeneity issues arising from three primary sources: omitted variable bias, measurement error, and reverse causality. Addressing these problems is critical to establishing credible causal relationships.

  - **Omitted Variable Bias:** In all three chapters, the outcome variables may be confounded by unobservable household and city characteristics, such as environmental awareness or preferences or city economic and social capital stock, which are correlated with both migration decisions and environmental practices. Hence, the models controlled for different fixed effects at the province, year, city, and household levels for time-invariant and time-variant unobserved heterogeneity at the regional or household level.
  - **Selection Bias:** Migration decision is often threatened by selection problems, since it is not randomly assigned. Migration decision and behaviour are easily influenced by economic, environmental and social factors. To address it, the regional out-migration rate and return migration rate, so-called migration network, is applied to split the selection bias.
  - **Measurement Error:** Pollution measures used in Chapters 3 and 4, such as $PM_{2.5}$ concentrations, may suffer from reporting inaccuracies or spatial aggregation errors. To mitigate these concerns, high-resolution satellite data from NASA and Washington University is employed, ensuring consistent measurement across locations and time.
  - **Reverse Causality:** In Chapter 3, the relationship between pollution and migration flows may suffer from reverse causality, as migration patterns themselves can influence pollution levels through increased urbanisation. Similarly, in Chapter 4, To address this, the study uses the relative pollution gap between origin and destination cities, treating it as an exogenous variable in a differential-based gravity model.

  In Chapter 4, pollution exposure may be endogenous to retirement expectations if individuals select locations based on anticipated health effects. To handle this, thermal inversion days are used as an instrumental variable for $PM_{2.5}$, exploiting natural variations in pollution unrelated to human activity.

- **Robustness Check and Complementary Analyses:**

  A range of strategies is employed to validate and deepen the findings:

  - **Robustness Checks:** The core results are re-estimated using alternative pollution measures, different instrumental variables, and varying index thresholds. In addition, a Heckman selection model is employed to assess potential sample selection bias. To account for possible macro-level confounders, policy changes during the study period are also incorporated into the analysis.
  - **Heterogeneity Analyses:** Subsample analyses are used explicitly to assess heterogeneity. These analyses compare rural versus urban households, gender groups, and education levels to examine how institutional and demographic contexts condition the behavioural effects of pollution and migration.
  - **Complementary Analyses:** Models incorporating lagged $PM_{2.5}$ values are estimated to investigate delayed and cumulative effects of pollution exposure. This approach identifies whether behavioural responses emerge immediately or persist over time, distinguishing longer-term impacts from short-term shocks.

Together, these strategies enhance the credibility of the empirical results by addressing omitted variable bias, reverse causality, and measurement errors, providing a solid foundation for causal inference across all three chapters.

1.5 Structure of the Thesis

The thesis is organised into five chapters:

Chapter 1: Introduction

Outlines the research background, motivations, and key questions, and introduces the Chinese institutional context.

Chapter 2: Migration and Green Behaviours

Examines how rural-to-urban migration affects environmental practices in origin households, distinguishing between current-migrant and return-migrant households.

Chapter 3: Pollution and Migration Flows

Investigates whether air pollution influences inter-city migration flows in China, using a gravity modelling approach.

Chapter 4: Pollution and Retirement Planning

Analyses how pollution exposure affects expected retirement age and labour supply among older adults, with emphasis on vulnerable groups.

Chapter 5: Conclusion

Summarises key findings, discusses policy implications, and proposes avenues for future research.

Chapter 2

Green Remittances: How does Migration shape Green Behaviours in Rural China

Abstract

This paper investigates the impact of internal rural-to-urban migration on the green behaviour of left-behind families in China, using willingness to recycle and fixed-place garbage disposal practices as indicators. The study examines the distinct effects of current migration and return migration, addressing selection bias, heterogeneity, and socio-economic gaps between urban and rural areas. The findings reveal that current migration significantly reduces recycling and fixed garbage placement probabilities by 71.9% and 21.8%, respectively, as the absence of household members disrupts routines. In contrast, return migration increases these behaviours by 79.0% and 46.3%, with returnees transferring urban-acquired norms to their families. Mechanism analysis highlights that better green infrastructure and social capital in destination cities amplify these positive effects, suggesting social remittance channels alongside financial remittances. This paper also found that younger, male-headed families with higher education and financial stability benefit most from return migration, while older families are more negatively affected by current migration. These findings underscore the importance of family capital in adopting sustainable practices and the challenges faced by vulnerable groups. This study is the first to apply social remittance theory to internal migration in China, offering insights into how migration shapes rural environmental behaviours.

Keywords: Rural-to-Urban Migration, Social Remittance, Green Behaviour

2.1 Introduction

The concept of a Green Economy is central to sustainable development, particularly in developing countries where rural-to-urban migration plays a pivotal role in economic growth. However, the impact of migration on green behaviours, such as recycling and waste disposal — remains underexplored. This paper seeks to address this gap by examining how migration, particularly current migration and return migration, influences the green behaviours of left-behind families in China. Specifically, it investigates whether migration acts as a channel for the diffusion of environmental norms and practices through knowledge spillovers.

Labour migration is often associated with financial remittances, which improve the living standards of left-behind families by supporting education, healthcare, and other necessities. Beyond financial transfers, migration also facilitates the movement of social remittances, including advanced skills, knowledge, and behavioural norms from urban to rural areas. When migrants are exposed to urban environments, their habits and preferences are shaped by new circumstances. Upon returning home, these "new norms" may diffuse to left-behind family members, potentially altering household behaviours. However, this process can have positive or negative effects, depending on the nature of the norms transmitted. Empirical evidence on these spillovers is limited, given the challenges of direct data collection and identification problems in microeconomic studies. This paper addresses these challenges by examining the impact of migration on green behaviours such as recycling willingness and fixed trash placement, using methods that control for double selection into current and return migration.

China provides an ideal context for this study due to its unique socio-economic conditions and large-scale internal migration. Over the past four decades, rapid urbanisation has dramatically reshaped China's social and economic landscape. The urbanisation rate surged from 18% in 1978 to 60.6% in 2019 (National Bureau of Statistics, 2020). As of 2021, China had 292.51 million migrant workers, with 171.72 million originating from rural areas (National Bureau of Statistics, 2021). This large-scale rural-to-urban migration has given rise to split-household families, where left-behind members often include children and elderly relatives. China's hukou system, which restricts labour mobility by tying social benefits to one's registered residence, further reinforces the prevalence of left-behind families, as migrants face challenges in permanently settling in urban areas.

China's case offers unique insights compared to international migration. First, the temporary nature of labour migration under hukou restrictions increases the likelihood of return migration, emphasising the role of urban exposure in shaping left-behind households' behaviours. Second, advancements in infrastructure, such as internet access and logistics systems, facilitate communication between migrants and their families, reinforcing the channels for norm diffusion. Finally, existing studies on migration in

China primarily focus on financial remittances, health, and education, with limited attention to the transmission of environmental norms or green behaviours.

Using data from the China Labour-Force Dynamics Survey (CLDS), which covers 29 provinces and 159 cities, this paper quantitatively examines how migration affects green behaviours among left-behind families. The study focuses on recycling willingness (e.g., classifying daily waste) and fixed trash placement (e.g., using public rubbish bins). Controlling for double selection into current and return migration, the findings reveal a dual effect: current migration disrupts green behaviours due to the absence of key family members, around 71.9% reduction in recycling probability, while return migration fosters positive changes by transmitting environmental norms acquired in urban areas, approximately 79.0% increase. Furthermore, the study explores the mechanisms of norm diffusion, showing that exposure to better urban green and information infrastructure and higher social capital can improve the green awareness of left-behind families. Also, according to the heterogeneity analysis, I found that the male-led, with younger family members, higher education, and more sustainable financial left-behind families are easier to accept environmental norms and minimise the serious effects caused by the absence of family members.

This paper contributes to the literature in two keyways. First, it is one of the first studies to quantitatively analyse the impact of migration on green behaviours, a green dimension largely neglected in the current literature. While prior studies have explored migration's effects on political preferences, welfare, and gender norms (Frey and Meier, 2004; Giulietti et al., 2013; Hildebrandt et al., 2005; Nikolova and Graham, 2015), few have examined environmental behaviours, particularly in the context of social remittances. Existing research on green economics often relies on qualitative approaches (Moran-Taylor and Taylor, 2010; Taylor et al., 2011) or focuses on narrow topics like fuel consumption (Alem et al., 2016; Wu et al., 2021b). This study fills this gap by examining green behaviours at the household level in a large-scale internal migration setting. Second, this paper expands the scope of social remittance literature, which has predominantly focused on international migration and its impact on political and social norms (Nikolova et al., 2017; Tuccio and Wahba, 2018; Wahba, 2015). By focusing on rural-to-urban migration in China, the study provides a new perspective on how internal migration shapes environmental behaviours in developing countries.

The structure of this paper is as follows: Section 2.2 reviews the relevant literature on migration and social remittances, highlighting gaps that this paper aims to address. Section 2.3 describes the data source, variable definitions, and descriptive statistics. Section ?? outlines the econometric model and identification strategy employed to address endogeneity and ensure robust estimates. Section 2.5 presents the empirical results and provides a detailed analysis of the findings. Section 2.6 provides the robustness check to the main results. Finally, Section 2.7 concludes the whole paper.

2.2 Literature Review

A substantial body of literature examines the determinants of migration and its effects on left-behind individuals, origin communities, and host regions. Migrants often leave their hometowns to improve their well-being and that of their families, primarily through remittances that enhance education, healthcare, and living standards (Abramitzky et al., 2012; Nikolova and Graham, 2015). While the economic effects of remittances have been a primary focus of early research, recent attention has turned to the social effects of migration. Existing remittance studies typically distinguish between two types: financial remittances and social remittances.

Financial remittances are among the most visible and extensively studied consequences of migration, illustrating its economic impact on origin countries. Significant progress has been made in understanding the drivers and consequences of financial remittances at both microeconomic and macroeconomic levels. Rapoport and Docquier (2006) provides a comprehensive review of the literature in this area. Within the context of migration and the left-behind, much of the research focuses on the effects of financial remittances on education, health, and well-being. For instance, Lee (2011) analyses the impact of migration on the health and education of left-behind children, while Yi et al. (2019) examine its effects on the health of elderly left-behind individuals in China. These studies highlight financial remittances as a critical channel through which migration affects left-behind households.

In addition to financial remittances, migration also produces non-financial impacts, such as shifts in social norms, political preferences, and attitudes. Social remittances, a concept introduced by Levitt (1998), refer to the transmission of cultural and social ideas, behaviours, and norms from host to origin regions through migration. This diffusion process encompasses normative structures, systems of practices, and social capital (Levitt, 1998; Levitt and Lamba-Nieves, 2011). Social remittances have been shown to influence societal norms positively, fostering entrepreneurship, gender equality, political preferences, and even environmental protection (Giulietti et al., 2013; Ljunge, 2012; López-Feldman and Chávez, 2017; Tuccio and Wahba, 2018).

Although the concept of social remittances was first introduced by Levitt (1998), its systematic analysis gained momentum with studies like Spilimbergo (2009), which brought more attention to this area. Compared to sociologists, economists are particularly interested in identifying causal relationships and addressing selection bias. Economists aim to determine not only whether migration changes behaviours or norms but also whether these changes are directly driven by migration. This distinction underpins the theoretical framework of this study.

Unlike financial remittances, which are easily measured and tracked by organisations like the World Bank and the IMF, social remittances are challenging to quantify. Most empirical studies rely on household surveys or field experiments, yet many social behaviours and norms remain complex to capture (Levitt, 1998; Levitt and Lamba-Nieves, 2011). While substantial evidence confirms the existence of social remittances, critical questions remain unanswered. These include how social remittances are formed, how they diffuse between origin and host regions, and how they transform within host communities (Tuccio and Wahba, 2020).

The impacts of migration on left-behind families are diverse, encompassing education, well-being, entrepreneurship, and social norms. For example, Giulietti et al. (2013) use data from rural China to show that return migration promotes self-employment among left-behind families and their communities. Hildebrandt et al. (2005) examine health knowledge spillovers in Mexico, finding that migration improves young mothers' health knowledge through indirect channels. In the domain of social norms,Wahba (2015) demonstrates that migration from Morocco to democratic European countries positively influences political preferences and social norms. However, Tuccio and Wahba (2018) find that return migration from conservative countries can reinforce traditional gender norms, illustrating that social remittances are not always positive.

Nikolova et al. (2017) identify two primary mechanisms through which social remittances are transmitted: (1) the absence effect, where the migration of socially proactive individuals weakens community networks and reduces pro-social behaviours among the left-behind Frey and Meier (2004); and (2) the diffusion effect, where migrants acquire language skills, cultural knowledge, and social norms in host regions and transfer these back to origin communities through communication, financial support, or physical visits (Levitt and Lamba-Nieves, 2011). This study builds on these mechanisms to explore how social remittances influence green behaviours among left-behind households.

The relationship between migration and green behaviours or environmental protection remains underexplored, creating a significant research gap that this paper aims to address. Existing literature on migration and the environment predominantly focuses on environmental migration or the health impacts of pollution on left-behind populations (Afifi, 2011). In China, the socio-economic divide between urban and rural areas provides a unique opportunity to study the transfer of green norms. Urban environments, with their superior infrastructure—including public rubbish bins, garbage collection stations, and sanitation workers—offer a stark contrast to rural conditions, potentially influencing the behaviours of rural-to-urban migrants.

The mechanism through which migration affects green behaviours is rooted in urban regulations and social norms. Chai et al. (2022) examine the influence of the China National Civilised City Award (NCCA), a programme designed to improve urban sustainability. They find that local governments' efforts to win the NCCA title significantly enhance urban infrastructure and environmental standards, creating a conducive environment for migrants to adopt green practices. Migrants exposed to such conditions are likely to transfer these norms to their left-behind families through channels such as financial support, communication, and life experiences.

2.3 Data

The data used in this paper is sourced from the China Labour Force Dynamics Survey (CLDS). Established in 2012 by Sun Yat-Sen University, the CLDS aims to conduct a biennial follow-up survey to observe changes and interactions between social structures, families, and the labour force in both urban and rural settings. The survey collects data across three dimensions: individual labour force, households, and communities (cities and villages). Spanning 29 provinces (excluding Hong Kong, Taiwan, Tibet, and Hainan), the survey provides comprehensive interdisciplinary data on topics such as education, work, migration, health, social participation, economic activities, grassroots organisations, and green behaviours.

To ensure representativeness, the CLDS employs a multistage cluster, stratified, and probability-proportional-to-size sampling approach, with survey weights adjusted to align with the distribution of the labour force population. These methodological strengths make the dataset particularly suitable for studying the relationship between migration and green behaviours. In 2014, the survey incorporated additional questions on green behaviours and environmental attitudes, offering unique opportunities to explore how labour migration influences family-level recycling behaviours.

This paper utilises data from the 2014 and 2016 waves of the China Labour-force Dynamics Survey (CLDS), constructing a repeated cross-sectional framework at the individual, household, and community levels. Since questions on environmental behaviours were not uniformly administered across all respondents, and because practices such as recycling and designated garbage disposal were compulsory in certain cities, the sample was restricted to ensure comparability and internal validity.

The initial dataset contains approximately 15,000 households. To obtain a clean and consistent analytic sample, several restrictions were applied. First, observations with missing values on key variables—such as migration status or recycling willingness—or with erroneous entries or extreme outliers (identified using standard deviation thresholds for continuous variables) were excluded. Second, households located in cities where local regulations during the survey period mandated both recycling and fixed trash placement were removed, as these policy requirements could mechanically inflate reported green behaviours and confound behavioural comparisons across locations; this policy-related exclusion affects less than 2% of the initial sample.

In addition, households reporting both current and return migrants were excluded to avoid conflating the effects of these two distinct migration types. After applying these criteria, the final estimation sample consists of 8,126 households, divided into three mutually exclusive categories: 3,679 non-migrant households (control group), 3,444 current migrant households, and 1,003 return migrant households (treatment groups).

These restrictions ensure that the estimation sample is free from measurement errors, unaffected by location-specific mandatory policies, and representative of household characteristics relevant to the study's objectives, thereby providing a consistent basis for estimating the impact of migration on environmental behaviours.

2.3.1 Dependent Variable

A key strength of the CLDS dataset lies in its inclusion of green behaviour-related variables. Since few field surveys or censuses capture data on environmental behaviours, this paper uses the questions "Are the households willing to recycle?" and "Are you willing to put your garbage in a fixed placement for disposal?" to generate the outcome variable. By using recycling willingness as a proxy for green behaviour, the study introduces an innovative approach to measuring family-level environmental awareness and action.

2.3.2 Migration Status

Using the detailed household and individual-level data provided by the CLDS, I categorised families based on the migration status of their members. Each individual's status was examined to determine whether they were living with the household or absent, and if absent, the reason for their absence. Families with at least one member currently absent for work or business were classified as current-migration families, while those with at least one member who had left the household for work for at least six months but had since returned were categorised as return-migration families. Families without any migration experience were identified as non-migration families. This classification provides a clear framework to distinguish between ongoing, past, and no migration experiences, enabling a comprehensive analysis of their impact on green behaviours at the family level.

2.3.3 Control variables

For covariates, current literature on explaining the factors affecting green life is mainly based on the Theory of Reasoned Action (TRA) and the Theory of Planned Behaviour (TPB) (Ramayah et al., 2012). Both theories split the motivation for green life into two aspects: internal factors involving self-attitude to green life and environment, and subjective norms affected by publicity, and external factors such as personal characteristics, and objective factors related to family, residence and even governance. Accordingly, apart from personal characteristics, I also add family-level characteristics, particularly in income, family size, the socioeconomic factors of the family head, and village-level factors, like the location, infrastructure condition and wealth (Otoma et al., 2013; Ojeda-Benítez et al., 2008; Xu et al., 2016) .

Table 2.1 shows the definition of variables, and Table 2.2 presents the summary statistics for the overall sample as well as subsamples defined by household migration status: non-migration, current migration, and return migration. Several key patterns emerge.

First, environmental behaviours vary significantly across groups. Households with return migrants report higher rates of recycling (86.1%) and fixed garbage disposal (96.2%) than both non-migrant and current migrant households. In contrast, current migrant households show lower environmental engagement, suggesting a potential disruption effect associated with ongoing migration.

Second, demographic and socioeconomic characteristics differ markedly across groups. Return migrant households tend to be younger (mean age 48.1), better educated (e.g., 54.2% with secondary education), and have higher annual incomes (45.8k RMB), compared to non-migrant households. They also exhibit larger family sizes (4.5 members on average), possibly reflecting multigenerational living arrangements or return of younger migrants.

Third, current migrant households are more likely to reside in lower-income villages (mean village income 7.8k RMB), and have less access to environmental infrastructure, as proxied by the number of recycling sites (30.4%). Meanwhile, non-migrant households are disproportionately located in suburban areas of metropolitan regions (11.5%).

These descriptive patterns suggest substantial heterogeneity in green behaviours and socioeconomic conditions across migration groups, supporting the hypothesis that migration status is closely associated with both household behaviour and environmental exposure.

TABLE 2.1: **Description of Variables**

Variable	**Description**
Dependent Variable	
Recycling	Are the households willing to do recycling? (Yes=1, No=0)
Fixed Disposal	Are the households willing to accept the fixed placement of garbage? (Yes=1, No=0)
Interest Variable	
Current Migration Household	The family with at least one current migrant. (Yes=1, otherwise=0).
Return Migration Household	The family with at least one return migrant. (Yes=1, otherwise=0).
Household Characteristics	
Family Size	The number of family members
Annual Income	Annual income of family in each year (10^3 CNY)
Near Recycling Sites	House is near the garbage bins or stations (Yes=1, No=0)
Individual Characteristics	
Male	Gender of family head (Male=1, Female=0)
Age	The age of family head
Marital Status	Married=1, otherwise=0
Employ status	Employed=1, otherwise=0
Education level	Illiteracy / Basic / Secondary / Post-secondary education
Party	Is the family head a member of the political party? (Yes=1, No=0)
Village Characteristics	
Village average income	The average income in village (10^3 CNY)
Own School	There is a primary school in the local administrative area. (Yes=1, No=0)
Suburb of metropolis	The village is in the Suburb of metropolis (Yes=1, No=0)

Note: Interest variable: The family with current migration and with return migration. Both variables are mutually exclusive, which means the sample does not involve any family with current and return migration at the same time.
Source: CLDS

Table 2.2: **Descriptive Statistics**

Sample	Overall Sample				Non-migration	Curretn Migration		Return Migration	
	Mean	Std. Dev.	Min	Max	Mean	Mean	t-test	Mean	t-test
Variable		(1)			(2)		(3)		(4)
Recycling	0.80	0.401	0	1	0.810	0.769	-4.206***	0.861	3.770***
Fixed Disposal	0.92	0.271	0	1	0.925	0.902	-3.351***	0.962	4.199***
Male	0.85	0.357	0	1	0.844	0.843	-0.056	0.898	4.347***
Age	53.63	12.767	10	108	53.91	54.96	3.463***	48.05	-12.224***
Illiteracy	0.13	0.334	0	1	0.141	0.131	-1.142	0.072	-5.839***
Basic Education	0.38	0.485	0	1	0.373	0.393	1.732*	0.360	-0.788
Secondary Education	0.46	0.499	0	1	0.448	0.454	0.474	0.542	5.327***
Post-Secondary Education	0.03	0.170	0	1	0.0380	0.022	-4.016***	0.026	-1.842*
Marriage	0.98	0.154	0	1	0.962	0.992	8.251***	0.968	0.841
Employ status	0.80	0.403	0	1	0.751	0.825	7.676***	0.860	7.374***
Political Party Member	0.09	0.282	0	1	0.0920	0.084	-1.057	0.080	-1.167
Annual income	39.84	66.224	-30	3000	37.90	40.18	1.440	45.78	3.694***
Family size	4.40	2.079	1	18	3.574	5.254	36.190***	4.526	14.600***
Recycling sites	0.34	0.472	0	1	0.358	0.304	-4.791***	0.362	0.246
Own School	1.61	0.488	1	2	1.563	1.642	6.831***	1.653	5.143***
Suburb of metropolis	0.09	0.285	0	1	0.115	0.061	-8.124***	0.090	-2.292**
Village average income	9.31	9.391	0	60	10.61	7.817	-12.840***	9.650	-2.651***
N		8126			3679		3444		1003

Note: This table is the descriptive statistics in different groups: Column 1 is the non-migration, which means the family will not have any member with migration experience; and Columns 2 and 3 are current migration and return migration groups respectively.
The t-test is for different means, where the control group is always non-migration group.
**, ** and * represent 1%, 5% and 10% significance levels, respectively.
Source: CLDS 2014,2016

2.4 Methodology

2.4.1 Baseline Specification

The baseline models are estimated using the Linear Probability Model (LPM) within a 2SLS framework because the coefficients are directly interpretable as percentage point changes in the outcome probability, facilitate the inclusion of multiple high-dimensional fixed effects, and allow for straightforward implementation of weak-instrument diagnostics. While IV-Probit models better reflect the bounded nature of probabilities, they are computationally more demanding in the presence of high-dimensional fixed effects and their coefficients are not directly interpretable without transformation. For robustness, I estimate IV-Probit models and report the corresponding Average Partial Effects (APEs) in Section 2.6 Robustness Check, which are consistent in sign and significance with the LPM estimates.

At the same time, standard errors are clustered at the household level to account for potential intra-household correlation in environmental behaviours and migration status. Since the analysis relies on repeated cross-sectional data, and the key explanatory variables—such as current and return migration—are defined at the household level, this clustering strategy ensures statistically robust inference and mitigates the risk of underestimating standard errors.

Although clustering at a higher level (e.g., village) was considered, the relatively small number of observations per village (typically 30–50) and the household-specific nature of the treatment variables make household-level clustering a more appropriate and conservative choice in this context.

The model of migration effect on left-behind family i in village j is specified as:

$$Y_{i,j} = \alpha_{i,j} + \beta M_{i,j} + \gamma X_{i,j} + \epsilon_{i,j} \tag{2.1}$$

where

$$M_{i,j} = \begin{cases} \mathrm{CM}_{i,j} & \text{if current migration} = 1, \text{otherwise } 0, \\ \mathrm{RM}_{i,j} & \text{if return migration} = 1, \text{otherwise } 0. \end{cases}$$

- $Y_{i,j}$ represents recycling (whether the family supports daily garbage recycling) and fixed disposal (whether the family uses fixed places for garbage disposal). Both variables are binary, with 1 indicating “yes” and 0 “no,” serving as proxies for green behaviour among left-behind families.

- M represents the migration status, with two mutually exclusive binary indicators: current migration (CM), indicating that at least one family member is a current migrant (1 if yes, 0 otherwise), and return migration (RM), indicating that at least one family member has returned after working away for six months or longer (1 if yes, 0 otherwise). Non-migrant households are the reference group.

- X is a vector of covariates at the individual, household, and village levels. At the individual level, the family head's characteristics—including age, gender, education, employment status, political party membership, and marital status—are considered, as they typically represent the most influential role in the household. Household-level covariates include annual income (a proxy for wealth) and family size (labour capital). Village-level variables capture the broader context, including the presence of a community primary school (1 if yes, 0 otherwise), location (whether the village is suburban, serving as a proxy for infrastructure), and whether there are garbage bins or carts in the village, as well as average village income (a measure of community wealth).

2.4.2 Identification Strategy

This study confronts two key identification challenges: reverse causality and selection bias. Reverse causality is less of a concern in this context, as environmental behaviours—such as recycling or waste sorting—are unlikely to influence migration decisions in China, where migration is primarily driven by economic incentives and structural disparities (Roback, 1982).

In contrast, selection bias poses a more significant threat to causal inference. Migration decisions are inherently non-random, shaped by a complex array of household and community-level characteristics. For example, households with greater financial or social resources may be more likely to send migrants, which would bias OLS estimates of migration's impact on environmental behaviours. In the case of return migration, additional selection concerns arise, as returnees may self-select based on unobserved individual preferences or macroeconomic shocks such as layoffs or urban policy changes (Li and Li, 2022; Deschenes et al., 2020).

To address these challenges, I employ two instrumental variables derived from city-level migration trends. For current migration, I use the change in outmigration rates from 2005 to 2015, capturing the strength of existing migration networks that lower the costs of moving. For return migration, I adopt a similar strategy using the change in return migration rates over the same period. These instruments reflect long-standing social ties that influence migration behaviour, following the approach of Lai et al. (2021), but are plausibly exogenous to current household environmental practices.

To enhance the credibility of the exclusion restriction, I also include village-level infrastructure indicators, such as the availability of waste collection facilities and public bins, to control for potential confounding factors. This setup ensures that the instruments capture variation in migration exposure while remaining uncorrelated with unobserved determinants of green behaviour, thereby supporting a robust identification strategy.

It is important to note that this study constructs migration network variables at the city level, rather than the village level, for two key reasons. First, the China Labourforce Dynamics Survey (CLDS) does not provide detailed village identifiers; the precise village address of each respondent is unavailable. As a result, village-level migration rates cannot be reliably calculated. Second, due to institutional constraints such as the hukou system, within-city migration (e.g., from one village or neighbourhood to another) is considerably easier and more frequent than cross-city migration. Relying on within-city movements would overstate the strength of migration networks and introduce measurement error into the instrumental variables. To mitigate this, the paper uses 1% samples of the China Census from 2010 and 2015 to compute cross-city outmigration and return migration rates, which serve as proxies for village-level migration trends while maintaining consistency and plausibility in measurement.

The identification of migration and return migration in the 1% China Census relies on two core questions: (1) "What is your hukou registration location?" and (2) "Where did you live five years ago?". Individuals whose current place of residence differs from their hukou location are classified as current migrants. Conversely, return migrants are defined as those whose current residence matches their hukou location but who lived in a different place five years prior.

Using individual-level responses, I first assign a migration status to each person in the sample according to these definitions. I then aggregate the number of current and return migrants by city, dividing by the total city population in the corresponding year to obtain city-level outmigration and return migration rates. These rates are calculated separately for the 2005 and 2015 Census waves. Comparing the two waves yields the change in migration intensity over time, which serves as a proxy for the strength of current and return migration networks.

This approach provides a consistent and comparable measure of cross-city migration dynamics, capturing both the scale and the directional shifts of mobility flows over a five-year interval. The construction algorithm is summarised as follows:

The first stage regression equations for current migration and return migration groups are specified as:

$$\mathrm{CM}_{i,j,c} = \theta_i + \lambda z_c + \tau X_{i,j} + \epsilon_{i,j,c}, \tag{2.2}$$

$$\mathrm{RM}_{i,j,c} = \rho_i + \phi s_c + \omega X_{i,j} + e_{i,j,c}, \tag{2.3}$$

where $\mathrm{CM}_{i,j,c}$ and $\mathrm{RM}_{i,j,c}$ represent the current and return migration status of household i in village j of city c. z_c is the change in outmigration in origin city c from 2005 to 2015, and s_c is the change in return migration in origin city c from 2005 to 2015. Both instrumental variables are calculated in logarithmic differences from Census 2005 and 2015:

$$z_c = \log(\textit{outmigration rate}_{c,2015}) - \log(\textit{outmigration rate}_{c,2005}), \tag{2.4}$$

$$s_c = \log(\textit{return rate}_{c,2015}) - \log(\textit{return rate}_{c,2005}). \tag{2.5}$$

The *outmigration rate* and *return migration rate* are computed based on the 1% China Census data. Specifically:

- The ***outmigration rate*** for city c is defined as the ratio of the number of *current migrants* whose *hukou* is registered in city c to the number of *non-migrant residents* currently living in city c:

$$\text{Outmigration Rate}_c = \frac{\text{Number of current migrants with hukou in city } c}{\text{Number of non-migrant residents in city } c} \quad (2.6)$$

- The ***return migration rate*** for city c is defined as the ratio of the number of *return migrants*—individuals whose current address matches their *hukou* in city c but who lived elsewhere five years ago—to the number of *non-migrant residents* in city c:

$$\text{Return Migration Rate}_c = \frac{\text{Number of return migrants in city } c}{\text{Number of non-migrant residents in city } c} \quad (2.7)$$

To ensure the instrumental variables effectively address selection bias, Figure 2.1 presents the distribution of the most recent migration years for current migrants and return migrants in the raw sample. Over 75% of migrants are concentrated in the post-2000 period, reflecting the significant migration boom after 1995. This trend aligns with the rise of labour networks that underpin the instrumental variables, ensuring robust coverage of migration dynamics.

The timeframe also coincides with China's increasing emphasis on environmental protection, which became a national priority following its accession to the WTO in 2001. For instance, policies such as the Environmental Law amendment in 2014 and the expansion of recycling infrastructure reflect this shift (Sun et al., 2019). Before 2000, environmental awareness and recycling behaviours were largely neglected at both the governmental and household levels. Thus, migration trends from 2000 to 2015 are not only representative of labour networks but also align with the initial development of green awareness and infrastructure. This ensures the instrumental variables effectively capture the key period for analysing the impact of migration on green behaviours.

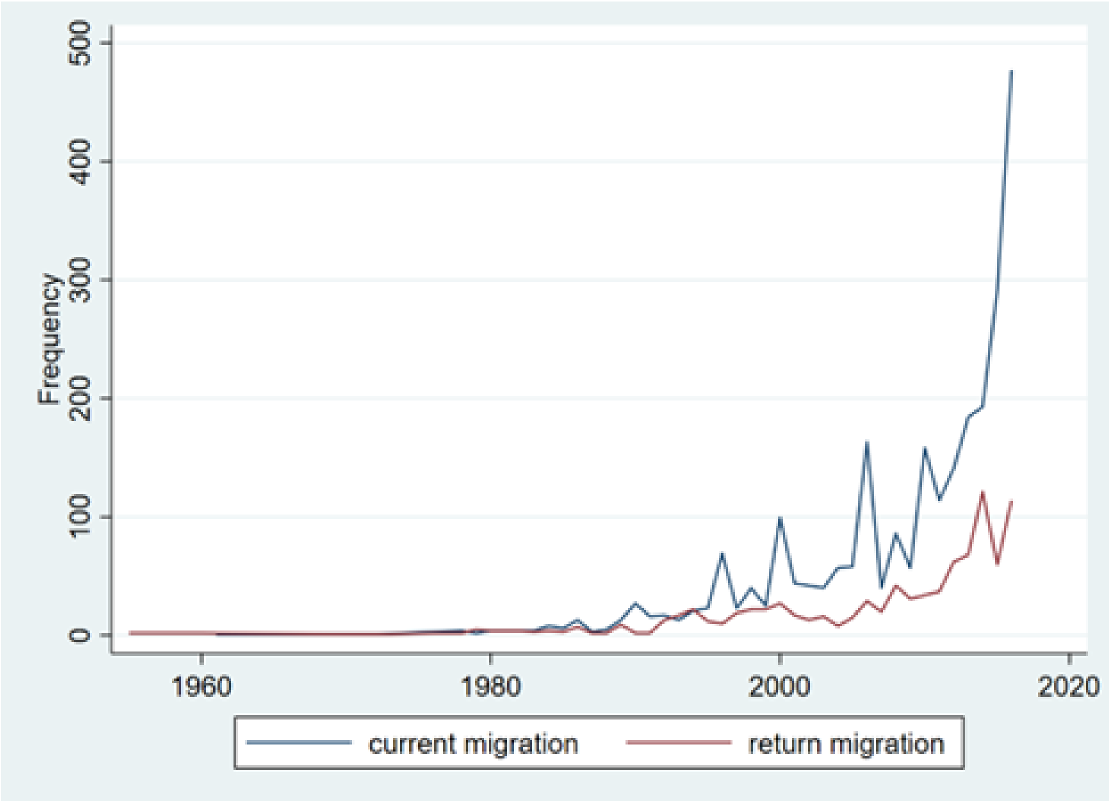


Figure 2.1: **Migration Frequency in the Sample**
Note: This Figure shows current and return migration frequency in the Sample from 1960 to 2018
Source: CLDS; Census

2.5 Empirical Results

2.5.1 Baseline Results - Pooled OLS and 2SLS

Following Equation 2.1, I first estimate pooled OLS models for the recycling outcomes of households with current migrants and those with return migrants. Table 2.3 presents the pooled OLS and 2SLS results for the impact of current and return migration on two measures of green behaviour—recycling willingness and fixed trash placement — using household-level clustering of standard errors.

The Pooled OLS results (Panel A, Columns 1–4) suggest that current migration is associated with a 4.3 percentage point decrease in recycling probability and a 2.2 percentage point decrease in fixed trash placement probability, both statistically significant at the 1% level.

In contrast, return migration shows a positive effect, increasing recycling willingness by 4.8 percentage points and fixed trash placement by 2.6 percentage points, both significant at conventional levels. These findings suggest that return migrants are more likely to engage in green behaviours, potentially due to greater environmental awareness or social norm internalisation.

However, since migration decisions may be influenced by unobserved household characteristics correlated with pro-environmental behaviours, these estimates may suffer from endogeneity bias..

The 2SLS results (Table 2.4) address endogeneity concerns by using city-level out-migration and return-migration rates as instruments. The first-stage F-statistics exceed conventional thresholds, confirming the relevance and strength of the instruments.

The second-stage estimates reveal substantially larger effects than the OLS results. Specifically, current migration reduces recycling probability by 71.9 percentage points and fixed trash placement by 21.8 percentage points, both statistically significant at the 1% level. These results indicate a severe disruption to green behaviours when household members migrate away, likely due to reduced available household labour and time for environmentally responsible tasks.

In contrast, return migration increases recycling probability by 79.0 percentage points and fixed trash placement probability by 46.3 percentage points, again statistically significant. These large effects suggest that returnees contribute positively to environmental behaviours, potentially by bringing back pro-environmental norms, experiences, and resources acquired in urban settings.

The considerable gap between the OLS and 2SLS estimates highlights potential biases in the naïve estimates. OLS appears to underestimate the negative effect of current migration and understate the positive effect of return migration, likely due to unobserved household characteristics correlated with both migration decisions and environmental behaviours. These findings support the interpretation that out-migration weakens green household practices due to the higher time and effort costs in the absence of migrants, while returnees help fill this gap by re-engaging in daily tasks and promoting environmentally friendly routines.

Interestingly, the magnitude difference between the OLS and 2SLS estimates—ranging from 10 to 20 times—is considerably larger than what is typically reported in studies on migration and behavioural outcomes, where IV corrections generally increase coefficient estimates by a factor of 3 to 7 (Giulietti et al., 2013; Tuccio et al., 2019) . Several interrelated factors may account for this pronounced divergence. First, the 2SLS strategy identifies a local average treatment effect (LATE), capturing the impact of migration among households whose migration decisions are influenced by exogenous variation in city-level migration trends between 2005 and 2015. These marginal households may be particularly sensitive to institutional or environmental factors, resulting in stronger behavioural responses than the average household captured in the OLS estimates.

Second, although the dependent variables—recycling and fixed garbage placement—are binary, they are not rare in this sample, with mean rates of 0.80 and 0.92, respectively. In the context of a linear probability model, this creates a setting where treatment effects can appear disproportionately large, especially when the treatment shifts the probability of an already common behaviour near its upper bound. Third, measurement error in self-reported migration status and omitted variables such as environmental preferences likely attenuate OLS estimates, biasing them toward zero. In contrast, the 2SLS approach corrects for these sources of endogeneity, leading to substantially larger and arguably more accurate effect sizes. Taken together, these factors—treatment effect heterogeneity, outcome scale effects in LPM, and endogeneity correction—help explain the unusually large gap between the OLS and 2SLS estimates in this context.

The contrasting effects can be explained by the dynamics of migration. Current migration disrupts household routines and reallocates labour, making it harder for left-behind families to maintain green behaviours. Additionally, left-behind households may prioritise economic survival over environmental practices, compounding the negative effects. In contrast, return migration introduces positive spillovers through exposure to better environmental norms, financial resource transfers, and social influence in origin communities. These findings align with migration theories emphasising the role of norm diffusion and resource flows (Tuccio and Wahba, 2020).

TABLE 2.3: **Baseline Reuslts - Pooled OLS**

Groups	Current Migration		Return Migration	
Dependent Variable	Recycling	Fixed Garbage Placement	Recycling	Fixed Garbage Placement
	(1)	(2)	(3)	(4)
Current Migration	-0.043***	-0.022***		
	(0.012)	(0.008)		
Return Migration			0.048***	0.026***
			(0.014)	(0.007)
Covariates Included	Yes	Yes	Yes	Yes
Fixed Province	Yes	Yes	Yes	Yes
Fixed Year	Yes	Yes	Yes	Yes
Fixed Year × Province	Yes	Yes	Yes	Yes
N	7122	7122	4681	4681
R2	0.059	0.075	0.073	0.086

Note: This table reports OLS results of regression on current migration and return migration. Columns 1 and 2 are the current migration group; columns 3 and 4 are the return migration group.
4. Clustering standard errors at the household level is in the parenthesis. ***, **,* represent the level of significance at 1%, 5%, 10%.
Source: CLDS 2014, 2016

TABLE 2.4: **Baseline Results - 2SLS**

Groups	Current Migration Group		Return Migration Group	
Dependent Variable	Recycling	Fixed Garbage Placement	Recycling	Fixed Garbage Placement
	(1)	(2)	(3)	(4)
First stage				
Out-migration rate	0.082***	0.082***		
	(0.013)	(0.013)		
Return-migration rate			0.052***	0.052***
			(0.015)	(0.015)
F-test	41.75	41.75	12.13	12.13
Second Stage				
Current migration	-0.719***	-0.218***		
	(0.170)	(0.074)		
Return migration			0.790***	0.463***
			(0.287)	(0.175)
Covariates Included	Yes	Yes	Yes	Yes
Fixed Province	Yes	Yes	Yes	Yes
Fixed Year	Yes	Yes	Yes	Yes
Fixed Year×Province	Yes	Yes	Yes	Yes
N	7122	7122	4681	4681

Note: This table reports 2SLS results of regression on current and return migration, with two-stage results; columns 1 and 2 are the current migration group and columns 3 and 4 are the return migration group;
Instrumental variables are the change of out-migration and return-migration rates in origin city c between 2005 and 2015, based on the data from the census survey data. This is calculated by the logarithm difference forms following Equations 2.4 and 2.5. Household-level clustering standard errors at the household level are in the parentheses. ***, **,* represent the level of significance at 1%, 5%, 10%.
Source: CLDS 2014 and 2016; Census 2005-2015

2.5.2 Joint Effect of Current and Return Migration

To jointly estimate the effects of current and return migration on green behaviours, we combine the two treatment indicators into a single specification. Both are entered as mutually exclusive dummy variables, with households without migrants serving as the reference group. This approach improves comparability across groups and reduces model complexity. The model is estimated using both OLS and Two-Stage Least Squares 2SLS, where the change in city-level outmigration and return migration rates from 2005 to 2015 serve as instrumental variables for current and return migration, respectively. The specification is as follows:

$$Y_{i,j} = \alpha_{i,j} + \beta CM_{i,j} + \gamma RM_{i,j} + \lambda X_{i,j} + \epsilon_{i,j} \tag{2.8}$$

Where CM and RM are mutually exclusive dummy variables that document the migration status of current or return migration.

Compared to Table 2.4, which estimates subgroup-specific models separately for current and return migration and does not allow for direct comparison between the two effects due to differing sample compositions and reference groups, Table 2.5 presents a joint regression that includes both migration variables in a single model. By pooling all households and using non-migrant families as the common reference group, the joint specification enables a direct comparison of the relative strength and direction of current and return migration effects. Moreover, it reduces the risk of omitted variable bias that may arise from estimating separate models on split samples.

Interestingly, the joint 2SLS estimates reported in Table 2.5 reveal that both the negative impact of current migration and the positive impact of return migration on green behaviours are substantially larger in magnitude than those found in the separate subgroup regressions (Table 2.4). Specifically, current migration is associated with a reduction of 113.9 percentage points in the probability of recycling in the joint model, compared to 71.9 percentage points in the separate model. Similarly, return migration increases the likelihood of recycling by 295.5 percentage points in the joint model, as opposed to 79.0 percentage points in the separate case.

Given that this is a Linear Probability Model, these large coefficients do not reflect literal probabilities exceeding 100%, but instead indicate strong marginal effects. The inflated magnitudes in the joint model likely result from improved identification due to a shared reference group (households without any migration experience) and the simultaneous estimation of both migration effects, which helps mitigate omitted variable bias that may arise in disjoint subsample specifications. This approach provides a clearer contrast between the two types of migration and their opposing associations with environmental behaviours.

In sum, the joint estimation framework improves comparability across migration types by using a unified reference group and model structure. The results suggest that current and return migration have significantly divergent effects on green behaviours, with the magnitude of these effects becoming more pronounced when estimated simultaneously. This highlights the importance of accounting for multiple migration pathways within a common analytical framework to capture their distinct and contrasting roles in shaping environmental outcomes.

Table 2.5: **Joint Effect of Current and Return Migration**

	OLS		2SLS			
			First stage		Second stage	
Dependent variable	Recycling	Fixed Garbage Placement	Current Migration	Return Migration	Recycling	Fixed Garbage Placement
	(1)	(2)	(3)	(4)	(5)	(6)
Out-migration rate			0.054***			
			(0.012)			
Return-migration rate				0.051***		
				(0.014)		
F-test			57.4	14.8		
Current Migraiton	-0.041***	-0.024***			-1.139***	-0.355***
	(0.012)	(0.008)			(0.197)	(0.106)
Return Migration	0.052***	0.031***			2.955***	1.399**
	(0.014)	(0.008)			(0.945)	(0.608)
Covariates Included	Yes	Yes			Yes	Yes
Fixed Province	Yes	Yes			Yes	Yes
Fixed Year	Yes	Yes			Yes	Yes
Fixed Cohort	Yes	Yes			Yes	Yes
Fixed Year × Province	Yes	Yes			Yes	Yes
R^2	0.059	0.073				
N	8124	8124			8124	8124

Note: This table reports OLS and 2SLS results of regression on current and return migration, with two-stage results; Household-level clustering standard errors at the household level are in the parentheses. ***, **,* represent the level of significance at 1%, 5%, 10%.
Source: CLDS 2014 and 2016; Census 2005-2015

2.5.3 Recovery Potential in Rural China

A better understanding of the composition and economic value of rural domestic waste helps contextualise the behavioural impact of migration. According to Han et al. (2019), rural areas in China generate an average of 0.521 kg of household waste per capita per day. While this is lower than the rural average in developed countries, it still accumulates into substantial volumes given China's vast rural population. This waste is largely composed of organic (39.05%) and inert components, with significant amounts of nutrients such as nitrogen (1.02%), phosphorus (0.50%), and potassium (1.42%), indicating high potential for agricultural reuse and composting.

Despite these properties, the proportion of recyclable waste remains modest, and infrastructure for waste sorting and recovery is limited in many rural areas. Nevertheless, the economic value of improved recycling remains notable. Official data show that in 2016, the recycling volume of ten major categories of recoverable waste reached 246 million tonnes in China, with an estimated market value of 515 billion yuan (Xiao et al., 2018). This implies an average value of approximately 2,093 yuan per tonne of recyclable waste.

Applying this benchmark, even small changes in rural recycling behaviour can yield meaningful returns. Based on our findings, current migration reduces the probability of proper recycling by 71.9%. Assuming a rural household of three people produces around 570 kg of domestic waste per year (0.521 kg $\times$ 365 $\times$ 3), a 71.9% behavioural loss corresponds to a waste of roughly 410 kg annually. Valued at 2,093 yuan per tonne, this results in an estimated annual economic loss of **859 yuan per household** due to disrupted recycling practices from current migration.

Conversely, return migration improves recycling probability by 79.0%, which would translate into a potential **gain of around 943 yuan per household** annually, based on the same assumptions. These back-of-the-envelope calculations underscore the tangible economic benefits associated with sustaining or improving green behaviours in rural households, especially those affected by migration.

In sum, the reduction in recycling observed among left-behind families is not only a behavioural setback but also represents forgone economic value and environmental efficiency. The positive impact of return migration—via social remittances and urban exposure—suggests that targeted policy support can amplify such gains in rural sustainability efforts.

2.5.4 Channel Analysis

This section explores how macro-level indicators, including information infrastructure, social capital, and green infrastructure, interact with migration status to influence recycling willingness among left-behind families. The inclusion of interaction terms in the analysis provides nuanced insights into how external structural factors amplify or mitigate the effects of migration on environmental behaviour.

In this section, the interaction related to the difference between destination and origin is added to check the possible channel affecting the left-behind family:

$$\text{Recycling}_{i,j,c} = \alpha_{i,j,c} + \beta M_{i,j,c} + \theta \text{Channel}_{i,c} + \gamma M_{i,j,c} \times \text{Channel}_{i,c} + \delta X_{i,j} + \epsilon_{i,j,c}. \tag{2.9}$$

Where,

- The interaction term $M_{i,j,c} \times \text{Channel}_{i,c}$ captures the potential influence of the difference between destination and origin.
- $M_{i,j,c}$ represents the dummy variables of interest (CM and RM),
- $M_{i,j,c}$ is instrumented by the migration network as the same as Equations 2.2 and 2.3. Interaction $M_{i,j,c} \times \text{Channel}_{i,c}$ is also instrumented by $IV \times \text{Channel}_{i,c}$
- $\text{Channel}_{i,c}$ denotes the macro indicator differential between the destination and origin by the ratio of the capital city of the Hukou province to home cities. This is because CLDS does not provide exact migrants' destinations, and then the capital city of *hukou* provinces is used to approximate the destination. This substitution is justified by the fact that the individual survey in CLDS recorded that 75% of migrants work within the province, indicating that within-province rural-to-urban migration dominates the sample. Moreover, according to the Chinese Rural Labour Report (National Bureau of Statistics, 2021), 58.8% of rural labour migration occurs within provinces, with eastern provinces having over 80% internal migration. Thus, provincial capital cities, which represent the highest levels of economic and environmental infrastructure, serve as reasonable proxies for migration destinations. Table 2.6 shows the channel analysis.

Information Infrastructure

The interaction between current migration and the scale of telecommunications (Column 1) yields a significant positive coefficient (0.013, $p < 0.1$), suggesting that information infrastructure mitigates the disruptive effects of migration on recycling behaviour. Enhanced access to communication technologies helps left-behind families stay informed about urban recycling norms, reducing the negative impact of migrant absence. This aligns with the theoretical framework of *information asymmetry reduction* Stigler (1961), as the improved flow of environmental knowledge strengthens behavioural continuity in left-behind households.

For return migration, the interaction term with telecommunications infrastructure (0.015, Column 2) is positive but insignificant. While return migrants benefit from urban information networks, their influence on recycling behaviour appears more reliant on direct social engagement and reintegration upon returning to rural areas. Notably, this contrasts with findings on *social capital*, which suggest that indirect influences, such as role models and norm carriers, are more impactful.

Social Capital

The interaction between current migration and social capital, proxied by the number of college students (Column 3), is significant and positive (0.014, $p < 0.05$). This indicates that higher human capital buffers the disruptive effects of migration, as college students in destinations often act as role models or carriers of urban norms. This finding complements the *information infrastructure* results, as it reinforces the idea that indirect mechanisms—whether through communication technologies or social networks—help sustain green behaviours in left-behind families.

However, the interaction with return migration (0.013, $p < 0.1$, Column 4) is weakly negative and significant. A plausible explanation for this counterintuitive result lies in the *destination-origin mismatch hypothesis*, where return migrants struggle to reconcile advanced environmental norms from highly educated destinations with the less developed social and cultural settings of their rural origins. This mismatch may undermine the adoption of recycling behaviours, especially when returnees' expectations exceed the practical realities of their home communities. This tension parallels findings in *green infrastructure*, where similar mismatches exacerbate behavioural challenges.

Green Infrastructure

The interaction between current migration and disparities in sewage disposal rates in destinations (Column 5) is negative but insignificant (-0.727). This suggests that green infrastructure differences do not directly mediate the effects of current migration on recycling willingness. However, this result contrasts with findings on *social capital* and *information infrastructure*, where indirect influences help sustain behaviours despite structural disparities.

For return migration, the interaction (-0.802, $p < 0.01$, Column 6) is significant and negative. This result supports the hypothesis of a *crowding-out effect*, where exposure to advanced green infrastructure in destinations diminishes individual responsibility for recycling upon return to less developed origins. Interestingly, this finding complements the *social capital* results: while highly educated destinations provide a human capital buffer for current migrants, they may inadvertently undermine the recycling practices of return migrants when destination-origin disparities are stark. These results collectively highlight the nuanced role of structural and social factors in shaping recycling behaviour.

The findings across these subsections reveal both intersections and contradictions. Information infrastructure and social capital both mitigate the negative effects of migration, suggesting that indirect mechanisms—whether technological or human capital-based—help sustain recycling behaviours. However, both mechanisms falter when structural mismatches, such as disparities in green infrastructure or destination-origin social norms, come into play.

For return migration, these mismatches are particularly pronounced, as evidenced by the weakly negative effect of social capital and the significant negative interaction with green infrastructure. Together, these results suggest that while indirect influences (e.g., college students and telecommunications) buffer the effects of migration, their efficacy is contingent on structural alignment between destinations and origins. Policymakers should focus on bridging these gaps by enhancing rural green infrastructure and fostering the contextual transferability of urban norms.

In addition to recycling, the analysis also examines fixed garbage placement as a complementary measure of green behaviour (in Table 2.7). While the recycling-related channels show stronger and more consistent patterns, the results for fixed garbage placement offer additional insights. Notably, the mitigating role of information infrastructure is weaker in this domain, as shown by the insignificant or negative interaction effects. This suggests that fixed disposal behaviours—likely governed more by local enforcement or habitual compliance—may be less sensitive to informational or social cues than voluntary recycling.

However, green infrastructure disparities still exert influence: the interaction between return migration and sewage disposal rate remains significantly negative, indicating that returnees may feel disillusioned or less motivated to comply with local waste placement norms after exposure to superior systems in urban destinations. These findings imply that institutional and infrastructural quality play a critical role in sustaining pro-environmental practices beyond attitudinal willingness, especially for habitual or regulated behaviours such as waste sorting and placement.

Table 2.6: **Channel Analysis I**

Dependent Variable	Recycling					
Channels	Information Infrastructure		Social Capital		Green Infrastructure	
	(1)	(2)	(3)	(4)	(5)	(6)
Current Migration	-0.879***		-1.133***		0.063	
	(0.202)		(0.346)		(0.672)	
Return Migration		0.679		1.742*		1.763***
		(0.424)		(0.898)		(0.589)
Current Migration×Scale of Telecommunications	0.013*					
	(0.007)					
Return Migration×Scale of Telecommunications		0.014				
		(0.021)				
Current Migration× No. of College Students			0.013**			
			(0.006)			
Return Migration×No. of College Students				-0.012		
				(0.008)		
Current Migration× Sewage Disposal Rate					-0.727	
					(0.610)	
Return Migration×Sewage Disposal Rate						-0.802***
						(0.286)
Covariates Included	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Province	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year × Province	Yes	Yes	Yes	Yes	Yes	Yes
N	6589	4285	6455	4170	6589	4285

Note: This table represents the channel analysis based on Equation 2.9 and 2SLS. Channels are information infrastructure, social capital and green infrastructure are captured by the provincial capital city to origin city ratio of scale of telecommunications, the number of college students and sewage disposal rate.
Interaction terms are used to capture the joint effect of channels and are also instrumented by $IV \times \text{Channel}_{i,c}$.
Migration Status is instrumented by the current/return migration trend.
Household-level clustering standard errors are in the parenthesis. ***, **,* represent the level of significance at 1%, 5%, 10%.
Source: CLDS 2014, 216; China City Yearbook, Census 2005-2015.

Table 2.7: **Channel Analysis II**

Dependent Variable	Fixed Garbage Placement					
Channels	Information Infrastructure		Social Capital		Green Infrastructure	
	(1)	(2)	(3)	(4)	(5)	(6)
Current Migration	-0.308***		-0.382***		0.211	
	(0.090)		(0.148)		(0.416)	
Return Migration		0.443**		0.337		0.748**
		(0.211)		(0.431)		(0.372)
Current Migration×Scale of Telecommunications	-0.02					
	(0.004)					
Return Migration×Scale of Telecommunications		-0.002				
		(0.006)				
Current Migration× No. of College Students			0.001			
			(0.003)			
Return Migration×No. of College Students				0.002		
				(0.005)		
Current Migration× Sewage Disposal Rate					-0.408	
					(0.391)	
Return Migration×Sewage Disposal Rate						-0.210
						(0.179)
Covariates Included	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Province	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year × Province	Yes	Yes	Yes	Yes	Yes	Yes
N	6589	4285	6455	4170	6589	4285

Note: This table represents the channel analysis based on Equation 2.9 and 2SLS. Channels are information infrastructure, social capital and green infrastructure are captured by the provincial capital city to origin city ratio of scale of telecommunications, the number of college students and sewage disposal rate.
Interaction terms are used to capture the joint effect of channels and are also instrumented by $IV \times \text{Channel}_{i,c}$.
Migration Status is instrumented by the current/return migration trend.
Household-level clustering standard errors are in the parenthesis. ***, **,* represent the level of significance at 1%, 5%, 10%.
Source: CLDS 2014, 216; China City Yearbook, Census 2005-2015.

2.5.5 Heterogeneity Analysis

2.5.5.1 Heterogeneity Analysis in Recycling

Table 2.8 presents the heterogeneity analysis of current migration on the recycling willingness of left-behind families, with results disaggregated by gender, age, education, and income in Columns (1) to (9). Figure 2.2 shows all results of heterogeneity in recycling and fixed garbage placement.

In Columns (1) and (2), the results show a significant negative effect of current migration on male-led households, with a reduction of around 78.0% at a 1% significance level, but no significant impact on female-led households. This aligns with studies suggesting that men, often the primary economic decision-makers, prioritize income-generating activities over environmental considerations (Shen et al., 2021). In contrast, women may maintain household routines, including recycling, despite disruptions caused by migration.

For different age subgroups, young families (aged 16–40) experience the largest but insignificant negative impact of current migration on recycling willingness, possibly reflecting the economic pressures and social adjustments they face during a key life stage. On the other hand, younger families may also adapt faster to urban environmental norms when one family member migrates. Middle-aged families (40–60) in Column (4) exhibit significant negative effects, with a 1% significant 53.3% reduction in willingness, while older families (60–80) face the strongest negative effect, approximately double that of the younger group (Column 5). This result implies that stronger reliance on habits and weakened community ties may diminish recycling efforts due to migration (Frey and Meier, 2004).

For different education levels, the adverse effect of current migration is not significantly pronounced for families with low education, but for those with higher education, the effect is more significant (-53.3%, Column 7). This finding supports the hypothesis that education fosters environmental awareness, enabling families to sustain recycling behaviours despite migration-induced disruptions (Nikolova and Graham, 2015).

While migration significantly and negatively impacts the recycling willingness of left-behind families, high-income families face stronger negative effects (89.7%, Column 9) than low-income families (-57.0%, Column 8). This result suggests that high-income households experience greater lifestyle adjustments during migration, which may deprioritize recycling efforts. This aligns with the research of Moran-Taylor and Taylor (2010), linking income shocks to reduced pro-social behaviours, including environmental practices.

Similarly, Table 2.9 reveals heterogeneity analysis of return migration across the same demographic and socio-economic groups. The effect of return migration is positive and significant for male-led families but insignificant for female-led families. Male heads often have broader social networks and community influence, which can amplify the adoption of recycling practices. Return migrants in these families might leverage these networks to promote environmental norms or participate in community recycling initiatives (Taylor et al., 2011).

Return migration positively and significantly influences the recycling willingness of younger families (31.6%, Column 3) and middle-aged families (51.1%, Column 4). By contrast, the effect on older families is not significant, potentially reflecting the influence of long-standing habits and experiences.

In terms of education levels, return migration significantly enhances recycling willingness among higher-educated families (38.5%, Column 7), supporting the view that education facilitates the internalization of environmental norms. Compared to the insignificant effect on the low-educated group, the results imply barriers to the transmission of environmental behaviours in this group.

Additionally, compared to low-income families (Column 8), high-income families (Column 9) show a significant positive effect of return migration, with a 1% significant 89.1% increase in recycling willingness. This is consistent with the hypothesis that economic stability supports pro-environmental investments (Wu et al., 2021b).

These findings highlight the complex relationship between migration and environmental behaviour in left-behind families, emphasizing the importance of subgroup characteristics. Regardless of whether families experience current migration or return migration, factors such as female leadership, higher educational levels, younger family members, and better financial situations play a positive role in transforming norms and offsetting the negative impacts caused by a lack of family labour.

TABLE 2.8: **Heterogeneity Analysis of Current Migration**

Dependent Variable	Recycling								
Subgroup	Gender		Age			Education		Income	
	Male	Female	[16,40]	(40,60]	(60,80]	Low schooling	High schooling	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Current Migration	-0.780***	-0.564	-1.600	-0.541***	-1.004**	-0.937**	-0.533***	-0.570**	-0.891***
	(0.202)	(0.385)	(2.373)	(0.198)	(0.398)	(0.398)	(0.181)	(0.247)	(0.308)
Covariates Included	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Province	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year×Province	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	6008	1113	952	3762	2405	3697	3422	3782	3338

Note: This table reports the heterogeneity analysis of the current migration group by socioeconomic characteristics.
Household-level clustering standard errors are in the parenthesis. ***, **,* represent the level of significance at 1%, 5%, 10%.
Source: CLDS 2014, 2016; Census 2005-2015

TABLE 2.9: **Heterogeneity Analysis of Return Migration**

Dependent Variable	Recycling								
Subgroup	Gender		Age			Education		Income	
	Male	Female	[16,40]	(40,60]	(60,80]	Low schooling	High schooling	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
return	0.485*	4.000	0.316**	0.511*	-3.324	8.568	0.375**	2.383	0.434**
	(0.253)	(4.059)	(0.151)	(0.275)	(4.330)	(29.002)	(0.160)	(2.245)	(0.182)
Covariates Included	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Province	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year×Province	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	4005	676	920	2332	1425	2322	2356	2175	2505

Note: This table reports the heterogeneity analysis of the return migration group by socioeconomic characteristics.
Household-level clustering standard errors are in the parenthesis. ***, **,* represent the level of significance at 1%, 5%, 10%.
Source: CLDS 2014, 2016; Census 2005-2015

2.5.5.2 Heterogeneity Analysis in Fixed Garbage Placement

In addition to recycling, Tables 2.10 and 2.11 report heterogeneity analyses for fixed garbage placement—a more habitual and rule-based form of green behaviour. Several patterns emerge that contrast with the recycling results.

First, the negative impact of current migration is consistently observed across subgroups, particularly among female-headed households and middle-aged individuals (40–60), with stronger effects among the low-educated and low-income groups. This suggests that fixed waste placement is more vulnerable to institutional detachment or a loss of routine, especially in vulnerable households. Compared to recycling, this behaviour appears to rely more on community norms and enforcement, which can deteriorate in the absence of key household members.

Second, the effect of return migration on fixed garbage placement is generally positive and significant for middle-aged, high-educated, and high-income households, mirroring patterns found in recycling. However, for female-headed households and those with low education or income, the coefficients are large and unstable, sometimes exceeding the LPM's realistic bounds. This reflects both statistical noise and potential behavioural mismatch: for example, returnees may attempt to transfer urban practices that clash with rural realities, leading to heterogeneity in response depending on household adaptability.

Overall, the heterogeneity findings suggest that while return migration may encourage both recycling and fixed waste placement, its effectiveness is conditional on socioeconomic resilience and infrastructural compatibility. In contrast, the disruptive effect of current migration appears more widespread in fixed garbage placement, pointing to the importance of institutional continuity and household stability in sustaining habitual green practices.

Table 2.10: **Heterogeneity Analysis of Current Migration**

Dependent Variable	Fixed Garbage Placement								
Subgroup	Gender		Age			Education		Income	
	Male	Female	[16,40]	(40,60]	(60,80]	Low schooling	High schooling	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Current Migration	-0.199**	-0.376*	0.367	-0.340***	-0.228	-0.365**	-0.152***	-0.201**	-0.277***
	(0.081)	(0.216)	(0.650)	(0.098)	(0.172)	(0.181)	(0.080)	(0.131)	(0.119)
Covariates Included	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Province	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year×Province	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	6008	1113	952	3762	2405	3697	3422	3782	3338

Note: This table reports the heterogeneity analysis of the current migration group by socioeconomic characteristics. Household-level clustering standard errors are in the parenthesis. ***, **,* represent the level of significance at 1%, 5%, 10%.
Source: CLDS 2014, 2016; Census 2005-2015

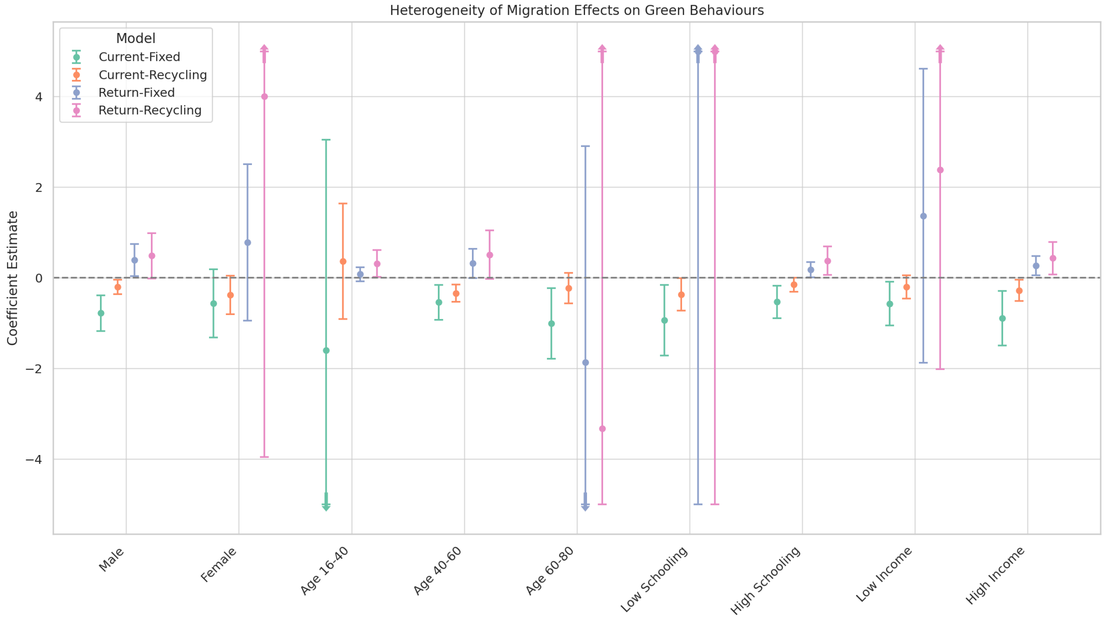


FIGURE 2.2: **Heterogeneity Analysis of Current Migration**

Note: This Figure shows the all heterogeneity analysis of current migration and return migration based on Tables 2.8, 2.9,2.10 and 2.11.

Source: CLDS; Census

Table 2.11: **Heterogeneity Analysis of Return Migration**

Dependent Variable	Fixed Garbage Placement								
Subgroup	Gender		Age			Education		Income	
	Male	Female	[16,40]	(40,60]	(60,80]	Low schooling	High schooling	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
return	0.391**	0.779	0.079	0.318*	-1.864	31.099	0.177**	1.369	0.269**
	(0.179)	(0.881)	(0.079)	(0.164)	(2.433)	(673.901)	(0.085)	(1.656)	(0.109)
Covariates Included	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Province	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year×Province	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	4005	676	920	2332	1425	2322	2356	2175	2505

Note: This table reports the heterogeneity analysis of the return migration group by socioeconomic characteristics.
Household-level clustering standard errors are in the parenthesis. ***, **,* represent the level of significance at 1%, 5%, 10%.
Source: CLDS 2014, 2016; Census 2005-2015

2.6 Robustness Check

2.6.1 Reduced Sample

The marginal effects reported in Table 2.13 are Average Partial Effects (APEs) from the IV-Probit model. These represent the average change in the probability of engaging in green behaviours—such as recycling or fixed garbage disposal—associated with changes in migration status, holding other covariates constant. APEs are expressed on the probability scale and are therefore directly interpretable.

The APEs are smaller in magnitude than the 2SLS coefficients in Table 2.4, which is expected given the differences between the model types. The Linear Probability Model (LPM) assumes constant marginal effects across the distribution and does not constrain predicted probabilities within [0,1], which can overstate effect sizes when the dependent variable has a high mean (here, 0.80 for recycling and 0.92 for fixed garbage disposal). In contrast, the IV-Probit model accounts for the bounded and non-linear nature of probabilities, producing more conservative—yet arguably more realistic—estimates.

The consistency in sign and statistical significance across LPM and IV-Probit results strengthens confidence in the robustness of the main findings.

Table 2.12: **Robustness Check - Reduced Sample**

Groups	Current Migration Group		Return Migration Group	
Dependent Variable	Recycling	Fixed Disposal	Recycling	Fixed Disposal
	(1)	(2)	(3)	(4)
Current Migration	-0.649***	-0.199**		
	(0.181)	(0.079)		
Return Migration			0.857***	0.489***
			(0.304)	(0.178)
Covariates included	Yes	Yes	Yes	Yes
Fixed Province	Yes	Yes	Yes	Yes
Fixed Year	Yes	Yes	Yes	Yes
K-P Wald F statistic	34.66	34.66	12.30	12.30
N	6253	6253	4121	4121

Note: This table reports the robustness check based on the reduced sample excluding the repeated households. All results are consistent with Table 2.4.
Household-level clustering standard errors are in the parentheses. ***, **,* represent the level of significance at 1%, 5%, 10%.
Source: CLDS 2014, 2016; Census 2005-2015

2.6.2 Alternative Specification

The reported marginal effects in Table 2.13 correspond to Average Partial Effects (APEs) derived from the IV-Probit model. These reflect the average change in the probability of engaging in green behaviours, such as recycling or using fixed garbage disposal—associated with changes in migration status, while holding other covariates constant. Compared to the latent-index coefficients from the IV-Probit estimation, APEs provide a more interpretable effect size on the actual probability scale.

It is worth noting that the APEs are smaller in magnitude than the 2SLS estimates reported in Table 2.4, which is expected due to differences in the underlying model structure. The Linear Probability Model (LPM) assumes a constant marginal effect across the entire distribution and does not impose bounds on the predicted probabilities. As a result, it may overstate effect sizes, especially when the dependent variable is already concentrated near the upper bound, as is the case here with recycling (mean = 0.80) and fixed garbage disposal (mean = 0.92). By contrast, the IV-Probit model accommodates the non-linear probability structure and the bounded nature of the outcome, thereby yielding more conservative, yet arguably more realistic, marginal effects.

Together, the consistency in the sign and significance across LPM and IV-Probit models provides additional confidence in the robustness of the main findings.

TABLE 2.13: **Heterogeneity Analysis of Return Migration**

Groups	Current Migration Group		Return Migration Group	
Dependent Variable	Recycling	Fixed Disposal	Recycling	Fixed Disposal
	(1)	(2)	(3)	(4)
Current Migration	-1.707***	-2.052***		
	(0.240)	(0.167)		
Marginal Effect	-0.222	-0.137		
Return Migration			2.052***	2.157***
			(0.318)	(0.360)
Marginal Effect			2.300	2.917
Covariates Included	Yes	Yes	Yes	Yes
Fixed Province	Yes	Yes	Yes	Yes
Fixed Year	Yes	Yes	Yes	Yes
Fixed Year×Province	Yes	Yes	Yes	Yes
N	7122	7122	4681	4681

Note: This table reports the robustness check for the specification by alternative methods, IV-Probit. All results are still consistent with Table 2.4.
Household-level clustering standard errors are in the parentheses. ***, **,* represent the level of significance at 1%, 5%, 10%.
Source: CLDS 2014, 2016; Census 2005-2015

2.7 Conclusion

This paper investigates the impact of migration on the green behaviour of left-behind families in China, focusing on internal rural-to-urban migration. Using two indices — willingness to recycle and fixed-place garbage disposal practices—the study examines the distinct effects of current migration and return migration on green behaviours. It incorporates selection bias adjustments, heterogeneity analysis, and interaction terms to explore mechanisms and socio-economic gaps between urban and rural areas.

The findings reveal significant differences in the effects of migration types. Current migration negatively impacts green behaviour by 71.9% and 21.8% reduction in recycling and fixed trash placement probability, as the absence of household members disrupts daily routines and weakens environmental practices. In contrast, return migration has a positive effect, around 79.0% and 46.3% increase in both, as returnees bring urban-acquired norms and practices back to their families.

To uncover the underlying mechanisms, this study examines the interaction between migration and the socio-economic gap between origin and destination cities. The results show that better green infrastructure, information networks, and higher social capital in destination cities amplify the positive effects of migration on recycling willingness. Migration not only provides financial remittances but also facilitates social remittance channels, enabling the transfer of urban environmental norms and practices to rural families. However, these benefits are likely to vary depending on the quality of infrastructure and socio-economic conditions in destination cities, which highlights the uneven potential for green behaviour transformation.

Heterogeneity analysis reveals that younger, male-headed families with higher educational levels and sustainable financial situations are better positioned to benefit from the positive effects of return migration. Enhanced family capital, such as higher education and financial stability, often plays a pivotal role in adopting and sustaining green practices. Conversely, families with older members are more vulnerable to the negative effects of current migration due to challenges in adapting to disrupted routines and responsibilities.

This paper contributes to the literature by being the first to explore the effects of migration on green behaviour in left-behind families in China, differentiating between current and return migration. It also pioneers the application of social remittance theory to the context of China's internal migration, differing from studies on international migration. Additionally, this study uniquely addresses the selection bias in return migration, providing robust insights into how migration influences rural sustainability. These findings are critical for informing policies aimed at bridging the widening urban-rural divide and promoting sustainable development.

The findings have several practical implications. First, strengthening green infrastructure in rural areas could magnify the positive effects of return migration by making it easier for left-behind families to adopt environmentally friendly practices. Second, targeted support for older families impacted by current migration, such as community waste management services or household support programsc— could mitigate the negative effects of migrant absence. Third, education and training programs for migrants could include modules on green behaviours to strengthen the transfer of environmental awareness through social remittance channels.

This study faces some limitations. The absence of detailed data on migrants' destinations restricts the analysis of cross-province migration, potentially underestimating the broader impacts of migration. While within-province migration accounts for 75% of cases, cross-province migration often involves larger socio-economic differences and could provide richer insights into the transformative effects of urban experiences. Additionally, the reliance on existing longitudinal datasets constrains the range of outcome variables, as green behaviour data is rarely collected alongside migration surveys. Future research should integrate more comprehensive migration destination data and expand the scope of environmental behaviours studied to provide a fuller picture of migration's impact on sustainability.

By addressing these challenges, this paper lays a solid foundation for future research and offers actionable insights into how migration influences rural environmental practices. It contributes to the broader discourse on sustainable development, social remittances, and rural modernisation in China.

Chapter 3

The Role of Air Pollution in Migration Decision: Evidence from China

Abstract

This paper examines the effect of air pollution on urban-to-urban migration flows in China. I utilize comprehensive and unique data, including the destination-to-origin ratio of $PM_{2.5}$ concentrations, city-level bidirectional urban migration flows, and city-level macro data, using a gravity model to analyse the impact of air pollution on migrants' destination choices. After addressing the endogeneity of air pollution using cumulative thermal inversion days as an exogenous instrument, my findings suggest that the attractiveness of destinations with higher comparative air pollution is significantly reduced for migrants from less polluted origins. Specifically, each doubling of the relative air pollution concentration between destination and origin cities reduces migration inflows by approximately 42%. Additionally, I examine the roles of green infrastructure, regulatory policies, welfare policies, and settlement costs as mechanisms influencing the impact of pollution on migration decisions. I also find that the effect of air pollution on migration flows initially intensifies and then diminishes as migration distance increases. Moreover, there is an asymmetric response to pollution across demographic groups, with older, middle-educated, married male migrants being more vulnerable to the effects of air pollution on migration decisions.

Keywords: Air Pollution, Urban-to-Urban Migration, Gravity Model, Environmental Economics

3.1 Introduction

3.1.1 Pollution and China

Industrialisation not only encourages the development of many developing countries but also causes serious pollution in some developing countries, increasing the difficulty of sustainable development. Meanwhile, the gradually increasing environmental contamination, like air pollution, threatens public health (Lelieveld et al., 2015). According to the WHO (2020), air pollution is responsible for nearly seven million premature deaths, with over four million cases of early death directly caused by ambient air pollution. Moreover, low- and middle-income countries face a higher level of exposure, particularly in Asia and Africa. In the face of air pollution, some short-term reactions, like particulate-filtering facemasks, pharmaceutical purchases, and medical interventions, are applied to mitigate the negative effects of air pollution (Deschenes et al., 2017; Zhang and Mu, 2018). However, in the long run, urban dwellers may choose to migrate, seeking a better residential environment (Banzhaf and Walsh, 2008). Therefore, does air pollution matter in migration? How does air pollution affect decision-making? This paper aims to study the effect of air pollution on urban-to-urban migration decisions using micro-survey aggregate data and macroeconomic data.

Compared to past studies on environmental migration caused by natural disasters, recent research has constructed a framework for understanding the pollution effect on migrants' behaviour and has verified the negative causal effect of pollution on migration in terms of health, welfare, and education (Chen et al., 2022; Deschenes et al., 2020; Shao et al., 2021; Zhang et al., 2022). However, studies on the effect of air pollution on migration decision-making are still limited due to limited access to reliable information for residents. While the role of information in decision-making has been emphasised for many years (Stigler, 1961), pollution information is often intentionally distorted, purposely kept secret, or not collected (Barwick et al., 2024). Hence, obtaining air pollution information is particularly difficult in developing countries, although many are suffering from air pollution. By 2020, among the 20 countries with the worst air quality, only four countries (Nepal, Saudi Arabia, India, and China) had installed pollution monitoring systems. Therefore, China provides an interesting case with abundant data to analyse the effect of air pollution on urban migrations. Particularly, the launching of the 2013 nationwide program of information disclosure on air pollution provided reliable access to air pollution information for the public.

In China, after decades of high-speed development, outdoor air quality has gradually and noticeably deteriorated, raising concerns for public health and economic development (Ebenstein et al., 2015). According to Forouzanfar et al. (2016), air pollution has been the fifth leading risk factor for all-cause mortality in China, causing 1.1 million early deaths in 2015. In response, the government launched a program in 2012 to undertake real-time monitoring and automated reporting of air pollution in three phases. From December 31, 2012, the first wave covered 74 major cities from the Economic Delta, province-level cities, and national environmental protection exemplary cities. In the second wave, 116 cities were added by the end of 2013, and the remaining 117 cities were included by October 2014 in the third wave. This initiative was eventually implemented across all cities, covering over 98% of the population in China by the end of 2015. The nationwide coverage and progressive roll-out provide a specific and ideal case to study the effect of air pollution on labour migration and economic development.

Based on this pollution information disclosure, labour migration could be concerned about the potential risk of air pollution at individual/household levels, and studies could check the role of air pollution in an array of health, productivity, economic, and financial behaviour (Deschenes et al., 2020; Dupuy, 2021; Heyes et al., 2016; Khanna et al., 2021).

3.1.2 Migration Mode in China

China provides an interesting case of migration analysis in terms of unique types of migration - rural-to-urban and urban-to-urban migration.

In the Chinese context, urban-to-urban and rural-to-urban migration differ markedly in institutional constraints, migration motivations, and human capital profiles. Rural-to-urban migrants—typically holding agricultural hukou—face limited access to urban welfare services and tend to occupy low-skilled, labour-intensive positions (Afridi et al., 2015). Their migration is largely driven by income differentials and surplus rural labour. In contrast, urban-to-urban migrants usually possess non-agricultural hukou, higher educational attainment, and greater integration capacity in urban systems. Their decisions are more selective and responsive to non-economic factors, including environmental quality, healthcare, and children's education opportunities.

It is important to clarify that the terms "rural" and "urban" in the Chinese context do not simply denote differences in geography or population density, as often assumed in Western studies. Instead, they reflect a deep-rooted institutional and administrative hierarchy shaped by China's hukou system. In particular, "rural" refers to lower-tier administrative areas—often underdeveloped, infrastructure-poor, and agriculturally based—where residents typically hold agricultural hukou. Migration from these areas to urban centres involves not only geographic mobility but also attempts to overcome systemic barriers to social and economic inclusion.

In contrast, urban-to-urban migration in China generally occurs between prefecture-level cities or higher-tier urban areas, and migrants involved in such flows often possess non-agricultural hukou, higher education, and better access to information (Cheng et al., 2014). This form of migration reflects a qualitatively different decision-making process, less constrained by institutional exclusion and more sensitive to environmental amenities and life quality.

Focusing on urban-to-urban migration serves two analytical purposes. First, it isolates the role of environmental factors—particularly air pollution—in shaping migration decisions among individuals with greater autonomy, information access, and selectivity. Second, it avoids conflating environmentally motivated migration with structural poverty-driven movements that dominate the rural-to-urban pattern. In the Chinese context, urban-to-urban migrants often represent skilled labour and human capital flows between similarly ranked administrative units (e.g., prefecture-level cities) but with differing environmental and economic endowments. These individuals are more likely to engage in strategic relocation based on quality-of-life considerations, making them more sensitive to marginal changes in air quality.

While the findings of this study may not be directly generalisable to rural-to-urban migration — where institutional constraints and economic necessity often dominate decision-making, existing research suggests that rural migrants also respond to environmental degradation, particularly when it affects health or agricultural viability (Yue et al., 2024). However, the strength and nature of this response tend to differ. Compared to the constrained and reactive nature of rural-to-urban flows, urban-to-urban migration provides a clearer lens into the marginal valuation of environmental quality in residential choice and sheds light on how air pollution affects the spatial allocation of human capital across cities.

In this paper, aiming to study the air quality effect on migration, the 2SLS and PPML models induced from the gravity model are constructed and controlled for the reverse causality by the instrumental variable. This paper found each double comparative air pollution concentration between destination and origin cities could reduce around 42% migration inflow at the 1% level of significance. On the other hand, this paper tries to analyse potential mechanisms between air pollution and migration decision-making through the physical distance, infrastructure, policy, and hedonic value between host and home cities and the education of migrants. The results found:

1. A U-shape relationship between distance and decision-making is affected, i.e., the negative effect of air pollution could be attenuated for longer than 1000km migration.

2. Positive green or welfare policies could offset the partial negative effect of air pollution, but a higher level of hedonic price in destinations could push potential migrants to give up their choice of destinations.

3. The compensating wage and job market made higher-educated groups less sensitive as low-skilled migrants to air pollution.

There are four key contributions of this paper.

First, it focuses on the pollution differential between destinations and origins as a proxy for the information gap perceived by migrants. Prior studies mostly examine net inflows or outflows at either origin or destination only, overlooking the comparative nature of migration decisions (Fu et al., 2021; Zhang et al., 2022). By constructing a destination-to-origin $PM_{2.5}$ ratio, this paper captures the relative pollution exposure that informs migrants' choices, accounting for both push and pull factors.

Second, it is among the first to analyse domestic bidirectional migration flows in relation to pollution at the macro level. While most existing studies focus on individual migration intentions (Cai and Wang, 2007; Liu et al., 2018; Thissen et al., 2010), and macro-level analyses tend to focus on international migration (Di Iasio and Wahba, 2024; Forouzanfar et al., 2016; Kaczan and Orgill-Meyer, 2020; Manning and Taylor, 2014; Piguet et al., 2011), this study uses aggregated city-to-city flows from the 2017 China Migrants Dynamic Survey (CMDS) to assess internal urban-to-urban movements, where pollution information is more salient and accessible.

Third, this paper goes beyond information disclosure mechanisms by exploring macro-level channels, including welfare policies, environmental regulations, and green infrastructure. These institutional and structural factors may condition the migration response to pollution and are often neglected in micro-level studies.

Finally, this is the first study to focus specifically on urban-to-urban migration affected by air pollution. While rural-to-urban migration has received considerable attention (Bierkamp et al., 2021; Shao et al., 2021), urban-to-urban flows involve more educated, skilled workers and play a vital role in human capital allocation. This paper highlights how environmental quality shapes the mobility of skilled labour within the urban system, offering new insights into spatial sorting in developing economies.

The structure of this paper is as follows: Section 3.2 is a literature review on migration and pollution based on the hedonic model and mechanism discussion. Sections 3.3 and 3.4 provide the explanation of data and methodology, particularly for the measurement of variables and empirical strategy. Section 3.5 shows the basic empirical results. Section 3.6 discusses the heterogeneity of the sample, and Section 3.7 concludes the paper.

3.2 Literature Review

3.2.1 Modern Migration and Pollution Effect

Migration can be perceived as costly in physical and psychological aspects. Regarding physical and economic motivation, migrants need to consider the probability of accessibility to a higher-quality and sustainable life at their destinations. On the other hand, long-distance migration, new community integration, and environment adaptation also increase the cost of mental health (Bayer et al., 2009; Luechinger, 2010; Piguet et al., 2011). The potential risk to health caused by pollution is also seen as a cost of migration, which affects both physical and psychic sides (Chen et al., 2013a; Manning and Taylor, 2014).

Whereas, what is the role of air pollution in modern migration? Past literature has studied both micro and macro levels and verified the negative effect of pollution on domestic or international migrants' well-being, staying willingness, and migration intentions (Cao et al., 2015; Zhang et al., 2022; Zhao et al., 2021). Moreover, Liu and Yu (2020) find that an increase in air pollution concentration has short-term positive production effects but long-term negative influences on personal income. Liu et al. (2020) also find that indoor air pollution from cooking reduces the capacity to deal with daily life and work. Similarly, some studies note the significant negative effect of air pollution on individuals' health, such as obesity, labour supply, and productivity (Borgschulte et al., 2022; Deschenes et al., 2020; Fu et al., 2021).

For studies on different groups of migrants, Sun et al. (2019) focus on urban pollution and low-income migrants, finding that urban pollution drives migrants to leave their current cities. Lai et al. (2021) study the migration response of high-talent individuals to air pollution. They discuss the attractiveness of the environment to graduates and find that $PM_{2.5}$ concentration pushes graduates away from their college cities, with a stronger effect on those from elite universities. This paper also examines the effect of air pollution on different migrant groups.

As for the potential mechanism of air pollution's effect on migration decisions, past literature found some determinants of air pollution affecting migration decision-making. The first factor is physical health concerns. Chen et al. (2022) examine migration decisions affected by physical health and sensitivity to air pollution based on individuals' subjective health evaluations. On the other hand, satisfaction with life and work also affects migration decisions. Wu et al. (2021a) found that pollution lowers residents' subjective well-being (SWB), while Li et al. (2014) found that pollution increases the risk of depressive symptoms during long-term urban work. Such declines in life or work satisfaction could cause migrants to leave (Luechinger, 2010). However, most literature focuses on micro-level individual reactions to pollution or economic factors, with few papers discussing the effect of macro public green information on decision-making, such as green infrastructure, environmental regulations, and migration policy. This paper also contributes to this point.

3.2.2 Environment Migration and Hedonic Model

Rosen (1979) and Roback (1982) introduced the Rosen-Roback hedonic model to measure the values of non-market amenities. In this model, it is assumed that with accurate information on the whole space, migration decisions reflect the trade-off between income, housing cost, and amenities. The hedonic price, measured by the geographical differential in income, housing price, or life quality, is often used to reflect the implicit value of recovering amenities, such as residential environment, green life, and air quality. In other words, the hedonic price can be seen as the cost of migration, which directly influences decision-making and can estimate individuals' marginal willingness to pay (MWTP) for clean and high-quality air and environment based on observed migration decisions. For instance, Chay and Greenstone (2005) applied the hedonic model to address aggregation problems and reflect pollution costs known to individuals. Similarly, Bayer et al. (2009) found that high-skilled migration responses to air pollution are more sensitive and have a higher MWTP. This allows the paper to use migration flow across regions to measure the effect of air pollution and interpret why migrants could be affected by information differentials in air pollution between city pairs.

Noticeably, the underlying assumption of the hedonic model is that "perfect information" and "imperfect information" on pollution could distort migration decisions. Furthermore, given the "imperfect information," relocation decisions are often distorted by regulations, agents, and market failure in developing countries. This phenomenon is particularly salient in China, where the residents' registration system, so-called "hukou," strongly restricts labor mobility and increases the threshold of migration costs (Kinnan et al., 2018). At the same time, the discretion of government officials and corruption in land transactions also distort the housing market and affect migration decisions through the hedonic path (Chen and Kung, 2019). Hence, the information asymmetry of pollution between the government and the public biases migration decisions.

For "imperfect" information, Bayer et al. (2009) developed an alternative discrete-choice approach to overcome distortions and bias caused by high moving costs. Additionally, Gao et al. (2023) augmented the hedonic model with imperfect information and found that air-pollution information disclosure induces urban residents to leave. Their work relaxes the initial assumption and helps this paper construct the framework to study the function of this "imperfect" information in migration decision-making.

3.3 Data

3.3.1 Bidirectional Migration Flow

To measure bidirectional migration flows within a gravity model framework, this paper uses data from the 2017 China Migrants Dynamic Survey (CMDS), conducted by the National Health Commission. The survey covers all 31 provinces and provides detailed city-to-city migration histories, allowing for the aggregation of flows between origin and destination cities. A key advantage of the 2017 wave is its rich detail on individual migration trajectories, enabling the identification of urban-to-urban migrants who have resided in the destination for over two years. This temporal filter ensures that the analysis focuses on long-term migrants, thereby excluding short-term or temporary movers whose decisions may be less sensitive to persistent environmental conditions. Moreover, the CMDS dataset features a large and nationally representative sample — covering approximately 20 million individuals — which enhances the precision and external validity of the empirical estimates.

3.3.2 Air Pollution

This paper uses $PM_{2.5}$ to proxy for air pollution level. Firstly, $PM_{2.5}$ is the major source of air pollution and is more typically and easily observed. According to China Environmental State Bulletin (2015), $PM_{2.5}$ accounted for over 60%-80% of annual days, by contrast, other indexes, like SO_2, CO and NO_2 are relatively low in China and their proportions of exceeding the standard are less than 1.5% of annual days. Secondly, $PM_{2.5}$ more seriously harms the public health. Compared to PM10, the particle size of $PM_{2.5}$ is much smaller, and it can easily penetrate the fine bronchial and alveoli, causing more serious illness and health issues. Particularly, even in the short-run exposure, high-density $PM_{2.5}$ could cause dyspnoea. In the long run, exposure to $PM_{2.5}$ disorders of the respiratory system and even lead to lung cancer and significantly increase the probability of death (Tanaka, 2015; Zhang et al., 2018). Finally, $PM_{2.5}$ comprehensively contains fossil fuel combustion and further chemical reactants in the air. Hence, this paper utilises the city-level $PM_{2.5}$ concentration to measure air pollution levels in each city.

Panel A of Figure 2 shows the national level of $PM_{2.5}$ concentration, and it kept a relatively stable trend between 2013 and 2017, ensuring no extra exogenous pollution shock. Considering that migrants need a longer time to respond to air pollution information, and short-run pollution information could be mitigated by the high cost of migration, this paper calculated a five-year average $PM_{2.5}$ density (2013-2017) for 278 cities to cover the migrant flow in 2017. On the other hand, aiming to capture pollution information differentials in migration decision-making, the destination-to-origin ratio of the five-year-average $PM_{2.5}$ concentrate is used as the interest variable, i.e., this ratio is calculated by $PM_{2.5}$ concentrate in the destination divided by that in the origin.

The data on $PM_{2.5}$ concentrate comes from Global Annual $PM_{2.5}$ Grids derived from satellite data by Van Donkelaar et al. (2016). Ground-level $PM_{2.5}$ is estimated by combining aerosol optical depth (AOD) retrievals from NASA's MODIS, MISR, and SeaWiFS. Subsequently, use geographically weighted regression (GWR) to calibrate to global ground-based $PM_{2.5}$. With the 0.01-degree resolution of the original raster, and a city-level grid is formed. Furthermore, to capture the differential of air pollution between origins and destination and then verify the push or pull effect, the ratio of destination-origin five-year-average $PM_{2.5}$ density is computed to measure the pull or push power (the ratio is calculated by destinations' $PM_{2.5}$ divided by corresponding origins' $PM_{2.5}$). Compared to the difference in $PM_{2.5}$ density between destination and origin cities, this ratio could avoid the negative-value interest variable and ease interpretation. If the ratio is less than 1 means the pollution is higher in origin, and corresponding migrants are searching for cleaner citie,s and the pollution is pushing migrants to leave. In terms of econometric meaning, each additional unit increase in this ratio means the destination pollution level is double that origin.

3.3.3 City Characteristics

Since this paper is derived from the gravity model (Equation 1) and focuses on bidirectional migration, all the symmetric characteristics of destination and origin cities are involved. Urban economic development is a crucial factor influencing migration (Kennan and Walker, 2011), thus the logarithm of GDP per capita, logarithm of urban wage, and the employment rate are used as proxies. Also, the urban population size affects migration choice (Zhang and Shunfeng, 2003). Considering the complex and imbalanced geographical conditions, the population density is used to replace the total population to proxy for the adsorption effect of population agglomeration. Moreover, the employment market, welfare and public services are correlated with migration selection (Kennan and Walker, 2010), thus the share of the second and third industries in urban GDP is used to proxy for the industrial structure. Furthermore, the spatial pattern of new employment is influenced by investments and infrastructure. Accordingly, the number of schools, hospitals, and urban water supply capacity are used. Finally, since the complex geographical and climate conditions also affect the settlement intention for individuals (Bardsley and Hugo, 2010), the annual rainfall, sunshine duration, humidity, and temperature are added to the model.

All these covariates of city-level characteristics are from the China City Statistic Yearbook conducted by the National Bureau of Statistics. The city yearbook is an abundant statistical dataset covering 294 cities and involves the population, economy, geographical conditions, industrial structure, and infrastructure conditions. The Climate conditions are from the China Meteorological Administration. Finally, after cleaning and dropping the missing observation, the dataset includes 288 cities and 25718 city-pairs.

Table 3.1 shows the variable definition as below, and Table 3.2 presents the descriptive statistics for the key variables used in the migration flow analysis. The migration flow dataset includes 25,718 city-pair observations, while city-level characteristics are based on 288 unique cities.

The average migration flow is 7.17 persons per city-pair, with a substantial standard deviation (50.47) and a maximum flow of 6,577, suggesting a highly skewed distribution. The log transformation (mean = 1.00, SD = 1.15) helps mitigate this skewness and facilitates regression analysis.

The mean destination-to-origin $PM_{2.5}$ ratio is 1.12, indicating that, on average, destination cities are slightly more polluted than origin cities. However, the wide range (from 0.068 to 14.73) reflects significant variation in pollution exposure across migration links, which underpins the identification strategy in this chapter.

City characteristics also display considerable heterogeneity. For instance, the second industry share averages 46.6%, highlighting the continued industrial orientation of many Chinese cities, while third industry shares remain lower (41.0%), suggesting uneven economic transformation. Urban unemployment rates are moderate (mean 3.07%), but city infrastructure indicators vary widely—for example, the number of hospitals ranges from 9 to over 1,500.

Weather variables, which serve as controls in the migration model, also vary significantly: the average temperature ranges from -0.35°C to 25°C, and rainfall and humidity differ substantially across regions. These descriptive patterns highlight the substantial heterogeneity across cities and migration links, reinforcing the need for a differential gravity model that accounts for relative conditions between origin and destination cities.

TABLE 3.1: **Description of Variables**

Variable	Description
Dependent variable	
ln Migration flow	Logarithm of the number of migrants from origin city to destination city
Interest variable	
Ratio of Destination-to-Origin $PM_{2.5}$	Ratio of Fiver-year-average $PM_{2.5}$ density of destination to origins
Control variables for origins and destinations	
- Economy level	
Ln GDP per capita	Logarithm of GDP per capita
Second industry share	The share of the second industry in GDP (%)
Third industry share	The share of the third industry in GDP (%)
Unemployment rate	Urban unemployment rate (%)
Population density	The differential in population density ($10^3/km^2$)
Ln Housing price	Logarithm of Housing prices (ě$/m^3$)
Ln Wage	Logarithm of Urban Average wage
- Infrastructure level	
Number of hospitals	Number of hospitals
Number of high schools	Number of high schools
Wastewater Disposal	Annual Wastewater Disposal (t)
- Geographical level	
Distance	The distance from origin to destination city ($10^3 km$)
Average temperature	Annual average temperature ($°C$)
Average sunshine duration	Annual average sunshine duration (h)
Average rainfall	Annual average rainfall (ml)
Average humidity	Annual average humidity ($\%rh$)

Note: This table reports all variables applied to check the pull-push factors affecting migration decisions.
Source: CMDS 2017, NASA EOSDIS 2016, China City Yearbook 2016, and China Meteorological Administration.

Table 3.2: **Descriptive Statistics**

Variable	N	Mean	Std. dev.	Min	Max
Bidirectional Variable					
Migration flow	25,718	7.172	50.47	0	6577
Ln Migration flow	25,718	0.999	1.151	0	8.791
Destination-to-Origin $PM_{2.5}$ Ratio	25,718	1.116	0.575	0.0679	14.73
Distance	25,718	9.950	4.609	9.49e-08	19.96
City Characteristics					
Population Density	288	0.43	0.339	.006	2.5
Ln GDP per capita	288	10.70	0.528	9.3	12
The Second Industry Share	288	46.62	9.570	15	71
The Third Industry Share	288	41.03	8.761	24	80
Ln Urban wage	288	3.96	0.231	1.6	4.7
Ln Housing Price	288	1.61	0.412	0.92	3.8
Urban Unemployment Rate	288	3.07	0.750	0.9	4.7
No. of hospitals	288	217.80	170.078	9	1568
No. of High School	288	209.24	134.220	5	1167
Wastewater Disposal	288	87.13	10.810	36	100
Weather Characteristics					
Average Temperature	288	14.80	5.199	-0.35	25
Average Sunshine Duration	288	5.35	1.342	2.9	8.8
Average Rainfall	288	3.32	1.862	0.23	7.5
Average humidity	288	70.82	10.040	41	84

Note: This table reports the descriptive statistics.The Destination-to-Origin $PM_{2.5}$ Ratio is obtained by the ratio of destination cities' five-year-average $PM_{2.5}$ density to origin cities' five-year-average $PM_{2.5}$ density. $PM_{2.5}$ data is obtained by ArcGIS. All Cities' Characteristics lagged for one year.
Source: CMDS 2017, NASA EOSDIS 2013-2017, China City Yearbook 2016, and China Meteorological Administration.

3.4 Methodology

3.4.1 Empirical Specification

The gravity model is commonly applied in studies on international trade and migration to analyse the effects of distance and key regional factors on flows between origin and destination areas. By incorporating characteristics from both regions, the model effectively examines potential "push" and "pull" factors. Vanderkamp (1977) reviewed the gravity model's application to migration decisions, addressing challenges in interpreting migration behaviour and demonstrating its utility in understanding migration dynamics.

This study builds on the gravity model framework to analyse how pollution differentials influence migration decisions, specifically examining the impact of pollution as a "push" factor in origin cities and a "pull" factor in destination cities:

$$M_{o,d} = k\frac{P_o^{\alpha} P_d^{\gamma}}{\text{distance}_{o,d}^{\beta}} \tag{3.1}$$

where:

- $M_{o,d}$ denotes the migration flow from origin city o to destination city d,
- P represents population size,
- $\text{distance}_{o,d}$ indicates the distance between the two cities,
- k, α, and β are parameters representing the influence of each factor.

From Equation (3.1), a linear form of the model can be derived (see the detailed process in Appendix A.1):

$$\begin{aligned} \ln(M_{o,d}) = \beta_0 + \beta_1 \frac{\text{PM}_{2.5d}}{\text{PM}_{2.5o}} + \beta_2 \ln(\text{distance}_{o,d}) \\ + \beta_3 \ln(P_o) + \beta_4 \ln(P_d) + X_d' \beta_5 + X_o' \beta_6 + \epsilon_{o,d} \end{aligned} \tag{3.2}$$

where:

- $\ln(M_{o,d})$ represents the logarithm of the migration flow from origin to destination, and to account for zero migration flows, $\ln(M_{o,d} + 1)$ is used,
- P denotes population density, and X is a matrix of city-level covariates, such as economic, infrastructure, and geographical characteristics.

The primary explanatory variable, $PM_{2.5}$, is the five-year average $PM_{2.5}$ concentration (2013–2017), serving as a proxy for air pollution levels. This study uses the destination-to-origin $PM_{2.5}$ ratio to capture pollution differentials, representing pollution as a “push” or “pull” force in migration decisions.

Given the focus on urban-to-urban migration and the fact that all migrants in the sample have already resided in their destination cities for more than two years, short-term exposure measures, such as a 12-month air pollution window, are not considered appropriate. Furthermore, in the Chinese context, the high institutional and financial costs of migration, stemming from the hukou system and related barriers to full settlement, imply that relocation decisions are rarely made in response to short-term environmental fluctuations. Rather, they are shaped by more sustained environmental conditions. The use of a 5-year exposure window thus better captures the medium-run air quality differential that households likely consider when making long-term relocation decisions. Moreover, the lagged nature of migration decisions, often tied to job mobility, school enrollment, or property investment, supports this longer horizon. While short-term effects are not the focus here, they are explicitly examined in the following chapter (see Chapter 4.4). There, I conduct a complementary analysis of short-term air pollution shocks and find that although their influence is statistically significant, the magnitude of the effect is limited, underscoring the importance of long-run exposure in shaping urban-to-urban migration flows.

Standard errors in Equation (3.2) are clustered at the destination city level to account for potential within-destination correlation in unobserved determinants of migration flows. Since city-level policies, environmental conditions, and unobserved amenities may affect all inflows to the same destination in a similar way, ignoring such intra-destination correlation could lead to underestimated standard errors and inflated significance levels. This clustering choice is particularly relevant when the outcome variable—the logarithm of migration flow—is observed across multiple origins for each destination city. Clustering at the destination level addresses heteroscedasticity and serial correlation across these grouped observations.

This approach follows the logic of common practice in gravity models and migration regressions, where one of the two nodes (either origin or destination) is selected for clustering to reflect the grouping structure in the error term (Santos Silva and Tenreyro, 2006). As this study focuses on the environmental and policy determinants of attractiveness at the destination, clustering by destination is the most theoretically consistent and empirically conservative strategy.

3.4.2 Identification Strategy

This study focuses on identifying the causal effect of pollution on migration flows, addressing several potential identification challenges:

1. **Omitted Variable Bias:** The gravity model includes several covariates to capture essential characteristics of origin and destination cities; however, the air pollution variable may still be endogenous. Factors like economic activities and industrial structures could influence both pollution levels and migration attractiveness. Including these city-level characteristics in the model helps reduce, but does not entirely eliminate, potential omitted variable bias.

2. **Reverse Causality:** Migration itself may influence local pollution levels, especially in cities experiencing large inflows of migrants. Urban migration increases pressure on infrastructure and the environment, potentially exacerbating pollution in destination cities. Additionally, developed cities with higher income and job opportunities often experience higher pollution due to economic activities, creating a potential feedback loop between migration and pollution. This study mitigates reverse causality by using province-pair fixed effects, which control for unobserved regional factors such as shared cultural, historical, and economic ties. Province-level fixed effects are also more appropriate than city-pair effects, as they avoid the risk of overfitting and noise absorption that might arise from city-specific effects. Within-province consistency in policy and socio-economic factors further supports the use of province-pair fixed effects. Additionally, migration restrictions under the *hukou* system primarily apply across provinces, adding a practical justification for using province pairs.

3. **Measurement Error:** While $PM_{2.5}$ is used as a proxy for general pollution exposure, actual individual exposure varies significantly due to factors like personal habits, indoor ventilation, and work environments (EPA, 2009). Furthermore, daily $PM_{2.5}$ values average across the day, and nighttime levels are typically lower due to reduced human activity. This averaging may underestimate actual exposure, particularly during peak activity hours.

To address endogeneity issues, this study uses thermal inversion as an instrumental variable, a meteorological phenomenon often applied in studies of air pollution's impact on health and productivity (Deschenes et al., 2020; Van Donkelaar et al., 2016; Arceo et al., 2016). Thermal inversion occurs when a layer of warm air traps pollutants near the ground, preventing them from dispersing and exacerbating pollution levels. Chen et al. (2022) used five-year average thermal inversion data as an instrument for air pollution in migration studies.

Since the main explanatory variable here is the five-year destination-to-origin $PM_{2.5}$ ratio, this study computes the five-year average of accumulated thermal inversion days to align with the pollution variable. The calculation process is as follows:

- **Step 1:** Split each day into four 6-hour periods to observe both daytime and nighttime temperature variations, capturing the influence of peak human activity hours on pollution.
- **Step 2:** Code each day as 1 if thermal inversion is observed during any of the four periods; otherwise, code it as 0.
- **Step 3:** Sum the thermal inversion days for each year and compute the five-year average to obtain the destination-to-origin ratio.

Figure 3.1 shows the correlation between the log-transformed thermal inversion days and the $PM_{2.5}$ ratio, indicating a significant relationship between the instrumental variable and the primary explanatory variable.

To confirm the exogeneity of thermal inversion, Figure 3.2 illustrates annual national thermal inversion days alongside migration flows from 1980 to 2020, showing no apparent association. Additionally, Figure 3.3 examines the relationship between thermal inversion and GDP at both the national level and in major economic cities (Beijing, Tianjin, Shanghai, Chongqing). In all cases, no significant correlation is observed, supporting the exclusive restriction is not violated.

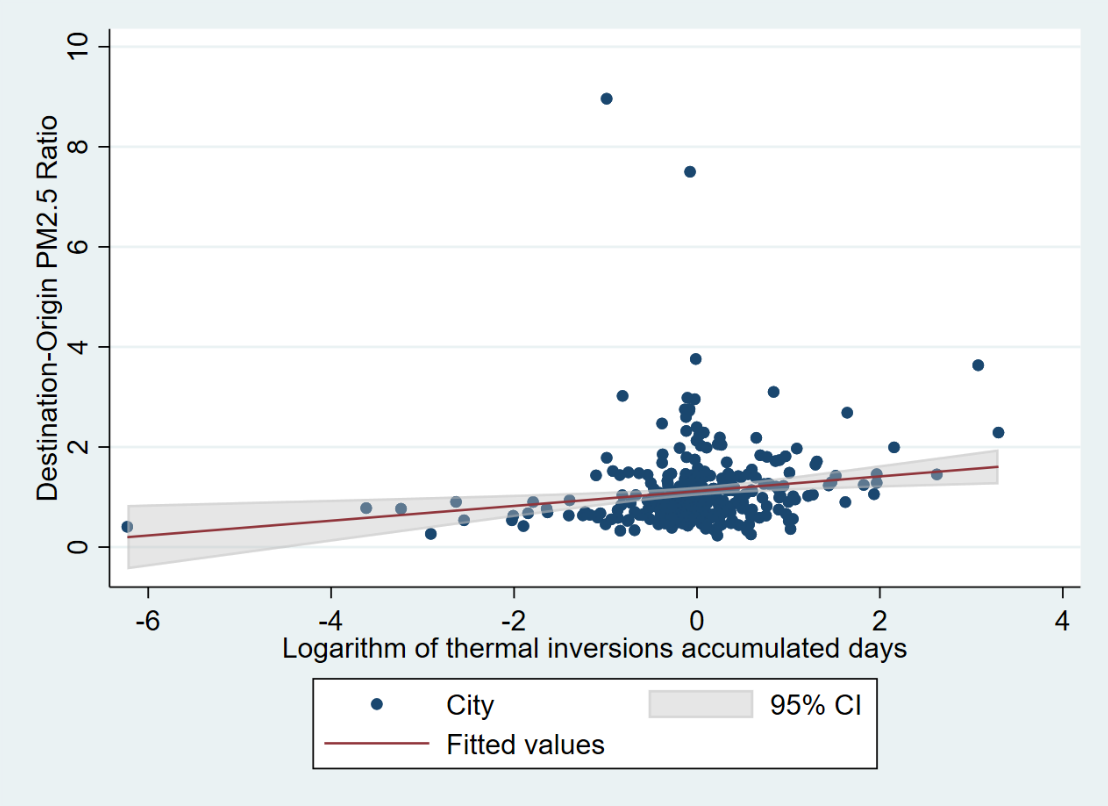


FIGURE 3.1: **The Correlation between destination-origin $PM_{2.5}$ ratio and destination-origin ratio of Thermal Inversion Accumulated Days**

Note: This figure shows the relationship between the $PM_{2.5}$ ratio and the logarithm of thermal inversions accumulated days for each destination and origin pair in the sample.
Source: CMDS 2017, NASA EOSDIS 2013-2017

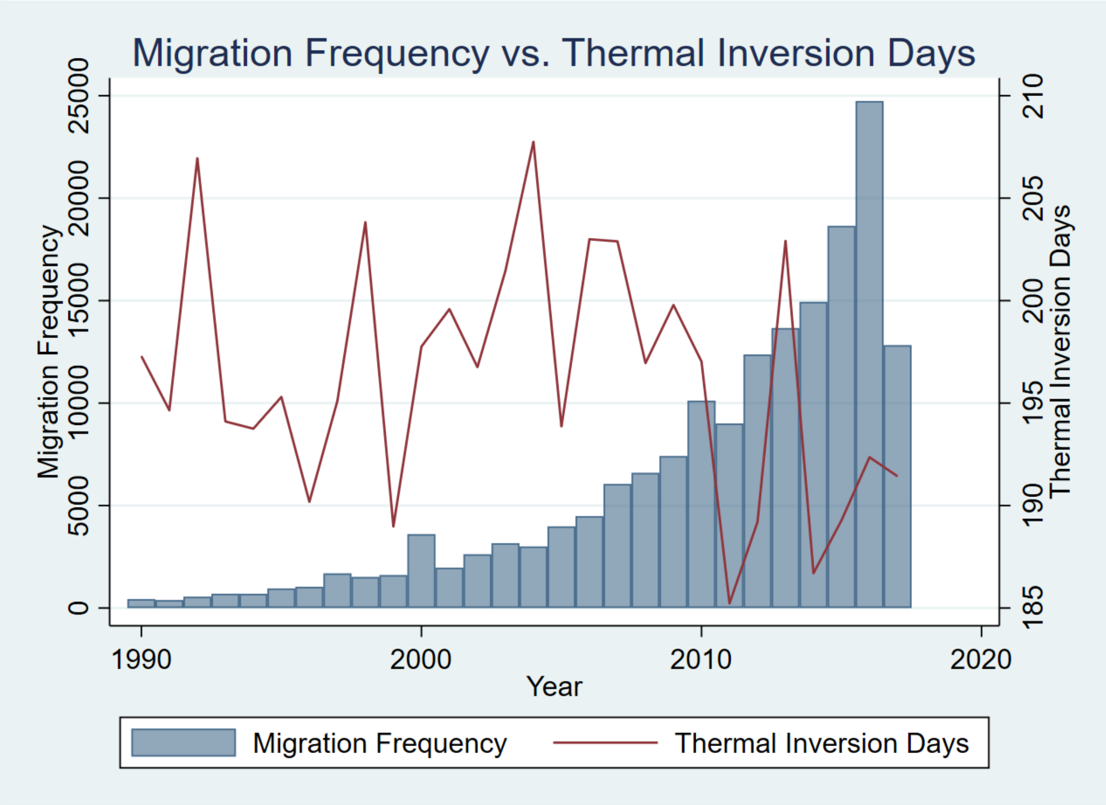
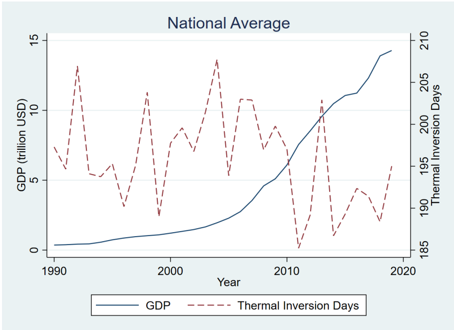


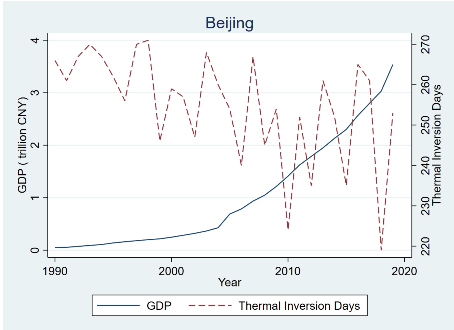
Figure 3.2: **Migration Flow and National Thermal Inversion Accumulated Days**

Note: This figure shows the time trend of migrants' frequency in the sample and national average thermal inversion accumulated days by year. The strong correlation between the thermal inversion days and migrants' frequency is not evident, i.e., the thermal inversion days are relatively exogenous to the mitigation

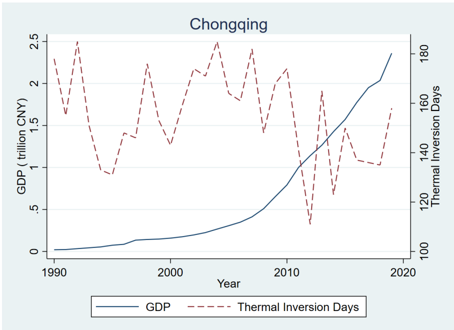
Source: CMDS 2017, NASA EOSDIS 2013-2017



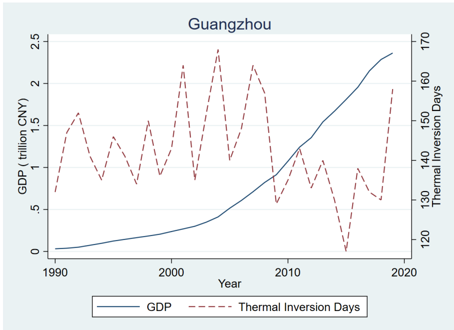
(A) Cross Nation



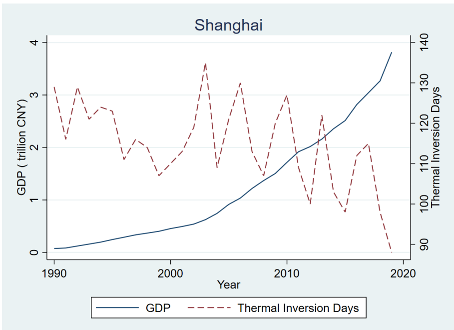
(B) Beijing



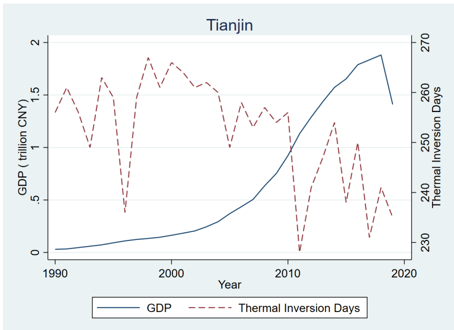
(C) Chongqing



(D) Guangzhou



(E) Shanghai



(F) Tianjin

FIGURE 3.3: **The Time Trend of Thermal Inversions Accumulated Days and GDP**

Note: To check the exclusive restriction of instrumental variables, i.e., IV is exogenous to potential unobserved factors, I check the time trend of thermal inversion days and GDP per capita at the national average and municipality levels. Subfigures **(A)** to **(F)** are the national average, Beijing, Chongqing, Guangzhou, Shanghai, and Tianjin. Noticeably, the fluctuation of thermal inversion does not show an association with GDP.

Source: NASA EOSDIS 1990-2020, World Bank, China City Yearbook.

3.5 Empirical Results

3.5.1 Baseline Results

Table 3.3 presents the baseline OLS results, analysing the effect of air pollution differentials—measured as the relative $PM_{2.5}$ concentration between destination and origin cities—on migration flows. Column (1) controls for fixed effects at the destination city level, while Column (2) controls for fixed effects at the origin city level. Column (3) includes fixed effects for both destinations and origins, capturing unobserved characteristics specific to each location. The results suggest that higher pollution at the destination has a "push" effect, reducing migration flows. Specifically, each doubling in the pollution differential corresponds to a 35.98% $(= e^{-0.446} - 1)$ decrease in migration flows.

To account for potential cultural and historical ties between provinces, such as shared borders, local dialects, and similar geographic conditions, Column (4) introduces province-pair fixed effects, which control for unobserved regional connections that may influence migration patterns. The results remain robust, showing a similar effect: a doubling in pollution differential is associated with a 7.02% $(= e^{-0.07} - 1)$ reduction in migration flows, though the coefficient on air pollution is slightly reduced.

Given the nature of the data — a cross-sectional, bidirectional migration flow dataset with many zero-valued observations — the OLS model with log transformation, ln(Migration flow+1), may lead to inconsistent estimates, as log transformations may bias estimates when zero outcomes are prevalent (Chen and Roth, 2023; Santos Silva and Tenreyro, 2022). To address this, a Poisson pseudo-maximum likelihood (PPML) estimator is applied as a robustness check. PPML accommodates zero migration flows without transformation, providing consistent estimates of migration elasticities. Since city-level fixed effects would offset city-specific covariates, province-level fixed effects are instead used to account for unobserved factors specific to each destination and origin province.

Table 3.3: **Baseline Results of OLS**

Dependent variable	Ln (Migration Flow+1)			
	(1)	(2)	(3)	(4)
Destination-origin ratio of $PM_{2.5}$	-0.139***	-0.137**	-0.446***	-0.0728**
	(0.0267)	(0.0605)	(0.108)	(0.0292)
Covarites Included	Yes	Yes	Yes	Yes
Destination City Fixed Effect	Yes	No	Yes	No
Origin City Fixed Effect	No	Yes	Yes	No
Province-pair Fixed Effect	No	No	No	Yes
$adj.R^2$	0.404	0.393	0.451	0.476
N	25717	25717	25717	25717

Note: This table reports the air pollution effect on cross-city migration flow between origin and destination cities based on OLS. The dependent variable is the logarithm of the number of cross-city migration flows plus one.
All control variables lagged one year and included bidirectional cities from origins to destinations.
Robust standard errors clustered at destination-city and are shown in parentheses.
***, **, and * representative the level of significance at 1%, 5% and 10%.
Source: CMDS 2017, NASA EOSDIS 2016, China City Yearbook 2016, and China Meteorological Data Sharing Service System (CMDSSS).

Table 3.4 reports the PPML results for the effect of air pollution on migration flows. Columns (1) and (2) apply fixed effects at the destination and origin province levels, respectively. Column (3) includes separate fixed effects for both provinces, and Column (4) applies province-pair fixed effects. The results confirm the robustness of the OLS findings: each doubling in the pollution differential leads to a 23.27% ($= e^{-0.265} - 1$) reduction in migration flows at a 5% significance level. When province-pair fixed effects are applied, the effect of air pollution is slightly smaller, reducing migration flows by approximately 2.64% ($= e^{-0.0264} - 1$).

TABLE 3.4: **Baseline PPML Results (Robustness Check to OLS)**

Dependent variable	Ln Migration Flow			
	(1)	(2)	(3)	(4)
Destination-origin ratio of $PM_{2.5}$	-0.238*	-0.298***	-0.265**	-0.0268
	(0.144)	(0.0963)	(0.114)	(0.0440)
Pseudo R^2	0.333	0.325	0.349	0.644
N	25717	25717	25717	25716
Destination Provinces Fixed Effect	Yes	No	Yes	No
Origin Provinces Fixed Effect	No	Yes	Yes	No
Province-pair Fixed Effect	No	No	No	Yes

Note: This table reports the air pollution effect on cross-city migration flow between origin and destination cities based on PPML. The dependent variable is the logarithm of the number of cross-city migration flows.
All control variables lagged one year and included bidirectional cities from origins to destinations.
Robust Standard errors clustered at destination-city and are shown in parentheses.
***, **, and * representative the level of significance at 1%, 5% and 10%.
Source: CMDS 2017, NASA EOSDIS 2013-2017, China City Yearbook 2016, and China Meteorological Administration.

3.5.2 2SLS and IV-PPML Results

To address potential reverse causality, accumulated thermal inversion days are used as an instrumental variable for pollution, as discussed in Section 4.2. Controlling for historical and cultural ties through province-pair fixed effects, Table 3.6 presents the 2SLS and IV-PPML results. Column (1) shows the baseline OLS results for comparison, while Columns (2) and (3) report the 2SLS results, with the first stage in Column (2) and the second stage in Column (3). The Kleibergen–Paap rk Wald F-statistic is 128.529, exceeding the critical value of 16.38, indicating a strong instrumental variable. Compared to the OLS results, the effect of air pollution on migration flows becomes more negative in the 2SLS model—approximately six times larger, suggesting an upward bias in OLS estimates. This bias likely arises because migration inflows tend to increase pollution in destination areas, while outflows decrease pollution in origin areas.

This finding aligns with prior literature (Chen et al., 2022; Khanna et al., 2021), which suggests that endogenous pollution variables may correlate with factors such as wages, job opportunities, and infrastructure that attract migration. Additionally, reverse causality may result in an overestimated pollution differential (destination-to-origin $PM_{2.5}$ ratio) as migration inflows worsen pollution in destination cities while outflows reduce it in origin cities.

Column (4) presents the IV-PPML model results, with standardised coefficients on the right. Similar to PPML, IV-PPML uses a Poisson pseudo-maximum likelihood approach to handle models with endogenous variables (Silva and Tenreyro, 2006). The IV-PPML results are consistent with the 2SLS findings, indicating a strong negative effect of air pollution on migration: each doubling in pollution between destination and origin cities reduces migration by approximately 42%.

In summary, the OLS, PPML, and 2SLS/IV-PPML results consistently indicate that air pollution significantly reduces migration flows, underscoring pollution's role as a "push" factor in migration decisions.

TABLE 3.5: **2SLS and IV-PPML Results**

	2SLS		IVPPML	
	First Stage	Second stage		Standardised
Dependent variable	$PM_{2.5}$ Ratio	Ln (Migration Flow+1)	Ln Migration Flow	Coefficients
	(2)	(3)	(4)	
Destination-origin ratio of $PM_{2.5}$		-0.465***	-0.545**	-0.420**
		(0.111)	(0.275)	
Destination-origin ratio of Thermal Inversion Days	0.147***			
	(0.013)			
Covariates Included	Yes	Yes	Yes	Yes
Province-pair Fixed Effect	Yes	Yes	Yes	Yes
Kleibergen–Paap rk Wald F statistic	128.529			
N	25512		25512	

Note: This table reports the air pollution effect on cross-city migration flow between origin and destination cities based on 2SLS and IV-PPML. Column (1) is the OLS results as baseline, and columns (2) and (3) are the first and second stages of the 2SLS results. Column 4 is IV-PPML results with standardised coefficients. All control variables lagged one year and included bidirectional cities from origins to destinations.
The province-pair effect is fixed. Standard errors clustered at destination city and are shown in parentheses. ***, **, and * representative the level of significance at 1%, 5% and 10%.
Source: CMDS 2017, NASA EOSDIS 2013-2017, China City Yearbook 2016, and China Meteorological Administration.

3.5.3 Channel Analysis

3.5.3.1 Channels of Air Pollution's Impact on Migration

Existing literature largely focuses on micro-level factors in migration decisions, such as individual health concerns, life satisfaction, and job opportunities in destination cities. However, the role of macroeconomic channels remains underexplored. This paper addresses this gap by examining how various macroeconomic indicators mediate the impact of air pollution on migration flows.

Guided by hedonic theory and the methods of Fu et al. (2021); Khanna et al. (2021); Wu et al. (2021a), I select a range of indicators representing amenities or costs in destination cities, including green infrastructure, environmental regulation strength, social welfare accessibility, and settlement costs. These proxies capture the potential mechanisms through which air pollution may affect migration flows.

The following model incorporates these channels:

$$\ln\left(M_{o,d}\right)=\gamma_0+\gamma_1\frac{\mathrm{PM}_{2.5d}}{\mathrm{PM}_{2.5o}}+\gamma_2\,\mathrm{channel}_d+\gamma_3\left(\frac{\mathrm{PM}_{2.5d}}{\mathrm{PM}_{2.5o}}\times\mathrm{channel}_d\right)+X'_{o,d}\delta+e_{o,d} \tag{3.3}$$

where $\mathrm{channel}_d$ denotes macroeconomic factors in destination cities, specifically green infrastructure, environmental regulation strength, social welfare accessibility, and settlement costs. The interaction term $\frac{\mathrm{PM}_{2.5d}}{\mathrm{PM}_{2.5o}}\times\mathrm{channel}_d$ captures the joint effect of air pollution and each macro factor on migration flows.

For channels, firstly, green infrastructure, represented by the harmless treatment rate of domestic waste, serves as a proxy for sustainable amenities and waste management capacity in destination cities. Also, since the wastewater is not likely associated with air pollution, the wastewater-related indicator can be seen as exogenous to green infrastructure. This variable captures the quality of urban green infrastructure, which can influence migration decisions by mitigating some negative impacts of air pollution.

Secondly, environmental regulation strength is proxied through the government's annual work reports for each destination. This approach captures the extent of regulatory focus on environmental issues and public awareness of such concerns. While past literature has widely applied different methods to measure the environmental regulation strength, e.g., Sanchez-Vargas et al. (2013) used the investment in the technical reduction of pollution and Kheder and Zugravu (2012) used the $Z-score$ to obtain the standardised value of the environment, because of the limited number of NGOs in China, this paper relies on the frequency of environment-related keywords in government reports as a proxy for environmental regulation strength.

As for the social welfare channel, the hukou threshold index is applied as a proxy for the probability of integrating and enjoying social welfare in destination cities. Hukou, the Chinese unique residents' registered system, restricts the migrants' welfare, such as local health insurance, education, and property management. Hence, acquiring the destinations' hukou determines if migrants can acquire the same level of welfare as residents. Following Zhang et al. (2019), this index is calculated and quantified by the government reports, policy text, local socioeconomic characteristics, registered information and relative hukou management. This index could show the difficulty of acquiring the destinations' hukou and imply the potential channels of migration cost in terms of social welfare.

Additionally, housing prices in destination cities serve as a proxy for settlement costs, capturing the trade-off between the cost of living and access to urban amenities. This variable represents the hedonic price of amenities, which may either attract or deter migrants depending on the perceived benefits relative to environmental quality.

Table 3.6 presents the results for each macroeconomic channel, with columns (1) through (4) examining the effects of green infrastructure, environmental regulation, social welfare, and settlement costs, respectively. All models are estimated using 2SLS, and interaction terms are also instrumented.

In Table 3.6, the results in column (1) indicate that green infrastructure positively influences migration flows, with a significant negative interaction term suggesting a substitutive relationship between air pollution and green amenities. This implies that high-quality green infrastructure in destination cities can partially offset the adverse effects of air pollution on migration.

For environmental regulation in column (2), stricter regulations appear (higher frequency of related keywords) to mitigate the negative impact of air pollution on migration, as indicated by the significant interaction term coefficients. This finding suggests that effective regulatory frameworks can alleviate migrant concerns about pollution and attract more migrants to regions with robust environmental policies.

In columns (3) and (4), the hukou threshold and housing prices both negatively impact migration flows, with positive coefficients of interaction terms. This indicates that social welfare restrictions and high settlement costs discourage migration inflows and complement the deterrent effect of air pollution. Notably, highly urbanised cities like Beijing and Shanghai impose stricter policies that limit long-term migration, particularly where pollution is a concern. Thus, social welfare barriers and high settlement costs are closely linked with air pollution as factors reducing migration flows.

Overall, these results suggest that while green infrastructure and strong environmental regulation can offset the negative effects of air pollution on migration, social welfare restrictions and high settlement costs exacerbate them. This analysis highlights the

TABLE 3.6: **Channel Analysis**

Dependent variable	Ln (Migration Flow+1)			
Channel	Green Infrastructure	Environmental Regulation	Social Welfare	Settlement Cost
	(1)	(2)	(3)	(4)
Air Pollution	-0.373***	-0.060	-0.975***	-0.894***
	(0.0999)	(0.303)	(0.257)	(0.170)
Green Infrastructure	0.279**			
	(0.142)			
Air Pollution × Green Infrastructure	-0.247**			
	(0.117)			
Green Regulation		0.021*		
		(0.010)		
Air Pollution × Green Regulation		-0.016*		
		(0.010)		
Hukou Threshold			-0.919**	
			(0.397)	
Air Pollution × Hukou Threshold			0.748***	
			(0.290)	
Housing Price				-0.0202**
				(0.00941)
Air Pollution × Housing Price				0.0463***
				(0.0151)
Covariates Included	Yes	Yes	Yes	Yes
Province-pair Fixed Effect	Yes	Yes	Yes	Yes
K-P F-test	42.13	19.97	69.93	65.54
N	25512	25063	25512	25512

Note: Based on 2SLS, this table discusses the channels of pollution effect in terms of infrastructure (column 1), environment regulation (column 2), social welfare (column 3) and housing prices (column 4).
All control variables are lagged one year and include bidirectional cities from origins to destinations.
The province-pair effect is fixed. Standard errors are clustered at destination-city and are shown in parentheses.
***, **, and * representative the level of significance at 1%, 5% and 10%.
Source: CMDS 2017, NASA EOSDIS 2013-2017, China City Yearbook 2013-2016, and China Meteorological Data Sharing Service System (CMDSSS), China City Government Annual Reports.

importance of considering both micro and macro channels when examining the effects of environmental factors on migration flows.

3.5.3.2 The Role of Distance in Pollution-Driven Migration

In environmental migration, the primary perceived benefits include escaping air pollution and protecting health. Migrants typically weigh the risks and benefits of moving to destinations with different pollution levels from their origins, and this consideration is influenced by migration distance and information on pollution in the destination city. On one hand, geographical distance directly impacts migration costs, with longer distances leading to higher expenses. On the other hand, greater distances weaken migrants' immediate perception of pollution in the destination, making environmental information about distant cities more uncertain compared to nearby ones.

This section analyses the role of distance in the relationship between air pollution and migration decisions. Migration distances in the sample range from 0 to 2000 km and follow a normal distribution. For analysis purposes, the distance range is divided into six intervals, each with a step length of 330 km. Table 3.7 presents regression results for the effect of air pollution on migration flows across these distance intervals, with Columns (1) to (6) representing different subgroups by migration distance.

The results indicate a U-shaped effect of air pollution on migration flows as distance increases. This means that at shorter distances, the negative effect of pollution on migration flow is more significant, but it diminishes beyond a certain distance. Figure 3.4 visualises the main regression results from Table 3.7, focusing on the effect of air pollution. When migration distance is between 660 and 990 km, this effect reaches its peak, reducing migration flows by approximately 49.49% ($= e^{-0.683} - 1$), statistically significant at the 1% level. Notably, when the migration distance exceeds 1650 km, the effect of air pollution remains negative but is no longer statistically significant.

These findings suggest that within 1000 km, distance amplifies the impact of air pollution on migration decisions, but this influence weakens beyond 1000 km. Migrants appear more sensitive to air pollution within a 1000-kilometre radius. Two possible explanations support this observation. First, perceived benefits related to lower pollution levels are more accessible within this distance, which also approximates the average distance between major economic hubs in China (e.g., Beijing to Shanghai, or Shanghai to Guangzhou). As distance increases, obtaining reliable information on the destination becomes more difficult, with pollution information likely attenuated over long distances Coombs (1978). Thus, beyond 1000 km, the clarity of pollution information decreases, affecting migrants' decision-making.

Second, economic factors, particularly migration costs, may overshadow the impact of air pollution at longer distances. Economic considerations remain the primary driver of migration decisions in China, and migration costs increase significantly with distance. For families deciding on long-distance moves, higher migration costs may prioritise economic opportunities, such as increased income or job prospects, over potential environmental risks. This helps explain why the effect of air pollution is negligible and statistically insignificant in Column (6).

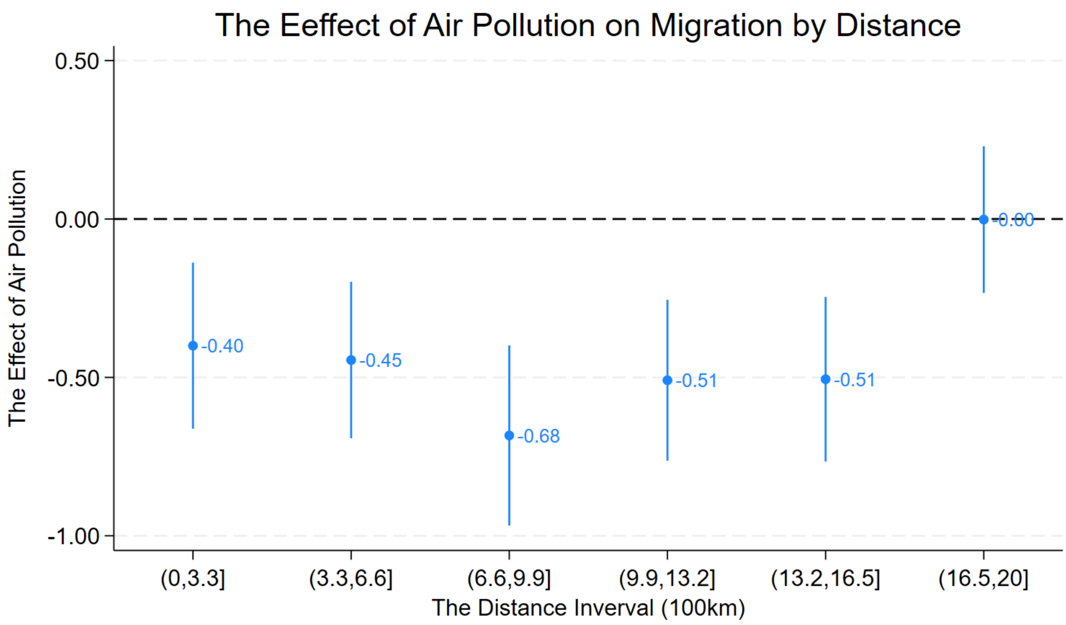


FIGURE 3.4: **The Effect of Air Pollution on Migration Flow by Distance**
Note: This figure shows the regression results of Table 3.7. Based on the distance between origins and destinations, the subsamples include six different intervals with a unit length of 330km
Source: CMDS 2017, NASA EOSDIS 2016, China City Yearbook 2016, and China Meteorological Data Sharing Service System (CMDSSS).

Table 3.7: **The Effect of Air Pollution on Migration by Distance**

Dependent variable	Ln (low Migration+1)					
Distance Interval (unit:100km)	[0, 3.3]	(3.3, 6.6]	(6.6, 9.9]	(9.9, 13.2]	(13.2, 16.5]	(16.5, 20]
	(1)	(2)	(3)	(4)	(5)	(6)
Destination-origin ratio of $PM_{2.5}$	-0.400***	-0.445***	-0.683***	-0.509***	-0.506***	-0.00176
	(0.134)	(0.126)	(0.145)	(0.130)	(0.133)	(0.118)
Covariates Included	Yes	Yes	Yes	Yes	Yes	Yes
Province-pair Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
N	2109	4710	5890	5934	4566	2303

Note: This table reports the air pollution effect on migration flow by different migration distance intervals.
All control variables lagged one year and included bidirectional cities from origins to destinations. The overall sample is split into six groups based on distance between origins and distance, and each interval length is 330km, covering all sample distances from 0 to 2000km. The province pairs are fixed to absorb the unobserved cultural and historical connection between destinations and origins. Robust Standard errors clustered at destination city and are shown in parentheses.
***, **, and * representative the level of significance at 1%, 5% and 10%.
Source: CMDS 2017, NASA EOSDIS 2016, China City Yearbook 2016, and China Meteorological Data Sharing Service System (CMDSSS).

3.5.4 Heterogeneity Analysis

Education

Do highly educated migrants exhibit a stronger aversion to air pollution? Does greater knowledge lead to “greener” choices? Research suggests that individuals with higher education are generally more sensitive to air pollution, tend to avoid highly polluted areas, and prefer cities with better air quality (Chen et al., 2022; Sun et al., 2019; Lai et al., 2021). However, other studies indicate that high-pollution cities often offer higher compensating wages, which may lessen the environmental concerns of highly educated individuals (Khanna et al., 2021; Wang et al., 2021). Additionally, high-income migrants often afford protective technologies, such as air purifiers or short-term travel, to mitigate pollution exposure (Chen et al., 2021; Gao et al., 2023; Ito and Zhang, 2020). This section examines whether the highly educated are more responsive to air pollution in migration decisions.

Migration flow is categorised by education level into four groups: primary (Grade 1 to 6), secondary (Grade 7 to 12), college (Bachelor), and postgraduate (Master and PhD). Table 3.8 presents the impact of air pollution on these groups. Column (1) represents individuals with primary education (grades 1 to 6), Column (2) includes those with secondary education (grades 7 to 12), Column (3) comprises college-educated individuals, and Column (4) represents those with postgraduate degrees (master’s and doctoral). The results are also shown in Figure 3.5

Results indicate that individuals with secondary education show the strongest aversion to air pollution, with migration decreasing by 54.5%. In contrast, the highly educated group shows a smaller reduction in migration at 27.8%, half the effect observed in the low-education group. This finding aligns with studies by Zhang et al. (2017) and Liu and Yu (2020). Compared to those with lower education, highly educated migrants often receive higher income, including compensating wages, and face lower migration costs, which may reduce their sensitivity to pollution (Bayer et al., 2009). Particularly, due to the unbalanced development and different industry structures, the highly polluted cities often have a high-quality job market which is more sticky for urban-to-urban migrants.

Highly educated migrants likely possess more knowledge about pollution prevention, particularly this paper focuses on urban-to-urban migrants. Their higher income and better working conditions, including air-conditioned and air-purifier environments, allow them to minimise pollution exposure, thus making pollution a less significant factor in their migration decisions. On the other hand, between highly educated groups, the principle "better knowledge more sensitive" still holds (postgraduates -27.8% vs graduates -15.6%)

Table 3.8: **Effect of Air Pollution on the Migrants Flow by Education**

Dependent Variable	Ln (Migration Flow+1)			
Subsample	Primary	Secondary	Graduate	Post-graduate
	(1)	(2)	(3)	(4)
Destination-origin ratio of $PM_{2.5}$	-0.328***	-0.545***	-0.156***	-0.278**
	(0.0734)	(0.0880)	(0.0739)	(0.111)
Covariates Included	Yes	Yes	Yes	Yes
Province-pair Fixed Effect	Yes	Yes	Yes	Yes
K-P F-test	100.56	107.71	103.63	87.27
N	7613	17129	9148	1013

Note: This table reports the air pollution effect on migration flow by different education levels. The overall sample is split into four groups. Column (1) is the group of migrants with primary or parenthesestion (grade 1 to 6), column (2) is the migrants with secondary education (grade 7 to 12); column (3) is the migrant group with a college degree (grade 10 to 12); and column (4) is migrants with master and over degrees. The province pairs are fixed to absorb the unobserved cultural and historical connection between destinations and origins.

All control variables lagged one year and included bidirectional cities from origins to destinations. Robust Standard errors clustered at destination-city and are shown in parenthesis.

***, **, and * representative the level of significance at 1%, 5% and 10%.

Source: CMDS 2017, NASA EOSDIS 2016, China City Yearbook 2016, and China Meteorological Data Sharing Service System (CMDSSS).

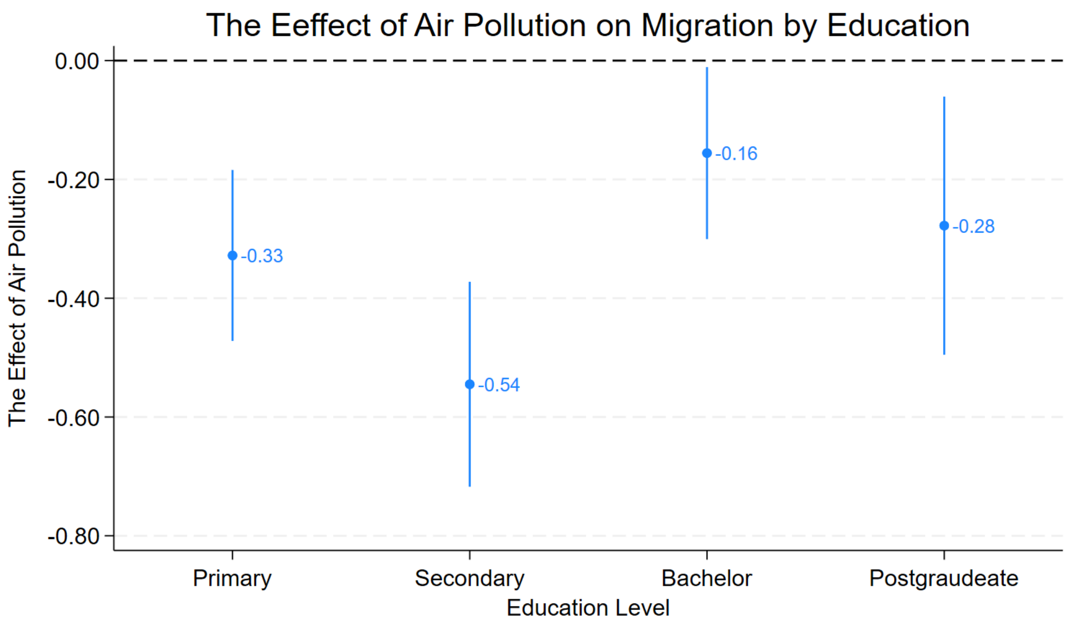


Figure 3.5: **The Effect of Air Pollution on Migration Flow by Education**

Note: This figure is the regression results of Table 3.8, showing the effect of air pollution on migration in different educated groups. The whole sample is split into four different education levels, and the point is the coefficients of air pollution and the line documents the 95% confidence interval.

Source: CMDS 2017, NASA EOSDIS 2016, China City Yearbook 2016, and China Meteorological Data Sharing Service System (CMDSSS).

Age

To explore how different age groups respond to air pollution, Table 3.9 presents migration behaviours across age groups. Columns (1) to (4) cover migrants aged 15 to 75, with each column representing a 15-year interval. Results indicate that as age increases, migration flow decreases more dramatically, suggesting that older migrants are more concerned with air pollution than younger ones. This finding is consistent with Guo et al. (2022), who notes that older individuals are more sensitive to health risks posed by pollution. Interestingly, Column (4) shows that although older groups have a relatively weaker aversion to pollution, they still exhibit a significantly higher aversion than the youngest group (ages 15-30). The results are also shown in Figure 3.6.

One potential explanation for the weaker migration response among less educated and older migrants is their overrepresentation in pollution-intensive sectors such as manufacturing, construction, and logistics (Fan, 2002). These industries often locate in more polluted areas, and the workers have limited job mobility across sectors or regions (Meng, 2012). Additionally, these groups may have lower income, weaker information access, and stronger social ties to their original residence, which jointly constrain their relocation options. As such, although they may suffer greater health risks from air pollution, economic necessity and job immobility often outweigh environmental considerations in their migration decisions. This occupational and structural constraint helps explain why their aversion to pollution is less strongly reflected in migration flows.

TABLE 3.9: **Effect of Air Pollution on the Migrants Flow by Age Cohorts**

Dependent Variable	Ln (Migration Flow+1)			
Subsamples	Age 15-29	Age 30-44	Age 45-59	Age 59-75
	(1)	(2)	(3)	(4)
Destination-origin ratio of $PM_{2.5}$	-0.195***	-0.424***	-0.412***	-0.373***
	(0.0814)	(0.0803)	(0.0800)	(0.0962)
Covariates Included	Yes	Yes	Yes	Yes
Province-pair Fixed Effect	Yes	Yes	Yes	Yes
K-P F-test	113.12	105.47	111.98	69.36
N	13284	14717	8641	3149

Note: This table reports the air pollution effect on migration flow by different age groups. The range of age in the sample is split into four intervals, each 15 years counts as a group. Column (1) is age "15-29"; column (2) is "30-44; column (3) is "45-59"; and column (4) is "59-75". The province pairs are fixed to absorb the unobserved cultural and historical connection between destinations and origins.
All control variables lagged one year and included bidirectional cities from origins to destinations.
Robust standard errors clustered at destination-city and are shown in parentheses.
***, **, and * representative the level of significance at 1%, 5% and 10%.
Source: CMDS 2017, NASA EOSDIS 2016, China City Yearbook 2016, and China Meteorological Data Sharing Service System (CMDSSS).

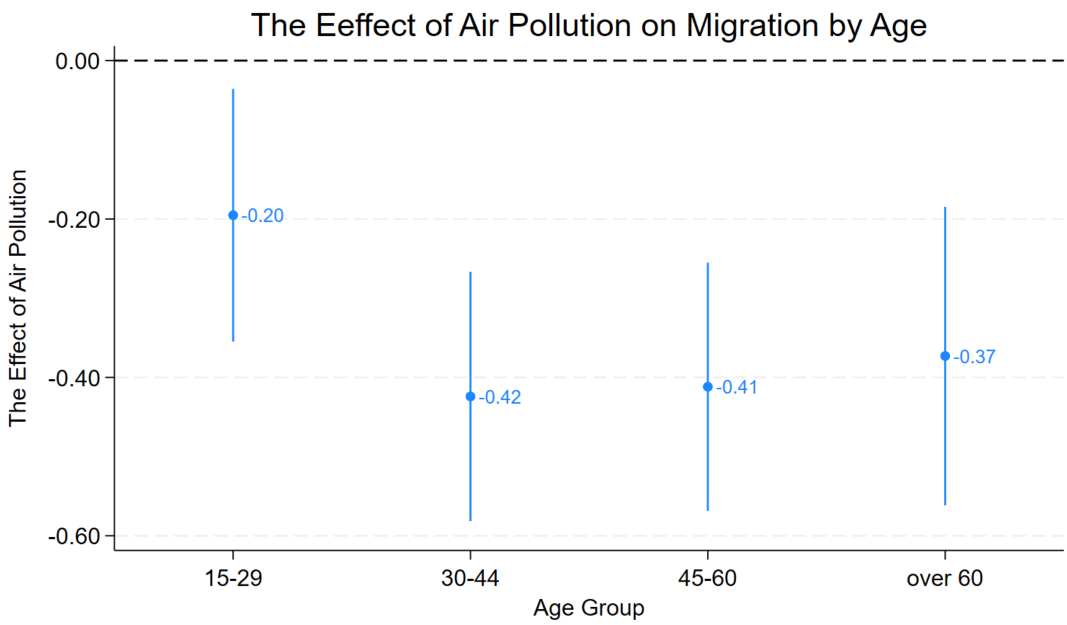


FIGURE 3.6: **The Effect of Air Pollution on Migration Flow by Age**
Note: This figure is the regression results of Table 3.9, showing the effect of air pollution on migration in different age groups. The whole sample is split into four different age groups. The point is the coefficients of air pollution, and the line documents the 95% confidence interval.
Source: CMDS 2017, NASA EOSDIS 2016, China City Yearbook 2016, and China Meteorological Data Sharing Service System (CMDSSS).

Gender Marital Status, and Cross-Province

This study also examines how marital status and gender influence migration decisions in response to air pollution. Table 3.10 presents the impact of air pollution on different groups by marital status (single in Column (1), non-single in Column (2)) and by gender (male in Column (3), female in Column (4)). Notably, non-single migrants demonstrate a stronger aversion to pollution compared to singles, with each unit increase in the destination-to-origin air pollution ratio reducing migration by an insignificant 10.9% for singles and by significant 41.5% for married individuals.

This difference likely stems from age and family responsibilities. Single migrants are typically younger and may have less life and work experience, which makes them less sensitive to pollution. In contrast, married migrants often consider the health of their families as well as their own, leading to greater sensitivity to air pollution (see Figure 3.7).

Interestingly, I did not find a significant difference in environmental attitudes between males and females. This insignificant result (t-test = -1.46) challenges traditional arguments that males exhibit greater concern for environmental issues (Chang et al., 2011). Both genders demonstrated similar aversion to air pollution, with a reduction of -36.4% among males and -34.8% among females. A plausible explanation lies in the composition of my research sample—urban migrants, who are typically more highly educated and experience greater gender equality in urban settings.

TABLE 3.10: **Effect of Air Pollution on the Migrants Flow by Gender and Marital Status**

Dependent Variable	Ln (Migration Flow+1)			
Subsample	Marital		Gender	
	Single	Married	Male	Female
	(1)	(2)	(3)	(4)
Destination-origin ratio of $PM_{2.5}$	-0.109	-0.499***	-0.364***	-0.348***
	(0.0762)	(0.0920)	(0.0798)	(0.0791)
Covariates Included	Yes	Yes	Yes	Yes
Province-pair Fixed Effect	Yes	Yes	Yes	Yes
K-P F-test	122.94	110.14	112.43	113.04
N	9197	18645	17228	16287
Coefficients Difference t-Test	7.44***		-1.46	

Note: This table reports the effects of air pollution on males and females with different marital statuses. Columns (1) and (2) are the single group; columns (2) and (3) are the married group. The coefficients difference test tests the coefficients of air pollution across different groups.
All control variables lagged one year and included bidirectional cities from origins to destinations.
The province pairs are fixed to absorb the unobserved cultural and historical connection between destinations and origins.
Robust standard errors clustered at destination-city and are shown in parentheses. ***, **, and * representative the level of significance at 1%, 5% and 10%.
Source: CMDS 2017, NASA EOSDIS 2016, China City Yearbook 2016, and China Meteorological Data Sharing Service System (CMDSSS).

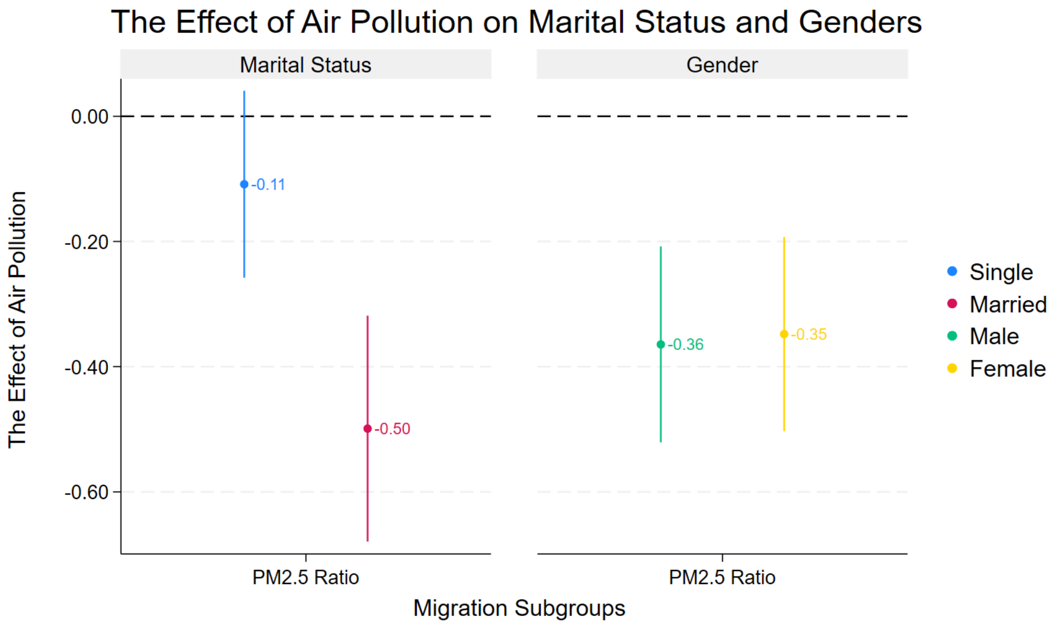


Figure 3.7: **The Effect of Air Pollution on Migration Flow by Marital and Gender**

Note: This figure is the regression results of Table 3.10, showing the effect of air pollution on migration by marital status and gender. Point is the coefficient of air pollution and the line documents the 95% confidence interval.

Source: CMDS 2017, NASA EOSDIS 2016, China City Yearbook 2016, and China Meteorological Data Sharing Service System (CMDSSS).

To further examine gender differences by marital status, we analysed subgroups based on both gender and marital status. Table 3.11 and Figure 3.8 present the impact of air pollution on males and females, with Columns (1) and (2) representing single individuals and Columns (3) and (4) representing married individuals. Among singles, males did not exhibit a significantly stronger aversion to pollution than females (t-test = -1.08). This finding aligns with the results in Table 3.10, which show that younger, single migrants are less likely to be affected by pollution.

In contrast, within the married group, males display a slightly stronger response to pollution. However, the difference between males and females remains statistically insignificant (t-test = -1.63). This suggests that marriage and family life may enhance individuals' sensitivity to pollution, which is consistent with 3.10.

TABLE 3.11: **Effect of Air Pollution on the Migrants Flow by Marital Status across Gender**

Dependent Variable	Ln (Migration Flow+1)			
Subsample	Single		Married	
	Male	Female	Male	Female
	(1)	(2)	(3)	(4)
Destination-origin ratio of $PM_{2.5}$	-0.103	-0.0254	-0.375***	-0.353***
	(0.0722)	(0.0811)	(0.0766)	(0.0743)
Covariates Included	Yes	Yes	Yes	Yes
Province-pair Fixed Effect	Yes	Yes	Yes	Yes
K-P F-test	131.00	119.00	115.03	115.51
N	6850	5773	16302	16232
Coefficients difference t-test	-1.08		-1.63	

Note: This table reports the air pollution effect on migration flow by marital situation and gender. All control variables lagged one year and included bidirectional cities from origins to destinations. Column (1) is a single group; column (2) is a married group; column (3) is the male group; and column (4) is the female group.
The province pairs are fixed to absorb the unobserved cultural and historical connection between destinations and origins.
The coefficients difference test tests the coefficients of air pollution across different groups. Robust standard errors clustered at destination-city and are shown in parentheses. ***, **, and * representative the level of significance at 1%, 5% and 10%.
Source: CMDS 2017, NASA EOSDIS 2016, China City Yearbook 2016, and China Meteorological Data Sharing Service System (CMDSSS).

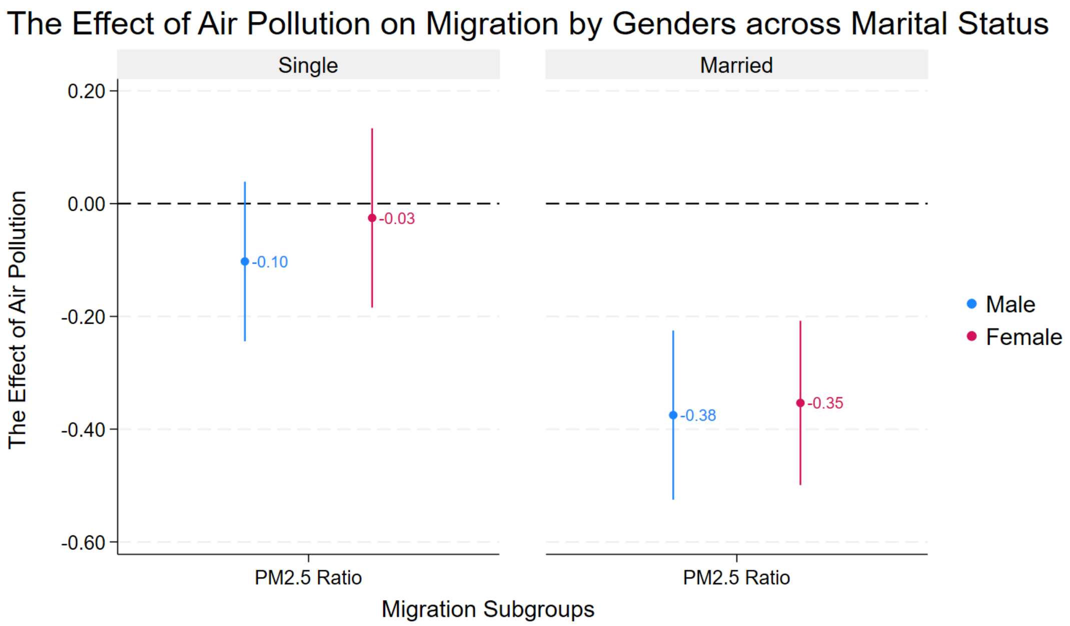


FIGURE 3.8: **The Effect of Air Pollution on Migration Flow by Marital Status across Gender**

Note: This figure is the regression results of Table 3.11, showing the effect of air pollution on migration by marital status across genders. Point is the coefficient of air pollution, and the line documents the 95% confidence interval.

Source: CMDS 2017, NASA EOSDIS 2016, China City Yearbook 2016, and China Meteorological Data Sharing Service System (CMDSSS).

On the other hand, cross-province migration in China entails additional costs and challenges. For cross-province migrants, moving outside their hukou province imposes restrictions on accessing social benefits such as education, healthcare, and insurance. Table 3.12 divides the sample into cross-province migration (Columns (1) to (3)) and within-province migration (Columns (4) to (6)), further analysing gender differences within these groups (also see Figure 3.9).

The results in Table 3.12 indicate that cross-province migrants exhibit a significant aversion to pollution, whereas within-province migrants do not. Cross-province migration entails higher costs due to greater distances, cultural and dialect differences, and hukou-based restrictions on social benefits, which likely increases concern about pollution exposure in new areas. Consistent with earlier findings, female migrants show relatively lower sensitivity to pollution, reinforcing the gender difference observed in Table 3.11.

Overall, marital status and the added complexities of cross-province migration significantly influence pollution-related migration decisions. While married, older, higher-educated, and cross-province migrants demonstrate the strongest aversion to pollution, there is no significant evidence of differing aversion to air pollution between male and female migrants.

TABLE 3.12: **Effect of Air Pollution on the Migrants Flow by Administrative Boundary and Gender**

Dependent Variable	Ln (Migration Flow+1)					
Subsample	Cross Province			Within Province		
	Overall	Male	Female	Overall	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)
Destination-origin ratio of $PM_{2.5}$	-0.483***	-0.356***	-0.343***	1.268	0.848	1.375
	(0.0899)	(0.0734)	(0.0706)	(1.000)	(0.926)	(0.959)
Covariates Included	Yes	Yes	Yes	Yes	Yes	Yes
Province-pair Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
N	17232	14749	14568	2589	2403	2444
Coefficients difference test		-0.54			-0.22	

Note: This table reports the air pollution effect on migration flow by migration routine (if migrating across the province) and gender. Columns (1), (2) and (3) are migration across the province; columns (4), (5) and (6) are migration within the province.
The province pairs are fixed to absorb the unobserved cultural and historical connection between destinations and origins. All control variables lagged one year and included bidirectional cities from origins to destinations.
Robust standard errors clustered at destination-city and are shown in parentheses. ***, **, and * representative the level of significance at 1%, 5% and 10%.
Source: CMDS 2017, NASA EOSDIS 2016, China City Yearbook 2016, and China Meteorological Data Sharing Service System (CMDSSS).

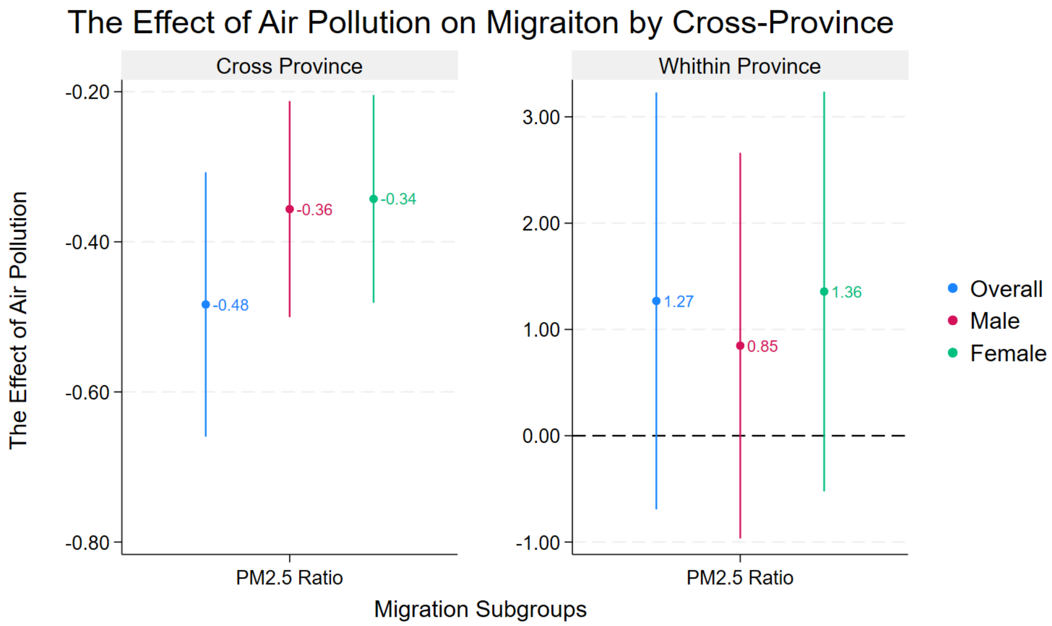


FIGURE 3.9: **Heterogeneity Analysis**
Note: This figure is the regression results of Table 3.12, showing the effect of air pollution on migration by marital status and gender. Point is the coefficient of air pollution and the line documents the 95% confidence interval.
Source: CMDS 2017, NASA EOSDIS 2016, China City Yearbook 2016, and China Meteorological Data Sharing Service System (CMDSSS).

3.5.5 Economic Cost of Air Pollution-Induced Migration Deterrence

To assess the direct economic cost of air pollution on migration decisions, I estimate the net benefit a potential migrant may forgo due to increased pollution exposure in the destination city. This calculation integrates three main components:

1. **Wage differentials:** The difference between the minimum wage in the destination and origin cities.
2. **Housing cost differentials:** Estimated as 2.5% of the annual housing price difference, following the rental-price ratio established by Tombe and Zhu (2019).
3. **Migration cost:** A one-time distance-based cost estimated at ¥0.4 per kilometre, based on average intercity transport costs from He and Ding (2019).

Formally, the one-year net benefit from migration is computed as:

$$\text{Net Benefit}_i = \text{Wage}_{\text{des}} - \text{Wage}_{\text{ori}} - 0.025 \times (\text{HP}_{\text{des}} - \text{HP}_{\text{ori}}) - 0.4 \times \text{Distance}_{\text{ori,des}} \quad (3.4)$$

Based on individual-level city-pair data and estimated parameters, I compute that the average net income gain forgone due to air-pollution-induced migration deterrence is approximately **¥10,179 per person per year**. This quantification underscores the substantial economic cost of air pollution in distorting labour mobility and hindering environmentally sensitive migration flows. Assuming an annual volume of 100,000 potential urban-to-urban migrants, a 42% reduction in migration flows—estimated from the gravity model—translates into a total annual economic loss of:

$$100{,}000 \times 10{,}179 \times 0.42 = \textbf{¥427.5 million}$$

This figure highlights the sizable welfare loss associated with pollution-driven misallocation of human capital in the urban labour market.

3.6 Robustness Check

This paper examines the robustness of the instrumental variable and the overall sample. Table 3.13 presents robustness check results for both the 2SLS and IVPPML models. Columns (1) to (3) use instrumental variables based on an alternative measure—the five-year average of the destination-to-origin thermal inversion strength ratio (2013–2017)—to validate the main results. These results remain consistent and significant, aligning closely with those in Table 3.4, thus confirming the reliability of the primary findings.

Additionally, Columns (4) and (5) use the logarithm of migration flows greater than zero as the dependent variable to test the robustness of the specification. Excluding all zero migration flows between cities reduces the sample size to 14,925. However, the results remain robust and statistically significant, supporting the strength of the original findings.

To further test the robustness of the air pollution effect, alternative pollution measures replace the $PM_{2.5}$ concentration. Table 3.14 reports results using different air quality indicators—Air Quality Index (AQI), PM10 (particulate matter with a diameter of 0.01 mm or smaller), sulfur dioxide (SO_2), and nitrogen dioxide (NO_2). All results remain consistent with those obtained using $PM_{2.5}$, reinforcing the main conclusions on pollution-driven migration.

Overall, the robustness checks across both instrumental variables and alternative pollution measures confirm the stability and reliability of the study's main results.

Table 3.13: **Robustness Check: Alternative IV Measurement and Non-zero Sample**

Robustness Check	Different IV Measurement			Reduced Sample	
	2SLS		IVPPML	2SLS	
	First stage	Second Stage		First stage	Second Stage
Dependent variable	$PM_{2.5}$ Ratio	Ln (Flow+1)	Flow	$PM_{2.5}$ Ratio	Ln Flow
	(1)	(2)	(3)	(4)	(5)
Destination-origin ratio of $PM_{2.5}$		-0.449***	-0.497*		-0.443***
		(0.119)	(0.264)		(0.135)
Ratio of Thermal Inversion Strength	0.260***			0.244***	
	(0.025)			(0.033)	
Covariates Included	Yes	Yes	Yes	Yes	Yes
Province-pair Fixed Effect	Yes	Yes	No	Yes	Yes
Kleibergen–Paap rk Wald F statistic	106.44			56.54	
N	25512		25512	14925	14925

Note: This table reports the robustness check for Table 2.4 and instrumental variables. Columns (1) and (2) are two-stage results of 2SLS. Column (3) is IV-PPML results with standardised coefficients. Columns (4) and (5) are two-stage results of 2SLS based on the non-zero migration flow subsample. All regressions applied the alternative instrumental variable - the average destination-origin ratio of thermal inversion strength covered the past five years (2013-2017)
All control variables are lagged one year and include bidirectional cities.
The province-pair effect is fixed. Standard errors clustered at destination city and are shown in parentheses.
***, **, and * representative the level of significance at 1%, 5% and 10%.

Source: CMDS 2017, NASA EOSDIS 2013-2017, China City Yearbook 2016, and China Meteorological Administration.

Table 3.14: **Robustness Check: Alternative Air Pollutants**

Dependent Variable	Ln (Migration Flow+1)			
	(1)	(2)	(3)	(4)
Destination-origin ratio of AQI	-0.700***			
	(0.164)			
Destination-origin ratio of $PM10$		-0.622***		
		(0.146)		
Destination-origin ratio of SO_2			-0.227***	
			(0.0580)	
Destination-origin ratio of NO_2				-0.805***
				(0.226)
Covariates Included	Yes	Yes	Yes	Yes
Province-pair Fixed Effect	Yes	Yes	Yes	Yes
K-P rk Wald F statistic	115.70	112.11	157.06	52.90
N	25512	25512	25512	25512

Note: This table reports a robustness check for alternative air pollution proxies based on 2SLS. Columns (1), (2,) (3,) and (4) used destination-origin ratios of AQI, $PM10$, SO_2 and NO_2, respectively.
All control variables are lagged one year and include bidirectional cities.
The province-pair effect is fixed. Standard errors clustered at destination city and are shown in parentheses.
***, **, and * representative the level of significance at 1%, 5% and 10%.
Source: CMDS 2017, NASA EOSDIS 2013-2017, China City Yearbook 2016, and China Meteorological Administration.

3.7 Conclusion

This study investigates the impact of air pollution on urban-to-urban migration in China, using bidirectional migration flow data from the China Migration Data Survey (CMDS) and the National Bureau of Statistics. The findings reveal that higher pollution levels in destination cities reduce migration inflows, with each doubling of the $PM_{2.5}$ concentration ratio between destinations and origins decreasing migration by approximately 42%.

In addition to quantifying the pollution effect on migration, this study explores the channels that influence migration decisions in response to air pollution. The results show that strict environmental regulations and high-quality infrastructure in destination cities can mitigate the adverse effects of pollution, whereas higher living costs and restrictive hukou policies exacerbate them. Interestingly, the study identifies a U-shaped relationship between pollution impact and migration distance, with pollution sensitivity declining for distances over 1,000 km. Demographic analysis reveals that older, less-educated, married, and cross-province migrants are less likely to choose highly polluted destinations.

These findings suggest important policy implications. Proactive environmental regulations and investments in green urban infrastructure could promote migration to cleaner cities, supporting sustainable urban development. Additionally, as air quality plays a significant role in attracting skilled migrants, improving air quality should be a priority for cities aiming to retain and attract talent. Finally, more flexible hukou policies could help reduce the negative impacts of pollution on migration and encourage the settlement of high-skilled workers, promoting human capital accumulation and long-term economic growth.

However, this study has certain limitations. The analysis relies on data from 2017, preventing an exploration of time trends. Additionally, the absence of prior migration histories may lead to an underestimation of pollution's long-term effects. Future research could incorporate longitudinal data to observe trends over time and explore the cumulative effects of pollution on migration patterns.

Overall, this study highlights air pollution as a significant factor in urban migration decisions, providing insights for policymakers on how environmental quality, infrastructure, and regulatory flexibility can enhance urban resilience and sustainable development.

Chapter 4

“Breathing Easy, Retiring Early?” Effects of Air Pollution on Retirement Age Expectations: Evidence from China

Abstract

This study investigates the impact of air pollution on retirement expectations in China, using panel data from the China Health and Retirement Longitudinal Study (CHARLS) and city-level PM2.5 concentration. Findings indicate that each 1% increase (0.12-1.04 $\mu g/m^3$) in PM2.5 concentration reduces the expected retirement age by 5.11 months, with more pronounced effects in rural areas compared to urban regions (9.03 months vs. 4.45 months). Also, the dynamic study of pollution shock implies that the sudden increase in pollution could dramatically aggregate the individuals’ adjustment to retirement expectations. The study also explores the mechanisms through which pollution affects retirement expectations, revealing that family financial support, green infrastructure, and knowledge capital mitigate pollution’s negative impact, while social welfare and insurance systems enable earlier retirement by enhancing financial security. A heterogeneity analysis shows that vulnerable groups—such as females, older adults, individuals with lower education levels, rural hukou holders, agricultural workers, and low-income earners—are particularly sensitive to pollution in terms of retirement planning.

Keywords: Air Pollution, PM2.5, Retirement Expectation, Aging Group, Health Economics, Environmental Economics

4.1 Introduction

4.1.1 Air Pollution and China

Rapid industrialisation and urbanisation, especially in developing economies, have led to deteriorating air quality. Air pollution has emerged as one of the most pressing environmental and public health challenges of the 21st century. Compared to soil and river pollution, air pollution exerts a more severe negative influence on individuals to communities because of its ability to spread and disperse widely (Feng et al., 2020). According to the estimation of the World Health Organisation, approximately 200,000 people die from air pollution-related diseases annually (Khuda, 2020) and related economic losses are around 60−270 million (Hossain et al., 2019). These alarming figures highlight not only the immediate threat to human health but also the far-reaching economic consequences for societies grappling with the burden of environmental degradation.

In China, the issue of air pollution is particularly acute due to rapid economic development. China has been facing quite serious air pollution problem for a long time, particularly, in some megacities, such as Beijing, Shanghai and Guangzhou (Li and Li, 2022). In 2015, air pollution in 96% of Chinese cities exceeded WHO air quality guidelines over four times higher than the safe level on average (Khanna et al., 2021). Since 2013, the Chinese government has implemented a national program to monitor and report air pollution, though the effectiveness of these controls remains inconsistent (Zeng et al., 2019).

Since the detrimental effects of prolonged exposure to air pollution on respiratory and cardiovascular health, chronic exposure to pollutants such as particulate matter ($PM_{2.5}$) and nitrogen dioxide (NO_2) has been linked to an increased risk of cardiovascular and respiratory diseases, which in turn impair workers' physical and cognitive capacities (Zivin and Neidell, 2012). Then, beyond these immediate health and economic influence, the cumulative impact of these health issues can reduce labour productivity and increase absenteeism, placing strain on both individual well-being and broader economic systems (Neidell, 2023). Furthermore, the pervasive impact of air pollution extends into multiple dimensions of human life, including outdoor activities, cognitive functioning, and long-term well-being (Khanna et al., 2021; Shen and Sun, 2023; Zhang et al., 2017), shaping broader socio-economic behaviours, including work decisions, retirement timing, and healthcare usage (Graff Zivin and Neidell, 2013). This creates a feedback loop where deteriorating health leads to lower labour force participation, exacerbating economic inequality and reducing household income, especially among older adults nearing retirement.

4.1.2 Retirement and Pension System in China

China's retirement and pension system is characterised by rigid statutory thresholds, fragmented institutional coverage, and wide disparities between formal regulations and actual labour market practices. The statutory retirement age—60 for men, 55 for female white-collar workers, and 50 for female blue-collar workers—has remained unchanged since the 1950s (Fang and Feng, 2018b). However, actual retirement behaviour is far more complex and heterogeneous, shaped by individual health conditions, family responsibilities, income adequacy, and sector-specific employment arrangements.

For instance, many urban residents in the informal sector or self-employed workers continue to work beyond the official retirement age due to insufficient pension benefits, while some civil servants retire early with generous entitlements. In rural areas, where formal employment is rare, retirement is often a gradual and informal transition rather than a clear-cut exit from the labour force (Feng et al., 2019).

China's pension system consists primarily of two pillars: (1) the Urban Employees' Basic Pension (UEBP), which is contribution-based and covers formal urban workers; and (2) the Resident-based Basic Pension Scheme, a flat-rate, non-contributory programme targeting rural and informal urban residents. The coverage and generosity of these schemes differ substantially. According to Ministry of Human Resources and Social Security of the People's Republic of China (2021), by the end of 2020, around 500 million people were covered by the urban employee scheme, while 460 million were covered by the resident scheme. However, the average monthly benefit under the resident scheme was less than ¥170 RMB (approximately $25 USD), compared to over ¥3000 RMB in major urban pension systems.

Further complicating retirement planning is the limited portability of pension benefits across regions due to the hukou-based administration and local fiscal disparities. Migrant workers often face interrupted contribution records and cannot easily consolidate their pension accounts when moving between cities (Liu and Sun, 2016).

These structural challenges result in significant variation in both retirement timing and expectations, particularly among individuals in non-standard employment or with weaker institutional attachment. Thus, the statutory retirement age is a poor proxy for actual labour supply decisions. Instead, this study focuses on expected retirement age, which better captures individual preferences and perceived capacity to remain in the workforce under various economic, health, and environmental conditions.

This paper examines the impact of long-run exposure to air pollution on retirement expectations, revealing that each 1% increase in $PM_{2.5}$ concentration (approximately 0.12 to 1.04 μg/m³) reduces the expected retirement age by an average of 5.11 months. This negative effect is more pronounced in rural regions, where it results in an approximate reduction of 9 months, compared to 4.45 months in urban areas. At the same time, this paper checks individuals' responses to dynamic pollution shocks, and the results illustrate that sudden pollution increases could lead to further adjustments and shorten retirement expectations, particularly in urban areas where pollution fluctuations are more noticeable.

This paper also checked the effect of short-run air pollution shocks on retirement expectation adjustment. The results of shocks in 2-year to 5-year time windows show that pollution-induced retirement adjustments are gradual rather than immediate, and their effects vary significantly across different socioeconomic groups.

Additionally, this paper explores the potential channels through which air pollution affects retirement expectations. The findings indicate that a lighter household burden, higher levels of green infrastructure, and social capital through knowledge spillover can help mitigate the negative impact of air pollution on retirement expectations. In contrast, a stronger social welfare and insurance system may enhance financial security, enabling individuals to retire earlier.

Finally, this paper investigates heterogeneity based on demographic and socioeconomic characteristics, finding that women, older individuals, those with lower levels of education, rural hukou holders, workers in the agricultural sector, individuals without pensions, and low-income individuals are more sensitive to the effects of air pollution on retirement expectations. This paper contributes to the literature by examining the causal relationship between air pollution and expected retirement age using panel data of CHARLS 2011-2020 from China and check the dynamic shock of air pollution's effect on individuals' adjustment of retirement expectations.

This paper contributes to three aspects. Firstly, this is the first paper studying the effect on retirement expectations of people nearing retirement. While there is a robust body of research on the immediate health and economic effects of air pollution, significant gaps remain, particularly concerning its impact on retirement behaviour (Bound et al., 2010). Few studies directly examine how air pollution—an environmental health shock—affects retirement age expectations. Most research has focused on productivity losses, absenteeism, or healthcare costs, rather than the long-term decision to retire (Khanna et al., 2021; Neidell, 2023).

Second, this paper is the first to investigate the effect of dynamic air pollution shocks on individuals' adjustments to their retirement expectations. The findings provide robust empirical evidence that improvements in air quality can extend labour supply in the long term. By integrating environmental and economic perspectives, this research offers valuable insights into how environmental stressors may reshape retirement decisions, with implications for public health, labour policy, and the sustainability of social security systems.

Third, this study contributes to the literature by jointly examining the short-run (2-year) and medium-run (5-year) effects of air pollution on retirement expectations, providing a dynamic perspective on behavioural and structural responses. While previous studies often focus on immediate or long-term exposure, our approach reveals how individuals adjust retirement planning over time, highlighting both short-term reactions to environmental shocks and persistent effects driven by cumulative pollution exposure.

This paper is organised into six sections. Section 4.2 reviews the existing literature. Sections 4.3 and 4.4 present the data sources, data structure, descriptive statistics, and the methodology. In Section 4.5, I examine both the macro and micro channels through which air pollution affects individuals nearing retirement, as well as conduct a heterogeneity analysis. Section 4.7 provides robustness checks for the main findings.

4.2 Literature Review

Air pollution, particularly in the form of fine particulate matter ($PM_{2.5}$) and nitrogen dioxide (NO2), has been extensively studied for its deleterious effects on human health. Long-term exposure to air pollutants has been linked to increased risks of respiratory diseases, cardiovascular problems, and premature mortality (Apte et al., 2015). Studies such as those by Burnett et al. (2014) have quantified the global burden of disease attributable to ambient air pollution, estimating millions of premature deaths annually. This body of research indicates that air pollution is one of the most significant environmental health risks globally, especially in rapidly industrializing regions like China.

Importantly, these health effects are not evenly distributed across populations. Vulnerable groups, such as the elderly, individuals with pre-existing health conditions, and those in lower-income brackets, experience more severe impacts. Older adults, in particular, face higher susceptibility to pollution-related illnesses due to declining immune systems and pre-existing health conditions (Shi et al., 2016). Lower-income populations are also disproportionately affected as they tend to live in areas with higher pollution exposure and have less access to healthcare services (Boyce, 2013).

Beyond its health consequences, air pollution has wide-ranging economic effects, particularly in terms of labour productivity, absenteeism, and economic growth. A study by Chang et al. (2019) demonstrated that high pollution levels significantly reduce worker productivity, especially in outdoor jobs such as construction and agriculture. The authors found that increasing $PM_{2.5}$ concentration by 10 g/m^3 decreases worker productivity by about 6%, which underscores the significant economic costs of pollution.

Similarly, Chen et al. (2020) explored how pollution affects cognitive performance, finding that prolonged exposure to air pollutants impairs cognitive function, particularly in older adults. This can hinder productivity in cognitive-intensive industries and exacerbate economic inequality, especially in developing economies like China. Furthermore, Neidell (2023) highlighted how pollution-induced health issues contribute to increased absenteeism, presenteeism (when employees are at work but underperform due to illness), and healthcare costs. These factors cumulatively lead to lower overall economic output and increased inequality as marginalised groups bear the brunt of these effects.

The relationship between air pollution and retirement decisions is complex and, at times, contradictory. On one hand, the adverse health effects of long-term pollution exposure often lead individuals to consider early retirement. Research by Maestas and Zissimopoulos (2010) indicated that health shocks, such as those caused by chronic respiratory or cardiovascular conditions — are significant predictors of early retirement. This is particularly relevant for older workers, who may be more susceptible to pollution-induced health deterioration.

On the other hand, the economic strain caused by pollution-related illnesses may compel individuals to delay retirement. Shen et al. (2021) noted that in regions with high pollution, workers might feel compelled to prolong their working lives to compensate for lost earnings or higher medical expenses. This creates a paradox: while health deterioration suggests earlier retirement, financial necessity encourages individuals to continue working despite declining health. Such conflicting incentives are particularly pronounced in developing countries like China, where pension systems may not provide sufficient income security for early retirees (Fang and Feng, 2018a). Thus, pollution creates opposing forces in retirement decisions. On the one hand, deteriorating health encourages early retirement, while on the other hand, financial insecurity pushes individuals to delay it. Few studies have explored these dual effects of air pollution on retirement, representing a significant gap in the literature.

The public health implications of air pollution in China have been extensively documented. For instance, Huang et al. (2019) found that prolonged exposure to high pollution levels in urban China has significantly reduced life expectancy, particularly in northern regions. Ebenstein et al. (2017) also noted that Chinese residents in more polluted areas suffer from higher rates of chronic illnesses and earlier mortality compared to those in cleaner regions.

China's ageing population adds another layer of complexity to the study of air pollution and retirement. As the country faces increasing pressure to reform its pension system, pollution-induced health issues among older workers could exacerbate financial strain on the government. Chen et al. (2023) argued that pollution-driven early retirements, combined with a rapidly ageing population, are likely to pose significant challenges to the sustainability of China's pension system. Therefore, understanding how pollution affects retirement decisions is crucial for both labour policy and social security planning.

4.3 Data

The individual and household data for this study are drawn from the China Health and Retirement Longitudinal Study (CHARLS), a nationally representative longitudinal survey focusing on Chinese adults aged 45 and above. CHARLS employs a four-stage, stratified, random sampling method to ensure that the data are representative of the Chinese population. CHARLS was initiated in 2011 and has conducted multiple biennial waves of data collection, with its most recent wave providing data through 2020, covering over 150 counties and 450 rural and urban committees with detailed demographic and socioeconomic variables. The survey has been widely studied in a bunch of topics, such as health status, physiological function, economic status, family structure, employment, and subjective recognition. including those that are vital for controlling confounding factors in the analysis. Therefore, CHARLS provides a suitable sample to study the people nearing retirement age and retirees. Based on this survey, this paper constructed a unique micro panel dataset to study the causality between air pollution and retirement decisions.

4.3.1 Retirement Expectation Data

In the CHARLS 2011 to 2020, there is a longitudinal question:

"*At what age do you plan to stop working? Stopping work in this context shall refer to having stopped all income-related activities, unpaid family business, and having no intention of engaging in anything more serious than small pastime work.*"

This question covers different job types, including agricultural and non-agricultural jobs, providing sufficient representativeness of data. The survey records the response of individuals aged 45 and above to this question, and the answer is used as the dependent variable to capture the expected age of retirement and the dynamic change in retirement expectation.

In the Chinese context, the statutory retirement age is a weak proxy for actual labour force participation due to both institutional constraints and behavioural diversity. First, the official retirement ages—60 for men, 55 for female cadres, and 50 for female workers—are low by global standards. As a result, many individuals continue working after retirement, especially in the informal sector. In the CHARLS Sample, only 48% of individuals aged 45 and above report some form of post-retirement economic activity, such as self-employment, casual labour, or helping with family businesses.

Second, the retirement age is primarily binding in the public sector and among employees in large state-owned enterprises. In the private sector and informal economy, where nearly half of China's urban employment now resides, retirement norms are more flexible. Workers in these sectors often retire later than the statutory age or continue working without formal retirement at all. This sectoral discrepancy renders the official retirement age less informative as an indicator of labour supply.

Third, for public-sector employees, retirement is mandatory once the statutory age is reached, regardless of health status or personal willingness to remain employed. This institutional rigidity makes actual retirement age an administratively imposed outcome rather than a reflection of individuals' preferences or constraints.

In contrast, expected retirement age serves as a forward-looking, subjective indicator that better captures individuals' labour supply intentions. It incorporates personal assessments of work capacity, financial need, health risk, and job satisfaction. This approach is particularly relevant in China's dual-track labour market. Recent survey evidence supports this view: a 2024 IZA study finds that the average expected retirement age among urban workers is 59.87 years — 61.03 for men and 57.17 for women, all above statutory thresholds. Moreover, 50.6% of respondents reported plans to delay retirement, and 15.9% intended to postpone retirement by at least five years due to financial or family pressures (Liu et al., 2024).

This measure is especially relevant in China's context of fragmented pension coverage. By the end of 2020, the Urban Employees' Basic Pension (UEBP) covered 500 million people, while the Resident-based Basic Pension covered another 460 million—but the latter provides only minimal income, and pension portability remains limited across regions and sectors (Ministry of Human Resources and Social Security of the People's Republic of China, 2021). These institutional frictions make long-term retirement planning more uncertain, especially for migrants and informal workers.

Therefore, this chapter uses expected retirement age as a more reliable measure of retirement behaviour. It is particularly suitable for capturing the effect of air pollution, which can alter perceptions of health risks and influence work-life planning, especially among mobile or vulnerable populations not fully protected by formal retirement structures.

4.3.2 Air Pollution Data

To capture air pollution exposure, this study uses satellite-derived $PM_{2.5}$ concentration data provided by Washington University in St. Louis. $PM_{2.5}$ — particulate matter with a diameter of 2.5 micrometres or less — is widely regarded as one of the most harmful air pollutants due to its ability to penetrate deep into the lungs and enter the bloodstream, increasing the risk of respiratory and cardiovascular diseases (Brook et al., 2010). It also serves as a key proxy for overall air quality, as it originates from various anthropogenic sources such as industrial processes, vehicle emissions, and fossil fuel combustion. Long-term exposure to $PM_{2.5}$ has been linked to elevated risks of mortality and chronic health conditions (Pope III and Dockery, 2006), making it particularly relevant for studying long-term outcomes like retirement expectations.

The $PM_{2.5}$ dataset integrates satellite observations (MODIS, VIIRS, MISR, SeaWiFS) processed using retrieval algorithms (Dark Target, Deep Blue, MAIAC), atmospheric simulations from the GEOS-Chem model, and calibration with ground-based monitors and AERONET data. A final statistical fusion yields high-resolution estimates of annual $PM_{2.5}$ concentrations at the city level across China, capturing both spatial and temporal variation.

In this study, pollution exposure is assigned at the city level, where each individual within the same city-year is assigned the same log $PM_{2.5}$ value, regardless of their urban or rural hukou status. This approach is consistent with the spatial resolution of the pollution data and standard practice in the literature using satellite-based or city-level administrative exposure measures. While this limits the ability to observe intra-city variation, the city-wide average reflects integrated exposure across the full municipal jurisdiction, including both urban districts and surrounding rural subregions. Given the spatial resolution and integrated nature of the pollution data, this assignment approach minimises systematic bias in estimated exposure between rural and urban residents within the same city.

It is crucial to note that China's urban–rural classification differs significantly from that in Western contexts. In China, the distinction between rural and urban areas is primarily based on administrative boundaries and household registration (hukou), not geographical remoteness. Most rural residents live within the same administrative cities as urban dwellers, often in close proximity to industrial clusters and traffic networks. Therefore, rural and urban populations are typically exposed to similar environmental conditions within a given city. The use of city-wide $PM_{2.5}$ averages, rather than imposing differential assignments, provides a consistent and policy-relevant measure of exposure.

Although this study does not model spatial heterogeneity in pollution within cities, it focuses on behavioural and institutional heterogeneity in response to a common environmental shock. In particular, the model allows for differential effects across urban and rural resident groups that may vary significantly in terms of economic vulnerability, healthcare access, and adaptive capacity, even when facing the same level of ambient pollution.

4.3.3 Control Variables

In this paper, covariates cover individual, household, city characteristics and climate data. For individuals and households, variables are from CHARLS, including age, gender, education levels, job type, marital and health status, individual and household socioeconomic characteristics, family structure and pension coverage. City characteristics, from the China Yearbook, include the city's economic development, employment market and infrastructure levels, such as GDP per capita, unemployment rate, and the number of hospitals and industry wastewater discharges. Also, since weather is associated with air pollution, the climate variables from the China Meteorological Administration are included.

Table 4.1 reports the description of the variable, and Table 4.2 presents descriptive statistics for the full sample and by urban-rural subsamples. The total sample includes 6,197 individuals aged 45 and above, with 2,582 urban and 3,615 rural residents. Significant urban–rural differences are observed across most variables.

On average, rural respondents report a substantially higher expected retirement age (65.7 years) than urban respondents (61.4 years), a gap of over four years that is statistically significant at the 1% level. This may reflect limited pension coverage and financial insecurity in rural areas, pushing individuals to plan for later retirement.

Individual-level characteristics also differ meaningfully across the two groups. Rural residents are older (58.6 vs. 54.8), have more children (2.46 vs. 1.77), are less likely to hold pensions (45.7% vs. 50.3%), and are more likely to engage in agricultural work. Educational and health indicators (e.g., chronic illness) also vary, highlighting socioeconomic disparities.

City-level variables show considerable heterogeneity but more moderate urban–rural gaps. Rural respondents are associated with lower GDP per capita, lower average wages, and weaker environmental infrastructure (e.g., lower wastewater discharge volumes), but only minor differences in temperature or population density.

The average log air pollution level is 3.79. Since pollution exposure is measured at the city level, all individuals within a city share the same $PM_{2.5}$ value, regardless of their urban or rural status. Therefore, no within-city variation in exposure exists across urban and rural groups in this dataset. However, their behavioural responses to the same pollution level may still differ, given disparities in health infrastructure, pension access, and socioeconomic vulnerability. This modelling strategy enables us to isolate the differential effects of a common environmental shock under different institutional and socioeconomic conditions, a relevant dimension for policy targeting and cross-country comparison.

These descriptive patterns underscore the importance of accounting for urban–rural heterogeneity in the analysis of retirement expectations and pollution exposure. The observed differences in socioeconomic characteristics and institutional access also support the relevance of subgroup analysis in the empirical sections that follow.

TABLE 4.1: **Description of Variables**

Variable	**Description**
Dependent variable	
Expected retirement age	At what age does individual plan to stop working?
Interest variable	
Ln Air Pollution	Logarithm of the past five-year average PM2.5 before the survey year
Control variables	
- Individual level	
Age	The age of individual
Male	The gender of individual (male=1, female=0)
Education level	Illiteracy / primary school/ Junior and Senior middle school /over college
Marital status	Weather individual is married (Yes=1, No=0)
Rural region	Weather individual work and lived in rural regions (Yes=1, No=0)
The number of children	The number of alive children
Chronic illness	Weather the individual has chronic illness (Yes=1, No=0)
Health insurance	Weather the individual has health insurance (Yes=1, No=0)
Pension	Weather the individual has pension (Yes=1, No=0)
Family size	The number of family members living together
Family expenditure	The average expenditure of all family members per month ($kRMB$)
Agriculture work	Whether the individual worked in agriculture (Yes=1, No=0)
- City level	
Ln GDP per capita	Logarithm of GDP per capita
Population density	Population density ($10^6/km^2$)
Second industry share	The share of the second industry in GDP (%)
Third industry share	The share of the third industry in GDP (%)
The number of hospitals	Number of hospitals (k)
Wastewater disposal capacity	The capacity of wastewater disposal ($10^6 kg\ /year$)
- Weather Controls	
Rainfall	Annual average rainfall (mm)
Average temperature	Annual average temperature ($°C$)

Note: This table shows the description of all variables.
Source: CHARLS 2010, 2012, 2014, 2018, and 2020; CFPS 2014, 2016, and 2018; China City Yearbook 1980-2020; WshU

Table 4.2: **Descriptive Statistics**

Sample	Overall Sample (1)				Urban (2)	Rural (3)	Rural vs. Urban (4)
Variable	Mean	SD	Min	Max	Mean	Mean	t-test
Expected retirement age	63.87	6.748	50	100	61.36	65.66	-26.643***
Ln air pollution	3.79	0.358	2.7	4.5	3.797	3.791	0.703
Male	0.65	0.478	0	1	0.631	0.659	-2.238**
Age	57.00	7.363	45	80	54.80	58.58	-21.191***
Age2/100	33.03	8.723	20	64	30.43	34.89	-21.244***
Marital status	0.94	0.234	0	1	0.953	0.934	3.283***
Hukou	0.71	0.452	0	1	0.437	0.913	-43.867***
Agricultural work	1.41	0.492	1	2	1.707	1.198	45.643***
No. of children	2.17	1.126	0	10	1.774	2.461	-25.541***
Chronic illness	0.72	0.449	0	1	0.699	0.737	-3.289***
Health insurance	0.97	0.181	0	1	0.963	0.968	-0.996
Own pension	0.48	0.499	0	1	0.503	0.457	3.563***
Family size	3.10	1.471	1	11	3.126	3.080	1.246
Family expenditure	1.43	2.237	0.012	84	1.852	1.135	11.404***
Ln GDP per capita	10.61	0.566	8.6	12	10.75	10.51	16.877***
Second industry share	45.82	9.430	12	72	45.85	45.79	0.244
Third industry share	40.77	9.000	20	72	42.27	39.70	10.990***
Ln urban wage	10.82	0.357	9.8	12	10.84	10.80	5.155***
Unemployment rate	3.11	0.748	1.2	4.5	3.093	3.123	-1.536
Ln population density	5.84	0.970	2.3	7.7	5.856	5.820	1.467
No. of hospitals	202.80	150.986	29	1175	205.9	200.6	1.364
Wastewater discharge	6615.11	8078.544	132	96501	7753	5802	9.349***
Average Rainfall	0.96	0.475	0.15	2.5	0.975	0.948	2.141**
Average Temperature	14.05	4.675	-1.2	24	14.04	14.05	-0.0880
N		6197			2582	3615	

Note: This table reports the descriptive statistics of all variables. Column (1) is the overall sample, including mean, standard deviation, minimum and maximum observations. Columns (2) and (3) are urban and rural subsamples respectively. Column (4) is the t-test for the urban-rural difference.
Source: CHARLS 2010, 2012, 2014, 2018, and 2020; China City Yearbook 1980-2020; NASA; WshU 2022

4.4 Methodology

4.4.1 Empirical Specification

This study employs two models to investigate the impact of air pollution on retirement expectations: a panel data model for long-term effects and a cross-sectional model for short-term adjustments to pollution shocks.

4.4.1.1 Long-term Exposure and Expectation

According to CHARLS 2011 to 2020 merging with air pollution data, this paper constructed the unbalanced panel data. To examine the effect of air pollution on the expected retirement age of individual i in city c in year t, the model is as below:

$$\begin{aligned}\text{Expected Retirement Age}_{i,c,t} = \beta_0 + \beta_1 \ln\left(\overline{\text{Air Pollution}}_{i,t,t-5}\right) + \mathbf{X}'_{i,t-1}\beta \\ + \mathbf{Z}'_{c,t}\theta + \alpha_i + \upsilon_c + \gamma_p + \theta_{\text{age}} + \lambda_{p,t} + \epsilon_{i,t}\end{aligned} \tag{4.1}$$

Where:

- $\ln\left(\overline{\text{Air Pollution}}_{i,c,t,t-5}\right)$ is the logarithm of the average $PM_{2.5}$ concentration over the past five years, i.e., from year $t-5$ to t. This approach better captures long-term effects and reduces measurement error from short-term volatility caused by single-year outliers.

- $\mathbf{X}$ represents individual/household-level covariates, which provide demographic background, including age, gender, marital status, education levels, rural residence, and whether the individual has a chronic illness, pension, health insurance, family size, the number of children, and family expenditure.

- $\mathbf{Z}$ denotes city-level control variables, reflecting the city's economic and infrastructure conditions, such as the logarithm of GDP per capita, population density, the share of the second and third industries, wastewater discharge capacity, the number of hospitals or clinics, and climate factors like rainfall and temperature. All city-level controls are lagged by one year to capture delayed effects and account for sufficient exposure to the objective group.

- α_i, υ_c, γ_p, θ_{age}, and $\lambda_{p,t}$ represent fixed effects for individual, city, province, time trend, age cohort, and province-by-year interactions, respectively. $\epsilon_{i,t}$ is the error term.

4.4.1.2 Dynamic Shock and Expectation Adjustment

Furthermore, this paper also examines the impact of dynamic changes in air pollution on adjustments to retirement expectations. Using data from CHARLS, I construct a cross-sectional and Panel dataset. First, based on the first difference, I obtained the difference in air pollution and retirement expectation age between the two rounds of the longitudinal survey. Second, the change in $PM_{2.5}$ concentration is used as a proxy for the shock of air pollution. This approach allows us to examine how individuals' expectations dynamically change in response to air pollution shocks.

Unlike Equation 4.1, which examines long-term effects, Equation 4.2 captures both macro and micro shocks by applying differencing to both individual- and city-level variables. This approach assumes that city-level changes, such as fluctuations in pollution levels or economic conditions, serve as external shocks affecting all residents, while individual-level changes, such as shifts in marital status or employment, reflect personal economic and social shocks (Ebenstein et al., 2017; Neidell, 2023).

By differencing both levels of variables, this methodology removes time-invariant unobserved heterogeneity, allowing for a more precise identification of how short-term economic and environmental fluctuations shape retirement expectations (Wooldridge, 2010). This approach ensures that broad external changes are captured while isolating individual-level behavioural adjustments, making it well-suited for studies analysing external shocks and their heterogeneous effects on retirement decisions (Deschênes and Greenstone, 2011; Graff Zivin and Neidell, 2013).

Hence, I apply repeated cross-section data to examine the effect of dynamic pollution shocks on individual i in year$t$ in city c is specified as:

$$\begin{aligned}\Delta\text{Expected Retirement Age}_{i,c,t} = \beta_0 + \beta_1 \Delta \ln\left(\text{Air Pollution}_i\right) + \mathbf{X}_i'\beta \\ + \Delta\mathbf{Z}_c'\theta + \upsilon_c + \gamma_p + \theta_{\text{age}} + \epsilon_i\end{aligned} \tag{4.2}$$

Where:

- $\ln\left(\text{Air Pollution}_i\right)$ is the logarithm of the average $PM_{2.5}$ concentration change over the past five years.
- $\mathbf{X}$ represents individual and family-level covariates.
- $\Delta\mathbf{Z}$ denotes the change in city-level covariates over the past five years to capture macroeconomic shocks.
- θ_{age} controls for differences across age cohorts.
- ϵ_i is the error term.

Also, despite the limited sample size in my first-difference model, I still estimate the fixed effects model based on Equation 4.2 for comparison, as follows:

$$\begin{aligned}\text{Expected Retirement Age}_{i,c,t} = \beta_0 + \beta_1 \ln\left(\text{Air Pollution}_{i,c,t}\right) + \mathbf{X}'_{i,c,t}\beta \\ + \mathbf{Z}'_{c,t}\theta + \mu_i + \upsilon_c + \gamma_p + \lambda_t + \epsilon_{i,c,t}\end{aligned} \tag{4.3}$$

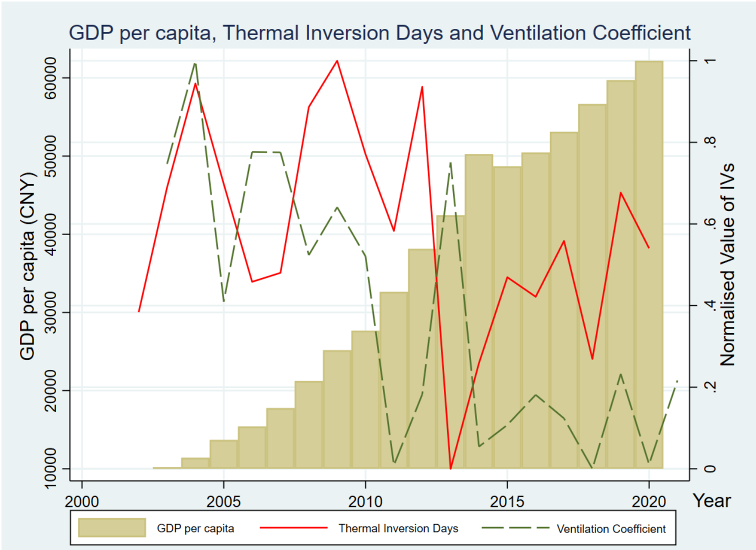
4.4.2 Identification

The identification strategy in this study leverages instrumental variables (IV) to account for the potential endogeneity between air pollution and retirement expectations. This suggests that pollution exposure might be endogenous, as individuals may move to less polluted areas due to declining health or other factors, causing reverse causality. For the long-run effect of air pollution, thermal inversion is applied. Thermal inversions are meteorological events that trap pollutants near the ground, leading to exogenous variations in pollution levels that are independent of human actions or local policies (Deschênes and Greenstone, 2011). Since these inversions affect pollution but not retirement decisions directly, they serve as a valid instrument for identifying the causal impact of long-term pollution exposure on retirement planning. By using this IV approach and lagging city-level variables to control for confounding factors, the analysis isolates the effect of pollution on expected retirement age. This approach has been applied successfully in studies examining environmental impacts on health and economic behaviour (Deschenes et al., 2020; Lavy et al., 2012), ensuring robust estimation of the causal relationship between air quality and retirement decisions.

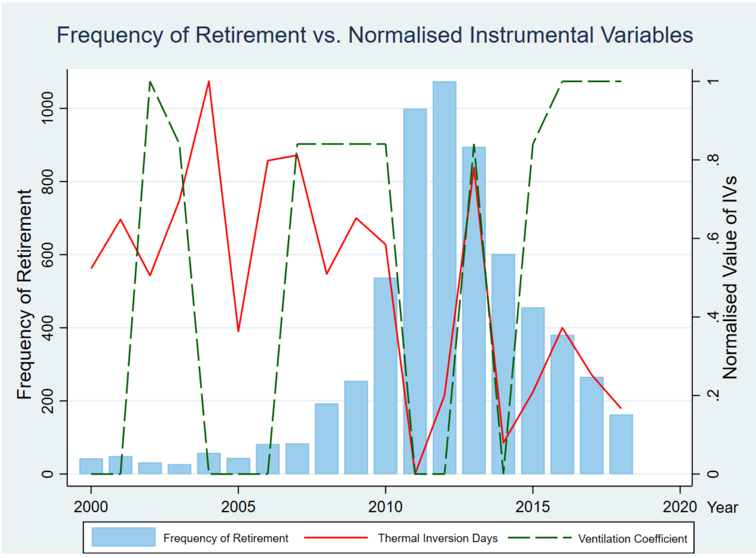
Furthermore, since thermal inversion data is recorded four times per day, its limited variation primarily captures long-run effects but is relatively insufficient for capturing short-term fluctuations and dynamic shocks. Therefore, I employ the ventilation coefficient (VC) as an instrumental variable to address potential reverse causality. The VC measures ventilation strength; a higher VC facilitates pollutant dispersion, results in cleaner air, and is widely used as an indicator in pollution studies, as discussed in Section 4.6.

Figure 4.1 presents the exclusion restriction check for both instrumental variables. Subfigure 1a illustrates the correlation between the IVs and national GDP, while Subfigure 1b shows the relationship between the distribution of retirement timing in the sample over the year and the IVs. Neither subfigure indicates a strong or significant correlation between the IVs and economic development, nor do they show any significant correlation between thermal inversion or ventilation coefficient and the distribution of retirement years.

Another identification threat is the potential selection bias from unbalanced panel data. In this case, individuals who leave the sample (missing in the following survey waves) may be due to deteriorating health, migration to different cities, or changes the employment status, raising selection bias. Table 4.12 shows the Heckman selection test to correct the selection bias. The insignificant inverse Mills ratio documents that selection bias is not a strong and main endogeneity problem.



(A) GDP vs. IVs



(B) Retirement Frequency vs. IVs

FIGURE 4.1: **The Time Trend of Thermal Inversions Accumulated Days, Frequency of Retirement and Economic Development**

Note: Subfigure (A) illustrates the relationship between the national average thermal inversion days, ventilation coefficients, and national GDP per capita over the years. The lack of a strong correlation between instrumental variables and GDP per capita suggests that thermal inversion phenomena and ventilation coefficients are largely independent of economic development.

Subfigure (B) presents the distribution of retirement flows in the sample alongside the average thermal inversion days and ventilation coefficients. Interestingly, no discernible pattern emerges to indicate a strong relationship between retirement behaviour and either thermal inversion days or ventilation coefficients.

Thermal inversion days and ventilation coefficients are normalised in all figures.

Source: CHARLS, NASA, WshU, China City Yearbook.

4.5 Empirical Results

4.5.1 Long-run Air Pollution Effect

Table 4.3 presents the long-run effect of air pollution on retirement expectations. Columns (1) to (3) report estimates from the fixed effects model. However, there may still be concerns about endogeneity, such as reverse causality or unobserved factors influencing both pollution exposure and retirement decisions. Columns (4) to (6) show 2SLS results with the instrumental variable - cubic transformation of thermal inversion days to address this and ensure a valid causal relationship. Both approaches provide results for the overall sample as well as for urban and rural subsamples. Fixed effects for individuals, provinces, years, age cohorts, and province-by-year interactions are included to ensure robust results, controlling for both time-invariant and regional heterogeneity.

The OLS results (columns (1) to (3)) indicate that increasing air pollution is significantly associated with a reduction in the expected retirement age, supporting the notion that long-term pollution exposure accelerates retirement planning. Specifically, a 1% increase in $PM_{2.5}$ concentration leads to a 0.121-year reduction (approximately 1.45 months) in expected retirement age for the overall sample, 0.111 years (1.33 months) for urban residents, and 0.2 years (2.4 months) for rural residents, all significant at the 1% level. These findings align with prior research and further support the notion that air pollution plays a critical role in shaping retirement decisions, particularly in rural areas where health impacts may be more severe and healthcare resources more limited (Chen et al., 2013b; Ebenstein et al., 2015).

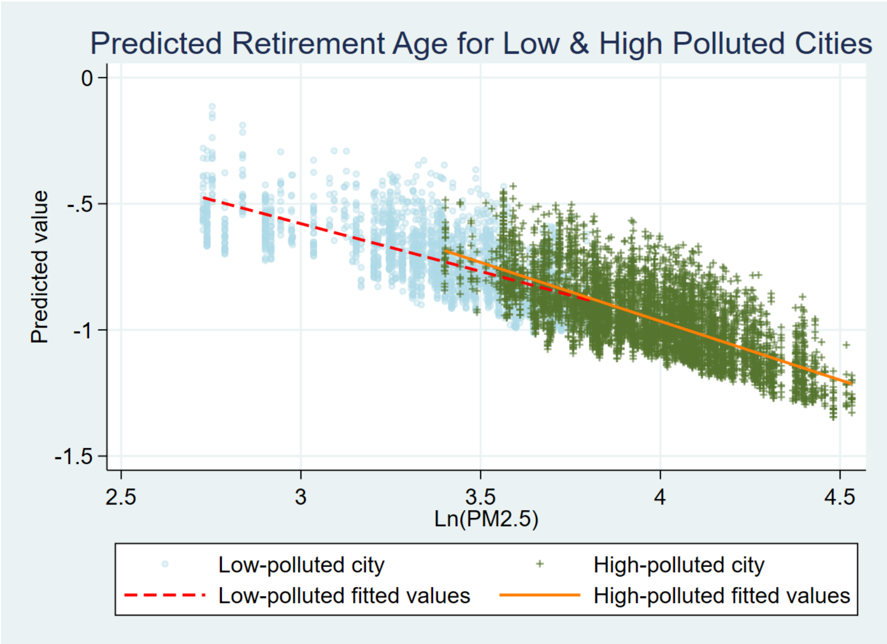
The 2SLS estimates (columns (4) to (6)) suggest an even stronger negative effect of long-run air pollution exposure on retirement expectations after addressing potential reverse causality. A 1% increase in air pollution is associated with a reduction of 0.426 years (5.12 months) in expected retirement age for the overall sample, 0.371 years (4.45 months) for urban areas, and a much larger reduction of 0.753 years (9.03 months) for rural areas, all statistically significant at the 1% level. These findings suggest that while air pollution significantly reduces expected retirement age across all groups, the effect is most pronounced in rural areas, likely due to higher exposure levels and lower healthcare accessibility.

Figure 4.2 illustrates the reductions in expected retirement age associated with varying levels of air pollution. Subfigure (a) presents fitted lines and patterns for high- and low-pollution cities, categorised according to the air quality index from the China Meteorological Administration. The steeper slope for high-pollution cities suggests that people living in these areas are more sensitive to air pollution in terms of retirement expectations. Subfigure (b) depicts the reactions of urban and rural residents to air pollution, showing that rural residents exhibit a more pronounced aversion to air pollution.

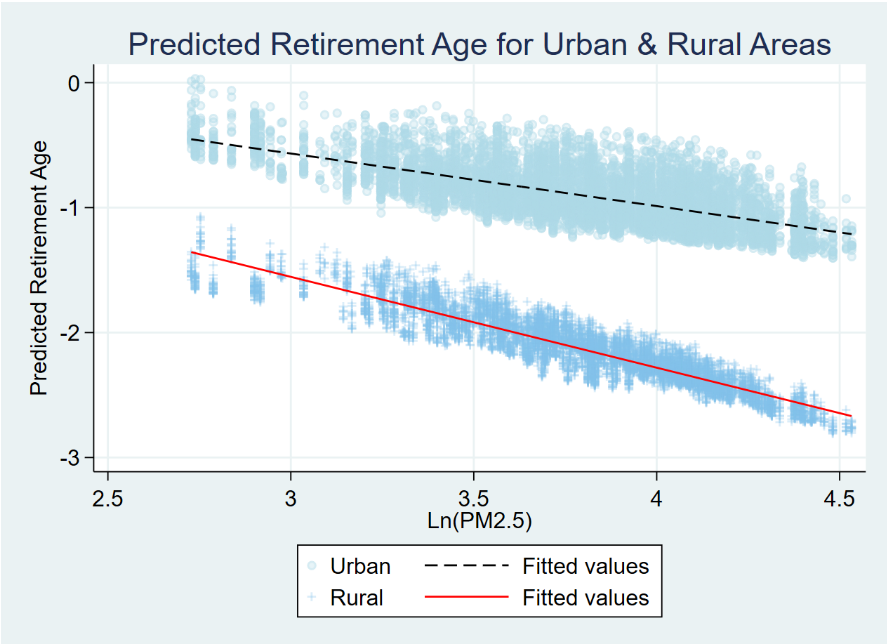
TABLE 4.3: **The Effect of Long-run Exposure to Air Pollution on Expected Retirement Age**

Dependent Variable	Expected Retirement Age					
Method	Fixed Effect			Fixed Effect + 2SLS		
Sample	Overall	Urban	Rural	Overall	Urban	Rural
	(1)	(2)	(3)	(4)	(5)	(6)
ln Air Pollution	-12.102***	-11.082***	-20.031***	-42.585***	-37.083***	-75.300***
	(2.984)	(3.014)	(5.471)	(12.882)	(10.384)	(21.792)
Fixed Individual	Yes	Yes	Yes	Yes	Yes	Yes
Fixed City	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Province	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Cohort	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year*Province	Yes	Yes	Yes	Yes	Yes	Yes
K-P F-statistic				12.16	19.64	16.99
R_2	0.848	0.848	0.829			
N	6128	2540	3588	6128	2540	3588

Note: This table reports the long-run air pollution exposure effect based on OLS baseline results and 2SLS results: the dependent variable is the expected retirement age. Both Panels report estimates based on log-transformed air pollution levels, where columns (1) to (3) and (4) to (6) present results for the overall sample, urban subsample, and rural subsample, respectively.
logarithm of $PM_{2.5}$ concentration is instrumented by the 5-year average annual thermal inversion days. The fixed effects models control for individual, province, year, cohort, and year-province interaction terms to account for unobserved heterogeneity at both the individual and regional levels.
The city-level clustering robust standard errors are in parentheses; ***, **, * represent the level of significance at 1%, 5% and 10% respectively.
Source: CHARLS 2011-2020, NASA, WshU, China City Yearbook



(A) The Predicted Value and Pollution by High and Low Polluted Cities



(B) The Predicted Value and Pollution by Rural and Urban Areas

FIGURE 4.2: **The Predicted Value against Pollution Level**
Note: This figure includes two subfigures: Subfigure (a) shows the predicted value and pollution by high- and low-polluted cities; Subfigure (b) shows the predicted value and pollution by rural and urban areas. All figures illustrate the relationship between the predicted retirement effect and air pollution by regions. In (a), the high-polluted individuals show a higher sensitivity to air pollution than the low-polluted individuals. The differing slopes of the lines in (b) indicate that air pollution has a greater impact on rural individuals compared to urban ones.
Source: CHARLS, NASA, WshU, China City Yearbook.

4.5.2 Short-run Dynamic Pollution Shock Effect

Tables 4.4 and 4.5 present the estimated effects of air pollution on changes in expected retirement age using Pooled OLS and 2SLS models for both 2-year and 5-year gaps, with results reported separately for the overall, urban, and rural samples. The Pooled OLS estimates reveal a significant negative relationship between pollution and expected retirement age, with the effects strengthening over time.

In the 2-year gap analysis, a 10% increase in air pollution reduces expected retirement age by $\frac{-0.064}{10} \times 365 = 2.3$ days for the overall sample, with a stronger effect in urban areas ($\frac{-0.088}{10} \times 365 = 3.2$ days) and an even larger reduction in rural areas ($\frac{-0.120}{10} \times 365 = 4.4$ days), significant at the 5% level. The 5-year gap analysis suggests an accumulation of the effect, where a 10% increase in pollution is associated with a $\frac{-0.215}{10} \times 365 = 7.8$-day reduction overall, with a particularly strong effect in urban areas ($\frac{-0.315}{10} \times 365 = 11.5$ days). These findings indicate that air pollution exerts both immediate and cumulative pressures on retirement planning.

The 2SLS estimates, which address potential endogeneity using the ventilation coefficient as an instrument, suggest that Pooled OLS may have underestimated the true impact of pollution. The IV estimates for the 2-year gap show a substantially larger reduction in retirement age, with a 10% increase in air pollution leading to an estimated decline of $\frac{-0.585}{10} \times 365 = 21.4$ days overall and an even stronger effect of $\frac{-1.037}{10} \times 365 = 37.9$ days in rural areas. The 5-year gap IV estimates remain large, with pollution-reducing retirement age by $\frac{-0.619}{10} \times 365 = 22.6$ days overall, and $\frac{-1.086}{10} \times 365 = 39.6$ days in rural areas. The larger IV coefficients suggest that omitted variable bias or measurement error may have led to downward-biased OLS estimates. The first-stage Kleibergen-Paap F-statistics, ranging from 10.43 to 20.99, indicate that the instrument is sufficiently strong, reducing concerns about weak instrument bias.

Taken together, these findings indicate that pollution not only triggers short-term adjustments in retirement expectations but also exerts an accumulating effect over time, reinforcing early retirement decisions. The stronger impact in rural areas suggests that limited access to healthcare and heightened exposure to pollution exacerbate the effect, necessitating targeted policy interventions. The significant difference between OLS and IV estimates underscores the importance of addressing endogeneity when evaluating pollution's labour market consequences. These results highlight the need for both short-term pollution control measures and long-term mitigation strategies to alleviate the economic burden of premature retirement induced by environmental degradation.

Table 4.4: The Short-Run Dynamic Effect of Air Pollution on Expected Retirement Age - Pooled OLS Results

Dependent Variable	Change in Expected Retirement Age					
Method	2-year Shock			5-year Shock		
Samples	Overall	Urban	Rural	Overall	Urban	Rural
	(1)	(2)	(3)	(4)	(5)	(6)
Δ ln air pollution	-0.064*	-0.088*	-0.120**	-0.121***	-0.215***	-0.176***
	(0.033)	(0.051)	(0.051)	(0.041)	(0.064)	(0.057)
Covariates Included	Yes	Yes	Yes	Yes	Yes	Yes
Fixed City	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Province	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Cohort	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year $\times$ Province	Yes	Yes	Yes	Yes	Yes	Yes
N	2384	1091	1288	1512	680	832
R^2	0.097	0.159	0.133	0.112	0.151	0.158

Note: The dependent variable is the change in expected retirement age. The interest variable is the annual average $PM_{2.5}$ concentration difference between different waves. Robust standard errors clustered at the city level are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.
Source: CHARLS 2011-2020, NASA, WshU, China City Yearbook.

Table 4.5: The Short-Run Dynamic Effect of Air Pollution on Expected Retirement Age - 2SLS Results

Dependent Variable	Change in Expected Retirement Age					
Method	2-year Shock			5-year Shock		
Samples	Overall	Urban	Rural	Overall	Urban	Rural
	(1)	(2)	(3)	(4)	(5)	(6)
Δ ln air pollution	-0.585***	-0.580***	-1.040***	-0.426***	-0.797***	-0.909***
	(0.191)	(0.177)	(0.353)	(1.392)	(2.692)	(2.613)
Covariates Included	Yes	Yes	Yes	Yes	Yes	Yes
Fixed City	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Province	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Cohort	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year × Province	Yes	Yes	Yes	Yes	Yes	Yes
K-P F-statistic	15.34	18.48	10.43	20.99	13.75	17.83
N	2384	1091	1288	1512	680	832

Note: The dependent variable is the change in expected retirement age. The interest variable, the annual average $PM_{2.5}$ concentration difference between different waves, is instrumented by the corresponding ventilation coefficient difference. All covariates are first-differenced. Robust standard errors clustered at the city level are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.
Source: CHARLS 2011-2020, NASA, WshU, China City Yearbook.

4.5.3 Channel Analysis

This paper also checks the potential channels influencing the effect of air pollution on retirement expectations. Specifically, I induce the four channels and their interaction terms with air pollution. The model is as below:

$$\begin{aligned}\text{Expected Retirement Age}_{i,c,t} = {} & \beta_0 + \beta_1 \ln\left(\overline{\text{Air Pollution}}_{i,t,t-5}\right) \\ & + \beta_2 \,\text{Channel}_{i,c} + \beta_3 \,\text{Channel}_{i,c} \times \ln\left(\overline{\text{Air Pollution}}_{i,t,t-5}\right) \\ & + X'_{i,t-1}\beta + Z'_{c,t}\theta + \alpha_i + \gamma_p + \theta_{\text{age}} + \lambda_{p,t} + \epsilon_{i,t}\end{aligned} \tag{4.4}$$

Where, apart from the same variables as Equation 4.1, Channel is the potential channel exogenous variable including family network, green infrastructure, social capital (knowledge spillover), and social welfare level:

- To capture the strength of the family network of individuals, I checked if they could receive money from their children and coded Yes as 1, and No as 0, covering individuals' comprehensive network in financial situation, feeding burden and family ties.
- The wastewater disposal rate captures green infrastructure because the wastewater disposal rate is not significantly associated with air pollution and precisely reflects a region's actual infrastructure level and disposal capacity, which is a long-run and consistent indicator.
- For social capital, the number of books in public libraries per hundred persons is used as an indicator. This measure aims to capture the spillover effect of knowledge (social capital) on residents, particularly regarding pollution awareness and green lifestyles, by emphasising the role of public libraries as a vital form of social capital. Libraries offer shared access to educational resources, enhancing community-level knowledge about environmental issues and sustainable practices. As repositories of information and interaction hubs, they foster knowledge transfer and social learning, enabling residents to make informed decisions about eco-friendly behaviours.

  This aligns with the concept of social capital discussed in economic literature, where institutions that promote trust and cooperation contribute to collective well-being (Ferguson, 2012; Vårheim et al., 2008). Moreover, recent research underscores that knowledge spillovers significantly influence societal outcomes, including environmental behaviours (Liang and Goetz, 2018). By democratizing access to information, libraries act as catalysts for building informed, environmentally conscious communities while encouraging greener lifestyles and reducing pollution. Thus, the indicator of books per hundred persons reflects not only access to knowledge but also the broader societal impacts fostered by public libraries as key contributors to social capital.

- Additionally, although I control the individuals' insurance status, the insurance coverage and investment in different regions are due to the government's financial status. Hence, the number of insured persons is used to capture the regional macro-insurance development level and insurance scale.

Table 4.6 represents the channel analysis results based on Equation 4.4. Columns (1) to (4) are family support, green infrastructure, knowledge spillover and capital and social welfare, respectively. Since our interest is the role of such channels in air pollution's effect on individuals' retirement expectations, the coefficients of interaction terms imply the substitute and complementary relationship between air pollution and channels.

The positive and significant in column (1), indicates that the stronger family network moderates the negative impact of pollution on retirement age. This finding is consistent with studies that show financial stability through family or otherwise allows individuals more freedom to avoid high-pollution or hazardous conditions, with more methods to avoid negative work and life conditions (Sun et al., 2019). In other words, family support provides a buffer that diminishes the worries of continuing to work under adverse environmental conditions.

In column (2), the green infrastructure mitigates the adverse effects of pollution on the expected retirement age. This finding is consistent with Kumar et al. (2019) – the high-quality green infrastructure could efficiently reduce the threats of air pollution.

The interaction between knowledge spillover and air pollution in column (3) suggests that the book numbers in libraries reduce the negative impact of pollution. The libraries, as one source of social capital, could facilitate awareness and adaptation to pollution risks, allowing individuals to adopt avoidance strategies that reduce pollution's impact on retirement age expectations. This is in line with the concept of knowledge spillover, where regions with greater access to information resources enable individuals to make more informed, more objective and useful methods to avoid the health issue (Hawe and Shiell, 2000; Parker and Kreps, 2005).

However, the negative interaction coefficient indicates that higher standards/scale of social welfare amplify the negative impact of pollution on retirement age. Individuals with access to better social support may feel financially secure enough to retire early when faced with high pollution levels. This phenomenon is consistent with studies highlighting that social safety nets can enable earlier retirement in response to adverse environmental conditions, as individuals are less reliant on continued employment for financial stability (Xie et al., 2023).

Table 4.6: **Channel Anslysis**

Dependent variable	Expected Retirement Age			
Channels	Family Networks	Green Infrastructure	Knowledge Spillover	Social Welfare
	(1)	(2)	(3)	(4)
Ln air pollution	-42.068***	-51.161***	-42.636***	-40.530***
	(13.101)	(16.003)	(12.541)	(12.533)
Children's financial supports	-0.484*			
	(0.291)			
Children's financial supports × Ln air pollution	0.131*			
	(0.079)			
Wastewater disposal rate		-0.673*		
		(0.402)		
Wastewater disposal rate × Ln air pollution		0.177*		
		(0.105)		
Social capital			-0.088**	
			(0.040)	
Social capital × Ln air pollution			0.022**	
			(0.010)	
Insurance scale				3.830
				(2.322)
Insurance scale × Ln air pollution				-1.011*
				(0.598)
Covariates Included	Yes	Yes	Yes	Yes
Fixed Individual	Yes	Yes	Yes	Yes
Fixed City	Yes	Yes	Yes	Yes
Fixed Province	Yes	Yes	Yes	Yes
Fixed Year	Yes	Yes	Yes	Yes
Fixed Cohort	Yes	Yes	Yes	Yes
Fixed Year ×Province	Yes	Yes	Yes	Yes
N	5879	6128	6128	6128

Note: This table reports channel analysis, and the dependent variable is the expected retirement age. Channels are children's financial support (family network), wastewater disposal rate (green infrastructure), number of books in public libraries per hundred persons (knowledge spillover), and the number of insured persons (insurance scale/social welfare coverage).
Both air pollution and interaction terms are instrumented by thermal inversion days[3], interacting with the channel variable.
The city-level clustering robust standard errors are in parentheses; ***, **, * represent the level of significance at 1%, 5% and 10% respectively.
Source: CHARLS 2011-2020, NASA, WshU, China City Yearbook.

4.5.4 Heterogeneity Analysis

This section explores the heterogeneity in the impact of air pollution on expected retirement age across demographic and socioeconomic groups, distinguishing between short-run dynamic shocks and long-run cumulative exposure. The long-run analysis considers sustained increases in $PM_{2.5}$, revealing how persistent pollution exacerbates vulnerabilities over time. The short-run analysis focuses on recent changes in air pollution ($\Delta \ln PM_{2.5}$ over the past five years), shedding light on individuals' immediate adjustments to sudden environmental shocks.

4.5.4.1 Heterogeneity Analysis in the Long-run Exposure to Air Pollution

Demographics– Gender, Age Education

Firstly, this paper checks the heterogeneity by demographic characteristics. Table 4.7 shows the results across gender, age, and education from columns (1) to (6). Since gender differences are critical in understanding the labour market's response to environmental hazards, columns (1) and (2) show the male and female reactions to air pollution. Women experience a larger reduction in expected retirement age due to a 1% increase in $PM_{2.5}$ (-0.70 years) compared to men (-0.33 years). While men are typically employed in more pollution-intensive jobs, women may be more vulnerable to the health effects of pollution due to differences in biological sensitivity and access to healthcare (Neidell, 2023).

Another crucial factor is age. While the sample focuses on ages over 45, the sensitivity to air pollution of groups closer to retirement age may be different. Hence, the overall sample is split into two subsamples based on the 55-age line, the legal retirement age in China. Younger workers (age $\leqslant 55$) exhibit a sharper decline in expected retirement age with a 1% increase in $PM_{2.5}$ (-0.57 years). Younger individuals are more likely to anticipate future health declines and adjust their retirement plans accordingly. However, older workers (age > 55) show an insignificant and smaller reduction (-0.29 years), as they are closer to retirement and less concerned with long-term health deterioration.

For different education levels, the sample is split into low schooling (primary school and below) and high schooling (secondary school and above). Workers with lower education levels experience a more significant reduction in expected retirement age, around 0.62 years in response to a 1% increase in $PM_{2.5}$. In contrast, individuals with higher education face a smaller and insignificant decline, around 0.28 years. This difference is likely due to the probability of exposure to manual labour and outdoor work. The highly educated group faces safer environments and has greater access to healthcare compared to the less educated group.

TABLE 4.7: **Heterogeneity by Demographics – Gender, Age & Education**

Dependent Variable	Expected Retirement Age					
Demographics	Gender		Age		Education	
Subsample	Male	Female	≤ 55	> 55	Low schooling	High schooling
	(1)	(2)	(3)	(4)	(5)	(6)
Ln air pollution	-33.086**	-70.339**	-56.921***	-29.060	-61.807***	-28.426*
	(14.001)	(34.775)	(16.943)	(18.597)	(20.642)	(15.634)
Covariates Included	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Individual	Yes	Yes	Yes	Yes	Yes	Yes
Fixed City	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Province	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Cohort	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year × Province	Yes	Yes	Yes	Yes	Yes	Yes
N	3987	2141	3448	2680	2931	3197

Note: This table represents the heterogeneity by demographics including gender, age, and education level.
The clustering robust standard errors are in parentheses. ***, **, * represent the significance levels at 1%, 5%, and 10%, respectively.
Source: CHARLS 2011-2020, NASA, WshU, China City Yearbook.

Socioeconomic– Hukou, Occupation, Pension, and Income

While this paper has split the rural and urban samples, the hukou status is often not consistent with residential addresses. Since hukou is the most significant threshold to access health care and other welfare resources, this paper also checks the hukou status. In Table 4.8 columns (1) and (2), rural hukou holders see a larger and more significant reduction in expected retirement age of 0.54 years compared to urban residents (insignificant 0.17 years). Again, the hukou system leads to significant disparities in access to healthcare and infrastructure between different groups.

Alternatively, occupations are highly associated with exposure to air pollution. Columns (3) and (4) in Table 4.8 show that agricultural workers are more sensitive to air pollution (0.57 years reduction) than non-agricultural workers (0.46 years), although both groups' reactions are 5% significant. In other words, the agricultural worker is more vulnerable in outdoor work conditions.

For financial status, columns (5) to (6) show the roles of pension and income in retirement plans. Pension coverage serves as a key buffer against the financial and health shocks associated with pollution – the individuals owning pensions (contribute and receive) are not as sensitive as the group without (insignificant -0.04 vs. 1% significant -0.62). Similarly, as a well-established determinant of health resilience, higher-income individuals have greater access to healthcare and protective resources. Hence, according to the average expenditure level, the low-income group shows a higher sensitivity to air pollution with a 10% significant 0.58 years of reduction, approximately double that high-income group with 0.33 years. Figure 4.3 shows the heterogeneity analysis based on the above analysis.

TABLE 4.8: **Heterogeneity by Socioeconomic – Hukou, Occupation, Pension & Income**

Dependent Variable	Expected Retirement Age							
Socioeconomic Category	Hukou		Occupation		Contribute to Pension		Income	
Subsample	Rural	Urban	Agriculture	Non-agriculture	Yes	No	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln air pollution	-53.061***	-17.166	-57.401**	-46.010**	-3.844	-61.552***	-58.163*	-33.420**
	(17.461)	(12.297)	(22.899)	(17.790)	(20.272)	(20.611)	(30.321)	(14.483)
Covariates Included	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Individual	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed City	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Province	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Cohort	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year × Province	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	4464	1660	3704	2422	1783	4345	3080	3048

Note: This table represents the heterogeneity by socioeconomic categories, including Hukou, occupation, pension contributions, and income levels. The clustering robust standard errors are in parentheses. ***, **, * represent significance levels at 1%, 5%, and 10%, respectively.
Source: CHARLS 2011-2020, NASA, WshU, China City Statistic Yearbook.

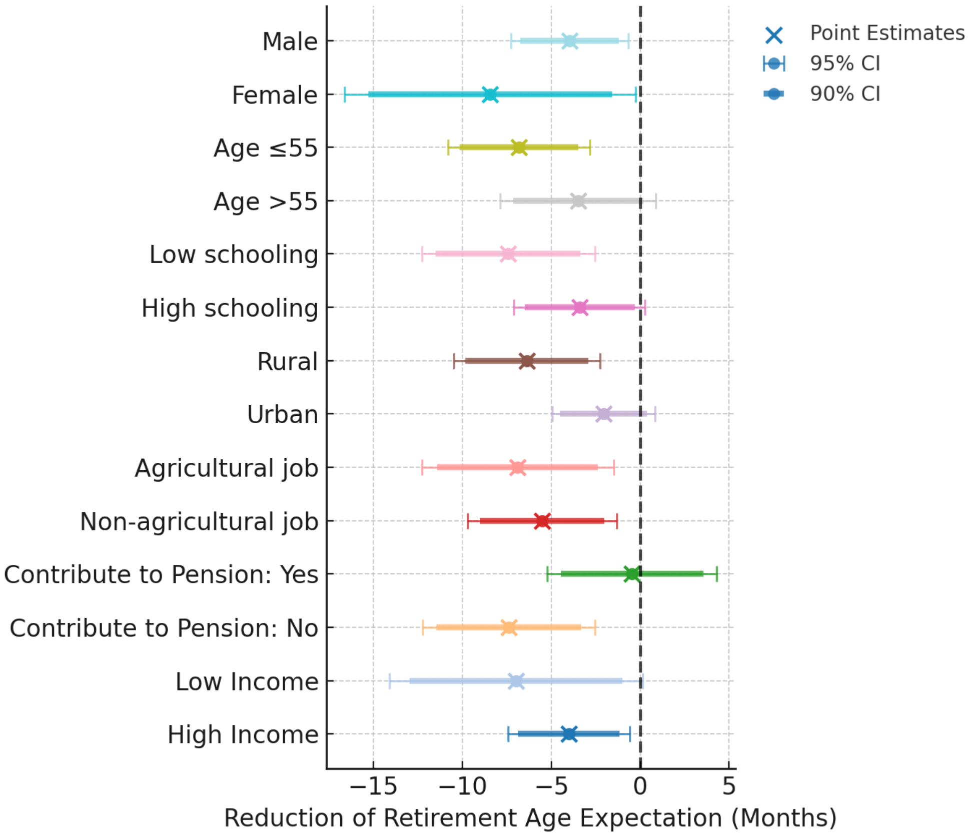


FIGURE 4.3: **Heterogeneity Analysis of Long-run Exposure to Air Pollution**
Note: This figure shows the heterogeneity analysis by demographic (gender, age, and education), and socioeconomic (hukou, occupation, pension, and income). The plots are standardized coefficients showing 90% and 95% confidence interval
Source: CHARLS, NASA, WshU, China City Yearbook

4.5.4.2 Heterogeneity Analysis in the Short-run Shock of Air Pollution

Tables 4.9 and 4.10 present the heterogeneous effects of short-run air pollution shocks on expected retirement age using two-year and five-year differencing approaches, respectively. Figures 4.4 and 4.5 are corresponding visualisations of heterogeneity results. The subsample analysis examines differences across age, education, occupation, pension contribution status, and income levels. While the general patterns remain consistent, notable differences arise in the magnitude and significance of estimated effects, suggesting that responses to pollution shocks vary across demographic and socioeconomic groups.

The results from the two-year differencing model in Table 4.9 indicate that younger individuals (≤ 55 years) exhibit a stronger response to pollution shocks compared to older individuals. A 10% increase in $PM_{2.5}$ concentration is associated with a 0.163-year (59.5-day) reduction in expected retirement age ($p < 0.1$) for younger individuals, while the estimated effect for those above 55 years is smaller and statistically insignificant. This suggests that younger workers may adjust their retirement expectations more readily in response to short-run pollution variations, whereas older workers, who are closer to actual retirement, may be less responsive to short-term environmental changes.

Education-based heterogeneity is not evident in the two-year model, as neither low nor high schooling subsamples show significant effects. Similarly, differences based on income level are minimal, with no statistically significant responses observed in either the low-income or high-income groups. However, occupational disparities emerge, as agricultural workers experience a significant reduction in expected retirement age of 0.079 years (28.8 days, $p < 0.1$), while non-agricultural workers do not exhibit a statistically significant response. This suggests that pollution-related health impacts may be more immediate for those engaged in physically demanding outdoor labour. Pension contribution status also appears to influence responsiveness to pollution, with non-contributors experiencing a much stronger, though imprecisely estimated, reduction in expected retirement age (0.310 years, or 113 days). The larger effect for those without pension contributions could reflect greater financial insecurity, making them more sensitive to pollution-induced health shocks that accelerate labour market exit.

Extending the time window to five years, the results in Table 4.10 provide a more stable and precise measure of the impact of air pollution on retirement expectations. The response among older individuals (> 55 years) becomes more pronounced, with a 0.203-year (74.2-day) reduction in expected retirement age ($p < 0.05$), while the effect on younger individuals declines to 0.097 years (35.4 days, $p < 0.1$). This shift suggests that while younger individuals react immediately to pollution, older individuals adjust their retirement expectations more gradually over time. Education-based differences also become clearer in the five-year model, as less-educated individuals experience a significant reduction in expected retirement age of 0.072 years (26.3 days, $p < 0.05$), while highly educated individuals remain largely unaffected. This aligns with previous findings suggesting that higher education enhances adaptability to environmental and economic shocks, possibly through better health awareness, financial planning, or job security.

Occupational differences are more pronounced in the five-year model, with non-agricultural workers showing a larger response (0.074 years $\approx$ 27 days, $p < 0.01$), which was not significant in the two-year model. This suggests that while pollution may immediately affect the health and retirement expectations of agricultural workers, urban workers experience a more gradual impact, potentially due to cumulative exposure in office-based environments or differences in healthcare access. The influence of pension contribution status remains consistent with the two-year results, as non-contributors experience a stronger reduction in expected retirement age, reinforcing the idea that individuals with weaker financial security are more vulnerable to environmental health risks.

Income-based heterogeneity becomes more apparent in the five-year model, where high-income individuals exhibit a stronger response (0.063 years $\approx$ 23 days, $p < 0.01$), while low-income individuals show no significant adjustment. This finding contradicts the common assumption that lower-income workers are more vulnerable to environmental shocks. A possible explanation is that low-income workers, despite experiencing greater health burdens, may lack the financial flexibility to adjust their retirement expectations, whereas higher-income workers, who can afford early retirement, react more strongly to pollution exposure.

Comparing the two time windows, the five-year differencing approach yields more stable and significant estimates, particularly among older individuals, less-educated workers, and non-agricultural occupations. The results suggest that pollution-induced retirement adjustments are gradual rather than immediate, and their effects vary significantly across different socioeconomic groups. The stronger response in the five-year model highlights the importance of considering longer-term adaptation mechanisms when assessing the economic consequences of environmental hazards.

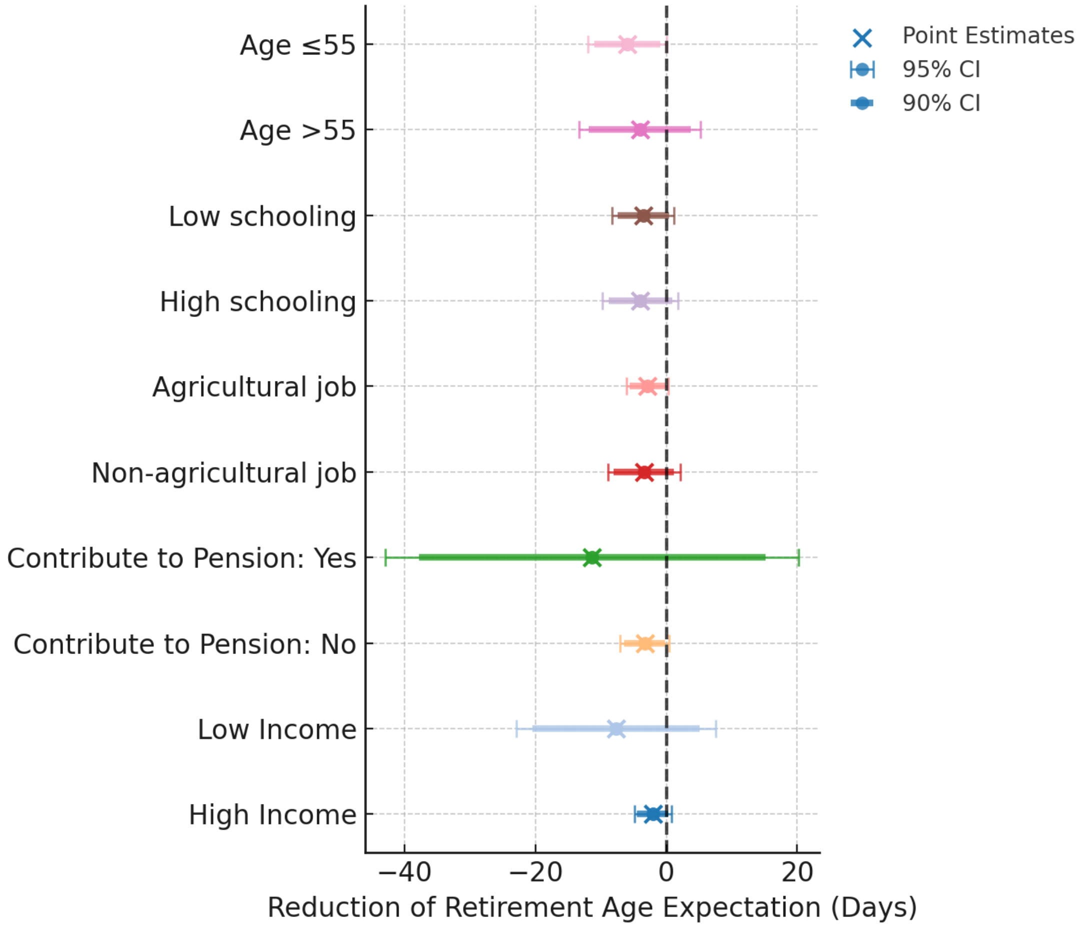


FIGURE 4.4: **Heterogeneity Analysis of Short-run Shock of Air Pollution (2-year gap)**

Note: This figure shows the heterogeneity analysis of Short-run Shock (2-year Gap) of Air Pollution. The plots are standardized coefficients showing 90% and 95% confidence interval

Source: CHARLS, NASA, WshU, China City Yearbook

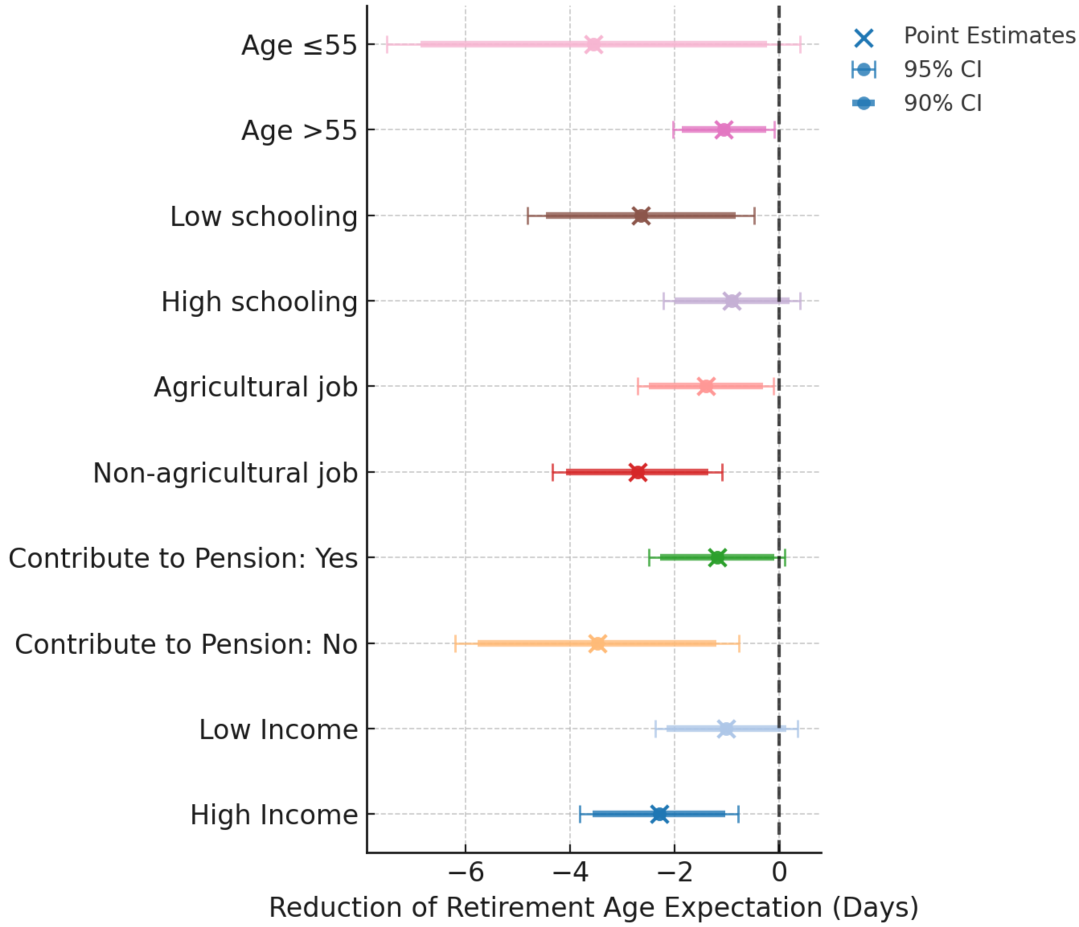


FIGURE 4.5: **Heterogeneity Analysis of Short-run Shock of Air Pollution (5-year gap)**

Note: This figure shows the heterogeneity analysis of Short-run Shock (5-year gap) of Air Pollution. The plots are standardized coefficients showing 90% and 95% confidence interval

Source: CHARLS, NASA, WshU, China City Yearbook

TABLE 4.9: **Heterogeneity of Short-run Dynamic Shock of Air Pollution (2-year Difference)**

Dependent variable	Expected Retirement Age									
Subsample	Age		Education		Occupation		Contribute to Pension		Income	
	≤ 55	> 55	Low schooling	High schooling	Agriculture	Non-agriculture	Yes	No	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ln air pollution	-1.632*	-1.106	-0.962	-1.088	-0.793*	-0.935	-3.098	-0.906*	-2.108	-0.557
	(0.838)	(1.291)	(0.660)	(0.806)	(0.449)	(0.770)	(4.409)	(0.524)	(2.123)	(0.399)
Covariates Included	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Province	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Cohort	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year × Province	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1065	1316	1030	1350	1315	1066	1109	1275	1097	1286

Note: This table represents the heterogeneity of the short-term dynamic (2-year time window) shock of air pollution.
The clustering robust standard errors are in parentheses; ***, **, * represent the level of significance at 1%,5% and 10% respectively.
Source: CHARLS 2011-2020, NASA, WshU, China City Yearbook

TABLE 4.10: **Heterogeneity of Short-run Dynamic Shock of Air Pollution (5-year Difference)**

Dependent variable	Change of Expected Retirement Age									
Subsample	Age		Education		Occupation		Contribute to Pension		Income	
	≤ 55	>55	Low schooling	High schooling	Agriculture	Non-agriculture	Yes	No	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ln air pollution	-0.973*	-0.290**	-0.724**	-0.247	-0.385**	-0.743***	-0.325*	-0.955**	-0.276	-0.629***
	(0.553)	(0.135)	(0.303)	(0.183)	(0.182)	(0.227)	(0.182)	(0.380)	(0.190)	(0.212)
Covariates Included	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Province	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Cohort	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year × Province	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	503	1007	350	877	811	699	1050	460	755	756

Note: This table represents the heterogeneity of the short-term dynamic shock (5-year time window) of air pollution.
The clustering robust standard errors are in parentheses; ***, **, * represent the level of significance at 1%,5% and 10% respectively.
Source: CHARLS 2011-2020, NASA, WshU, China City Yearbook

4.6 Robustness Check

4.6.1 Exclusive Restriction Check

To ensure the exclusive restriction makes sense, check the relationship between economic development and thermal inversion, and the correlation between the actual retirement age distribution of retirees in our sample and thermal inversion. Figure 4.1 shows the check for exclusive restriction. No matter the economic level, or retirement flow distribution is not associated with thermal inversion days, and exclusive restriction is not violated.

4.6.2 Short-Run Effect Check

To ensure the robustness of our findings, we employ an individual fixed effects (FE) model as a robustness check. While our main analysis relies on pooled OLS and 2SLS to estimate the short-run (2-year) and long-run (5-year) effects of air pollution on retirement expectations, the inclusion of individual FE helps control for time-invariant unobserved heterogeneity, such as inherent health conditions, risk preferences, or work motivations. However, using FE in the main analysis is not ideal due to the significant reduction in sample size, particularly in an unbalanced panel setting, where individuals may not be consistently observed across time. Additionally, FE models absorb within-individual variation, which can over-filter short-term pollution shocks, reducing statistical power. By applying individual FE as a robustness check rather than a primary specification, we ensure that our results remain consistent even after accounting for unobserved heterogeneity, while retaining sufficient variation in the main analysis to capture the dynamic short-run effects of air pollution. Table 4.11 shows the robustness check results with fixed effects at the individual level, which is consistent with the Pooled OLS and 2SLS results in Tables 4.4 and 4.5.

TABLE 4.11: **Fixed Effect and 2SLS Results**

Dependent Variable	Change in Retirement Age					
Method	Fixed Effect			Fixed Effect + 2SLS		
Samples	Overall	Urban	Rural	Overall	Urban	Rural
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Ln air pollution	0.007	-0.136	-0.026	-0.618**	-1.046**	-1.158*
	(0.070)	(0.090)	(0.158)	(0.249)	(0.457)	(0.664)
Covariates Included	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Individual	Yes	Yes	Yes	Yes	Yes	Yes
Fixed City	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Province	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Cohort	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year × Province	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.351	0.377	0.469			
K-P F statistic				14.04	9.09	7.34
N	1099	555	529	1099	555	529

Note: The dependent variable is the change in expected retirement age between five years, consistent with Table 4.3.
The interest variable, air pollution, changed over five years (2011-2016, 2015-2020)
The Kleibergen-Paap F-statistics are reported to assess the instrument's strength, with values above the general threshold of 10, and most are above 16.38.
The city-level clustering robust standard errors are in parentheses; ***, **, * represent the level of significance at 1%,5% and 10% respectively.
Source: CHARLS 2011-2020, NASA, WshU, China City Yearbook

4.6.3 Heckman Selection Test

The analysis is based on an unbalanced panel constructed from CHARLS, covering the period from 2011 to 2020. Of the 8,015 individual-wave observations in the initial dataset, only 6,128 are retained in the final regression sample due to non-response or attrition in specific survey waves. This implies that approximately 24% of the observations are missing from the outcome equation, primarily due to sample attrition or the inability to recontact participants across waves.

To assess whether the unbalanced nature of the panel introduces selection bias, I implement a Heckman two-step selection model. In the first-stage selection equation, a Probit model estimates the probability that an observation appears in multiple waves, using demographic characteristics and an owning-home-fixed-phone indicator as explanatory variables. The "Home With Fixed Phone" variable serves as an exclusion restriction, as it plausibly affects the likelihood of remaining in the panel but is not expected to influence retirement expectations directly. The second-stage outcome equation estimates the effect of air pollution on expected retirement age while incorporating the inverse Mills ratio derived from the first step to correct for any non-random selection.

The coefficient on the inverse Mills ratio is statistically insignificant, suggesting that sample selection bias is not a major concern in this context. Therefore, the results from the main analysis are unlikely to be driven by systematic differences between those who remain in the panel and those who do not.

TABLE 4.12: **Robustness Check - Heckman Selection Test**

Variable	First Step	Second Step
	Selection group	Expected retirement age
With Home Fixed Phone	0.0362	
	(0.062)	
Ln Air Pollution		-42.509***
		(12.877)
Inverse Mills Ratio (λ)		-0.882
		(4.506)
$\chi^2(24) = 219.34$	Prob $>\chi^2 = 0.0000$	
Controls Variable	Included	Included
N	8015	6128

Note: This table shows the Heckman selection test. The inverse Mills Ratio is insignificant, i.e., sample selection bias is not a serious issue in the model. In this case, the individuals selected in the sample are not systematically different from those who were not, in terms of the factors affecting the outcome.
***, **, * represents the level of significant at 1%,5% and 10% respectively.
Source: CHARLS 2011-2020, NASA, WshU, China Yearbook

4.6.4 Instrumental Variable Robustness Check

This paper also examines the robustness of the instrumental variable by using an alternative. Similar to thermal inversion, the ventilation coefficient (VC) is widely used in pollution studies. On one hand, VC is highly associated with air pollution as it determines the concentration of pollutants. For instance, regions with higher VC values indicate better ventilation, allowing pollutants to disperse more easily and resulting in cleaner air compared to regions with low VC values. On the other hand, the VC, derived from altitude and wind speed, is inherently exogenous as it reflects natural atmospheric conditions. These factors are unlikely to be directly influenced by human behaviours, economic conditions, or production patterns, ensuring the validity of the exclusion restriction.

Following the methods of Broner et al. (2012) and the investigation of Lai et al. (2021), I incorporate the VC as an alternative instrumental variable. The VC, which combines altitude and wind speed, has been extensively applied in atmospheric research and is constructed as follows:

$$\mathrm{VC}_{c,t} = \mathrm{WS}_{c,t} \times \mathrm{PBLH}_{c,t} \tag{4.5}$$

where, $\mathrm{VC}_{c,t}$ is the ventilation coefficient in city c in year t, WS is wind speed, and PBLH is the planetary boundary layer height.

The data is from the ERA5 atmospheric dataset from the Copernicus Climate Change Service (C3S) (2017) and calculated by ArcGIS. Figure 4.6 shows the strong positive correlation between the two instrument variables, and Table 4.13 reports the robustness check to IV by using VC as the new IV. While the size changes a little, the scale and significance are consistent with the baseline results in Table 4.3.

I also examined the exclusivity restriction of ventilation coefficients (VC). In Figure 4.1, I investigated the potential correlation between VC and economic or productive activities and retirement frequency. Additionally, I analysed the relationship between actual VC and the distribution of retirement flows within the sample. None of these figures exhibit strong patterns or associations. Therefore, based on the existing literature and robustness checks, I conclude that the exclusivity restriction holds.

TABLE 4.13: **Robustness Check – Alternative Instrumental Variable**

Dependent Variable	Expected Retirement Age					
2*Sample	Overall Sample		Urban		Rural	
	IV1	IV2	IV1	IV2	IV1	IV2
Variable	(1)	(2)	(3)	(4)	(5)	(6)
First stage						
Thermal Inversion Days	0.009***		0.015***		0.009***	
	(0.0027)		(0.0034)		(0.0022)	
Ventilation Coefficient		-0.003***		-0.005***		-0.002***
		(0.0005)		(0.0006)		(0.0004)
Second Stage						
Ln Air Pollution	-42.585***	-26.593***	-37.083***	-24.241***	-75.300***	-36.072**
	(12.882)	(6.584)	(10.384)	(5.863)	(21.792)	(16.721)
Fixed Individual	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Province	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Cohort	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Year × Province	Yes	Yes	Yes	Yes	Yes	Yes
K-P F-statistics	12.16	30.73	19.64	74.69	16.99	22.02
N	6128		2540		3588	

Note: This table checks the robustness of the instrumental variable.
The interest variable, air pollution, is instrumented using the ventilation coefficient.
ventilation coefficient = wind speed × planet boundary layer height
The Kleibergen-Paap F-statistics are reported to assess the strength of the instrument, with values above 16.38.
The city-level clustering robust standard errors are in parentheses; ***, **, * represent the level of significance at 1%,5% and 10% respectively.
Source: CHARLS 2011-2020, WshU 2022, China City Yearbook, C3S 2017

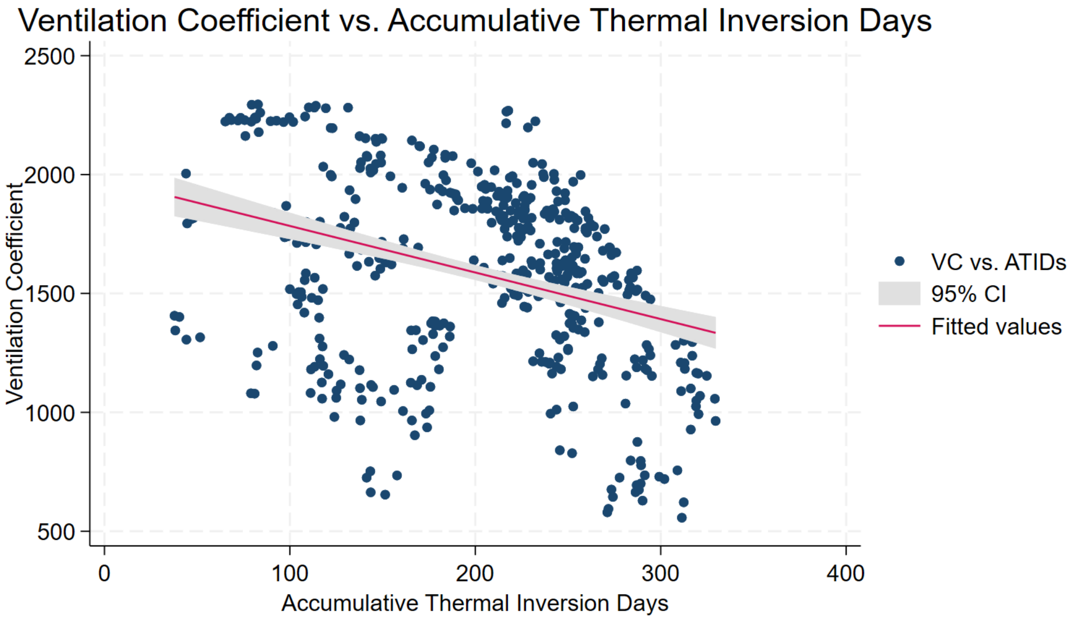


FIGURE 4.6: **The correlation between Instrumental Variables**
Note: This figure shows the correlation between thermal inversion days and ventilation coefficient. The red line shows the correlation with 95% confidence interval.
Source: NASA, Copernicus Climate Change Service

4.6.5 Different Lagged City-level Controls

I use different lagged city-level controls to check the robustness of the results, as shown in Table 4.14. Including lagged controls helps mitigate the influence of sudden shocks, such as unexpected policy changes or environmental disasters, that might distort city-level variables in the same period. Additionally, retirement decisions and behavioural changes often respond to conditions over time rather than immediately. By incorporating 3-year and 5-year lagged controls (columns (1) and (2)), the analysis captures these delayed effects and better reflects how residents process and act upon historical information.

Using lagged variables also avoids simultaneity bias, ensuring that the explanatory variables precede the outcome and reducing potential endogeneity concerns. This aligns with the idea that residents rely on accumulated and past knowledge of city-level conditions, such as infrastructure development or pollution trends, rather than reacting instantaneously. Importantly, the results remain consistent with the baseline estimates, demonstrating the robustness of the findings across different temporal frameworks.

TABLE 4.14: **Robustness Check – Different Lagged Controls**

Dependent variable	Expected Retirement Age	
Variables	(1)	(2)
Ln air pollution	-41.358***	-47.017***
	(12.505)	(16.580)
5-year lagged city controls	Yes	
3-year lagged city controls		Yes
Covariates Included	Yes	Yes
Fixed Individual	Yes	Yes
Fixed City	Yes	Yes
Fixed Province	Yes	Yes
Fixed Year	Yes	Yes
Fixed Cohort	Yes	Yes
Fixed Year × Province	Yes	Yes
K-P F-statistics	34.403	20.068
N	6128	6128

Note: This table shows a robustness check to the main results with different lagged years of pollution. Columns (1) and (2) are the 5-year and 3-year lagged city-level controls. All results are consistent with Table 4.3, showing the robustness of the main results. ***, **, * represents the level of significant at 1%,5% and 10% respectively.
Source: CHARLS 2011-2020, NASA, WshU, China Yearbook

4.6.6 Information Disclosure Policy Impact

Since information can influence individual optimal decision-making (Stigler, 1961), there is a concern that the sudden disclosure of air pollution data may lead to exaggerated perceptions of pollution. The Chinese government implemented a $PM_{2.5}$ concentration information disclosure policy in 2012, which could have prompted behavioural overreactions to pollution. To examine potential biases from this policy, I introduce an interaction term between air pollution and a policy treatment dummy variable, where the dummy indicates data after 2012 (Yes=1, No=0). This interaction tests whether information disclosure significantly altered behaviour.

Table 4.15 shows that the effect of air pollution remains consistent with the baseline estimates, and the interaction term is statistically insignificant. This suggests that individuals may not have substantially adjusted their behaviour in response to the disclosed information. However, the lack of significance could reflect structural barriers, such as limited access to avoidance options or scepticism toward government data, rather than an absence of awareness. The findings suggest that while the policy increased public awareness of pollution, it did not lead to behavioural changes substantial enough to influence retirement expectations.

TABLE 4.15: **The Effect of Air Pollution on Expected Retirement Age with the Effect of Information Disclosure**

Variable	Expected Retirement Age		
Samples	Overall	Urban	Rural
	(1)	(2)	(3)
Ln air pollution	-53.947***	-39.514***	-88.692***
	(15.923)	(11.882)	(21.402)
Policy Post	-2397.920	-2423.529	-2297.145
	(4.2×10^8)	(2.2×10^8)	(3.3×10^8)
Ln air pollution × Policy Post	8.881	2.555	6.046
	(7.969)	(4.203)	(6.545)
Covariates Included	Yes	Yes	Yes
Fixed Individual	Yes	Yes	Yes
Fixed City	Yes	Yes	Yes
Fixed Province	Yes	Yes	Yes
Fixed Year	Yes	Yes	Yes
Fixed Cohort	Yes	Yes	Yes
Fixed Year × Province	Yes	Yes	Yes
N	6128	2540	3588

Note: This table shows a robustness check of the main results combined with the effect of air pollution information disclosure. All results are consistent with Table 4.3, and the insignificant treatment effect shows information disclosure did not significantly affect individuals' retirement expectations in this sample.
***, **, * represents the level of significant at 1%,5% and 10% respectively. City-level clustering robust standard errors are in the parentheses.
Source: CHARLS 2011-2020, NASA, WshU, China Yearbook

4.7 Conclusion

This study examines the impact of air pollution on the expected retirement age in China, constructing panel data from the China Health and Retirement Longitudinal Study (CHARLS) and city-level $PM_{2.5}$ concentration data from the University of Washington in St Louis. The findings demonstrate a significant relationship between increased pollution levels and a reduction in expected retirement age, highlighting the health and economic implications of prolonged pollution exposure on labour supply, particularly among older individuals.

Our results suggest that each 1% increase in $PM_{2.5}$ concentration (around 0.12-1.04 µg/m³) correlates with a reduction in expected retirement age, with more pronounced effects in rural areas where individuals experience heightened vulnerability due to limited access to healthcare resources. Furthermore, dynamic analysis of pollution shocks illustrates those sudden increases in pollution lead to further adjustments in retirement expectations, particularly in urban areas where pollution fluctuations are more noticeable. These results underscore the role of environmental stressors in retirement decision-making, contributing to an evolving understanding of how pollution impacts economic behaviour and retirement planning.

Also, this paper suggests that pollution-induced retirement adjustments are gradual rather than immediate, and their effects vary significantly across different socioeconomic groups. The stronger response in the five-year model highlights the importance of considering longer-term adaptation mechanisms when assessing the economic consequences of environmental hazards.

Additionally, the channel analysis reveals that certain socioeconomic factors can mitigate or exacerbate the effects of pollution on retirement expectations. Family financial support, green infrastructure, and knowledge capital serve as buffers, allowing individuals to maintain higher retirement ages despite environmental challenges. Conversely, robust social welfare and insurance systems provide financial security, facilitating earlier retirement in polluted areas.

By integrating perspectives from environmental health and labour economics, this study underscores the critical need for policies that address air quality improvement and support for vulnerable populations. The findings provide valuable insights for policymakers aiming to strengthen social security frameworks, improve public health infrastructure, and encourage sustainable development in regions affected by high pollution levels. Future research could further explore the interplay between pollution, retirement, and labour supply, particularly in other developing economies facing similar environmental and demographic challenges.

Admittedly, there are some limitations to this study. First, data limitations affect the scope of the analysis. The use of city-level $PM_{2.5}$ data as a proxy for individual exposure may not capture the precise pollution levels each participant experiences, especially in rural areas where air quality can vary significantly within the same city. Additionally, the CHARLS dataset provides limited geographic coverage for certain rural and remote areas, which may impact the generalizability of the findings. Furthermore, self-reported retirement expectations may differ from actual retirement decisions. Relying on self-reported expected retirement ages introduces potential biases, as these expectations may be shaped by personal aspirations, social pressures, or inaccuracies in anticipating future circumstances. Consequently, actual retirement ages could diverge significantly from stated expectations. Lastly, this study employs a static analysis of welfare systems, assuming that welfare and social security provisions remain stable over time. However, evolving policies and changes in welfare provisions may alter the effects of pollution on retirement expectations. Future research could incorporate dynamic welfare variables to account for policy shifts that impact retirement behaviour in response to environmental factors.

This work adds to the limited literature on the long-term economic impacts of air pollution, demonstrating the need for comprehensive approaches to labour policy that consider environmental factors affecting retirement and overall workforce participation.

Chapter 5

Conclusions

This thesis has examined the intricate relationships between air pollution, migration, and socioeconomic behaviours, shedding light on how environmental stressors influence human decision-making and economic outcomes. By combining empirical evidence from three distinct but interconnected studies, it provides a comprehensive understanding of the mechanisms through which migration shapes green behaviours, pollution impacts urban mobility, and environmental factors alter retirement expectations. This concluding chapter synthesises the key findings, highlights their broader implications, and outlines directions for future research to address the challenges of sustainable development in rapidly transforming societies.

5.1 Introduction

This thesis investigates the complex interactions between air pollution, migration, and socioeconomic behaviours, focusing on their dynamic relationships within China. As one of the world's most rapidly industrialising and urbanising countries, China faces significant challenges related to environmental degradation, demographic transitions, and labour market adjustments. This study adopts an interdisciplinary approach, drawing insights from environmental economics, migration studies, and labour economics, to examine how environmental stressors influence migration decisions, reshape social norms, and affect retirement expectations.

The research is structured around three key questions: 1. How does rural-to-urban migration influence environmental behaviours in origin households? 2. What roles do air pollution disparities play in shaping urban-to-urban migration flows? 3. How do long-run exposure of air pollution and short-run pollution shock affect retirement expectations?

To address these questions, the thesis employs rigorous empirical methods, leveraging micro-level survey data and macro-level datasets on pollution and economic indicators. It combines fixed-effects regressions, gravity models, and instrumental variable approaches to establish causal relationships and disentangle complex dynamics.

In this chapter, Section 5.2 reviews the synthesis of key findings. Section 5.3 highlights the main academic contributions of this thesis. Section 5.4 outlines the corresponding policy implications. Finally, Section 5.5 discusses current limitations and suggests directions for future research.

5.2 Synthesis of Findings

This section provides a comprehensive summary of the key findings from the three empirical chapters, highlighting their connections, theoretical implications, and policy relevance. It demonstrates how each chapter contributes to answering the research questions and advancing the understanding of the interactions between air pollution, migration, and socioeconomic behaviours.

5.2.1 Migration and Green Behaviours

This chapter investigates the impact of migration on the green behaviours of left-behind families in China, focusing on internal rural-to-urban migration. Using two indices—willingness to recycle and fixed-place garbage disposal practices—the study examines the distinct effects of current migration and return migration on green behaviours. It incorporates selection bias adjustments, heterogeneity analysis, and interaction terms to explore mechanisms and socio-economic gaps between urban and rural areas.

The findings reveal significant differences in the effects of migration types. Current migration negatively impacts green behaviour by 71.9% and 21.8% reduction in recycling and fixed trash placement probability, as the absence of household members disrupts daily routines and weakens environmental practices. In contrast, return migration has a positive effect, around 79.0% and 46.3% increase in both, as returnees bring urban-acquired norms and practices back to their families.

To uncover the underlying mechanisms, this study examines the interaction between migration and the socio-economic gap between origin and destination cities. The results show that better green infrastructure, information networks, and higher social capital in destination cities amplify the positive effects of migration on recycling willingness. Migration not only provides financial remittances but also facilitates social remittance channels, enabling the transfer of urban environmental norms and practices to rural families. However, these benefits are likely to vary depending on the quality of infrastructure and socio-economic conditions in destination cities, which highlights the uneven potential for green behaviour transformation. Heterogeneity analysis reveals that younger, male-headed families with higher educational levels and sustainable financial situations are better positioned to benefit from the positive effects of return migration. Enhanced family capital, such as higher education and financial stability, often plays a pivotal role in adopting and sustaining green practices. Conversely, families with older members are more vulnerable to the negative effects of current migration due to challenges in adapting to disrupted routines and responsibilities. This heterogeneity suggests that policies aimed at strengthening household resilience and improving education and income levels could further enhance the adoption of green practices.

5.2.2 Pollution and Migration Flows

This chapter investigates the impact of air pollution on urban-to-urban migration in China, using bidirectional migration flow data from the China Migration Data Survey (CMDS) and the National Bureau of Statistics. The findings reveal that higher pollution levels in destination cities reduce migration inflows, with each doubling of the PM2.5 concentration ratio between destinations and origins decreasing migration by approximately 42%.

In addition to quantifying the pollution effect on migration, this study explores the channels that influence migration decisions in response to air pollution. The results show that strict environmental regulations and high-quality infrastructure in destination cities can mitigate the adverse effects of pollution, whereas higher living costs and restrictive hukou policies exacerbate them. Interestingly, the study identifies a U-shaped relationship between pollution impact and migration distance, with pollution sensitivity declining for distances over 1,000 km. Demographic analysis reveals that older, less-educated, married, and cross-province migrants are less likely to choose highly polluted destinations.

5.2.3 Pollution and Retirement Planning

This chapter examines the impact of air pollution on the expected retirement age in China, constructing panel data from the China Health and Retirement Longitudinal Study (CHARLS) and city-level $PM_{2.5}$ concentration data from the University of Washington in St Louis. The findings demonstrate a significant relationship between increased pollution levels and a reduction in expected retirement age, highlighting the health and economic implications of prolonged pollution exposure on labour supply, particularly among older individuals.

Results suggest that each 1% increase in $PM_{2.5}$ concentration (around 0.12-1.04 $\mu g/m^3$) correlates with a reduction in expected retirement age by approximately 5.11 months. The effects are more pronounced in rural areas, where individuals face higher vulnerability due to limited healthcare resources, leading to an average reduction of 9.03 months in retirement age. In contrast, urban areas experience a smaller reduction of about 4.45 months, reflecting better access to infrastructure and services.

Dynamic analysis of pollution shocks illustrates that sudden increases in pollution lead to further adjustments in retirement expectations. Specifically, a 1% increase in pollution over five years results in an additional reduction of 1.45 months on average. The short-run effect is more pronounced in urban areas, with reductions of 2.58 months, compared to 2.11 months in rural areas. These results underscore the role of environmental stressors in retirement decision-making, contributing to an evolving understanding of how pollution impacts economic behaviour and retirement planning.

Channel analysis reveals that socioeconomic factors can mitigate or exacerbate the effects of pollution on retirement expectations. Family financial support, green infrastructure, and knowledge capital serve as buffers, allowing individuals to maintain higher retirement ages despite environmental challenges. Conversely, robust social welfare and insurance systems provide financial security, facilitating earlier retirement in polluted areas.

This study underscores the critical need for policies that address air quality improvement and support for vulnerable populations. The findings provide valuable insights for policymakers aiming to strengthen social security frameworks, improve public health infrastructure, and encourage sustainable development in regions affected by high pollution levels. Future research could further explore the interplay between pollution, retirement, and labour supply, particularly in other developing economies facing similar environmental and demographic challenges.

5.3 Contribution and Significance

This thesis contributes to the broader academic discourse by advancing theoretical frameworks in migration studies, environmental economics, and labour economics. Through its interdisciplinary approach, it integrates insights across these domains to address the complex relationships between air pollution, migration, and socioeconomic behaviours. The findings not only refine existing theories but also introduce new perspectives to better understand the role of environmental stressors in shaping human decisions and behaviours.

1. **Contributions to Migration Theory**

   This research extends social remittance theory by incorporating environmental behaviours as part of migration outcomes, which has been largely overlooked in previous studies. While prior research has focused on economic remittances and cultural transfers, this thesis highlights how migration can facilitate the diffusion of green behaviours from urban to rural households. Specifically, the identification of both positive impacts of return migration and negative effects of current migration provides nuanced insights into how migration influences environmental practices, depending on context and channels of transmission.

   This research extends social remittance theory by incorporating environmental behaviours as part of migration outcomes, which has been largely overlooked in previous studies. While prior research has focused on economic remittances and cultural transfers, this thesis highlights how migration can facilitate the diffusion of green behaviours from urban to rural households. Specifically, the identification of both positive impacts of return migration and negative effects of current migration provides nuanced insights into how migration influences environmental practices, depending on context and channels of transmission.

   Additionally, the thesis breaks new ground by focusing on internal migration in a developing economy, addressing gaps in migration literature that have traditionally emphasized international migration. By examining rural-to-urban flows and bidirectional urban migration, it offers a novel perspective on how migration dynamics within countries shape environmental adaptation and sustainability practices.

2. **Contributions to Environmental Economics**

   This thesis bridges environmental economics with migration and labour studies, highlighting the role of pollution shocks in influencing mobility and retirement behaviours. It provides empirical evidence that pollution acts as both a push factor driving migration and a constraint affecting long-term labour supply decisions.

   The research also demonstrates how environmental quality disparities shape migration patterns, contributing to debates on hedonic valuation models and urban resilience. By incorporating dynamic pollution shocks, it expands traditional economic models to capture short- and long-term responses to environmental stressors, offering a richer understanding of adaptation mechanisms in polluted environments.

Furthermore, the integration of green remittance channels into migration theory advances the discourse on how urban experiences influence rural sustainability, highlighting the broader role of migration in promoting environmental adaptation.

3. **Contributions to Labour Economics**

   The thesis makes substantial contributions to understanding how environmental stressors, such as air pollution, influence retirement expectations and labour supply decisions. Prior research has primarily focused on health shocks or economic incentives as drivers of retirement, neglecting the role of environmental factors. By demonstrating how pollution accelerates retirement planning, this thesis introduces environmental stressors as a distinct category affecting life-cycle decisions.

   It also explores heterogeneity across urban and rural contexts, revealing that vulnerable groups, such as older adults, low-income households, and rural residents, are disproportionately affected. This emphasis on inequality highlights the distributional impacts of pollution, adding depth to existing theories on labour economics and welfare.

4. **Methodological Contributions**

   Methodologically, this research advances the use of instrumental variable (IV) approaches to address endogeneity concerns in migration and pollution studies. By employing thermal inversion days and migration network rates as instruments, it provides robust identification strategies that strengthen causal interpretations.

   The use of differential gravity models to capture bidirectional migration flows represents another methodological innovation, enabling a more comprehensive analysis of migration dynamics based on pollution gaps.

   Additionally, the integration of dynamic pollution shock models in retirement studies provides a framework for analysing short- and long-term responses to environmental stressors, which can be applied to other contexts and developing economies.

5. **Bridging Disciplines**

   A key contribution of this thesis lies in its interdisciplinary approach, bridging insights from environmental economics, migration studies, and labour economics. By linking environmental stressors with migration and labour supply decisions, it creates a unified framework for studying socioeconomic adaptation to environmental changes.

This synthesis offers a more holistic understanding of the interconnected challenges posed by urbanisation, environmental degradation, and demographic transitions. It also lays the groundwork for future research that further integrates environmental, social, and economic dimensions to address sustainability and inequality.

5.4 Policy Implications

The findings of this thesis provide substantial insights into the socioeconomic and behavioural impacts of air pollution and migration, offering several policy implications for environmental governance, sustainable development, and social welfare systems. By addressing challenges related to pollution, migration dynamics, and retirement planning, this research informs integrated policies that promote resilience and sustainability in rapidly urbanising and ageing societies like China.

1. **Promoting Environmental Sustainability and Infrastructure Development**

   The results highlight the role of migration in transferring environmental norms and practices through social remittances. Return migration, in particular, promotes green behaviours in rural households, suggesting that policies supporting green infrastructure development in rural areas could amplify these positive effects. Investments in public waste management systems, recycling facilities, and environmental education programs can create enabling environments for the adoption of sustainable practices.

   For urban regions, stricter enforcement of pollution control regulations and expanded investment in green technologies can reduce pollution levels, making cities more attractive destinations and retaining skilled labour. Strengthening data transparency through public disclosure of pollution levels can further influence migration decisions, empowering households to make informed choices based on environmental quality.

2. **Bridging Urban-Rural Socioeconomic Disparities**

   The findings underscore the persistent inequalities between urban and rural areas, especially in access to healthcare, social infrastructure, and environmental amenities. Policies should address these disparities by extending urban-level services to rural areas. Improving healthcare infrastructure in rural regions can mitigate the adverse effects of pollution on vulnerable populations, reducing disparities in retirement planning outcomes.

   Targeted support programs for left-behind families, such as subsidies for green technologies or community waste management initiatives, can counterbalance the negative effects of migration disruptions. Additionally, migration-related education programs should focus on strengthening environmental awareness among rural and urban populations to enhance the transfer of green practices.

3. **Strengthening Labour Market Resilience and Social Security Systems**

Pollution-induced health deterioration accelerates retirement planning, particularly in rural areas with weaker healthcare systems. This finding highlights the need for labour policies that encourage flexible retirement planning, allowing older individuals to remain economically active while managing health concerns. Expanding access to employer-based health insurance and retirement savings plans can provide stronger financial security, reducing the pressure to exit the workforce prematurely.

Additionally, reforms to pension systems should focus on addressing regional inequalities and adapting to the needs of ageing populations affected by environmental stressors. Policymakers should also consider incentives for delayed retirement, such as enhanced pension benefits or subsidies for older workers in high-risk environments.

4. **Managing Migration and Urbanisation Pressures**

   Air pollution emerges as both a push and pull factor in migration decisions, influencing urbanisation patterns and labour mobility. Policies that improve air quality in urban centres can enhance cities' attractiveness for skilled workers while relieving population pressures on less-developed regions. Flexible migration policies, including reforms to the hukou system, can support the integration of migrants into urban labour markets, reducing inequalities and enhancing human capital accumulation.

   Moreover, migration management should incorporate environmental planning to balance economic development with sustainability. Regional coordination between cities can create incentives for green urban planning, reducing pollution disparities and promoting balanced growth across regions.

5. **Supporting Vulnerable Populations**

   The heterogeneity analysis highlights how vulnerable groups, such as older adults, low-income households, and rural migrants—are disproportionately affected by environmental stressors and migration-related disruptions. Policies should focus on targeted interventions, such as housing subsidies, community support programs, and mobile healthcare services, to enhance the resilience of these populations.

   Special attention should be given to rural communities that rely heavily on migration for economic stability. Programs encouraging sustainable agriculture and eco-friendly industries can help rural areas diversify their economies while reducing dependence on environmentally damaging practices.

5.5 Limitations & Directions for Future Research

While this thesis provides new insights into the relationships between air pollution, migration, and socioeconomic behaviours, it also acknowledges several limitations that offer avenues for further research. These limitations primarily stem from data constraints, methodological approaches, and the scope of generalizability, each of which suggests opportunities to deepen and broaden future inquiry.

1. **Data and Measurement Limitations**

   A key limitation of this research is its reliance on city-level pollution data as a proxy for individual exposure to environmental stressors. Although this approach captures broader patterns, it may overlook intra-city variations in pollution levels, particularly in rural areas where disparities in air quality are more pronounced. Future studies could integrate remote sensing data or personal monitoring devices to achieve more granular measurements of pollution exposure.

   Additionally, the use of self-reported data for green behaviours and retirement expectations introduces potential reporting biases. Perceptions about environmental quality or retirement decisions may reflect aspirations rather than actual behaviours, which could affect the interpretation of results. Incorporating behavioural experiments or administrative data in future research could help address these biases.

   Moreover, the longitudinal datasets employed in this thesis, while robust, are limited in their ability to capture long-term dynamics and spillover effects. Expanding these datasets to include multi-year tracking or integrating cross-national comparisons could provide a more comprehensive understanding of how environmental stressors shape socioeconomic behaviours over time.

2. **Methodological Constraints**

   The econometric approaches used in this thesis, including instrumental variable (IV) techniques and dynamic panel models, address endogeneity concerns but rely on several assumptions. For example, the use of thermal inversion days as an instrument for pollution depends on its plausibility as an exogenous shock, which may vary across regions. Future research could test alternative instruments or use spatial econometric techniques to account for spillover effects between neighbouring cities.

   Additionally, the differential gravity model employed in studying migration flows provides insights into push and pull factors based on relative pollution levels but does not fully capture network effects or migration chains that influence mobility decisions. Incorporating social network data or agent-based simulations could enhance understanding of these dynamics.

3. **Generalisability of Findings**

   While this research focuses on China—a rapidly urbanising and developing economy—it is important to acknowledge that the findings may have limited generalisability to countries with different institutional structures, migration regimes, or environmental policies. For instance, China's hukou system imposes unique constraints on mobility and access to services, which may amplify the observed effects of pollution and migration in ways that are not applicable in other contexts.

   Moreover, the country's rapid urbanisation and large-scale internal migration, combined with disparities in green infrastructure and pension systems, shape specific behavioural responses to pollution shocks, particularly in retirement planning and the transmission of environmental norms. These structural features differ markedly from those in many developed economies, where more robust welfare systems and environmental regulations may alter both exposures and responses.

   Additionally, while the datasets used (e.g., CHARLS, CLDS, CMDS) are nationally representative and well-suited for causal analysis, they are embedded in specific social and institutional contexts. Self-reported expectations—such as retirement planning—may also reflect cultural and informational biases that constrain international comparability.

   To address these concerns, future research could extend the empirical framework to other developing economies facing similar environmental and demographic pressures or conduct cross-country comparative studies. Such work would help distinguish context-specific mechanisms from more generalisable patterns, thereby deepening our understanding of how environmental stressors shape human behaviour globally.

4. **Extensions and Future Research**

   Several avenues for future research emerge from this thesis:

   - **Longitudinal Analysis**: Extending the datasets to track individuals over time can shed light on the cumulative effects of pollution and migration, as well as life-cycle transitions in green behaviours and retirement planning.
   - **Spatial and Network Effects**: Exploring spatial dependencies and migration networks could improve understanding of spillover dynamics, particularly how migration patterns influence environmental and economic outcomes in neighbouring regions.
   - **Climate Change and Extreme Events**: Integrating dimensions such as climate-induced migration and disaster shocks would expand the scope of environmental stressors studied, offering insights into adaptation strategies under different scenarios.
   - **Gender and Demographic Differences**: Investigating how gender, education, and generational differences mediate responses to pollution could help design more targeted policies for vulnerable populations.

- **Policy Evaluations**: Future studies could assess the effectiveness of existing environmental regulations, green infrastructure investments, and pension reforms to determine how they mitigate the effects of pollution and migration disruptions.

In summary, this thesis offers valuable insights into the complex interactions between environmental stressors, migration, and socioeconomic behaviours, providing a foundation for future research and policymaking. By examining how air pollution shapes migration flows, influences green behaviours, and affects retirement planning, it highlights the far-reaching impacts of environmental challenges on individual and societal outcomes. The findings emphasise the need for integrated policies that address environmental governance, labour market resilience, and social inequalities to promote sustainable and inclusive development. While acknowledging certain limitations, this research paves the way for further exploration of environmental adaptation strategies in diverse contexts, contributing to the broader discourse on sustainability and socioeconomic transformation in a rapidly changing world.

Appendix A

Appendix

A.1 Deriving the Linear Model from the Gravity Equation

The gravity model is commonly used in studies on international trade and migration, capturing the effect of distance and regional characteristics on flows between an origin and a destination. Following this framework, we begin with the basic gravity model equation for migration flows:

$$M_{o,d} = k \frac{P_o^\alpha P_d^\gamma}{\text{distance}_{o,d}^\beta} \tag{A.1}$$

where:

- $M_{o,d}$ is the migration flow from the origin city o to the destination city d,
- P_o and P_d denote the population sizes of the origin and destination cities,
- $\text{distance}_{o,d}$ is the distance between the origin and destination cities,
- k is a constant of proportionality,
- α, γ, and β are parameters capturing the elasticity of migration flow concerning population sizes and distance.

To linearize this equation, we take the natural logarithm of both sides:

$$\ln(M_{o,d}) = \ln \left(k \frac{P_o^\alpha P_d^\gamma}{\text{distance}_{o,d}^\beta} \right) \tag{A.2}$$

Applying the properties of logarithms, $\ln(xy) = \ln(x) + \ln(y)$ and $\ln(x^a) = a \ln(x)$, we can rewrite the equation as:

$$\ln(M_{o,d}) = \ln(k) + \alpha \ln(P_o) + \gamma \ln(P_d) - \beta \ln(\text{distance}_{o,d}) \tag{A.3}$$

Letting $\ln(k) = \beta_0$, we get:

$$\ln(M_{o,d}) = \beta_0 + \alpha \ln(P_o) + \gamma \ln(P_d) - \beta \ln(\text{distance}_{o,d}) \tag{A.4}$$

To extend this model for migration-specific analysis, we incorporate additional explanatory variables:

- The pollution differential between destination and origin, represented as the ratio $\frac{\text{PM}_{2.5d}}{\text{PM}_{2.5o}}$,
- A set of other city-level characteristics for the destination and origin cities, denoted by X_d and X_o, which include economic, infrastructure, and geographical variables.

The extended model then becomes:

$$\ln(M_{o,d}) = \beta_0 + \beta_1 \frac{\text{PM}_{2.5d}}{\text{PM}_{2.5o}} + \beta_2 \ln(\text{distance}_{o,d}) + \beta_3 \ln(P_o) + \beta_4 \ln(P_d) + X_d\beta_5 + X_o\beta_6 + \epsilon_{o,d} \tag{A.5}$$

where:

- β_1 represents the effect of the pollution differential on migration flow,
- β_2, β_3, and β_4 correspond to the elasticity of migration flow with respect to distance and population sizes,
- $X_d\beta_5$ and $X_o\beta_6$ capture the effects of additional destination and origin city characteristics,
- $\epsilon_{o,d}$ is the error term.

To account for zero migration flows, we use $\ln(M_{o,d} + 1)$ in the model, resulting in the following final linear model:

$$\begin{aligned} \ln(M_{o,d} + 1) = \beta_0 + \beta_1 \frac{\text{PM}_{2.5d}}{\text{PM}_{2.5o}} + \beta_2 \ln(\text{distance}_{o,d}) \\ + \beta_3 \ln(P_o) + \beta_4 \ln(P_d) + \mathbf{X}_d\beta_5 + \mathbf{X}_o\beta_6 + \epsilon_{o,d} \end{aligned} \tag{A.6}$$

This linearised model (Equation 3.2) is now suitable for regression analysis, where each parameter represents the effect of a specific factor on migration flows between origin and destination cities.

References

R. Abramitzky, L. P. Boustan, and K. Eriksson. Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration. *Am Econ Rev*, 102(5):1832–1856, 2012. ISSN 0002-8282 (Print) 0002-8282 (Linking). . URL https://www.ncbi.nlm.nih.gov/pubmed/26594052. Abramitzky, Ran Boustan, Leah Platt Eriksson, Katherine eng R24 HD041022/HD/NICHD NIH HHS/ 2012/08/01 Am Econ Rev. 2012 Aug;102(5):1832-1856. doi: 10.1257/aer.102.5.1832.

Tamer Afifi. Economic or environmental migration? the push factors in niger. *International Migration*, 49:e95–e124, 2011. ISSN 0020-7985.

Farzana Afridi, Sherry Xin Li, and Yufei Ren. Social identity and inequality: The impact of china's hukou system. *Journal of public economics*, 123:17–29, 2015.

Yonas Alem, Abebe D. Beyene, Gunnar Köhlin, and Alemu Mekonnen. Modeling household cooking fuel choice: A panel multinomial logit approach. *Energy Economics*, 59: 129–137, 2016. ISSN 0140-9883. . URL https://www.sciencedirect.com/science/article/pii/S0140988316301748.

Joshua S Apte, Julian D Marshall, Aaron J Cohen, and Michael Brauer. Addressing global mortality from ambient pm2. 5. *Environmental science technology*, 49(13): 8057–8066, 2015. ISSN 0013-936X.

Eva Arceo, Rema Hanna, and Paulina Oliva. Does the effect of pollution on infant mortality differ between developing and developed countries? evidence from mexico city. *The Economic Journal*, 126(591):257–280, 2016. ISSN 0013-0133.

H Spencer Banzhaf and Randall P Walsh. Do people vote with their feet? an empirical test of tiebout's mechanism. *American economic review*, 98(3):843–863, 2008. ISSN 0002-8282.

Douglas K Bardsley and Graeme J Hugo. Migration and climate change: examining thresholds of change to guide effective adaptation decision-making. *Population and Environment*, 32:238–262, 2010. ISSN 0199-0039.

Panle Jia Barwick, Shanjun Li, Liguo Lin, and Eric Zou. From fog to smog: The value of pollution information. Report, American Economic Review, 2024.

Patrick Bayer, Nathaniel Keohane, and Christopher Timmins. Migration and hedonic valuation: The case of air quality. *Journal of Environmental Economics and Management*, 58(1):1–14, 2009. ISSN 0095-0696.

Sina Bierkamp, Trung Thanh Nguyen, and Ulrike Grote. Environmental income and remittances: Evidence from rural central highlands of vietnam. *Ecological Economics*, 179: 106830, 2021. ISSN 0921-8009. . URL https://www.sciencedirect.com/science/article/pii/S0921800920305711.

Mark Borgschulte, David Molitor, and Eric Yongchen Zou. Air pollution and the labor market: Evidence from wildfire smoke. *Review of Economics and Statistics*, pages 1–46, 2022. ISSN 0034-6535.

John Bound, Todd Stinebrickner, and Timothy Waidmann. Health, economic resources and the work decisions of older men. *Journal of econometrics*, 156(1):106–129, 2010. ISSN 0304-4076.

James K Boyce. *Economics, the environment and our common wealth*. Edward Elgar Publishing, 2013. ISBN 1782547673.

Fernando Broner, Paula Bustos, and Vasco M Carvalho. Sources of comparative advantage in polluting industries. Report, National Bureau of Economic Research, 2012.

Robert D Brook, Sanjay Rajagopalan, C Arden Pope III, Jeffrey R Brook, Aruni Bhatnagar, Ana V Diez-Roux, Fernando Holguin, Yuling Hong, Russell V Luepker, and Murray A Mittleman. Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the american heart association. *Circulation*, 121(21):2331–2378, 2010. ISSN 0009-7322.

Richard T Burnett, C Arden Pope III, Majid Ezzati, Casey Olives, Stephen S Lim, Sumi Mehta, Hwashin H Shin, Gitanjali Singh, Bryan Hubbell, and Michael Brauer. An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. *Environmental health perspectives*, 122(4): 397–403, 2014. ISSN 0091-6765.

He Cai and Jin Wang. A study on migrant workers' permanent migration intentions. *Sociological Studies*, 6:86–113, 2007.

Ruohong Cai, Shuaizhang Feng, Michael Oppenheimer, and Mariola Pytlikova. Climate variability and international migration: The importance of the agricultural linkage. *Journal of Environmental Economics and Management*, 79:135–151, 2016. ISSN 0095-0696.

Guangzhong Cao, Ming Li, Yan Ma, and Ran Tao. Self-employment and intention of permanent urban settlement: Evidence from a survey of migrants in china's four major urbanising areas. *Urban Studies*, 52(4):639–664, 2015. ISSN 0042-0980.

Kuang-Cheng Chai, De-Cong Xie, Chin-Piao Yeh, Hao-Ran Lan, and Zhen-Xin Cui. Chinese national civilized city and corporate social responsibility: will civilized city promote corporate social responsibility? *Applied Economics Letters*, 29(7):593–596, 2022. ISSN 1350-4851.

Hongqin Chang, Xiao-yuan Dong, and Fiona MacPhail. Labor migration and time use patterns of the left-behind children and elderly in rural china. *World Development*, 39 (12):2199–2210, 2011. ISSN 0305-750X.

Tom Y Chang, Joshua Graff Zivin, Tal Gross, and Matthew Neidell. The effect of pollution on worker productivity: evidence from call center workers in china. *American Economic Journal: Applied Economics*, 11(1):151–172, 2019. ISSN 1945-7782.

Kenneth Y Chay and Michael Greenstone. Does air quality matter? evidence from the housing market. *Journal of political Economy*, 113(2):376–424, 2005. ISSN 0022-3808.

Jiafeng Chen and Jonathan Roth. Logs with zeros? some problems and solutions. *The Quarterly Journal of Economics*, page qjad054, 2023. ISSN 0033-5533.

Juan Chen, Shuo Chen, and Pierre F. Landry. Migration, environmental hazards, and health outcomes in china. *Social Science Medicine*, 80:85–95, 2013a. ISSN 0277-9536. . URL https://www.sciencedirect.com/science/article/pii/S0277953612008064.

Shuai Chen, Yuyu Chen, Ziteng Lei, and Jie-Sheng Tan-Soo. Impact of air pollution on short-term movements: evidence from air travels in china. *Journal of Economic Geography*, 20(4):939–968, 2020. ISSN 1468-2702.

Shuai Chen, Yuyu Chen, Ziteng Lei, and Jie-Sheng Tan-Soo. Chasing clean air: Pollution-induced travels in china. *Journal of the Association of Environmental and Resource Economists*, 8(1):59–89, 2021. ISSN 2333-5955.

Shuai Chen, Paulina Oliva, and Peng Zhang. The effect of air pollution on migration: evidence from china. *Journal of Development Economics*, 156:102833, 2022. ISSN 0304-3878.

Ting Chen and James Kai-sing Kung. Busting the "princelings": The campaign against corruption in china's primary land market. *The Quarterly Journal of Economics*, 134 (1):185–226, 2019. ISSN 0033-5533.

Wei Chen, Xiaoyu Wang, Jing Chen, Chao You, Lu Ma, Wei Zhang, and Dong Li. Household air pollution, adherence to a healthy lifestyle, and risk of cardiometabolic multimorbidity: results from the china health and retirement longitudinal study. *Science of The Total Environment*, 855:158896, 2023. ISSN 0048-9697.

Yuyu Chen, Avraham Ebenstein, Michael Greenstone, and Hongbin Li. Evidence on the impact of sustained exposure to air pollution on life expectancy from china's huai river policy. *Proceedings of the National Academy of Sciences*, 110(32):12936–12941, 2013b. ISSN 0027-8424.

Zhiming Cheng, Ingrid Nielsen, and Russell Smyth. Access to social insurance in urban china: A comparative study of rural–urban and urban–urban migrants in beijing. *Habitat International*, 41:243–252, 2014.

China Environmental Awareness Index. 2021 china environmental awareness index report, 2021. URL https://www.chinagreen.org.cn. Accessed August 2025.

Gary Coombs. Opportunities, information networks and the migration-distance relationship. *Social Networks*, 1(3):257–276, 1978. ISSN 0378-8733.

Olivier Deschenes, Michael Greenstone, and Joseph S Shapiro. Defensive investments and the demand for air quality: Evidence from the nox budget program. *American Economic Review*, 107(10):2958–2989, 2017. ISSN 0002-8282.

Olivier Deschenes, Huixia Wang, Si Wang, and Peng Zhang. The effect of air pollution on body weight and obesity: Evidence from china. *Journal of Development Economics*, 145:102461, 2020. ISSN 0304-3878.

Olivier Deschênes and Michael Greenstone. Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the us. *American Economic Journal: Applied Economics*, 3(4):152–185, 2011. ISSN 1945-7782.

Valentina Di Iasio and Jackline Wahba. The determinants of refugees' destinations: Where do refugees locate within the eu? *World Development*, 177:106533, 2024. ISSN 0305-750X.

Arnaud Dupuy. Migration in china: To work or to wed? *Journal of Applied Econometrics*, 36(4):393–415, 2021. ISSN 0883-7252.

Avraham Ebenstein, Maoyong Fan, Michael Greenstone, Guojun He, Peng Yin, and Maigeng Zhou. Growth, pollution, and life expectancy: China from 1991–2012. *American Economic Review*, 105(5):226–231, 2015. ISSN 0002-8282.

Avraham Ebenstein, Maoyong Fan, Michael Greenstone, Guojun He, and Maigeng Zhou. New evidence on the impact of sustained exposure to air pollution on life expectancy from china's huai river policy. *Proceedings of the National Academy of Sciences*, 114 (39):10384–10389, 2017. ISSN 0027-8424.

US EPA. Final report: Integrated science assessment for particulate matter. *Washington, DC: US Environmental Protection Agency*, 2009.

C Cindy Fan. The elite, the natives, and the outsiders: Migration and labor market segmentation in urban china. *Annals of the association of American geographers*, 92 (1):103–124, 2002.

C Cindy Fan. *China on the Move: Migration, the State, and the Household*. Routledge, 2007. ISBN 0203937376.

C Cindy Fan and Tianjiao Li. Familization of rural–urban migration in china: evidence from the 2011 and 2015 national floating population surveys. *Area Development and Policy*, 4(2):134–156, 2019. ISSN 2379-2949.

Hanming Fang and Jin Feng. The chinese pension system. Report, National Bureau of Economic Research Cambridge, MA, USA:, 2018a.

Hanming Fang and Jin Feng. The chinese pension system. 2018b.

Qiushi Feng, Wei-Jun Jean Yeung, Zhenglian Wang, and Yi Zeng. Age of retirement and human capital in an aging china, 2015–2050. *European Journal of Population*, 35: 29–62, 2019.

Tong Feng, Huibin Du, Zhongguo Lin, and Jian Zuo. Spatial spillover effects of environmental regulations on air pollution: Evidence from urban agglomerations in china. *Journal of Environmental Management*, 272:110998, 2020. ISSN 0301-4797.

Stuart Ferguson. Are public libraries developers of social capital? a review of their contribution and attempts to demonstrate it. *The Australian Library Journal*, 61(1): 22–33, 2012.

Mohammad H Forouzanfar, Ashkan Afshin, Lily T Alexander, H Ross Anderson, Zulfiqar A Bhutta, Stan Biryukov, Michael Brauer, Richard Burnett, Kelly Cercy, and Fiona J Charlson. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the global burden of disease study 2015. *The lancet*, 388(10053):1659–1724, 2016. ISSN 0140-6736.

Bruno S Frey and Stephan Meier. Social comparisons and pro-social behavior: Testing" conditional cooperation" in a field experiment. *American economic review*, 94(5): 1717–1722, 2004. ISSN 0002-8282.

Shihe Fu, V Brian Viard, and Peng Zhang. Air pollution and manufacturing firm productivity: Nationwide estimates for china. *The Economic Journal*, 131(640):3241–3273, 2021. ISSN 0013-0133.

Xuwen Gao, Ran Song, and Christopher Timmins. Information, migration, and the value of clean air. *Journal of Development Economics*, 163:103079, 2023. ISSN 0304-3878.

Corrado Giulietti, Jackline Wahba, and Klaus F Zimmermann. *Entrepreneurship of the left-behind*. Emerald Group Publishing Limited, 2013. ISBN 1781907560.

Joshua Graff Zivin and Matthew Neidell. Environment, health, and human capital. *Journal of economic literature*, 51(3):689–730, 2013. ISSN 0022-0515.

Qingbin Guo, Yong Wang, Yao Zhang, Ming Yi, and Tian Zhang. Environmental migration effects of air pollution: Micro-level evidence from china. *Environmental Pollution*, 292:118263, 2022. ISSN 0269-7491.

Zhiyong Han, Changwen Ye, Yu Zhang, Zeng Dan, Zeyan Zou, Dan Liu, and Guozhong Shi. Characteristics and management modes of domestic waste in rural areas of developing countries: a case study of china. *Environmental Science and Pollution Research*, 26(9):8485–8501, 2019.

Penelope Hawe and Alan Shiell. Social capital and health promotion: a review. *Social science medicine*, 51(6):871–885, 2000. ISSN 0277-9536.

Yongming He and Baiqun Ding. Environmental and economic evaluation of superhighway based on travel cost. *Ekoloji Dergisi*, (107), 2019.

Anthony Heyes, Matthew Neidell, and Soodeh Saberian. The effect of air pollution on investor behavior: Evidence from the sp 500. Report, National Bureau of Economic Research, 2016.

Nicole Hildebrandt, David J McKenzie, Gerardo Esquivel, and Ernesto Schargrodsky. The effects of migration on child health in mexico. *Economia*, 6(1):257–289, 2005. ISSN 1529-7470.

Md Maruf Hossain, KA Majumder, Mahmuda Islam, and AA Nayeem. Study on ambient particulate matter (pm2. 5) with different mode of transportation in dhaka city, bangladesh. *American Journal of Pure and Applied Biosciences*, 1(4):12–19, 2019.

Keyong Huang, Fengchao Liang, Xueli Yang, Fangchao Liu, Jianxin Li, Qingyang Xiao, Jichun Chen, Xiaoqing Liu, Jie Cao, and Chong Shen. Long term exposure to ambient fine particulate matter and incidence of stroke: prospective cohort study from the china-par project. *Bmj*, 367, 2019. ISSN 1756-1833.

IIPS. Internal migration in india and its impact on development, 2020. Accessed August 2025.

International Labour Organization. India: Social protection for all - policy brief, 2021. URL `https://www.ilo.org`. Accessed August 2025.

Koichiro Ito and Shuang Zhang. Willingness to pay for clean air: Evidence from air purifier markets in china. *Journal of Political Economy*, 128(5):1627–1672, 2020. ISSN 0022-3808.

David J Kaczan and Jennifer Orgill-Meyer. The impact of climate change on migration: a synthesis of recent empirical insights. *Climatic Change*, 158(3-4):281–300, 2020. ISSN 0165-0009.

John Kennan and James R Walker. Wages, welfare benefits and migration. *Journal of Econometrics*, 156(1):229–238, 2010. ISSN 0304-4076.

John Kennan and James R Walker. The effect of expected income on individual migration decisions. *Econometrica*, 79(1):211–251, 2011. ISSN 0012-9682.

Gaurav Khanna, Wenquan Liang, Ahmed Mushfiq Mobarak, and Ran Song. The productivity consequences of pollution-induced migration in china. Report, National Bureau of Economic Research, 2021.

Sonia Ben Kheder and Natalia Zugravu. Environmental regulation and french firms location abroad: An economic geography model in an international comparative study. *Ecological Economics*, 77:48–61, 2012. ISSN 0921-8009.

KE Khuda. Air pollution in the capital city of bangladesh: its causes and impacts on human health. *Pollution*, 6(4):737–750, 2020. ISSN 2383-451X.

Cynthia Kinnan, Shing-Yi Wang, and Yongxiang Wang. Access to migration for rural households. *American Economic Journal: Applied Economics*, 10(4):79–119, 2018. ISSN 1945-7782.

Prashant Kumar, Angela Druckman, John Gallagher, Birgitta Gatersleben, Sarah Allison, Theodore S Eisenman, Uy Hoang, Sarkawt Hama, Arvind Tiwari, and Ashish Sharma. The nexus between air pollution, green infrastructure and human health. *Environment international*, 133:105181, 2019. ISSN 0160-4120.

Wangyang Lai, Hong Song, Chang Wang, and Huanhuan Wang. Air pollution and brain drain: Evidence from college graduates in china. *China Economic Review*, 68:101624, 2021. ISSN 1043-951X.

Victor Lavy, Avraham Ebenstein, and Sefi Roth. The impact of air pollution on cognitive performance and human capital formation. *Unpublished. http://www2. warwick. ac. uk/fac/soc/economics/staff/academic/lavy/*$text_and_tables_air_pollution_draft_20_09_12.pdf$, 2012.

Ming-Hsuan Lee. Migration and children's welfare in china: The schooling and health of children left behind. *The Journal of Developing Areas*, 44(2):165–182, 2011. ISSN 0022037X. URL `http://www.jstor.org/stable/23215246`.

Jos Lelieveld, John S Evans, Mohammed Fnais, Despina Giannadaki, and Andrea Pozzer. The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature*, 525(7569):367–371, 2015. ISSN 0028-0836.

Peggy Levitt. Social remittances: Migration driven local-level forms of cultural diffusion. *International Migration Review*, 32(4):926–948, 1998. . URL `https://journals.sagepub.com/doi/abs/10.1177/019791839803200404`.

Peggy Levitt and Deepak Lamba-Nieves. Social remittances revisited. *Journal of ethnic and migration studies*, 37(1):1–22, 2011. ISSN 1369-183X.

Shi Li and Björn Gustafsson. China's social security system: Current situation and challenges. *China Economic Review*, 68:101626, 2021. .

Xiaoqin Li and Yonghui Li. The impact of perceived air pollution on labour supply: Evidence from china. *Journal of Environmental Management*, 306:114455, 2022. ISSN 0301-4797.

Zhengtao Li, Henk Folmer, and Jianhong Xue. To what extent does air pollution affect happiness? the case of the jinchuan mining area, china. *Ecological Economics*, 99: 88–99, 2014. ISSN 0921-8009.

Jiaochen Liang and Stephan J Goetz. Technology intensity and agglomeration economies. *Research Policy*, 47(10):1990–1995, 2018.

Bing Liu, Xiangquan Li, and Dan Wang. Retirement age expectations of chinese workers: Evidence from a national survey. Technical Report DP No. 17642, IZA – Institute of Labor Economics, 2024. URL https://repec.iza.org/dp17642.pdf. IZA Discussion Paper Series.

Tao Liu and Li Sun. Pension reform in china. *Journal of aging & social policy*, 28(1): 15–28, 2016.

Ying Liu, Wei Deng, and Xueqian Song. Influence factor analysis of migrants' settlement intention: Considering the characteristic of city. *Applied Geography*, 96:130–140, 2018. ISSN 0143-6228.

Ziming Liu and Lu Yu. Stay or leave? the role of air pollution in urban migration choices. *Ecological Economics*, 177:106780, 2020. ISSN 0921-8009.

Ziming Liu, Jia Li, Jens Rommel, and Shuyi Feng. Health impacts of cooking fuel choice in rural china. *Energy economics*, 89:104811, 2020. ISSN 0140-9883.

Martin Ljunge. Cultural transmission of civicness. *Economics Letters*, 117(1):291–294, 2012. ISSN 0165-1765.

Simon Luechinger. Life satisfaction and transboundary air pollution. *Economics Letters*, 107(1):4–6, 2010. ISSN 0165-1765.

Alejandro López-Feldman and Estefanía Chávez. Remittances and natural resource extraction: Evidence from mexico. *Ecological Economics*, 132:69–79, 2017. ISSN 0921-8009. . URL https://www.sciencedirect.com/science/article/pii/S0921800916301902.

Nicole Maestas and Julie Zissimopoulos. How longer work lives ease the crunch of population aging. *Journal of Economic Perspectives*, 24(1):139–160, 2010. ISSN 0895-3309.

Dale T. Manning and J. Edward Taylor. Migration and fuel use in rural mexico. *Ecological Economics*, 102:126–136, 2014. ISSN 0921-8009. . URL https://www.sciencedirect.com/science/article/pii/S0921800914000925.

Xin Meng. Labor market outcomes and reforms in china. *Journal of economic perspectives*, 26(4):75–102, 2012.

Ministry of Human Resources and Social Security of the People's Republic of China. 2020, 2021. URL https://www.mohrss.gov.cn/wap/fw/rssj/202107/t20210726_419319.html. Accessed: 2025-06-04.

MOHRSS. Statistical bulletin on human resources and social security development in 2022, 2023. URL http://www.mohrss.gov.cn. Accessed August 2025.

Michelle J Moran-Taylor and Matthew J Taylor. Land and leña: linking transnational migration, natural resources, and the environment in guatemala. *Population and Environment*, 32(2-3):198–215, 2010. ISSN 0199-0039.

National Bureau of Statistics of China. Communiqué on the national migrant population development (2023), 2023. URL http://www.stats.gov.cn. Accessed August 2025.

Matthew Pestel Neidell. Air pollution and worker productivity. *IZA World of Labor*, 2023.

Milena Nikolova and Carol Graham. In transit: The well-being of migrants from transition and post-transition countries. *Journal of Economic Behavior Organization*, 112:164–186, 2015. ISSN 0167-2681.

Milena Nikolova, Monica Roman, and Klaus F Zimmermann. Left behind but doing good? civic engagement in two post-socialist countries. *Journal of comparative economics*, 45 (3):658–684, 2017. ISSN 0147-5967.

Sara Ojeda-Benítez, Carolina Armijo-de Vega, and Ma Ysabel Marquez-Montenegro. Household solid waste characterization by family socioeconomic profile as unit of analysis. *Resources, Conservation and Recycling*, 52(7):992–999, 2008. ISSN 0921-3449.

Suehiro Otoma, Hai Hoang, Hai Hong, Izumi Miyazaki, and Ricardo Diaz. A survey on municipal solid waste and residents' awareness in da nang city, vietnam. *Journal of Material Cycles and Waste Management*, 15:187–194, 2013. ISSN 1438-4957.

Ruth Parker and Gary L Kreps. Library outreach: overcoming health literacy challenges. *Journal of the Medical Library Association*, 93(4 Suppl):S81, 2005.

Pew Research Center. Environmental concerns and civic engagement in india, 2018. URL https://www.pewresearch.org. Accessed August 2025.

Etienne Piguet, Antoine Pécoud, and Paul De Guchteneire. Migration and climate change: An overview. *Refugee Survey Quarterly*, 30(3):1–23, 2011. ISSN 1471-695X.

C Arden Pope III and Douglas W Dockery. Health effects of fine particulate air pollution: lines that connect. *Journal of the air waste management association*, 56(6):709–742, 2006. ISSN 1096-2247.

Thurasamy Ramayah, Jason Wai Chow Lee, and Shuwen Lim. Sustaining the environment through recycling: An empirical study. *Journal of environmental management*, 102: 141–147, 2012. ISSN 0301-4797.

Hillel Rapoport and Frédéric Docquier. The economics of migrants' remittances. *Handbook of the economics of giving, altruism and reciprocity*, 2:1135–1198, 2006. ISSN 1574-0714.

Jennifer Roback. Wages, rents, and the quality of life. *Journal of political Economy*, 90 (6):1257–1278, 1982. ISSN 0022-3808.

Sherwin Rosen. Wage-based indexes of urban quality of life. *Current issues in urban economics*, pages 74–104, 1979.

Armando Sanchez-Vargas, Ricardo Mansilla-Sanchez, and Alonso Aguilar-Ibarra. An empirical analysis of the nonlinear relationship between environmental regulation and manufacturing productivity. *Journal of Applied Economics*, 16(2):357–371, 2013. ISSN 1514-0326.

J.M.C. Santos Silva and Silvana Tenreyro. The log of gravity. *The Review of Economics and Statistics*, 88(4):641–658, 2006. .

JMC Santos Silva and Silvana Tenreyro. The log of gravity at 15. *Portuguese Economic Journal*, 21(3):423–437, 2022. ISSN 1617-982X.

Shuai Shao, Baoli Li, Meiting Fan, and Lili Yang. How does labor transfer affect environmental pollution in rural china? evidence from a survey. *Energy Economics*, 102: 105515, 2021. ISSN 0140-9883. . URL https://www.sciencedirect.com/science/article/pii/S0140988321003972.

Wei-Teng Shen, Xuan Yu, Shun-Bin Zhong, and Hao-Ran Ge. Population health effects of air pollution: fresh evidence from china health and retirement longitudinal survey. *Frontiers in Public Health*, 9:779552, 2021. ISSN 2296-2565.

Yu Shen and Wenkai Sun. Information and avoidance behaviour: The effect of air pollution disclosure on labour supply in china. *International Labour Review*, 162(4): 665–686, 2023. ISSN 0020-7780.

Liuhua Shi, Antonella Zanobetti, Itai Kloog, Brent A Coull, Petros Koutrakis, Steven J Melly, and Joel D Schwartz. Low-concentration pm2. 5 and mortality: estimating acute and chronic effects in a population-based study. *Environmental health perspectives*, 124 (1):46–52, 2016. ISSN 0091-6765.

JMC Santos Silva and Silvana Tenreyro. The log of gravity. *The Review of Economics and statistics*, 88(4):641–658, 2006. ISSN 0034-6535.

Antonio Spilimbergo. Democracy and foreign education. *American economic review*, 99 (1):528–43, 2009. ISSN 0002-8282.

George J Stigler. The economics of information. *Journal of political economy*, 69(3): 213–225, 1961. ISSN 0022-3808.

WZ Sun, XN Zhang, and SQ Zheng. Air pollution and spatial mobility of labor force: Study on the migrants' job location choice. *Econ. Res. J*, 54:102–117, 2019.

Shinsuke Tanaka. Environmental regulations on air pollution in china and their impact on infant mortality. *Journal of health economics*, 42:90–103, 2015. ISSN 0167-6296.

Matthew J. Taylor, Michelle J. Moran-Taylor, Edwin J. Castellanos, and Silvel Elías. Burning for sustainability: Biomass energy, international migration, and the move to cleaner fuels and cookstoves in guatemala. *Annals of the Association of American Geographers*, 101(4):918–928, 2011. ISSN 0004-5608 1467-8306. .

Frans Thissen, Joos Droogleever Fortuijn, Dirk Strijker, and Tialda Haartsen. Migration intentions of rural youth in the westhoek, flanders, belgium and the veenkoloniën, the netherlands. *Journal of Rural Studies*, 26(4):428–436, 2010. ISSN 0743-0167.

Trevor Tombe and Xiaodong Zhu. Trade, migration, and productivity: A quantitative analysis of china. *American Economic Review*, 109(5):1843–1872, 2019.

Michele Tuccio and Jackline Wahba. Return migration and the transfer of gender norms: Evidence from the middle east. *Journal of Comparative Economics*, 46(4):1006–1029, 2018. ISSN 0147-5967. . URL https://www.sciencedirect.com/science/article/pii/S0147596718302518.

Michele Tuccio and Jackline Wahba. *Social Remittances*. Springer, 2020. ISBN 3319573659.

Michele Tuccio, Jackline Wahba, and Bachir Hamdouch. International migration as a driver of political and social change: evidence from morocco. *Journal of Population Economics*, 32:1171–1203, 2019. ISSN 0933-1433.

UNESCAP. Urbanisation and internal migration in indonesia, 2019. URL https://www.unescap.org. Accessed August 2025.

Aaron Van Donkelaar, Randall V Martin, Michael Brauer, N Christina Hsu, Ralph A Kahn, Robert C Levy, Alexei Lyapustin, Andrew M Sayer, and David M Winker. Global estimates of fine particulate matter using a combined geophysical-statistical method with information from satellites, models, and monitors. *Environmental science technology*, 50(7):3762–3772, 2016. ISSN 0013-936X.

John Vanderkamp. The gravity model and migration behaviour: An economic interpretation. *Journal of Economic Studies*, 4(2):89–102, 1977. ISSN 0144-3585.

Andreas Vårheim, Sven Steinmo, and Eisaku Ide. Do libraries matter? public libraries and the creation of social capital. *Journal of documentation*, 64(6):877–892, 2008.

Jackline Wahba. Selection, selection, selection: the impact of return migration. *Journal of Population Economics*, 28(3):535–563, 2015. ISSN 1432-1475.

Li Wang, Yunhao Dai, and Dongmin Kong. Air pollution and employee treatment. *Journal of Corporate Finance*, 70:102067, 2021. ISSN 0929-1199.

Deborah Williams and Krishna P Paudel. Migration, remittance, and adoption of conservation practices. *Environmental Management*, 66(6):1072–1084, 2020. ISSN 0364-152X.

Jeffrey M Wooldridge. *Econometric analysis of cross section and panel data*. MIT press, 2010. ISBN 0262296799.

World Bank. Indonesia: Toward a circular economy for plastic waste, 2019. URL https://www.worldbank.org/en. Accessed August 2025.

World Bank and Institute for Health Metrics and Evaluation. *The Cost of Air Pollution: Strengthening the Economic Case for Action*. World Bank Group, Washington, DC, 2016. URL https://openknowledge.worldbank.org/handle/10986/25013. Accessed: 2025-06-02.

World Bank. The future of social protection in east asia and the pacific, 2020. URL https://www.worldbank.org. Accessed August 2025.

World Health Organization. Air pollution. https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health, 2021. Accessed: 2025-06-02.

Wenjie Wu, Yanwen Yun, Jingtong Zhai, Yeran Sun, Guanglai Zhang, and Ruoyu Wang. Residential self-selection in the greenness-wellbeing connection: a family composition perspective. *Urban Forestry Urban Greening*, 59:127000, 2021a. ISSN 1618-8667.

Y. Wu, B. Entwisle, C. Sinai, and S. Handa. Migration and fuel use in rural zambia. *Popul Environ*, 43(2):181–208, 2021b. ISSN 0199-0039 (Print) 0199-0039 (Linking). . URL https://www.ncbi.nlm.nih.gov/pubmed/34924664.

Shijiang Xiao, Huijuan Dong, Yong Geng, and Matthew Brander. An overview of china's recyclable waste recycling and recommendations for integrated solutions. *Resources, Conservation and Recycling*, 134:112–120, 2018.

Tingting Xie, Ye Yuan, and Hui Zhang. Information, awareness, and mental health: Evidence from air pollution disclosure in china. *Journal of Environmental Economics and Management*, 120:102827, 2023. ISSN 0095-0696.

Lilai Xu, Tao Lin, Ying Xu, Lishan Xiao, Zhilong Ye, and Shenghui Cui. Path analysis of factors influencing household solid waste generation: a case study of xiamen island, china. *Journal of Material Cycles and Waste Management*, 18:377–384, 2016. ISSN 1438-4957.

Fujin Yi, Chang Liu, and Zhigang Xu. Identifying the effects of migration on parental health: Evidence from left-behind elders in china. *China Economic Review*, 54:218–236, 2019. ISSN 1043-951X.

Qian Yue, Yan Song, Ming Zhang, Xueli Zhang, and Longke Wang. The impact of air pollution on employment location choice: Evidence from china's migrant population. *Environmental Impact Assessment Review*, 105:107411, 2024.

Yingying Zeng, Yuanfei Cao, Xue Qiao, Barnabas C Seyler, and Ya Tang. Air pollution reduction in china: Recent success but great challenge for the future. *Science of the Total Environment*, 663:329–337, 2019. ISSN 0048-9697.

Chenglei Zhang, Minzhe Du, Liping Liao, and Wenxiu Li. The effect of air pollution on migrants' permanent settlement intention: Evidence from china. *Journal of Cleaner Production*, 373:133832, 2022. ISSN 0959-6526.

Jipeng Zhang, Ru Wang, and Chong Lu. A quantitative analysis of hukou reform in chinese cities: 2000–2016. *Growth and Change*, 50(1):201–221, 2019. ISSN 0017-4815.

Junjie Zhang and Quan Mu. Air pollution and defensive expenditures: Evidence from particulate-filtering facemasks. *Journal of Environmental Economics and Management*, 92:517–536, 2018. ISSN 0095-0696.

Kevin Honglin Zhang and SONG Shunfeng. Rural–urban migration and urbanization in china: Evidence from time-series and cross-section analyses. *China economic review*, 14(4):386–400, 2003. ISSN 1043-951X.

Xin Zhang, Xiaobo Zhang, and Xi Chen. Happiness in the air: How does a dirty sky affect mental health and subjective well-being? *Journal of environmental economics and management*, 85:81–94, 2017. ISSN 0095-0696.

Xin Zhang, Xi Chen, and Xiaobo Zhang. The impact of exposure to air pollution on cognitive performance. *Proceedings of the National Academy of Sciences*, 115(37): 9193–9197, 2018. ISSN 0027-8424.

Zhihao Zhao, Xin Lao, Hengyu Gu, Hanchen Yu, and Ping Lei. How does air pollution affect urban settlement of the floating population in china? new evidence from a push-pull migration analysis. *BMC Public Health*, 21:1–15, 2021.

Yu Zhu, Wenfei Winnie Wang, Liyue Lin, Jianfa Shen, and Qiang Ren. Return migration and in situ urbanization of migrant sending areas: Insights from a survey of seven provinces in china. *Cities*, 115:103242, 2021. ISSN 0264-2751.

Joshua Graff Zivin and Matthew Neidell. The impact of pollution on worker productivity. *American Economic Review*, 102(7):3652–3673, 2012. ISSN 0002-8282.

Index