THE IMPACT OF MIGRATION ON NATURAL AND HUMAN CAPITAL IN DEVELOPING COUNTRIES

by

Tae Hyun Yoon

(Under the Direction of Mateusz Filipski)

ABSTRACT

This dissertation delves into the socio-economic impact of migration in Mongolia and Mexico, focusing on urbanization and demographic shifts. The first two chapters assess Mongolia's urbanization, examining the environmental degradation from informal settlements and human capital accumulation influenced by in-migration. The environmental implications of these slum areas are explored via a unique demographic dataset, synthesized using a deep learning model and high-resolution satellite imagery. In the second chapter, the impact of in-migration on local students' education is investigated, analyzing the balance between the crowd-out and crowd-in effects in education and labor markets. In the last chapter, we have focused on the international migration's effects on children's welfare in Mexico's agricultural households. This study probes the correlation between cash transfers, children's health, and migration decisions, aiming to address the heterogeneity between cash transfer beneficiaries and non-beneficiary's households.

This dissertation underscores the correlation between the number of Mongolia's *'ger'* settlements and pollution levels, providing insights into urbanization's environmental impact. It reveals the nuanced relationship between in-migration and educational investment in Mongolia's youth, suggesting higher immigration rates generally enhance human capital development. The final chapter elucidates the intricate interplay of migration, cash transfer programs, and children's nutritional status in rural Mexico. Overall, this research highlights the complex environmental and socio-economic effects of migration, offering invaluable insights for future policy development and research. It employs unconventional datasets and computational methodologies, paving the way for future explorations of migration in developing nations.

INDEX WORDS: Development, Human Migration, Urbanization, Air Pollution, Environmental Degradation, Human Capital, Nutrition

THE IMPACT OF MIGRATION ON NATURAL AND HUMAN CAPITAL IN DEVELOPING COUNTRIES

by

Tae Hyun Yoon

B.A. in Economics, Soongsil University, 2014

M.A. in Economics, Soongsil University, 2017

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2023

THE IMPACT OF MIGRATION ON NATURAL AND HUMAN CAPITAL IN

DEVELOPING COUNTRIES

by

Tae Hyun Yoon

Major Professor: Mateusz Filipski
Committee: Susana Ferreira
Yukiko Hashida

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
August 2023

DEDICATION

The journey of acquiring my Ph.D., laden with academic exploration and personal growth, would not have been navigable without the unwavering support of my family. My profound gratitude extends to my father, Jaechool Yoon, my mother, Kyungsook Lee, and my sister, Dahye Yoon, who never once faltered in their belief in me, my aspirations, and my dreams. Even when I deviated from traditional paths, be it when I dropped out of high school, when I initially was not considering going to college, or when I was wrestling with my health condition, Tourette syndrome, their eyes held a guiding light, directing me towards unwavering determination and purpose. Their steadfast faith and support have been my strongest allies in Athens and will continue to be as I prepare for the next chapter of my journey in the U.S.

Above all, the greatest fortune in my life has been meeting my wife, Jaeseon Song, an extraordinary woman whose wisdom and intellect know no bounds. She has been more than just my life partner, she has been my unwavering supporter, sharing in my challenges and serving as a trusted confidante in my research. Her constant reassurances reinforced my belief in my uniqueness and potential to make a significant contribution. Remarkably, despite the fact that my research fell outside her own area of interest and expertise, she always listened attentively to my research challenges, continually encouraging me to yield results of immense value.

And so, with the deepest sense of gratitude and the highest regard, I dedicate this dissertation to my cherished family and my beloved wife.

ACKNOWLEDGEMENTS

I extend my deepest gratitude to my advisor, Dr. Filipski. His consistent encouragement and insightful guidance were instrumental in evolving my initial, naïve ideas into a polished, valuable dissertation. His patience and wisdom not only guided my development as a researcher, but also ensured that I maintained my enthusiasm for exploring and expanding upon new concepts and areas.

Additionally, when I first arrived at UGA, my programming and statistical software skills were practically nonexistent, which often led to feelings of inadequacy. However, under Dr. Filipski's mentorship, I honed these skills and integrated advanced technologies into my research. This not only added depth to my dissertation, but also equipped me with essential competencies that will undoubtedly prove invaluable throughout my career.

I remember my arrival at UGA six years ago, managing both Tourette syndrome and ADHD, conditions I had been medicating for over five years before I came to the US from South Korea. The initial phase of my Ph.D. program was challenging so that I would often find myself nodding off in class and struggling with coursework. But thanks to my wife's unwavering support and the impending pressure of preliminary exams, I decided to discontinue my medication at the beginning of 2018, a decision I have upheld ever since.

I believe Dr. Filipski was likely to aware of my condition or the symptoms of my Tourette's syndrome, but he never brought it up, and I never felt nervous or uncomfortable discussing my academic challenges with him. His acceptance and support enabled me to overcome the obstacles posed by my condition, and instead of using it as an excuse, I used it as a catalyst to work harder and successfully complete my Ph.D. program.

I have long awaited the moment when I could express my heartfelt appreciation to my professor, and it brings me immense joy to finally pen down these words of acknowledgement. His guidance has been nothing short of transformational, and for that, I am truly grateful.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

This dissertation provides a comprehensive analysis of the complex implications of both internal and external migration on natural and human capital within developing countries. Previous migration studies have largely focused on the influence of international migration or remittances on economic improvement or poverty reduction within migrants' origin countries (Acosta et al., 2007; Adams et al., 2008; Amuedo-Dorantes & Pozo, 2006; Cox-Edwards & Rodríguez-Oreggia, 2009; López Córdova, 2018; Semyonov & Gorodzeisky, 2008; Yang & Choi, 2007). However, there has been a recent shift towards the study of internal migration, now exceeding international migration volumes (Lucas, 2015; Selod & Shilpi, 2021). These shifts often result not only from traditional migration patterns seeking better economic opportunities, but also from factors such as conflict, violence, or natural disasters.

As a result, the urban population has grown rapidly, with the percentage of the global population residing in urban areas increasing from 43% in 1990 to an estimated 60% by 2030, concentrating about 90% of future urban growth in low- and middle-income countries (UN-HABITAT, 2016, 2018). This rapid urban expansion, often unplanned, has triggered various environmental and socio-economic issues that affect both original residents and immigrants, thus emerging as a primary concern for government long-term urban planning and economic growth strategies.

To address the multifaceted natural and human capital impact of demographic changes, this study focuses on Mongolia, thereby broadening our understanding of the implications of accelerated urbanization in developing nations. The first two chapters of this dissertation highlight environmental degradation and human capital accumulation in response to rapid influx of migrants to urban areas within a Mongolian context.

The first chapter investigates the influence of urbanization, specifically the proliferation of informal settlements, on air quality. Using advanced computational image processing technologies and unconventional satellite imagery data, we generated a detailed demographic dataset at the sub-district level. This enabled us to accurately estimate the causal relationship between the number of informal settlement households near the 14 monitoring stations and six major air pollutant concentration levels, highlighting potential health risks posed by increased informal settlements.

The second chapter explores the impact of immigration rates on the educational investment decisions of native youth across 22 urban areas in Mongolia, including Ulaanbaatar and 21 provincial capitals. It considers the complex interplay between the educational resource market and labor market, influenced by immigrant students and low-skilled immigrant workers respectively, addressing the intricate trade-off between the crowd-out effect driven by increased educational demand and the crowd-in effect due to changes in wage and unemployment rates.

In the final chapter, we return to the conventional perspective of migration studies, examining the international migration of household members from Mexico to the US. We assess the indirect welfare effects resulting from labor loss in agricultural households due to migration, investigating whether the migration of one or more household members has positive or negative effects on children from agricultural households in Mexico.

Each of these chapters has been meticulously and strategically designed to estimate the causal impact of internal and external migration on natural and human capital in developing countries. This enriches the academic discourse on human mobility and demographic dynamics by emphasizing their complex relationship. Moreover, the empirical results provide potential policy suggestions and propose strategies to mitigate the negative effects of rapid urbanization.

CHAPTER 2

THE IMPACT OF URBAN SPRAWL ON AIR QUALITY: THE CASE OF ULAANBAATAR, MONGOLIA[1]

[1] Yoon, T., Filipski, M., Ferreira, S., and Hashida, Y. To be submitted to *Ecological Economics*

Abstract

The process of rapid urbanization, driven by both voluntary and involuntary factors, is particularly pronounced in less-developed countries. This rapid urban expansion is fueling the growth of informal settlements. These incoming migrants often suffer from restricted access to basic services, serving as catalysts for environmental degradation and various health hazards. Air pollution stands out as the primary concern linked to these living conditions in Mongolia. In an attempt to better understand this issue, this study harnesses cutting-edge image processing techniques. They have been employed to create high-frequency demographic data, representing informal settlements, and to delineate study areas at the neighborhood level by integrating deep learning and satellite images. The focus of this investigation is the relationship between the slum population and air pollution levels in Ulaanbaatar, the capital city of Mongolia. For this purpose, a fixed-effect regression model was applied. This model treats six air pollution indicators ($PM_{2.5}$, PM_{10}, CO, SO_2, NO_2, O_3) as dependent variables, and the number of '*gers*' (traditional Mongolian herder's portable tents) within a half-mile radius of a monitoring station as an independent variable. The empirical findings of the study shed light on the relationships between the number of informal settlers in Ulaanbaatar and levels of air pollution concentration. Specifically, there is a positive association with $PM_{2.5}$, PM_{10}, SO_2, and CO, a negligible association with NO_2, and a negative association with O_3.

Introduction

In 1990, urban areas were home to 43% of the global population, accounting for approximately 2.3 billion people. This figure rose to 54% by 2015, which represented roughly 4 billion people. Forecast suggests that by 2030, the proportion of the worldwide population living in urban regions will reach 60%, with approximately 90% urban growth expected to take place in low- and middle-income countries (UN-HABITAT, 2016, 2018).

Conventional economic theory posits that the concentration of economic activities in urban areas induces the migration of the workforce from rural to urban regions in search of job prospects and improving living standards such as access to higher income, education, and healthcare services. However, in recent decades, especially in developing countries, the growth of urban population has been driven not only by voluntary migration in searching better economic opportunities, but also by the involuntary migration of people who has been forced to move due to income shocks resulting from such as violent conflicts or natural disasters, and migration has become a direct adaptation strategy for rural communities affected by the impact of climate change (Mastrorillo et al., 2016; McLeman & Smit, 2006; Selod & Shilpi, 2021). As a result, around 7 million people have been displaced from their homes due to natural disasters since the end of 2020 (*Environmental Migration*, 2021).

The rapid urbanization and population growth rates resulting from the significant influx of migrants and higher fertility rates in developing countries have created a situation where the available adequate and affordable housing is being outpaced. As a result, low-income population, mostly migrants from rural areas, often have limited financial resources, have been finding it challenging to access adequate housing, leading to a substantial increase in the number of individuals living in slum or informal settlements. In 2019, over 1 billion people reside in informal settlements, implying that roughly one in every eight people in the world lives in inadequate environments, enduring harsh living conditions and lacking access to basic infrastructure. This situation is deeply concerning as the United Nations predicts that the number of individuals residing in informal settlements will rise to three billion by 2050 (GSG, 2022; López et al., 2019; UN-HABITAT, 2016; UNSD, 2019; West et al., 2021).

The proliferation of slums and informal settlements in developing countries, which are often lacking in basic infrastructure and services, such as heating, sanitation, clean water, sewage, electricity, cooking fuels, and public transportation, has led to environmental degradation and health hazards among urban populations. One of the most pressing issues that arises from unplanned urban sprawl is air pollution, which poses an unavoidable threat to public health (GSG, 2022; López et al., 2019; Pandey et al., 2018; Porio, 2011; UN-HABITAT, 2018). In 2016, 90 % of urban dwellers were exposed to polluted air that failed to meet the World Health Organization's (WHO) air quality guidelines, with low- and middle-income countries being disproportionately affected. Of the cities with population over 100,000 inhabitants, 97 % of those in low- and middle-income countries failed to meet air quality standards, compared to only 49 % in high-income nations. The WHO estimates that outdoor air pollution is responsible for seven million deaths worldwide annually, with more than 90 % of fatalities caused by air pollution occurring in low- and middle-income countries (UNEP, 2019; UNSD, 2019; WHO, 2021a). The households, primarily located in slums or informal settlements, that burn wood, charcoal, or other inappropriate fuels for cooking, heating, or lighting purposes are significant contributors to both indoor and outdoor pollution. Consequently, the harmful effects of environmental pollution caused by population growth disproportionally affect the well-being of impoverished households in developing countries (UNEP, 2019). The rapid expansion of urban areas, characterized by the emergence of slums and informal settlements, resulting in heightened air pollution levels, has evolved into a significant concern for both local communities and policymakers.

This research aims to add to the current knowledge and help in policy making by exploring the connection between recent trends in internal migration, the increase in informal settlements, and environmental degradation, specifically air pollution. This study constructs and utilizes the innovative dataset on informal settlements in Ulaanbaatar, Mongolia. Specifically, the study examines the causal relationship between the concentration of slum population in urban areas of developing countries and six major air pollutant concentrations, $PM_{2.5}$, PM_{10}, SO_2, NO_2, CO, O_3, at the sub-district level. While previous research has primarily focused on examining the relationship between population or immigrant growth and environmental degradation at broader scale, such as the global, national, or city levels, this study focuses on the neighborhood level, providing a more nuanced understanding of how informal settlements contribute to local air pollution. The study's findings have potential implications for policy interventions aimed at improving the living conditions and health outcomes of slum residents while mitigating environmental damage. It is particularly important because many megacities in developing world suffer from exceptionally high levels of air pollution, which is highly associated with the health conditions of residents, and this adverse effect disproportionately falls on poor people.

Furthermore, another significant contribution of this research lies in its innovative data collection methodology. For the first time, state-of-the-art image processing techniques are being applied to quantify demographic factors within the current global migration trend, particularly in the developing world. This research focuses on Ulaanbaatar and identifies the number of *'gers'*, which are considered informal settlers in the scope of this study. The number of informal settlers is measured at the neighborhood level, with the count of *'gers'* within a half-mile radius of each of the 14 monitoring stations scattered throughout Ulaanbaatar, encompassing both non-slum urban centers and slum areas. Using high-resolution satellite imagery combined with cutting-edge deep learning techniques, we have quantified the number of informal settlements. This approach not only distinguishes informal settlers from urban areas but also estimates the demographic impact on environmental degradation over time, with special focus on the human migration trends mostly seen in developing countries.

Unlike previous studies that have primarily relied on infrequent data collection methods that cover larger areas, like census or household surveys, our methodology allows for the construction of more accurate and frequent dataset, even for smaller study regions. Between September 2015 and May 2019, we gathered 172 observations for the 14 monitoring stations. We factored in both high-density and low-density *'ger'* areas within a half-mile radius of each monitoring station to effectively represent Ulaanbaatar, a city noted for its rapid immigration and significant pollution levels. Leveraging this accurate dataset, the findings from the fixed effect econometric specification for '*gers*' within a half-mile radius of 14 monitoring stations in Ulaanbaatar and six air pollution sources ($PM_{2.5}$, PM_{10}, SO_2, CO, NO_2, and O_3) indicate a significant relationship. The empirical results show that on average, the presence of one additional '*ger*' in Ulaanbaatar was found to significantly increase the levels of $PM_{2.5}$, PM_{10}, SO_2, and CO by 0.0517 μg/m³, 0.140 μg/m³, 0.0215 μg/m³, and 0.989 μg/m³, respectively. However, we did not find a significant association between the number of informal settlers and the levels of NO_2. Interestingly, there was a negative relationship between the number of informal settlement households and O_3 levels. However, it is important to note that this negative relationship does not imply that informal settlers reduce O_3 levels. Instead, it is highly plausible to explain that O_3 levels can be influenced by various external factors and chemical reactions with other air pollutant sources, such as nitrogen oxides (NO_x) or volatile organic compounds. These pollutants can be increased by incoming population from rural areas, especially under the influence of heat and sunlight.

The structure of this paper is organized as follows: Section 1 provides an introduction, while Section 2 describes the study areas, Ulaanbaatar, the capital city of Mongolia. In Section 3, we review previous literature, considering three main aspects related to the relationship between immigration and environmental degradation. Section 4 describes our data, including the introduction of the new and innovative demographic dataset created using satellite images and deep learning-based image processing techniques. In Section 5, we present our empirical strategies and results, which involve various specifications to ensure the reliability and consistency of our results. Finally, in Section 6, we conclude the paper and provide further discussion points for future research.

Study Area

Mongolia has experienced a significant rise in its urban population in recent years, largely driven by rural to urban migration. This trend was initiated by a series of policy reforms that were implemented after the fall of the Soviet Union in 1989, which fostered the movement of individuals from rural areas to urban centers in search of improved economic prospects. Additionally, extreme climate shocks such as '*dzud*[2]', have been a significant contributing factor to this migration trend. For instance, during the winter of 2009-2010, the nation lost almost 22 % of its livestock, which has been the primary livelihood source of Mongolian herders, due to the extreme weather event, resulting in many herders seeking refuge in urban areas, Ulaanbaatar(UB), the capital city of

[2] A compounded weather shock of severe cold and snowy winter followed by drought in summer kills animals on a massive scale in Mongolian terms.

Mongolia (Byambadorj et al., 2011; Cui et al., 2019; Dore & Nagpal, 2006; Mayer, 2016; Park et al., 2019; Xu et al., 2021). As a result, the urbanization rate in Mongolia has increased significantly, with 70 % of the total population residing in urban areas, exceeding the Asian average urbanization rate of 50 % (ADB, 2018). Furthermore, the issue of informal settlement in Ulaanbaatar is compounded by the migration of rural populations who bring with them the use of *'gers'* – portable felt tent commonly used as dwelling by Mongolian nomads – which they continue to inhabit in urban centers.

However, the large-scale migration of people from rural areas to UB has raised concerns as most individuals live in large-scale informal areas, locally known as the '*ger district* [3]', that lack planning and regulation. In contrast to other countries that experience increasing slum areas, residents of Mongolia's *'ger districts'* are able to legally obtain ownership of land due to the country's Law on Land and Law on Allocation of Land to Mongolia Citizens for Ownership, which allows citizens to privatize land for residential purposes. For example, when migrants move to Ulaanbaatar, they can legally acquire 0.07 hectares of land for a cost of $2 and install their portable traditional nomadic houses, known as *'ger'* (Bauner & Richter, 2006; Jun, 2021; Miller, 2017; Park et al., 2019). Therefore, informal settlements in Ulaanbaatar are considered legally formalized under these laws. However, it is important to note that the definition of slums and informal settlements is not only determined by land tenure, but also other operational and conditional factors.

In line with SDG Target 11.1, which seeks to ensure access to affordable, safe, and adequate housing and basic services for all individuals, UN-Habitat defines slums and informal settlements as areas where residents face one or more following difficulties. First, slum households

[3] The *'ger'* is a traditional Mongolian round-shaped portable house used for a nomadic life on the steppe. The 'ger district' means the informal settlement of multiple *'gers'* characterized by large areas without public services and infrastructures.

are identified by operational criteria such as lack of access to improved water sources, sanitation facilities, sufficient living space, stable housing, or tenure security. Also, informal settlements are characterized by high population densities and low-quality physical housing. Additionally, areas can be considered informal settlements if inhabitants lack security of tenure, basic services, and formal city infrastructure or housing that does not adhere to present planning and construction rules, situated in hazardous areas, and lack a municipal permit (Fekade, 2000; Park et al., 2019; UN-HABITAT, 2018).

The expansion of these small migrant parcels has led to the creation of extensive '*ger districts.*' These areas primarily established by '*gers*' located in Ulaanbaatar's outskirts, are characterized as slums or informal settlements according to the UN-Habitat's definition. These districts currently cover 75 % of Ulaanbaatar's territory, with 65 % of the city's residents living in informal areas as of 2018. The current population of 1.5 million city residents in Ulaanbaatar is nearly three times its original capacity of 500,000. As a result, at least two out of three residents are excluded from basic infrastructure and public services, pointing to the serve lack of essential amenities in the city's informal settlements (Jun, 2021).

The present study focuses on the context of Mongolia's internal migration, which provides unique opportunities to explore the causal relationship between informal settlements and air pollutant concentration in developing countries. This study offers two distinct advantages in identifying informal settlement variables. Firstly, in contrast to other countries where the precise identification of informal settlements can be challenging, the use of portable traditional housing known as *'ger'* in Mongolia allows for a more accurate measurement of informal settlement population. Since *'gers'* have a consistent size and shape, they can be more precisely identified and measured using satellite imagery and deep learning object detection modeling.

Secondly, in other countries, it can be difficult to determine whether informal settlers have moved back to rural areas and abandoned their housing or are still residing in urban slum areas. However, in Mongolia, residents have a long nomadic history and often move with their housing, either within urban areas or back to rural areas. As a result, measurement error can be minimized, and the population of informal settlements can be more accurately determined using Mongolia's conventional lifestyle.

Literature Review

Three primary arguments support the question of whether increased migration rates can exacerbate environmental degradation of host communities. Firstly, some researchers have proposed that population growth and urban expansion, driven by an influx of newcomers, directly contribute to local environmental harm. This debate revolves around the potential significant impact of a rise in migrant influx rate at city, state, or national levels on local environmental conditions. Numerous researchers contend that a substantial increase in migration can boost population growth beyond an urban areas' sustainable limit, consequently straining local infrastructure and natural resources. These resources may be inadequate to sustain increased demands for water and food, manage congestion and pollution, and handle waste production (Bartlett & Lytwak, 1995; Daily et al., 1995; Dedeoğlu et al., 2021; Price & Feldmeyer, 2012). Various studies have examined the impact of migration, both internal and international, on environment degradation and found a significant negative relationship between immigration and air pollution. Qin and Liao (2016) and Rafiq et al. (2017) conducted an empirical investigation into the relationship between internal migration and environmental degradation in China, concluding that internal migration was positively correlated with increased levels of air and water pollution. Meanwhile, Alola and Kirikkaleli (2019) and Dedeoğlu et al. (2021) focused on the relationship between international immigration and CO_2 emission in the United States, finding a significant causal relationship and consistent evidence that increased incoming population elevates CO2 emissions.

Secondly, other researchers argue that the population growth induced by incoming migration can lead to more damaging ecological effects compared to other types of population growth. Incoming settlers tend to cluster geographically based on ethnic networks or economic conditions, with informal settlements serving as an example. These situations lead to a disproportionate increase in local population density, consequently intensifying environmental issues (Dedeoğlu et al., 2021; Price & Feldmeyer, 2012). For instance, Komatsu et al. (2013) investigated the effect of internal migration on residential energy consumption and CO_2 emissions in Hanoi, Vietnam's capital city. Their study found that while migration between urban areas had no significant impact on CO_2 emissions, rural-to-urban migration was positively and significantly associated with air pollution. Similarly, in the context of high urbanization rates in China, the study conducted by Long et al. (2022) revealed that internal migration had a positive impact on national carbon emissions, and magnitude of internal migration effects on emission was greater than emissions caused by net population growth.

Lastly, contrary to the second argument, some studies have suggested that incoming migrants exert a minimal impact on local environmental degradation. This theory's justification stems from the *ecological footprint perspective*, which proposes that although population influx contributes to local population growth, it could result in lesser ecological damage compared to other forms of local population growth. This is attributed to the fact that incoming settlers, particularly those from rural or less developed regions, usually have smaller ecological footprints than native-born residents and adopt lifestyles that are less consumptive and place a lesser burden on the environment (Dietz & Rosa, 1997; Price & Feldmeyer, 2012). Squalli (2010) examined the link between immigrants and air pollution in the United States and found that states with a larger proportion of US-born residents had higher levels of polluted gas emissions, whereas states with a greater proportion of foreign-born residents had lower levels of levels of NO_2 and SO_2 emissions. In a similar vein, Price and Feldmeyer (2012) also study employed spatial panel analysis to examine the relationship between population dynamics and air pollution in the United States. The study found that international immigration, mainly from developing countries, was not associated with air pollution levels in the U.S. In contrast, other forms of population growth, such as domestic migration and natural population growth, had significantly negative effects on air quality.

Despite thorough research exploring the impact of the incoming population on urban environmental degradation from three unique theoretical viewpoints, a definitive consensus remains elusive concerning whether the influx of people contributes to adverse environmental outcomes. Furthermore, there is a notable dearth of research studying the relationship between the rising count of households in informal settlement zones and their effect on air pollution in developing countries context. Past studies have yielded conflicting outcomes, suggesting that a surge in informal settlers could either significantly exacerbate or marginally affect air pollution. On the one hand, settlers residing in slum areas could potentially contribute to environmental degradation due to their inadequate resources and population pressure. However, it is also feasible that inhabitants of informal settlements, who often consume less and exert less pressure on the environment, could potentially have a smaller ecological footprint than established urban residents.

The current study aims to estimate that relationship between the number of *'gers'* at the sub-district level in Ulaanbaatar and air pollution concentration. By identifying and counting the number of *'gers'* within a half-mile radius of each of the 14 monitoring stations, covering not only slum district but also urban center, in Ulaanbaatar and analyzing data using a deep learning technique with high-resolution satellite imagery, we explore the causal relationship between informal settlements and air pollution. Through our analysis, this study contributes to narrowing the knowledge gap in the field of environmental research, particularly in the context of the impact of informal settlements on air pollution concentration in developing countries.

Data

To investigate the impact of the influx of inhabitants, specifically the quantity of '*gers*' in Mongolia's context, on the air quality of Ulaanbaatar, we compiled three essential datasets: (1) air pollution concentration levels collected from 14 monitoring stations in Ulaanbaatar, sourced from the OpenAQ API[4]; (2) meteorological data retrieved from the World Weather Online API[5]; and (3) the number of '*gers*' located around the 14 monitoring stations in Ulaanbaatar. These data sources allowed us to explore the connection between the population of informal settlements and air pollution concentration of 14 monitoring stations in Ulaanbaatar, thus providing a more thorough understanding of the problem.

Emission Data

We obtained subdistrict-level hourly concentration of six key air pollutant sources ($PM_{2.5}$, PM_{10}, CO, SO_2, NO_2, O_3) for 14 monitoring stations in Ulaanbaatar between 2015 and 2019 using the OpenAQ platform. OpenAQ is non-profit organization that publicly shares air quality data. Given that all six air pollution concentration data were not available for each of the 14 monitoring

[4] https://openaq.org/
[5] https://www.worldweatheronline.com/

stations throughout the entire study period, we collected an unbalanced panel dataset, utilizing as much data as was accessible for each monitoring station.

The dependent variables in this study are six air pollutant concentrations, namely particulate matter ≤2.5 mg ($PM_{2.5}$), particulate matter ≤10 mg (PM_{10}), sulfur dioxide (SO_2), nitrogen dioxide (NO_2), Carbon Monoxide (CO), and Ozone (O_3), all measured in micrograms per cubic meter (mg/m^3). $PM_{2.5}$ and PM_{10} are types of airborne particulate matter consisting of small solid and liquid droplets containing acid, organic chemicals, metals, soil particles, dust, and biological elements like pollen and fungal spores. While PM_{10} has been commonly used as an air pollution indicator in developing countries, researchers have increasingly focused on $PM_{2.5}$ due to its more severe impact on respiratory health (Sanbata et al., 2014). Inhalation of $PM_{2.5}$ can lead to respiratory diseases, asthma, chronic bronchitis, and other lung-related illnesses, as it can penetrate the bronchial and alveolar regions of the human body. The primary sources of particulate matter are vehicle emissions, biomass burning, combustion-derived carbon containing ultrafine nitrates and sulfate, air dust, and endotoxins from biological particles (Pires et al., 2008a, 2008b; Rahman et al., 2019; WHO, 2021b). Additionally, Carbon monoxide (CO), Sulfur dioxide (SO_2), Nitrogen dioxide (NO_2), and Ozone (O_3) are hazardous atmospheric gases. The primary anthropogenic sources of these emissions predominantly result from the combination of fossil fuel combustion, domestic heating, power generation, and transportation (Sharma et al., 1999). Traffic, specifically from cars, trucks, and buses, is the primary contributor to CO and NO_2 emissions, while electricity generation and coal combustion represent the largest sources of SO_2 emissions.

Airborne pollutants, comprising both particulate matter and gases, can be produced either through direct emission from sources or as a result of chemical reactions involving other pollutants. For example, nitrogen dioxide (NO_2) can be generated by a chemical reaction between nitrogen monoxide (NO) and Ozone (O_3). Conversely, nitrogen dioxide (NO_2) can react with oxygen (O_2) to generate ozone (O_3). For instance, during weekends when traffic volume and industrial activities decrease, the level of nitrogen monoxide (NO) declines, leading to an increase in ozone (O_3) levels due to a reduced production of nitrogen dioxide (NO_2). Table 2.1 presents the specific emission sources of these major air pollutants and their associated health impacts within the context of Mongolia.

Weather Data

The meteorological data includes temperature, wind speed, wind direction, precipitation levels, and atmospheric pressure. We then matched the subdistrict-level hourly pollution data with the city-level hourly meteorological data to create a comprehensive dataset for our analysis. The air quality in urban areas is primarily determined by the quantity and chemical composition of emissions, as well as meteorological factors that can impact the concentration and dispersion of pollutants. The dispersion of pollutants can be influenced by wind velocity and direction, while chemical reactions of pollutants may be influenced by temperature, precipitation, and pressure (Buchholz et al., 2010; Chen et al., 2008). However, a challenge arises as weather conditions may also be associated with population density and air pollution levels. Moreover, the effect of weather on observed air pollution trends may not follow a linear pattern and could be subject to interaction effects (Cole et al., 2020). As a result, our empirical model incorporates quadratic terms of meteorological factors to account for these complexities.

Identifying informal settlements in Ulaanbaatar

In this study, we measure and quantify the number of *'gers'* within a half-mile radius of monitoring stations by integrating data from very high-resolution (VHR) remote sensing images with advanced deep learning technology.

Typically, studies identify migration based on census or household survey data, although the availability of such data can be inconsistent. The use of these data sources is subject to several limitations. Firstly, household surveys are resource-intensive and time-consuming, making them infrequent, especially in developing countries. Secondly, survey data generally represent national or regional levels rather than local or subdistrict levels. Lastly, while georeferenced data is collected during surveys, the release of geographic information typically lags, creating a gap in real-time data for identifying information of interest (Burke et al., 2021).

In recent years, geospatial data and advanced image processing technologies have gained popularity across various research fields, offering new opportunities for researchers to overcome limitations imposed by data constraints. Remote sensing data, in particular, provides several advantages over traditional data collection methods. For example, satellite images offer consistent spatial and temporal coverage, enabling real-time, dynamic, cost-effective monitoring of the Earth's surface over extensive areas at a relatively low cost. Furthermore, this allows researchers to compare different regions at different times more effectively. Consequently, using remote sensing data can significantly reduce the cost of measuring targets and creating datasets more efficiently (Paganini et al., 2018; Persello et al., 2022).

Deep learning, a breakthrough technology inspired by the structure of the human brain, has significantly improved image classification and object detection tasks in image data. A subset of machine learning and artificial intelligence algorithms, deep learning methods, particularly

Convolutional Neural Networks (CNNs), are widely recognized as highly successful architectures. They offer numerous advantages for analyzing satellite images, enabling the identification of spatial, spectral, and temporal patterns hidden within image data. Researchers have successfully applied CNNs to a variety of tasks, including the detection of palm oil trees and vehicles in satellite images, demonstrating high levels of accuracy and reliability. These successful applications have motivated us to develop a deep-learning-based model for our research, aiming to leverage the power of advanced technologies to detect and quantify the number of *'gers'* in a given area using remote sensing data.

Many studies have used deep learning techniques for predicting and quantifying the number of oil palm trees in satellite imagery, with some achieving exceptional performance. These studies demonstrated that their model could detect palm oil trees with an accuracy level of over 90 % when compared to ground truth data (Aliero et al., 2014; Mubin et al., 2019; Srestasathiern & Rakwatin, 2014). Vehicle detection and counting, a more challenging task due to the small size and varied types and orientations of vehicles, has also achieved significant accomplishments with high levels of accuracy (Liu et al., 2010; Sang et al., 2018; Tayara et al., 2017).

The process of building object detection models using Convolutional Neural Networks (CNNs) involves identifying key patterns of the target object through iterative training. Initially, the detection model requires an annotated reference for training, consisting of sample images labeled with ground truth objects. During training, the CNNs optimize the model based on the gradient descent algorithm, minimizing the loss function by comparing the parameterized model with the annotated training samples. This process involves updating the model weights and biases based on gradient, evaluating the model performance and facilitating its improvement. Eventually, the model achieves a global minimum of the loss function in the high-dimensional feature space (Kattenborn et al., 2021).

To develop a CNN model for the detection of *‘gers’*, high-resolution satellite images were obtained from the Google Earth Pro platform and a manually annotated reference was created. Due to the iterative learning nature of the model, based on various features of *‘gers’* such as size, shape, color, and other characteristics, it was necessary to provide an adequate number of labeled images with informative features. Consequently, we manually created image chips of the training samples containing more than 50,000 labeled objects. The optimized object detection model demonstrated a globally minimized loss during the training process.

In order to ensure temporal consistency with the available air pollution concentration data, we collected all available satellite images for 14 monitoring stations in Ulaanbaatar, covering a period from the second half of 2015 to the first half of 2019. Our study areas, as depicted in Figure 2.1, include both urban centers—specifically non-ger districts—as well as slum areas identified as 'ger-districts' within Ulaanbaatar. For this study, we obtained 172 observations of unbalanced panel data for the number of *‘gers’*, which represents the number of informal settlements across the 14 monitoring stations. To align this variable with the hourly air pollution concentration and meteorological data, we operated under the assumption that the number of *‘gers’* around the 14 monitoring stations remains constant for a 20-day window, and we matched this data from 10 days prior to 10 days after the reference date of the immigrant data obtained from the specific dates of satellite images.

Table 2.2 presents the average number of detected *’gers’* and available air pollution concentration data for the 14 monitoring stations. The number of informal settlement household variable reflects the number of detected *’gers’* within a half-mile radius of each monitoring station.

Among the 14 stations, the *Bayankhoshuu* monitoring station has the highest density of detected *'gers'*, with an average of 1248 '*gers*' over time, while the *Bukhiin urguu* monitoring station has the lowest density with an average of 35.8 '*gers*' over time. Notably, *Bayankhoshuu* is typical '*ger district*' located far from the urban center, while *Bukhiin urguu* is situated in the urban center of Ulaanbaatar.

To facilitate a straightforward comparison of statistics across different areas, we employed a method whereby the 14 monitoring stations were evenly divided based on the density of immigrants. Specifically, seven monitoring stations were located in high-density areas with an average of more than 600 *'gers'*, while the other seven monitoring stations were situated in low-density areas with an average of less than 600 *'gers'*. However, it should be noted that the density of *'gers'* in a particular area does not necessarily indicate whether it is a slum or non-slum area defined by UN-Habitat. For instance, the *Urgakh naran* areas, despite a relatively low average number of '*gers*' within a half-mile radius from the monitoring station (less than 40 '*gers*'), is distanced from the urban center and lacks public infrastructure. In contrast, the *100-ail* monitoring station is located in the urban core where high-rise buildings and public facilities, which could potentially emit pollutants, are found. This area also presents a dense concentration of over 600 '*gers*'.

Table 2.3 presents statistics for six air pollutant sources, categorized by year and *'ger'* density. The values in parentheses represent the average month of available datasets for each year and variable. However, please note that only half a year's data is available for 2015 and 2019, as denoted by the numbers above nine (from September to December) and below two (January to April), respectively. Analysis of the average air pollution concentration data for 2016, 2017, and 2018 reveals that areas with a higher density of *'gers'* tend to experience higher levels of air pollution, with the exception of Nitrogen Dioxide (NO_2). Prior studies have suggested that NO_2 emissions are closely linked to traffic emissions. During the winter season, residential heating sources may combine with these traffic emissions, contributing to NO_2 levels. As Figure 2.5 illustrates, the difference in NO_2 levels between areas of low *'ger'* density and high *'ger'* density has remained relatively constant over the years, even during the winter season. However, the discrepancy in levels of other air pollutants between areas of low *'ger'* density and high *'ger'* density appears to be influenced by seasonality. This suggests that NO_2 is predominantly generated by traffic emissions, a source that remains consistent throughout the year (Briggs et al., 2000; Enkh-Undraa et al., 2019; Huang et al., 2013; Meng et al., 2008).

Table 2.4 presents the pairwise correlations among *'ger'* density, air pollution concentrations, and weather factors. The results suggest that *'ger'* density is positively associated with $PM_{2.5}$, PM_{10}, SO_2, and O_3, but negatively associated with NO_2 and CO. In addition, the six variables measuring air pollution concentration levels display strong positive correlations with one another, except for O_3 which shows a negative correlation with the other pollutants. As previously discussed, certain air pollutants can form as a result of chemical reactions involving other pollutants. For example, when the level of Nitric Oxide (NO) decreases during weekends, it reacts less with Ozone (O_3) to form Nitrogen Dioxide (NO_2). Consequently, more O_3 remains unreacted, leading to an increase in its concentration. This could explain why the weekday patterns of O_3 show a trend opposite to the other five pollutants. Table 2.4 also reveals that meteorological factors generally have negative associations with air pollution concentration levels, while atmospheric pressure displays a positive correlation.

Empirical Strategies and Results

Fixed-Effect Identification Strategies

Our study constitutes the first empirical exploration into the impact of informal settlements, specifically the number of *'gers'* around monitoring stations, on air pollution concentration levels in Ulaanbaatar, a city known for having one of the highest levels of air pollution worldwide.

$$Y_{i,y,m,t} = \beta_1 G_{i,y,m,t} + \beta_2 W_{y,m,t} + \delta_i + \theta_y + \pi_m + \varepsilon_{i,t} \quad (1)$$

$Y_{i,y,m,t}$, as the dependent variable, denotes the six air pollution indicators, namely $PM_{2.5}$, PM_{10}, CO, SO_2, NO_2, and O_3, measured in micrograms per cubic meter at monitoring station i, during year y, month m, and at time t. The independent variable $G_{i,y,m,t}$ reflects the number of *'gers'* located within a half-mile radius from the i monitoring station, matched correspondingly for the $Y_{i,y,m,t}$. The variable $W_{y,m,t}$ indicates the weather conditions in Ulaanbaatar, including temperature, wind speed, wind direction, precipitation, pressure. We also account for location, year, and month fixed effects through δ, θ and π, respectively. Figure 2.4 shows a monthly trend of the six dependent variables. $PM_{2.5}$, PM_{10}, CO, SO_2, and NO_2 reveal a decrease during the summer season and an increase during the winter season, likely due to increased coal burning for heating purposes. In contrast, O_3 concentrations increase during the summer season due to the crucial role played by heat and sunlight in the chemical reaction that generates this pollutant.

$$Y_{i,y,m,t} = \beta_1 G_{i,y,m,t} + \beta_2 W_{y,m,t} + \delta_i + \theta_y + \pi_m + \mu_w + \lambda_t + \delta_i \cdot \lambda_t + \varepsilon_{i,t} \quad (2)$$

In equation (2), we incorporate weekday variables and interaction terms of location and time. Figures 2.6 and 2.7 depict the levels of air pollution concentration by weekday and the number of *'gers'* in Ulaanbaatar. The figures indicate that, excluding ozone, the levels of the other five air pollutant concentrations tend to decline sharply on weekends due to reduced industrial and traffic activities. Figures 2.8 and 2.9 illustrate the variations in air pollution concentration by time of day and *'ger'* density. These figures reveal that, except for ozone, the pattern of air pollution concentration level movement is similar, while the rate of fluctuation differs according to *'ger'* density. For instance, during peak hours, the concentration of SO_2 in areas with high *'ger'* density exceeds that in areas with low *'ger'* density, while at the bottom line, they are lower. Consequently, to account for time-of-day fixed effects heterogenous by locations, interaction terms between time and location (e.g., location 1 * 03:00:00) are also included.

$$Y_{i,y,m,t} = \beta_1 G_{i,y,m,t} + \beta_2 W_{y,m,t} + \beta_3 W^2_{y,m,t} + \delta_i + \theta_y + \pi_m + \lambda_t + \mu_w + \delta \cdot \lambda_{it} + \varepsilon_{i,t} \quad (3)$$

Equation (3) extends equation (2) by introducing quadratic terms for weather variables. The concentration of air pollution is jointly influenced by both population growth and meteorological conditions. However, accounting for the effects of weather on air pollution concentration can be complex as these effects are unlikely to follow a linear pattern. To tackle this challenge, we conducted a regression analysis of hourly air pollution concentration data on hourly weather data, incorporating a broad array of weather variables and their interaction terms. Specifically, the equation includes variables for temperature, precipitation, wind speed, wind direction, atmospheric pressure, and quadratic terms for these factors—temperature, wind speed, and precipitation—considering their potential non-linear relationship with air pollution concentration.

Results

We present the results of our fixed-effect models for six dependent variables in Table 2.5 to 1.10. Column (1) displays the estimated results for the random-effect specification model with controlling meteorological variables. Column (2) reports the regression results of equation (1), which controls for location, year, and month fixed effects. Column (3) shows the results of equation (1) with the inclusion of weekday and time-of-day dummies. Column (4) presents the estimated results of equation (2), which includes location by time-of-day interaction terms. Finally, column (5) displays the results of equation (3), which includes quadratic terms for weather effects.

A glance at the results indicates that all four fixed-effect specification models, from column (2) to column (5), demonstrate a consistent pattern of coefficients for $PM_{2.5}$, PM_{10}, SO_2, and CO, indicating a statistically significant positive association with the number of *'gers'* at a 1 % level of significance. Conversely, the relationship between the number of *'gers'* and NO_2 shows a statistically insignificant effect and relatively smaller magnitude of coefficients. Additionally, our analysis reveals a negative and statistically significant relationship between the number of *'gers'* and concentration of Ozone level in Ulaanbaatar, also at a 1 % level of statistical significance.

The results of this study suggest that the number of *'gers'* has a significant positive effect on the average levels of $PM_{2.5}$, PM_{10}, SO_2, and CO in Ulaanbaatar. Specifically, the addition of a single *'ger'* within a half-mile radius from one of the 14 monitoring stations increases the average air pollution concentration level of Ulaanbaatar by 0.0517 μg/m³, 0.140 μg/m³, 0.0215 μg/m³, and 0.989 μg/m³ for $PM_{2.5}$, PM_{10}, SO_2, and CO, respectively, at a statistically significant level of 1%. However, the relationship between the number of *'gers'* and average NO_2 level in Ulaanbaatar is statistically insignificant, indicating a negligible effect of informal settlements on this air pollutant. Moreover, the findings indicate that the addition of one *'ger'* within a half-mile radius from one of the 14 monitoring stations leads to a decrease in the average level of ozone in Ulaanbaatar by 0.0491 μg/m³. This result is consistent with previous studies that suggest a negative association between population growth and density and Ozone growth (Borck & Schrauth, 2021; Cramer, 1998).

According to the American Cancer Society, an increase of 10 μg/m³ in PM2.5 concentration corresponds to a 4%, 6%, and 8% increase in overall mortality, cardiopulmonary disease mortality, and lung cancer, respectively (Xing et al., 2016). A straightforward extrapolation from these findings and our own results reveals the significant health implications of additional *'gers'* in Ulaanbaatar. Our research shows that each additional *'ger'* households contributes to an increase of 0.00517 μg/m³ in PM2.5. Consequently, every addition of 2,000 ger households can potentially lead to a 4%, 6%, and 8% increase in the probability of overall mortality, cardiopulmonary disease mortality, and lung cancer, respectively.

In addition, research conducted in China indicates that for every 10 μg/m³ increase in PM10 concentration, there is a 3.4% to 13.6% increase in lung cancer mortality (Zhou et al., 2017). Considering our own findings, where the coefficient of the relationship between PM10 concentration and the number of *'gers'* in Ulaanbaatar is 0.14 μg/m³, it can be inferred that the addition of every 60 *'gers'* potentially leads to an increase in the probability of lung cancer mortality by 3.4% to 13.6%, considering the context of China.

Given these potential health risks related to the number of *'ger'* population in Ulaanbaatar, it is critical to consider policy implications, particularly in the context of Ulaanbaatar's *'ger districts.'* Currently estimated to house 800,000 inhabitants—a 70% increase over the last 20 years—these districts are likely to continue expanding due to ongoing rural-to-urban migration trends.

Furthermore, the, the negative relationship between the number of informal settlers and ozone concentration does not necessarily indicate that increased *'ger'* households reduced air pollution concentration in Ulaanbaatar. Ozone is considered a secondary pollutant that is formed by chemical reactions between nitrogen oxides (NO_x) and volatile organic compounds (VOC) with heat and sunlight. In contrast, the O_3 is chemically decomposed into NO_2 and Oxygen (O_2) by photochemical reactions with Nitrogen oxides (NO) (Wang et al., 2020). For this reason, if the number of *'gers'* is highly correlated with NO concentration, it is possible to decrease O_3 level, resulting in negative causation between O_3 and the number of *'gers'* (Wang et al., 2020). Hence, if the number of *'gers'* is highly correlated with NO concentration, it is possible that the increase in *'ger'* households lead to a decrease in O3 level, resulting in negative association between O3 and the number of *'gers'*.

Conclusion and Discussion

In conclusion, this study sought to investigate the relationship between the number of informal settlements, referred to as *'gers'*, and air pollution, characterized by six major air pollutant concentrations prevalent in developing countries. Our focus was on the rapidly urbanizing city of Ulaanbaatar, Mongolia. Using sophisticated deep learning techniques and high-resolution satellite imagery, we obtained more precise and accurate data regarding the number of *'gers'*. These traditional Mongolian dwellings, commonly inhabited by herders, were counted within a half-mile radius of each of the 14 monitoring stations in Ulaanbaatar.

The findings of this research demonstrate a positive correlation between air pollution concentration levels and the number of *'gers'* for $PM_{2.5}$, PM_{10}, SO_2, and CO, while the effect on NO_2 is negligible. Furthermore, we found a negative correlation between *'ger'* population and O3. This study offers unique insights into the causal impact of informal settlements on air pollution in developing nations, thereby filling a significant gap in existing literature. The results indicate that an increasing number of households in slums or informal settlement areas can contribute to environmental degradation, thus necessitating policy interventions aimed at improving living conditions and health outcomes for residents, while mitigating environmental damage.

Moreover, these findings should serve as a warning of the potential human cost of unplanned expansion of informal settlements in Ulaanbaatar, particularly in light of the rapid rural-to-urban migration. It is imperative to devise and implement policies that address this emerging health crisis regarding the large influx of incoming migrants. There is a critical need for urban planning strategies that mitigate increased air pollution concentration levels attributable to '*ger'* proliferation. These may include planning for cleaner energy solutions, improving infrastructure to reduce slum expansion, and fostering greater public awareness about the health risks associated with increased air pollution.

This study contributes significantly to our understanding of the impact of slum and informal settlement expansion on environmental degradation and potential health implications in developing nations, with a particular focus on air pollution in Ulaanbaatar, Mongolia. The findings provide a nuanced understanding of the phenomenon, emphasizing the importance of addressing the living conditions of slum dwellers, representing the health risks regarding the increased number of informal settlements. This study's relevance extends to policymakers, urban planners, and researchers, providing insights and guidance for future research and policy interventions aimed at creating more sustainable and livable cities in developing nations.

In this study, we employed fixed-effect models by adding location and time dummy variables to control unobserved heterogeneity across the study areas and time periods. This model allows us to capture the effects of location-specific factors over constant time effects and those of time-specific factors over constant location effects. Despite our data being more robust than previous studies that relied on less frequent survey data at the provincial or national level, fixed-effect models are still subject to potential omitted variable bias that may arise from time-varying confounding factors that are correlated with both the dependent and independent variables.

Nevertheless, Mongolia's unique historical and political context reduces the likelihood of endogenous migration decisions by immigrants, thus minimizing potential omitted variable bias in our research. By law, immigrants from rural regions possessing portable tents have the right to acquire land parcels at equal prices, regardless of their proximity to urban centers in Ulaanbaatar. If economic opportunities were the primary driver for internal movement within Ulaanbaatar, our study area, households would prefer to settle closer to the city center. Moreover, the growth of informal settlements in Ulaanbaatar has been documented to occur progressively from the city center to its peripheries. Given these factors and this assumption, potential omitted variable bias can be minimized due to Mongolia's immigrant policy.

However, we did not collect satellite imagery during winter due to the potential for measurement error and the issue of reverse causality. Image quality inconsistencies during winter, caused by snow, diminish the accuracy of *'ger'* detection, leading to measurement errors. Moreover, the consumption of coal and other unsuitable fuel sources intensifies during colder months, exacerbating air pollution concentration levels. As air pollution worsens in winter, Ulaanbaatar residents, who mostly live in 'ger districts', might reconsider urban living and decide to return to rural areas, either temporarily or permanently (Kuo, 2018). Although we used temperature and monthly dummies to control for seasonal effects, the potential for residual omitted variables and measurement errors necessitated the exclusion of winter data when estimating the marginal effect of an additional informal settler on average air pollution concentration levels. By omitting winter data, this study seeks to mitigate the risks of reverse causality and potential omitted variable bias, thereby enhancing the reliability and accuracy of the results. In addition, these considerations strengthen the internal validity of our research design and its findings.

<Table 2.1. Description of Sources and Health Effects of Six Air Pollutants>[6]

<table>
<tr><th>Pollutant
(μg/m³)</th><th>Source</th><th>Effects</th></tr>
<tr><td>Particulate Matter≤2.5mg ($PM_{2.5}$)</td><td rowspan="2">• Traditional low-efficiency stoves (incomplete combustion)
• Coal combustion, biomass burning, and motor vehicle emissions and brick industry
• Forest fires
• Power plant ash ponds
• Dry ground conditions
• PM10 Source Contribution: Soil > Road Dust > Combustion
• PM2.5 Source Contribution: Combustion > Soil > Road Dust > Motor Vehicles > Biomass Burning</td><td rowspan="6">• Annually in Mongolia, approximately 3,010 deaths occur due to disease linked to indoor air pollution, and another 1,123 due to outdoor air pollution-related illnesses, including heart diseases, stroke, lung cancer, pneumonia, and COPD.
• In Ulaanbaatar in 2011, outdoor air pollution was responsible for 29% of heat and lung disease deaths and 40% of lung cancer deaths.
• The prevalence of asthma among children in Ulaanbaatar exceeds both global and Asia Pacific averages.
• A study in 2011 revealed that the rate of spontaneous abortions was 3.6 times higher in December, when air pollution levels peak, compared to May.</td></tr>
<tr><td>Particulate Matter≤10mg (PM_{10})</td></tr>
<tr><td>Sulfur Dioxide (SO_2)</td><td>• Ger stove combustion
• Traffic and power plant emissions
• Increase in fuel consumption requirements in cold weather (coal combustion)
• Ger area sites (higher concentrations compared to non-ger areas)</td></tr>
<tr><td>Nitrogen Dioxide (NO_2)</td><td>• Emissions from old vehicles
• Higher concentrations detected at roadside sites compared to urban sites</td></tr>
<tr><td>Carbon Monoxide (CO)</td><td>• Raw coal combustion for heating
• Power plant emissions
• Improper sealing of stoves and chimneys causes CO poisoning</td></tr>
<tr><td>Ozone (O_3)</td><td>• Secondary pollutant, formed by photochemical reaction between oxides of nitrogen (NOx) and volatile organic compounds (VOC)
• Associated with levels of sunshine and water vapor
• Possibly related to high traffic emission density on weekdays</td></tr>
</table>

[6] Dickinson-Craig, E., Bartington, S. E., Watts, R., Mandakhbayar, O., Khurelbaatar, E.-O., Ochir, C., Boldbaatar, D., Warburton, D., Thomas, G. N., & Pope, F. D. (2022). Carbon monoxide levels in households using coal-briquette fuelled stoves exceed WHO air quality guidelines in Ulaanbaatar, Mongolia. *International journal of environmental health research*, 1-12. , Soyol-Erdene, T.-O., Ganbat, G., & Baldorj, B. (2021). Urban air quality studies in Mongolia: pollution characteristics and future research needs. *Aerosol and Air Quality Research*, *21*(12), 210163. , WHO. (2018). World Health Organization issues recommendations to tackle health impacts of air pollution in Mongolia.

<Table 2.2. Summary Statistics of 14 Monitoring Stations>

Site Name	Avg. Num of Gers	Latitude	Longitude	Available AQ Indicators	Data Span
Bayankhoshuu	1248.0	N47.95756	E106.822752	$PM_{2.5}$, PM_{10}, SO_2	2016/05/25 - 2019/03/13
ub-11	1213.1	N47.9514305	E106.9040722	PM_{10}, SO_2, NO_2	2016/12/23 - 2019/12/30
Tolgoit	1077.9	N47.922495	E106.794805	$PM_{2.5}$, PM_{10}, SO_2, NO_2, CO, O_3	2015/08/31 - 2019/03/13
MNB	1018.8	N47.929732	E106.888629	$PM_{2.5}$, PM_{10}, SO_2, NO_2, CO, O_3	2015/08/31 - 2019/03/13
ub-6	809.0	N47.91345	E106.972030	SO_2, NO_2	2017/05/23 - 2019/12/30
Nisekh	767.2	N47.863943	E106.779094	$PM_{2.5}$, PM_{10}, SO_2, NO_2, CO, O_3	2015/08/31 - 2019/03/13
100 ail	653.1	N47.9329056	E106.9213833	PM_{10}, SO_2, NO_2, CO, O_3	2015/08/31 – 2019/03/13
Amgalan	576.5	N47.913429	E106.997907	$PM_{2.5}$, PM_{10}, SO_2, NO_2, CO, O_3	2015/09/03 - 2019/03/13
ub-9	531.3	N47.9819138	E106.9405111	SO_2, NO_2	2016/12/24 – 2019/11/29
Baruun 4 zam	212.7	N47.9153833	E106.8941944	$PM_{2.5}$, PM_{10}, SO_2, NO_2, CO	2015/08/31 - 2019/03/13
Mongol gazar	89.3	N47.9035389	E106.8507083	PM_{10}, SO_2, NO_2, CO	2015/08/31 - 2019/03/13
Misheel expo	47.3	N47.8943389	E106.8824722	PM_{10}, SO_2, NO_2, CO, O_3	2015/08/31 - 2019/03/13
Urgakh naran	39.1	N47.8664611	E107.1180194	PM_{10}, SO_2, NO_2, CO, O_3	2015/09/01 - 2019/03/13
Bukhiin urguu	35.8	N47.9176056	E106.9373611	$PM_{2.5}$, PM10, SO_2, NO_2, CO, O_3	2015/08/31 - 2019/03/13

<Table 2.3. Summary Statistics of Six Air Pollutants by Year and Population Densities>

	Full sample	high density areas	low density areas	Difference (Low)-(High)
$PM_{2.5}$				
2015 (N= 2304)	27.94	28.57	24.20	-4.37^{***}
	(9.06)	(9.05)	(9.10)	
2016 (N= 35042)	87.48	96.83	68.21	-28.62^{***}
	(7.76)	(7.37)	(8.57)	
2017 (N= 39720)	89.12	111.21	60.15	-51.07^{***}
	(5.58)	(5.73)	(5.38)	
2018 (N= 25093)	54.88	58.77	48.07	-10.69^{***}
	(5.39)	(5.36)	(5.44)	
2019 (N= 11008)	149.35	172.13	118.23	-53.90^{***}
	(1.80)	(1.79)	(1.80)	
PM_{10}				
2015 (N= 2309)	108.37	110.73	94.20	-16.54^{**}
	(9.06)	(9.05)	(9.10)	
2016 (N= 49707)	140.99	169.46	109.20	-60.26^{***}
	(8.10)	(7.67)	(8.57)	
2017 (N= 60249)	140.32	172.67	112.84	-59.83^{***}
	(5.44)	(5.66)	(5.26)	
2018 (N= 34243)	113.78	116.28	110.51	-5.77^{***}
	(5.39)	(5.43)	(5.34)	
2019 (N= 17374)	175.93	199.08	156.78	-42.30^{***}
	(1.79)	(1.78)	(1.80)	
SO_2				
2015 (N= 2317)	11.39	10.08	19.05	8.97^{***}
	(9.06)	(9.05)	(9.10)	
2016 (N= 47995)	38.31	41.43	34.41	-7.02^{***}
	(7.95)	(7.70)	(8.25)	
2017 (N= 55016)	32.90	39.19	26.99	-12.20^{***}
	(5.46)	(5.48)	(5.44)	
2018 (N= 28628)	16.36	15.26	17.53	2.27^{***}
	(5.54)	(5.94)	(5.11)	
2019 (N= 17537)	35.95	46.73	26.84	-19.89^{***}
	(1.84)	(1.89)	(1.80)	
NO_2				
2015 (N= 2285)	18.23	17.02	25.99	8.97^{***}
	(9.06)	(9.05)	(9.10)	
2016 (N= 43611)	40.52	30.73	52.48	21.75^{***}
	(7.91)	(7.43)	(8.48)	
2017 (N= 49917)	47.84	33.68	59.85	26.17^{***}
	(5.44)	(5.66)	(5.25)	

2018 (N= 30491)	36.65	26.06	47.56	21.50***
	(5.41)	(5.44)	(5.37)	
2019 (N= 11954)	39.13	27.85	45.31	17.45***
	(1.95)	(2.19)	(1.82)	
CO				
2015 (N= 2982)	529.97	459.48	669.05	209.57***
	(9.05)	(9.05)	(9.06)	
2016 (N= 40857)	1173.01	1353.68	958.68	-395.00***
	(7.22)	(7.38)	(7.02)	
2017 (N= 42541)	1315.39	1556.41	1036.80	-519.61***
	(5.86)	(5.67)	(6.09)	
2018 (N= 27636)	993.64	1104.70	882.70	-222.00***
	(5.04)	(4.87)	(5.20)	
2019 (N= 8490)	1142.06	1854.70	818.34	-1036.36***
	(1.73)	(1.59)	(1.79)	
O_3				
2015 (N= 3004)	35.17	35.49	34.57	-0.92
	(9.05)	(9.05)	(9.06)	
2016 (N= 39641)	33.24	36.35	29.41	-6.94***
	(7.35)	(7.59)	(7.05)	
2017 (N= 37619)	35.31	39.53	29.42	-10.11***
	(5.41)	(5.43)	(5.38)	
2018 (N= 22247)	27.98	33.96	19.37	-14.59***
	(5.39)	(5.44)	(5.31)	
2019 (N= 7737)	13.46	20.31	9.02	-11.29***
	(1.75)	(1.76)	(1.75)	

<Table 2.4. Pairwise Correlation Matrix>

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) Ger Density	1.000											
(2) $PM_{2.5}$	0.161***	1.000										
	(0.000)											
(3) PM_{10}	0.165***	0.661***	1.000									
	(0.000)	(0.000)										
(4) SO_2	0.052***	0.519***	0.337***	1.000								
	(0.000)	(0.000)	(0.000)									
(5) NO_2	-0.450***	0.213***	0.226***	0.286***	1.000							
	(0.000)	(0.000)	(0.000)	(0.000)								
(6) CO	-0.023***	0.701***	0.404***	0.435***	0.476***	1.000						
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)							
(7) O_3	0.389***	-0.282***	-0.055***	-0.146***	-0.380***	-0.394***	1.000					
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)						
(8) Temperatures	0.007	-0.324***	-0.162***	-0.309***	-0.138***	-0.315***	0.126***	1.000				
	(0.186)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)					
(9) Precipitation	0.001	-0.059***	-0.052***	-0.044***	-0.032***	-0.039***	0.013*	0.032***	1.000			
(mm)	(0.909)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.073)	(0.000)				
(10) Pressure	0.004	0.205***	0.032***	0.153***	0.084***	0.227***	-0.145***	-0.779***	-0.085***	1.000		
	(0.470)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)			
(11) Wind direction	0.024***	-0.025***	-0.078***	-0.004	-0.081***	-0.045***	-0.048***	-0.081***	-0.055***	0.066***	1.000	
(degree)	(0.000)	(0.000)	(0.000)	(0.467)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		
(12) Windspeed	0.011**	-0.129***	-0.107***	-0.091***	-0.142***	-0.155***	0.088***	0.076***	0.065***	-0.146***	0.413***	1.000
(kmph)	(0.034)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

<Table 2.5. Linear Regression Analysis: PM2.5 *(N = 21890)* >

	(1)	(2)	(3)	(4)	(5)
# Gers	0.0139^{***}	0.0515^{***}	0.0510^{***}	0.0510^{***}	0.0517^{***}
	(0.000547)	(0.00791)	(0.00760)	(0.00744)	(0.00742)
Temperature	-1.686^{***}	-0.196^{***}	-0.277^{***}	-0.291^{***}	-0.626^{***}
	(0.0392)	(0.0500)	(0.0589)	(0.0577)	(0.0754)
Precipitation	-7.759^{***}	-5.133^{***}	-4.881^{***}	-5.144^{***}	-11.03^{***}
	(0.996)	(0.915)	(0.888)	(0.871)	(1.572)
Pressure	-0.577^{***}	-0.438^{***}	-0.490^{***}	-0.492^{***}	-15.02^{**}
	(0.0406)	(0.0400)	(0.0403)	(0.0395)	(4.664)
Wind Direction	-0.00259	-0.00518^{*}	-0.00923^{***}	-0.00915^{***}	0.00240
	(0.00253)	(0.00233)	(0.00228)	(0.00224)	(0.00998)
Wind Speed	-0.605^{***}	-0.404^{***}	-0.470^{***}	-0.476^{***}	-0.748^{***}
	(0.0342)	(0.0326)	(0.0355)	(0.0348)	(0.0958)
Temperature^2					0.0193^{***}
					(0.00277)
Precipitation^2					2.375^{***}
					(0.495)
Pressure^2					0.00711^{**}
					(0.00230)
Wind Direction^2					-0.0000292
					(0.0000242)
Windspeed^2					0.00728^{**}
					(0.00262)
Constant	628.8^{***}	452.2^{***}	516.4^{***}	518.1^{***}	7939.9^{***}
	(41.45)	(41.29)	(41.70)	(40.86)	(2367.8)
Location	NO	YES	YES	YES	YES
Year, Month	NO	YES	YES	YES	YES
Weekday/Daytime	NO	NO	YES	YES	YES
Location*Time	NO	NO	NO	YES	YES
Weather Term^2	NO	NO	NO	NO	YES
R^2		0.273	0.329	0.357	0.362

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

< Table 2.6. Linear Regression Analysis: PM10 *(N = 30794)* >

	(1)	(2)	(3)	(4)	(5)
# Gers	0.137***	0.125***	0.146***	0.143***	0.140***
	(0.0100)	(0.0195)	(0.0187)	(0.0183)	(0.0183)
Temperature	-3.569***	-1.183***	1.612***	1.534***	1.105***
	(0.0864)	(0.117)	(0.136)	(0.133)	(0.171)
Precipitation	-23.85***	-16.22***	-11.89***	-12.04***	-22.00***
	(2.277)	(2.230)	(2.162)	(2.116)	(3.785)
Pressure	-2.899***	-2.573***	-1.617***	-1.666***	-60.86***
	(0.0888)	(0.0927)	(0.0931)	(0.0911)	(10.49)
Wind Direction	-0.0548***	-0.0481***	-0.0188***	-0.0177***	-0.0169
	(0.00561)	(0.00550)	(0.00538)	(0.00526)	(0.0235)
Wind Speed	-1.141***	-1.193***	0.0899	0.0490	-0.843***
	(0.0752)	(0.0760)	(0.0824)	(0.0807)	(0.224)
Temperature^2					0.0273***
					(0.00632)
Precipitation^2					4.167***
					(1.211)
Pressure^2					0.0290***
					(0.00516)
Wind Direction^2					-0.000000936
					(0.0000568)
Windspeed^2					0.0243***
					(0.00613)
Constant	3023.3***	2682.4***	1688.4***	1741.0***	31911.6***
	(91.18)	(95.50)	(96.18)	(94.09)	(5329.1)
Location	NO	YES	YES	YES	YES
Year, Month	NO	YES	YES	YES	YES
Weekday/Daytime	NO	NO	YES	YES	YES
Location*Time	NO	NO	NO	YES	YES
Weather Term^2	NO	NO	NO	NO	YES
R^2		0.132	0.202	0.238	0.242

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

< Table 2.7. Linear Regression Analysis: SO2 *(N = 28735)* >

	(1)	(2)	(3)	(4)	(5)
# Gers	0.0331^{***}	0.0206^{***}	0.0208^{***}	0.0214^{***}	0.0215^{***}
	(0.00185)	(0.00282)	(0.00277)	(0.00275)	(0.00274)
Temperature	-0.825^{***}	-0.153^{***}	-0.0959^{***}	-0.0940^{***}	-0.180^{***}
	(0.0144)	(0.0184)	(0.0219)	(0.0218)	(0.0286)
Precipitation	-2.628^{***}	-1.392^{***}	-1.730^{***}	-1.674^{***}	-3.876^{***}
	(0.371)	(0.349)	(0.347)	(0.344)	(0.614)
Pressure	-0.437^{***}	-0.353^{***}	-0.326^{***}	-0.323^{***}	5.582^{**}
	(0.0146)	(0.0146)	(0.0151)	(0.0150)	(1.730)
Wind Direction	0.00104	0.000337	-0.000252	-0.000318	0.0343^{***}
	(0.000927)	(0.000872)	(0.000874)	(0.000868)	(0.00383)
Wind Speed	-0.189^{***}	-0.137^{***}	-0.0816^{***}	-0.0830^{***}	-0.304^{***}
	(0.0123)	(0.0120)	(0.0133)	(0.0133)	(0.0364)
Temperature^2					0.00481^{***}
					(0.00106)
Precipitation^2					0.931^{***}
					(0.196)
Pressure^2					-0.00291^{***}
					(0.000852)
Wind Direction^2					-0.0000843^{***}
					(0.00000927)
Windspeed^2					0.00717^{***}
					(0.000988)
Constant	443.7^{***}	361.9^{***}	331.4^{***}	328.6^{***}	-2672.3^{**}
	(15.09)	(15.01)	(15.53)	(15.42)	(878.3)
Location	NO	YES	YES	YES	YES
Year, Month	NO	YES	YES	YES	YES
Weekday/Daytime	NO	NO	YES	YES	YES
Location*Time	NO	NO	NO	YES	YES
Weather Term^2	NO	NO	NO	NO	YES
R^2		0.259	0.284	0.297	0.301

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

< Table 2.8. Linear Regression Analysis: NO2 *(N = 26312)* >

	(1)	(2)	(3)	(4)	(5)
# Gers	0.000409	-0.00171	0.000183	0.000371	0.000705
	(0.00305)	(0.00362)	(0.00338)	(0.00328)	(0.00327)
Temperature	-0.558***	0.136***	0.470***	0.463***	0.405***
	(0.0182)	(0.0233)	(0.0265)	(0.0258)	(0.0338)
Precipitation	-3.578***	-2.989***	-2.126***	-2.008***	-2.523***
	(0.463)	(0.437)	(0.412)	(0.400)	(0.723)
Pressure	-0.303***	-0.337***	-0.233***	-0.237***	0.525
	(0.0184)	(0.0185)	(0.0180)	(0.0175)	(2.044)
Wind Direction	-0.00915***	-0.00838***	-0.00651***	-0.00593***	0.0355***
	(0.00117)	(0.00111)	(0.00105)	(0.00102)	(0.00455)
Wind Speed	-0.354***	-0.298***	-0.165***	-0.175***	-0.418***
	(0.0157)	(0.0153)	(0.0162)	(0.0157)	(0.0430)
Temperature^2					0.00349**
					(0.00124)
Precipitation^2					0.272
					(0.224)
Pressure^2					-0.000372
					(0.00101)
Wind Direction^2					-0.000101***
					(0.0000110)
Windspeed^2					0.00790***
					(0.00117)
Constant	345.4***	378.7***	275.5***	279.3***	-113.6
	(19.97)	(18.96)	(18.57)	(18.05)	(1037.5)
Location	NO	YES	YES	YES	YES
Year, Month	NO	YES	YES	YES	YES
Weekday/Daytime	NO	NO	YES	YES	YES
Location*Time	NO	NO	NO	YES	YES
Weather Term^2	NO	NO	NO	NO	YES
R^2		0.191	0.297	0.338	0.342

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

< Table 2.9. Linear Regression Analysis: CO *(N = 23389)*>

	(1)	(2)	(3)	(4)	(5)
# Gers	0.708***	0.977***	0.927***	0.975***	0.989***
	(0.0866)	(0.130)	(0.124)	(0.123)	(0.123)
Temperature	-23.65***	-0.445	-7.430***	-7.400***	1.070
	(0.641)	(0.818)	(0.950)	(0.940)	(1.237)
Precipitation	-73.73***	-50.44**	-45.60**	-43.94**	-105.4***
	(17.16)	(15.89)	(15.30)	(15.15)	(26.71)
Pressure	-5.328***	-4.503***	-7.645***	-7.608***	-122.2
	(0.654)	(0.649)	(0.648)	(0.641)	(74.22)
Wind Direction	-0.133**	-0.123**	-0.231***	-0.225***	0.116
	(0.0417)	(0.0387)	(0.0376)	(0.0372)	(0.166)
Wind Speed	-9.489***	-5.908***	-10.51***	-10.42***	-12.04***
	(0.560)	(0.535)	(0.581)	(0.576)	(1.578)
$Temperature^2$					-0.486***
					(0.0455)
$Precipitation^2$					22.83**
					(8.743)
$Pressure^2$					0.0573
					(0.0366)
Wind $Direction^2$					-0.000810*
					(0.000402)
$Windspeed^2$					0.0793
					(0.0430)
Constant	6007.5***	4859.8***	8350.5***	8281.1***	65443.5
	(671.6)	(664.6)	(665.2)	(657.9)	(37675.9)
Location	NO	YES	YES	YES	YES
Year, Month	NO	YES	YES	YES	YES
Weekday/Daytime	NO	NO	YES	YES	YES
Location*Time	NO	NO	NO	YES	YES
Weather $Term^2$	NO	NO	NO	NO	YES
R^2		0.257	0.325	0.342	0.347

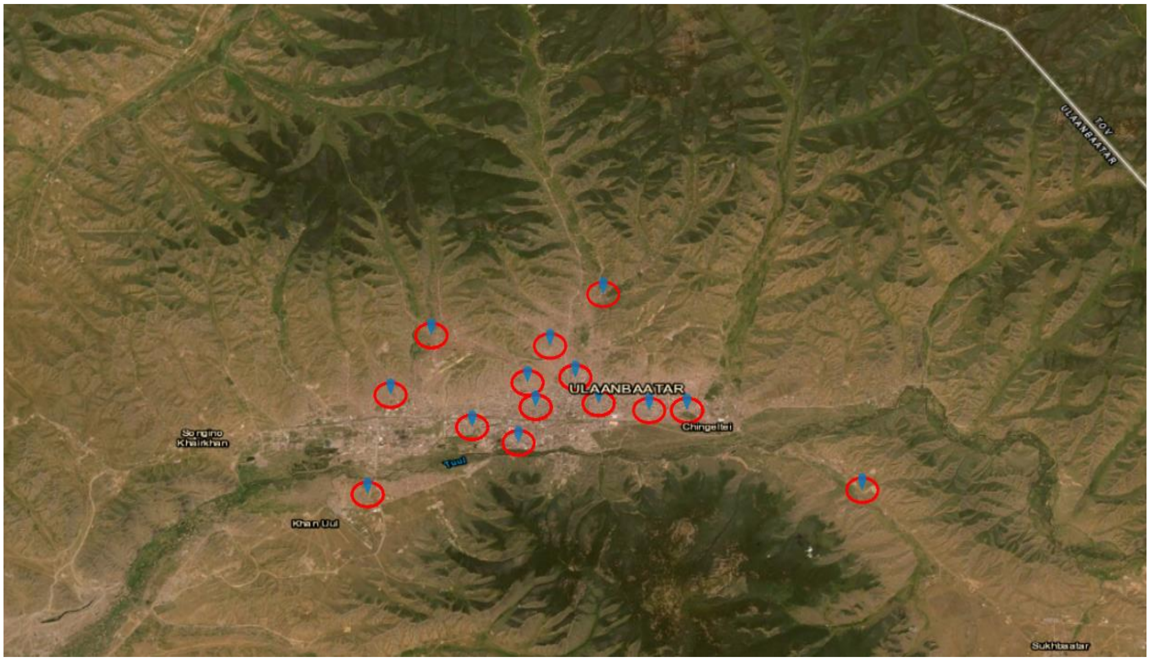
t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

< Table 2.10. Linear Regression Analysis: O3 *(N = 20105)*>

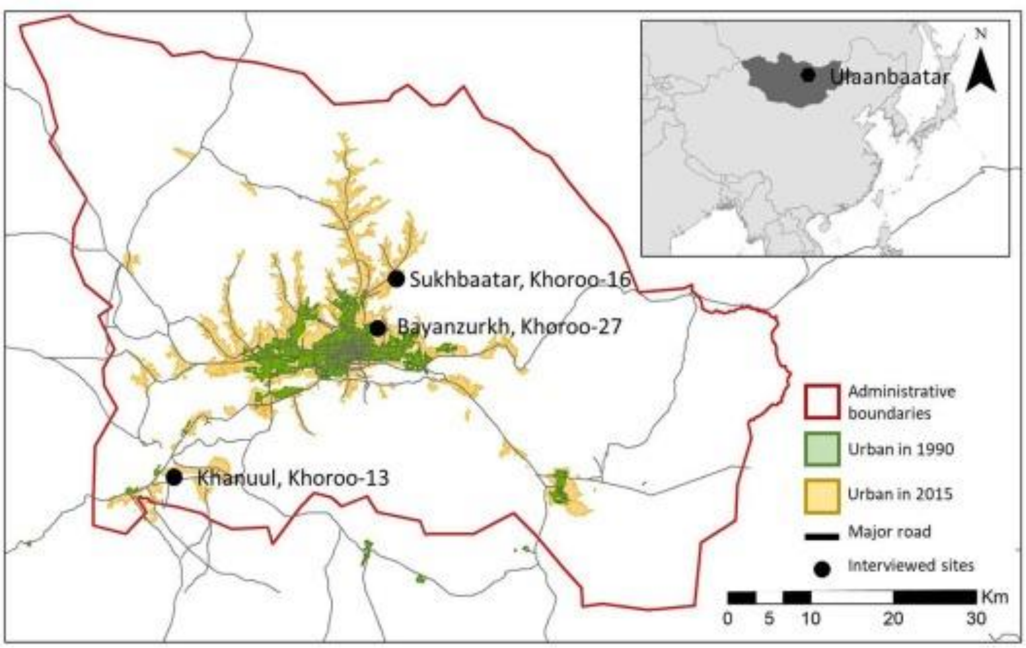
	(1)	(2)	(3)	(4)	(5)
# Gers	-0.0118**	-0.0518***	-0.0462***	-0.0469***	-0.0491***
	(0.00438)	(0.00506)	(0.00452)	(0.00435)	(0.00434)
Temperature	0.0550*	-0.689***	-0.116**	-0.110**	-0.436***
	(0.0278)	(0.0336)	(0.0367)	(0.0354)	(0.0468)
Precipitation	-1.400*	0.562	0.700	0.489	6.090***
	(0.669)	(0.600)	(0.542)	(0.522)	(0.961)
Pressure	-0.305***	-0.117***	0.132***	0.121***	2.766
	(0.0280)	(0.0269)	(0.0252)	(0.0243)	(2.838)
Wind Direction	-0.0263***	-0.0179***	-0.0111***	-0.0111***	-0.0168**
	(0.00179)	(0.00161)	(0.00146)	(0.00141)	(0.00626)
Wind Speed	0.336***	0.00450	0.353***	0.349***	0.257***
	(0.0237)	(0.0219)	(0.0223)	(0.0215)	(0.0588)
Temperature^2					0.0199***
					(0.00176)
Precipitation^2					-1.930***
					(0.287)
Pressure^2					-0.00133
					(0.00140)
Wind Direction^2					0.0000139
					(0.0000151)
Windspeed^2					0.00167
					(0.00158)
Constant	349.0***	212.4***	-60.97*	-48.85	-1358.6
	(30.22)	(27.59)	(25.91)	(24.98)	(1440.5)
Location	NO	YES	YES	YES	YES
Year, Month	NO	YES	YES	YES	YES
Weekday/Daytime	NO	NO	YES	YES	YES
Location*Time	NO	NO	NO	YES	YES
Weather Term^2	NO	NO	NO	NO	YES
R^2		0.242	0.397	0.441	0.447

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

<Figure 2.1. Study Area: Ulaanbaatar and 14 Monitoring Stations >



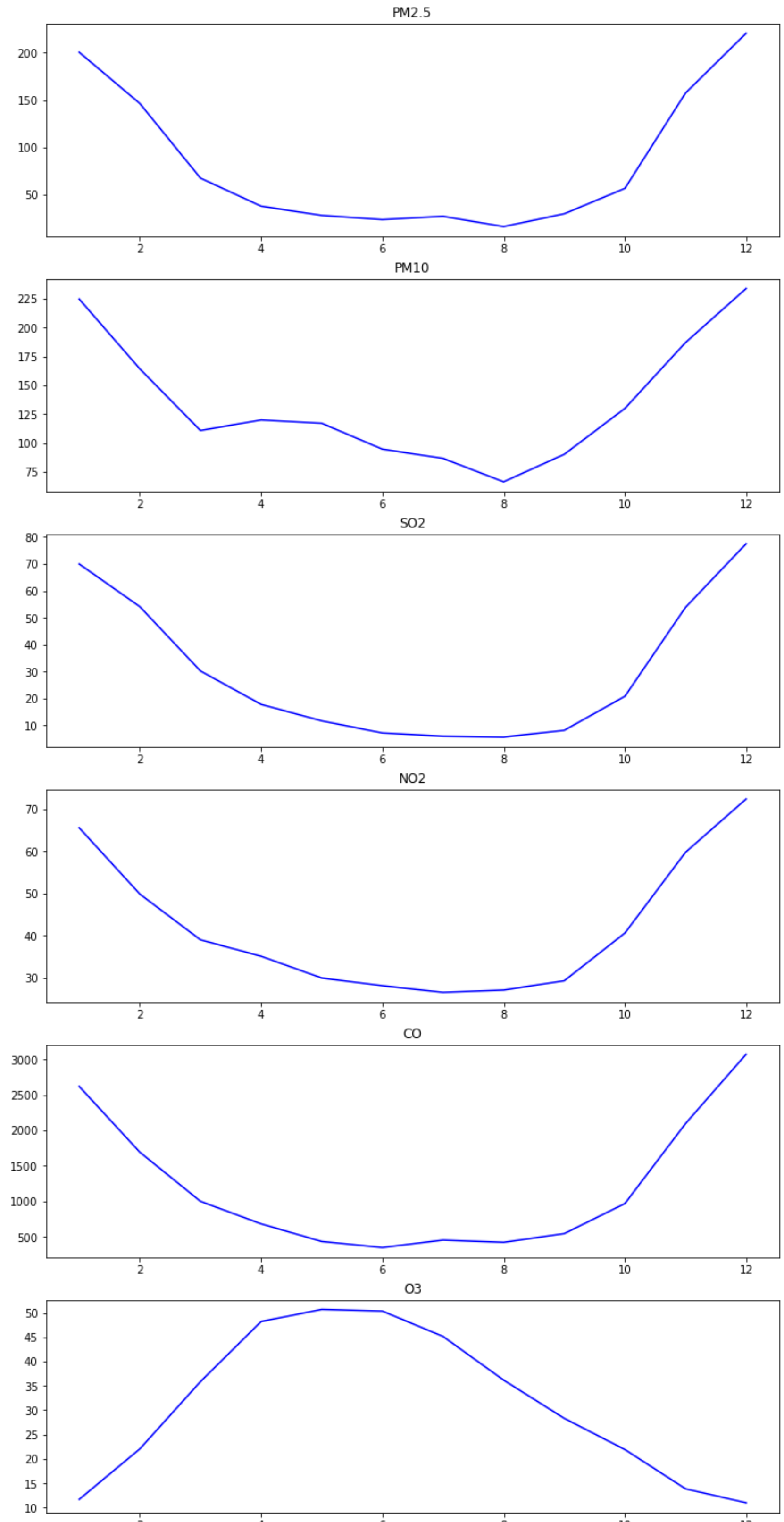
< Figure 2.2. The Growth of Ulaanbaatar from 1990 to 2015[7]>



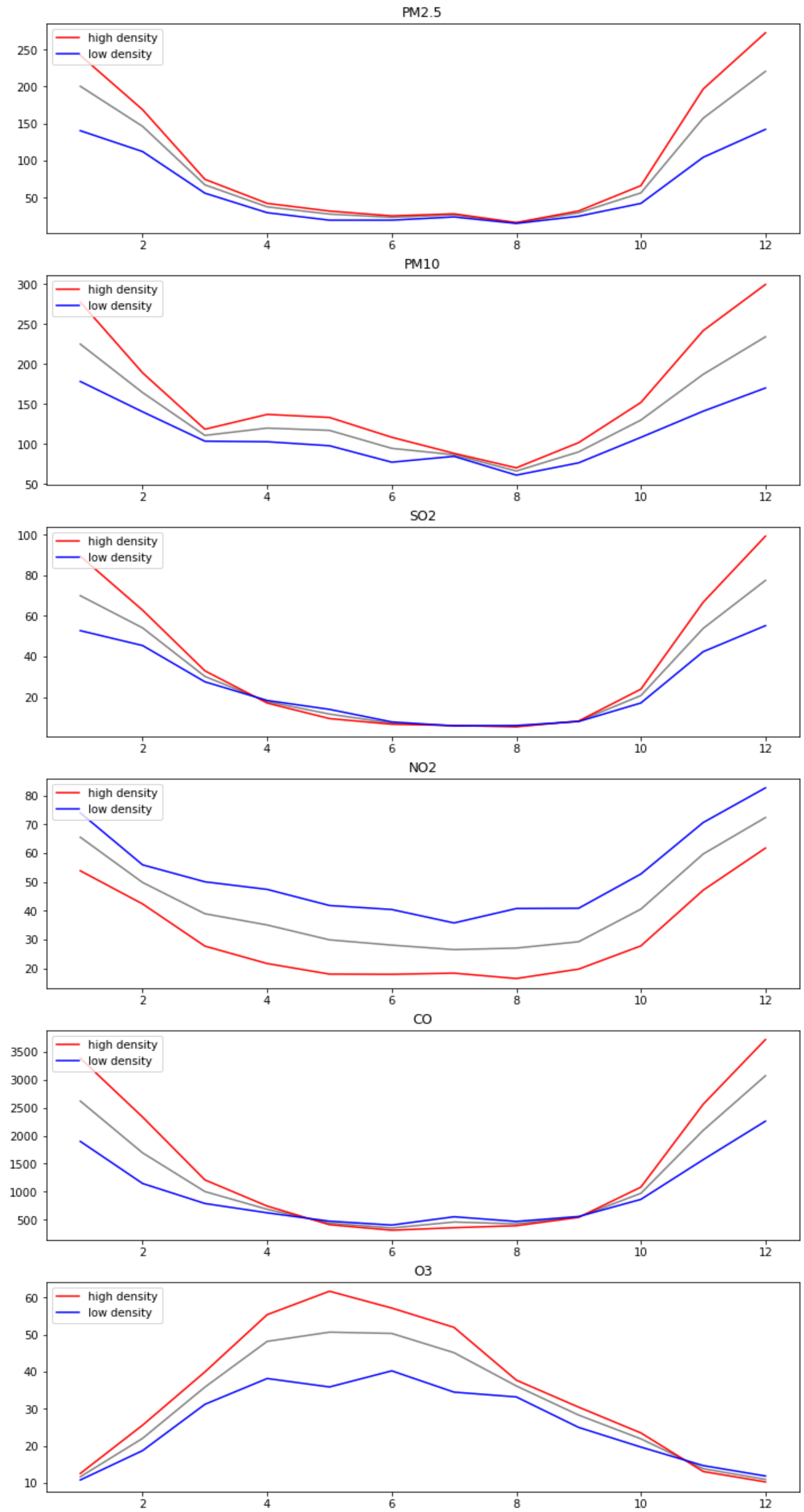
[7] This figure indicates the urban growth in Ulaanbaatar from 1990 to 2015. Adapted from Park, H., Fan, P., John, R., Ouyang, Z., & Chen, J. (2019). Spatiotemporal changes of informal settlements: Ger districts in Ulaanbaatar, Mongolia. *Landscape and urban planning*, *191*, 103630.

<Figure 2.3. Deep Learning Results Sample: Detected ‘Gers’ in UB >

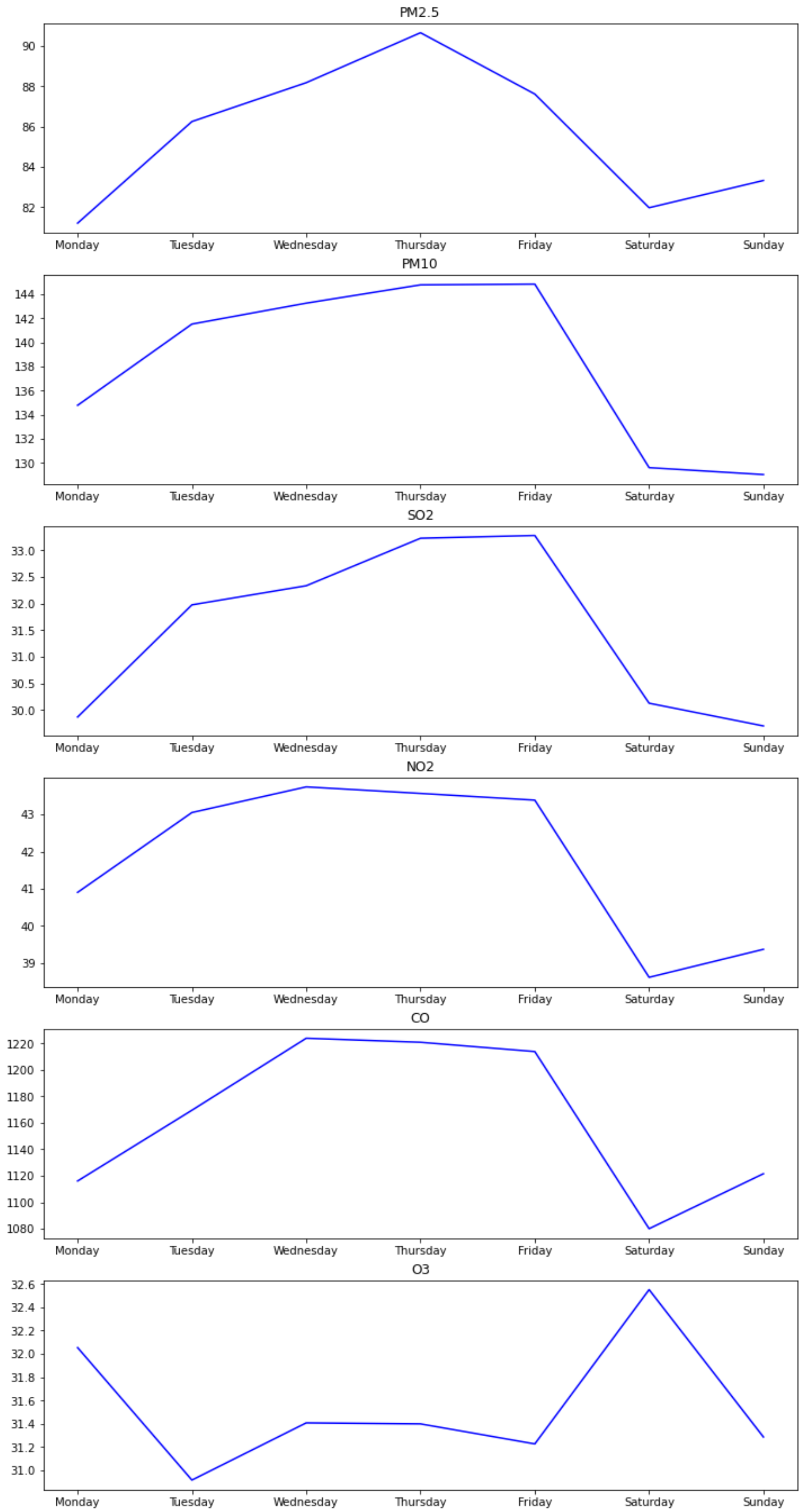
<Fig 2.4. Air Pollution Trends by Month>



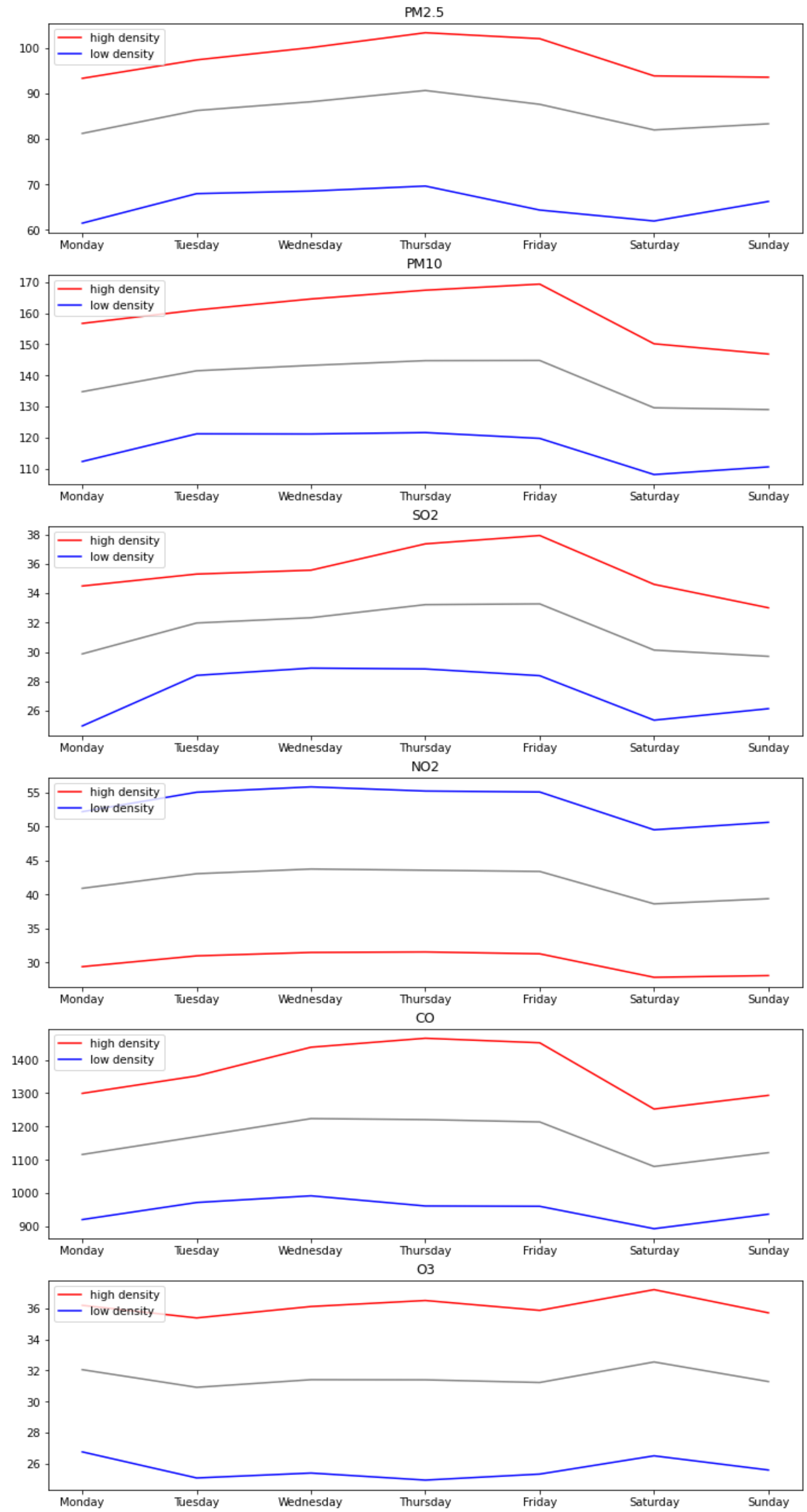
< Fig 2.5. Air Pollution Trends by Month and Density of Informal Settlements >



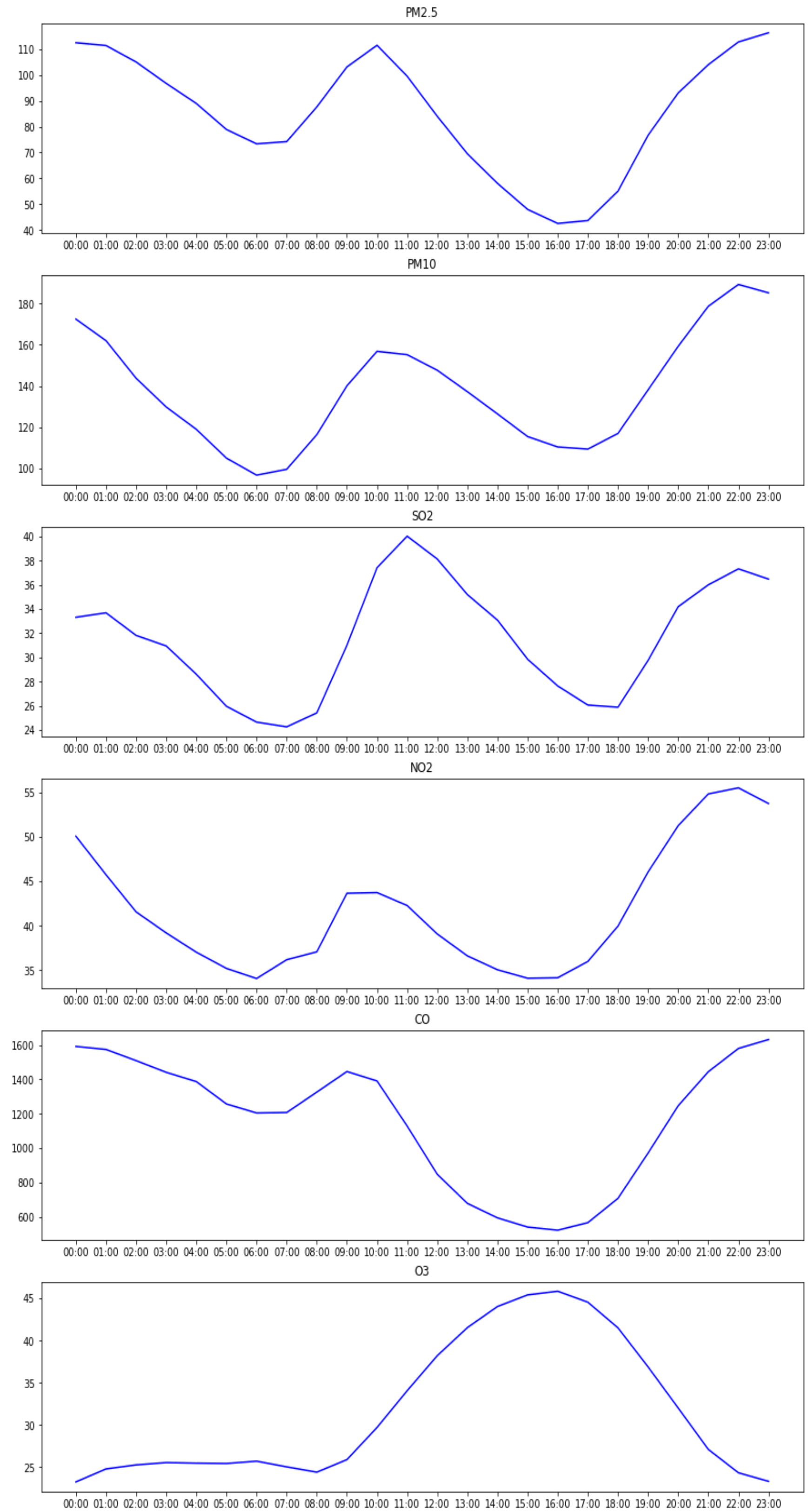
< Fig 2.6. Air Pollution Trends by Weekday >



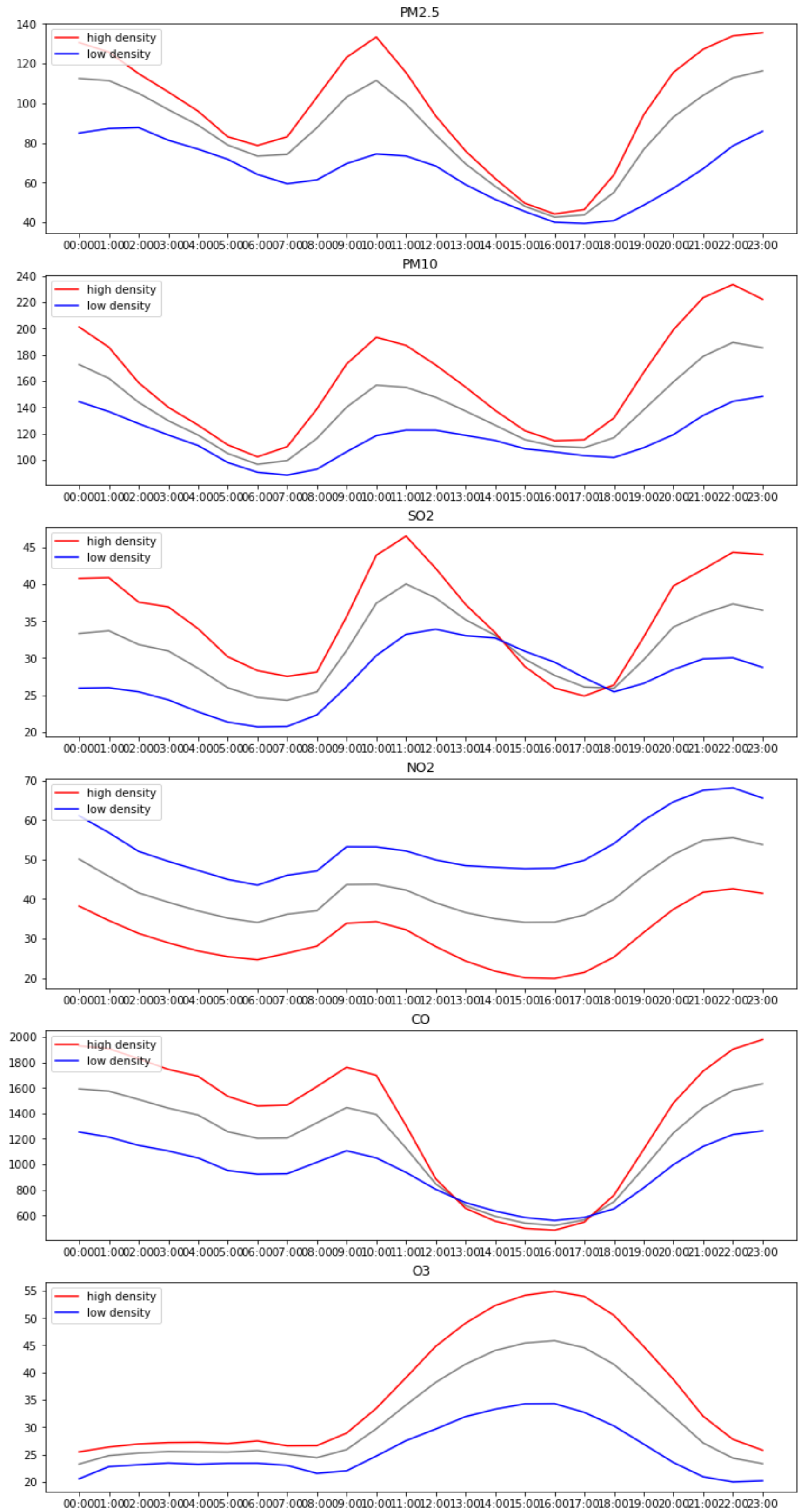
< Fig 2.7. Air Pollution Trends by Weekday and Density of Informal Settlements >



< Fig 2.8. Air Pollution Trends by Daytime >



< Fig 2.9. Air Pollution Trends by Daytime and Density of Informal Settlements >



CHAPTER 3

THE IMPACT OF IMMIGRATION ON EDUCATION ATTAINMENT OF URBAN YOUTH IN DEVELOPING COUNTRIES: THE CASE OF MONGOLIA[8]

[8] Yoon, T., and Filipski, M. To be submitted to *Development Economics*

Abstract

This chapter delves into the educational impacts of substantial internal migration from rural to urban areas. It raises concerns about potential negative effects on educational outcomes and human capital accumulation in the receiving communities in urban areas. While previous research suggests negative impacts on the education of native students, our work introduces a discourse arguing that this migration could encourage a higher pursuit of education among native individuals. This could be due to changes in the labor market following the immigration of low-skilled rural workers. This study used repeated cross-sectional data from the Mongolia Labor Force Survey (MLFS) spanning 2010 to 2018 to analyze the causal relationship between the immigration rate in urban areas and the likelihood of native youth completing secondary education. As an additional check for robustness, we incorporated historical weather data in instrumental variable identification. We hypothesized that lower precipitation levels significantly contribute to driving the rural population towards urban areas for survival and economic opportunities, by applying a Two-Stage Residual Inclusion (2SRI) specification. Our empirical findings show that whether we consider the total migration rate or the exclusive rural-to-urban migration rate as the independent variable, both yield significant and positive results for educational attainment. These results show a motivational trend among the local younger generation in urban areas, aged between 21 and 30, to pursue further education, particularly general secondary education as opposed to technical or vocational certification.

Introduction

In recent decades, there has been a growing focus on internal migration, which refers to the movement of individuals or households within a country's borders, typically from rural to urban areas, and is more pronounced among developing countries. Additionally, it is important to note that internal migration represents a much larger volume of migration compared to international migration, estimated to be three to four times greater (Lucas, 2015; Selod & Shilpi, 2021). In terms of pull and push factors, developing countries are experiencing rapid urbanization due to both voluntary migration in search of economic opportunities and involuntary migration, forced displacement from rural areas resulting from intensifying climate change and violent conflicts. As a result, these demographic changes in urban areas of developing countries have raised concerns among both local residents and policymakers regarding socio-economic impacts.

The growing influx of migrants from less developed to more developed regions, whether on an international or internal scale, raises concerns about the educational performance of native students. This issue has gained significant attention from both host communities and policy makers, particularly in light of the potential negative effects of immigration. Notably, education is a vital component in the process of long-term human capital accumulation and economic growth in developing nations. In response, there has been a significant accumulation of research over the past few decades probing into the correlation between the prevalence of immigrant concentration and the educational achievement of local students.

Previous studies primarily aimed to examine the impact of immigrants flowing from less developed areas, typically marked by inadequate educational infrastructure, into more urbanized regions. In particular, there are concerns that a significant portion of immigrant students in a classroom could alter the class composition, which is considered a public asset, and subsequently affect the learning and behavior of local students (Lazear, 2001; Wang et al., 2018). For instance, when students from rural areas move to urban areas and they are less educated compared to native students, even if they share the same language, teachers may adjust their teaching approach to accommodate these immigrant students. This, in turn, could limit teachers' ability to give individualized attention to each student in a large class, leading to a reduction in the overall educational inputs received by students.

There is a vast number of studies examining the impact of increased rates of immigrant concentration on native students' access to education and school performance. These studies support Lazear (2001) hypothesis that higher immigrant concentration can negatively impact native students (Betts, 1998; Betts & Lofstrom, 2000; Borjas, 2004; Brunello & Rocco, 2013; Farre et al., 2018; Gould et al., 2009; Hoxby, 1998; Jensen & Rasmussen, 2011). Additionally, some studies, such as Betts and Fairlie (2003), Farre et al. (2018), and Murray (2016), provide empirical evidence to explain that some native parents may respond to the presence of immigrant students by choosing to enroll their children in private schools in search of better educational quality.

Although there is limited research on the impact of immigrant peers on native students in the context of developing countries, the available literature found results align with findings from studies conducted in developed country settings (Contreras & Gallardo, 2022; Padilla-Romo & Peluffo, 2023; Tumen, 2018). Empirical evidence from developing countries indicates that the presence of immigrant students, predominantly from rural regions or other countries undergoing conflicts and violence, can have detrimental has the potential to lower educational quality and decrease returns on investments in schooling. This may discourage native students from pursuing higher education degrees, representing crowd-out effect (Hunt, 2017).

However, it is too early to conclude that immigrant households have a negative impact on educational attainment and human capital accumulation in host communities. The decision of native students to invest more in education is influenced not only by the educational resources but also by the dynamics of the labor market. Changes in wage and employment probability structures for native workers, affected by the influx of lower-skilled immigrant labor from less developed regions, may serve as a push factor for native youth to pursue higher education instead of taking low-skilled jobs.

For instance, when there is a large influx of low-skilled immigrant workers and assuming that laborers cannot move across states smoothly, the average wages among low-skilled natives with similar skills to immigrant workers decrease. Meanwhile, some wage-sensitive low-skilled workers may choose unemployment rather than accepting lower wages in an imperfectly inelastic labor market (Sarzin, 2021; Dustmann, Glitz, & Frattini, 2008, Schonberg, & Stuhler, 2016). Furthermore, the negative impact of immigration on wages and unemployment is not homogeneous across different skill groups[9] (Borjas, 2003; Ottaviano & Peri, 2012). The increased proportion of low-skilled labor in the host labor market is likely to reduce the relative low-skilled wage to high-skilled wage, decrease the marginal productivity of low-skilled labor, and increase the marginal productivity of high-skilled labor, assuming that low-skilled and high-skilled labor are complementary inputs (Borjas, 2003; Borjas, 2017; Dustmann et al., 2008; Dustmann et al., 2016; Edo, 2019; Ottaviano & Peri, 2012; Peri, 2016; Sarzin, 2021; Verme & Schuettler, 2021).

The relative proportion of low-skilled workers to high-skilled workers increases in host communities, leading to a decrease in wages for low-skilled native workers and a rise in wages for the high-skilled labor group. The reduced return for the low-skilled labor group, which competes

---

[9] In general, skills in the labor market can refer to an individual's educational attainment, occupation, and other abilities, such as their proficiency in performing manual or analytical tasks.

mainly with less-educated native youth, and the increased return for the high-skilled labor group incentivize native students to enroll in higher education (Betts, 1998; Hunt, 2017; Tumen, 2021). As a result, increased competition in the labor market, with reduced wages and employment probabilities, could encourage native youth to enroll in higher education services.

There is extensive empirical evidence on the effects of larger inflows of immigration, particularly low-skilled labor from developing countries, on the labor markets of developed countries. The authors argue that a mass influx of immigration has an adverse impact on specific skills groups, mainly comparable to low-skilled migrant (Altonji & Card, 1991; Angrist & Kugler, 2003; Boustan et al., 2010; Card, 1990, 2005; Carrington & De Lima, 1996; Friedberg, 2001; Friedberg & Hunt, 2018; Hanson, 2009; Hunt, 1992; Kugler & Kugler, 2009; LaLonde & Topel, 1991; Lemos & Portes, 2014; Longhi et al., 2005; McIntosh, 2008; Pischke & Velling, 1997; Schoeni, 1997; Suen, 2000). In a similar vein, empirical results from the contexts of internal migration and migration between developing countries show a significantly negative relationship between immigration and the labor market of host communities (El Badaoui et al., 2017; Kleemans & Magruder, 2018; Strobl & Valfort, 2015).

Therefore, two interrelated factors can significantly influence the educational investment decisions of indigenous populations: the education market and the labor market. The interplay between these markets shapes the choices of the native population regarding further investment in education. The effect of migration to urban areas is significant but under-researched, lacking consensus, particularly with the surge in internal migration trends in developing countries. Recognizing the importance of human capital accumulation for economic growth, this study evaluates the effects of immigration rates in 22 urban areas of Mongolia on the likelihood of secondary education completion among indigenous students.

To achieve this, we utilized fixed-effect logistic regression analysis using data from the Mongolia Labor Force Survey (MLFS), which provides repeated cross-sectional information. As a robustness check, and to account for potential remaining endogeneity concerns regarding the correlation between educational decisions and urban immigration rate, we performed an instrumental variable analysis. We adopted an empirical approach used in previous studies (Boustan et al., 2010; Kleemans & Magruder, 2018; Munshi, 2003; Strobl & Valfort, 2015), using weather shocks, specifically regional level lagged rainfall variables, as an instrumental variable.

Our study found a positive correlation between immigration rate and completion of general secondary education among urban youth. Specifically, a 1% rise in immigration rates corresponded to a 1% to 2.5% increased likelihood of general secondary education completion, considering different specifications with individual, household, and province-level control variables. However, a slightly negative correlation, albeit statistically insignificant, was observed when outcome variable includes technical/vocational certification or diplomas as a secondary education.

The first stage of the Two-Stage Residual Inclusion (2SRI) procedure indicated a negative associate between previous years' precipitation levels and immigration rate of reference years, consistent with our hypothesis. The second stage confirmed the positive correlation between immigration rates and general secondary education completion but was statistically significant when we consider rural-to-urban immigration as an independent variable.

The structure of this paper is as follows: Section 1 offers the introduction. Section 2 furnishes a background of the study areas, and Section 3 proposes a conceptual framework exploring the causal relationship between educational outcomes and incoming migration. Section 4 reviews the previous literature, and Section 5 provides a description of the utilized data and the control variables. The empirical strategies and results are presented in Section 6. Furthermore, Section 7 details the instrumental variable identification and the results from the Two-Stage Residual Inclusion (2SRI) estimate. Finally, Section 8 concludes the paper, discussing potential avenues for future research.

Background and Study Area

Mongolia has a long history of reliance on herding in rural areas. However, in recent times, approximately 70% of Mongolia's population now resides in urban areas, which is one of the highest urbanization rates among developing countries[10] (Dore & Nagpal, 2006; Xu et al., 2021). The country has undergone a significant increase in urbanization over the last few decades. The political and economic transition, along with weather-related losses in pastoral areas, have been major drivers of migration from rural to urban areas, acting as both push and pull factors (Cui et al., 2019; Mayer, 2016).

At the beginning of the 1990s, as shown in Figure 3.1, Mongolia experienced a decrease in urbanization due to economic difficulties in urban areas, as the government was unable to provide support during the post-Soviet Union transition period. However, the emergence of economic opportunities in urban areas in the mid-1990s resulted in a rapid increase in migration from rural to urban areas (Fernandez-Gimenez & Batbuyan, 2004; Gioli et al., 2014; Park et al., 2017; Sneath, 2006; Xu et al., 2021). Furthermore, climate change has increased the frequency and intensity of extreme weather phenomena, which has directly driven more migration from rural to urban areas. For instance, the "*dzud*," a Mongolian term for a harsh winter following a dry summer, which occurred in 2009-2010, led to the loss of 20% of Mongolia's entire livestock population. The combination of heavy snowfall, cold waves, and summer drought during this

10 Average urbanization rate in Asia is 50% ADB. (2018). Mongolia: Urban Sector Fact Sheet. Asian Development Bank. https://www.adb.org/publications/mongolia-urban-sector-fact-sheet

weather event caused a shortage of hay for winter and exacerbated the economic hardship of herders, leading to increased migration to urban areas (Chatty & Sternberg, 2015; Du et al., 2018; Mayer, 2016; Tugjamba et al., 2021; Xu et al., 2021).

Regarding the skill groups in urban labor market, immigrants from rural areas in Mongolia are often categorized as low-skilled laborers. Following Mongolia's transition to a market economy in the early 1990s, there has been a decline in the quality of boarding schools that primarily serve students from herder households. This decline can be attributed to the policy change implemented during this period. An illustrative example is the increase in school dropout rates and decrease in enrollment rates observed in 1995. These changes were primarily a consequence of heightened dormitory fees, which were imposed due to a lack of government financing (Batkhuyag & Dondogdulam, 2019; MECSS, 2015). Batkhuyag and Dondogdulam (2019) highlights that as a result of the aforementioned changes, the 2017 National Quality Assessment Survey in Mongolia revealed significant disparities in academic performance between rural and urban students at the primary level. Specifically, children from herder families scored 15% lower in mathematics compared to their urban counterparts. This disparity underscores the lower educational achievements of rural immigrant students in comparison to those from urban areas. Furthermore, since the majority of rural immigrants have backgrounds in agriculture, they are unlikely to fill the position of highly skilled laborers in urban settings. Consequently, migration tends to have a more adverse impact on the employment and wage proposed of low-skilled labelers in urban areas, where less-educated youth may also be seeking employment opportunities.

Conceptual Framework

The substantial influx of immigrant students from less developed regions can have a significant impact on the demand for education in host communities. If the supply of secondary educational resources is not perfectly elastic, this increased demand can potentially give rise to both direct and indirect costs, leading to negative peer effects. One direct cost is the increased demand for education brought about my immigration, which can lead to rising expenses such as tuition fees and dormitory charges. Additionally, there are indirect costs associated with reduced teaching quality and negative externalities resulting from the presence of immigrant students. This can potentially lead to the displacement of native students from secondary educational institutions, a phenomenon referred to as the crowd-out effect in the education market.

Nevertheless, it is important to consider that the increased presence of immigrant workers with low skill level can potentially result in a decrease in the relative wages of low-skilled workers to that of high-skilled workers. These dynamics could significantly encourage native students to aspire for wages associated with higher-skilled or educated labor roles, steering clear from the decreased wages of low-skilled work. Consequently, this can promote their educational attainments, exemplifying a crowd-in effect. As such, when considering the marginal benefits associated with secondary education for native students, it is crucial to consider the wages earned by skilled workers who have attained higher levels of education. On the other hand, the marginal costs of pursuing secondary education encompass both the direct expenses of education and the opportunity costs associated with lower wages typically earned by individuals with low skills.

The decision to invest in secondary education is determined by the trade-off between the marginal benefits of acquiring further education (crowd-in effect) and the marginal cost associated with pursuing additional years of schooling (crowd-out effect). When the marginal benefit of obtaining secondary education outweighs the marginal cost, individuals are more likely to choose to invest in higher levels of education. Conversely, if the marginal cost of pursuing secondary education surpasses the marginal benefits, individuals may opt not to pursue further schooling.

The crowd-in and crowd-out effects arising from the influx of migrant students and low-skilled workers are profoundly reliant on the secondary education resources and labor market dynamics within host communities. Consequently, the magnitude of these effects relies on the demand and supply elasticities present in the labor and secondary education markets, which dictate how wages and the expenses associated with higher education are impacted by increased immigrant concentration (Jackson, 2015).

Jackson (2015) posits that, in equilibrium, the variable β indicates the sensitivity of native enrollment demand to inflows of relatively unskilled immigrant labor, while α denotes the responsiveness of native enrollment demand to the arrival of immigrant students.

$$\beta = (-\epsilon_{wL})(\eta^N) + \left(-\epsilon_{fL}\right)(\phi^N) \in [0, \infty) \qquad \text{[Crowd-in] (1)}$$

$$\alpha = (-\epsilon_{wE})(\eta^N) + \left(-\epsilon_{fE}\right)(\phi^N) \in [-1, 0] \qquad \text{[Crowd-out] (2)}$$

η^N: Secondary education demand elasticities of native students for relative unskilled wage

ϕ^N: Secondary education demands elasticities of native students for cost of secondary education.

ϵ_{wL}: Relative low-skilled group wage elasticities to increased share of low-skilled immigrant labor.

ϵ_{fL}: Cost of secondary education elasticities increased share of low-skilled immigrant labor.

ϵ_{wE}: Relative low-skilled group wages elasticities to increased share of immigrant students.

ϵ_{fE}: Cost of secondary education elasticities to increased share of immigrant students.

The comparison of the magnitudes of $|\alpha|$ and $|\beta|$ reveals the impact of immigration on the secondary education investment decision in host communities. If $|\alpha|$ is greater than and $|\beta|$, this suggests that a large influx of immigration has a negative effect on secondary education investment. On the other hand, if $|\alpha|$ is less than and $|\beta|$, it represents a positive relationship between the secondary education completion rate and immigrant concentration in host communities (Jackson, 2015).

Figure 3.2 delineates two distinct market conditions: one for the education market and the other for the labor market. Sections (A) and (B) illustrate the effects of a demand shock in the educational resource market. Diagram (A) signifies a scenario with a perfectly inelastic supply of educational resources, while diagram (B) reflects a scenario with a perfectly elastic supply of educational resources. In contrast, diagram (C) and (D) detail the dynamics of relative wage reduction for low-skilled labor compared to high-skilled labor in two labor market scenarios: one with perfectly inelastic labor demand (C) and another with perfectly elastic labor demand (D).

Considering the various elasticity conditions of labor market demand and educational resource supply, several scenarios determine the lower and upper bounds of β and α. If we assume that labor cannot readily relocate to other labor markets in different states, the lower bound of β occurs when the supply of educational resources is perfectly inelastic, as illustrated in Figure 3.2, diagram (A), indicating no crowd-in effect from the labor market. Conversely, the upper bound of β is represented when labor demand is perfectly inelastic, as in diagram (C), and the educational resource supply is perfectly elastic, as in diagram (B).

In terms of the response of local students to an increase in the number of immigrant students, α approaches its upper bound when both the supply of educational resources and the local labor demand are perfectly elastic, as demonstrated in diagrams (B) and (D). Lastly, when the supply of educational resources is perfectly inelastic, as depicted in diagram (A), α reaches its lower bound of -1 (Jackson, 2015).

However, it is important to note that β and α are determined by the sensitivities of the labor market and secondary education supply to immigrant labor and students, respectively. Thus, differences in the characteristics of the labor market and educational supply can result in varying outcomes. Therefore, it would not be appropriate to generalize empirical evidence from developed countries to the developing world.

First, in developing countries, the labor market is not as well-integrated as it is in developed countries, and there is a high proportion of informal and unregulated sectors with low-skilled workers. This can lead to a depression in local wages and employment opportunities for native workers, as immigrant workers are able to take jobs in these unregulated sectors and work for lower wages. In addition, the exclusion of the informal sector from minimum wage regulations exacerbates the downward pressure on wages. Furthermore, the constrained capacity of labor markets in developing nations to adapt to labor supply shocks driven by a substantial influx of immigrants can intensify the consequences experienced by indigenous workers.

Second, the cost of internal migration, which is the most common migration pattern in the developing world, is lower compared to international migration, and those who migrate internally do not face language barriers or legal constraints. This makes it easier for immigrant households to integrate into the local labor market and potentially increases competition with native workers, contributing to negative effects on the labor market for natives (Bryant & Rukumnuaykit, 2013; Calderón-Mejía & Ibáñez, 2016; Kleemans & Magruder, 2018; Selod & Shilpi, 2021).

Moreover, the impact of reduced schooling resource distribution, which generally plays a role in the crowding-out effect, on educational decision-making in developing countries may differ from that in developed countries due to differences in their educational market characteristics. For instance, among higher-income countries, the quality of school instruction and student performance play a crucial role in gaining entry to higher education, such as a bachelor's or graduate degree. However, in developing countries, the influx of less-educated immigrant students may actually make it easier for native students to obtain a secondary education diploma, potentially leading to an increase in the number of native students who complete secondary education (Hunt, 2017).

Literature Review

While the research on the impact of immigration on human capital investment and educational outcomes in developing countries is limited, there is a modest body of literature that has explored this topic, primarily focusing on developed countries such as the United States. These studies have examined the effects of immigration on educational investment choices, taking into account both the crowd-out effect and the crowd-in effect.

Betts (1998) examines the relationship between state-level immigration and native students' high school completion using data from the U.S. Census. The study provides evidence of a negative impact of immigration on the likelihood of high school graduation for native-born black and Hispanic students in the U.S., supporting the "*educational crowding-out*" hypothesis. However, Hunt (2017) also analyzes the effects of immigration on native students' high school completion in the U.S. and finds different results, showing that state-level immigration rates increase the probability of high school completion among native-born black students. Hunt suggests that immigration may motivate native black students to attain higher levels of education, reducing competition in the labor market, despite a decrease in per capita school resource allocation. In a similar vein, McHenry (2015) uses National Education Longitudinal Study and U.S. Census data to examine the impact of low-skilled immigrant influx on U.S. native students' educational outcomes. The study finds that low-skilled immigration is positively associated with native students' high school attendance, grades, test scores, and enrollment in academically rigorous curricula. Jackson (2015) also examines the effect of immigration on native students' college enrollment using U.S. Census data. The study finds that an increase in the number of low-skilled workers at the state level raises native students' college enrollment, but the number of immigrant college students is not significantly associated with the probability of college enrollment for U.S. natives. Smith (2012) focuses on the impact of immigration on human capital in the U.S., specifically youth employment rather than schooling outcomes. The study finds evidence of a negative impact of immigration on employment outcomes for native youth. Brunello et al. (2020) use Italian provincial data to examine whether the increased share of low-skilled immigrants is associated with the likelihood of native students' high school completion. The study finds that the increased number of low-skilled immigrants is positively associated with both high school completion and higher education enrollment decisions.

There is also some empirical evidence from developing countries, particularly in the context of Syrian refugees displaced by conflict and seeking refuge in Middle Eastern countries. Tumen (2018) assess the impact of Syrian refugees on the enrollment of native students in high school and uncovers a positive correlation. The author argues that the presence of refugees, who predominantly consist of low-skilled laborers with limited access to the formal job market, creates competition with native youth for low-paying informal jobs. This increased competition has prompted native students to pursue higher levels of education. However, contrasting findings are reported by Assaad et al. (2018), who did not find a significant impact on the educational outcomes of Jordanians as a result of the presence of Syrian refugees.

These studies offer valuable insights into the complex relationship between immigration and educational investment, highlighting the importance of considering contextual factors and region-specific dynamics. However, the existing empirical evidence, although available for both developed and developing country contexts, is still limited, and there is a lack of consensus among studies. As a result, our comprehensive understanding of the intricate interaction between incoming migrant households and the human capital accumulation of native students remains incomplete and insufficient. Especially in developing countries, this interaction should be further and more diversely investigated, given the recent rapid trend of incoming migration to urban areas and long-term economic growth strategies centered on human capital accumulation. Therefore, it is imperative to conduct further research in developing countries to enhance our understanding of this intricate relationship and its implications for educational outcomes and the development of human capital.

Data and Descriptive Statistics

Mongolia Labor Force Survey Data

To analyze the causal relationship between immigration and human capital investment decisions in the 22 urban areas, including 21 province capital and Ulaanbaatar, we utilized data from the Mongolia Labor Force Survey (MLFS) from 2010 to 2018, excluding the years 2012 and 2015 due to missing data and sample size imbalances. The MLFS, conducted annually by the National Statistics Office in collaboration with the International Labor Organization, aims to generate official labor statistics for Mongolia. The survey covers 21 Aimags, 311 soums, and 9 districts of UB, providing national representativeness. The MLFS data is internally comparable, following the methodology definition of the International Labor Organization. The survey collects information on respondents' migration status over the past five years, including their current and previous locations, as well as their educational status. Specifically, respondents were asked to report their highest level of educational attainment. By leveraging the MLFS data, we could examine the relationship between immigration and human capital investment decisions, particularly in the context of secondary education completion rates among native youth in Mongolia's capital cities.

Human capital investment indicator

In this study, we investigated the causal relationship between the rate of incoming migration and the educational attainment of the native young population in urban Mongolia. Specifically, we explored whether individuals between the ages of 21 and 30 have completed secondary education. The education indicator variable was measured in two ways. First, individuals were assigned a value of 1 if they had completed general secondary education or higher-level degrees, and 0 otherwise. Second, a comparable binary classification was employed, in which individuals were assigned a value of 1 if they had completed secondary education or higher-level degrees, including technical or vocational diplomas and certificates, and 0 otherwise. It is important to note that the sample solely comprises individuals residing in the survey location during the survey period. Household members who were temporarily away from home for purposes such as studying or working are not included in the sample. Hence, our dataset consists of individual-level observations of respondents aged 21 to 30, who are residing in either one of the 21 Aimag centers (province capitals) or Ulaanbaatar at the time of the survey. Consequently, we collected a total of 43,352 observations from 22 urban areas spanning across 7 distinct time periods, ranging from 2010 to 2018[11].

There are inherent limitations in the decision to observe the 21 to 30 age group, which stem from the data used. The data does not include the specific year when respondents completed their secondary education, which is particularly relevant to the study's objective of assessing whether native youth pursue further education in response to immigrant inflow. For example, the immigration rate variable was defined based on the number of immigrants tracked over the past five years. This means that respondents who are 30 years old could have completed their education without considering the current immigration rate. Additionally, even if some individuals finished their secondary education within five years, their response to immigration may have been delayed, indicating a lagged time between immigration impact and educational investment decisions.

Despite these potential limitations arising from the data, it is still necessary to define our observation between the ages of 21 to 30 for several reasons. First, unlike in developed countries where students typically begin and complete their general education at the same age, various conditions and household constraints can hinder children from starting and completing their education in the same year. Our data also confirms that among respondents aged 21 to 30, more than 13% of them reported currently attending an educational institution at the time of the survey. Furthermore, the attainment of vocational and technical certificates and diplomas, included in our second education outcome variable, is not dependent on age.

Nevertheless, these data limitations highlight the need for a more strategic selection of the study group in future research. Incorporating data that captures the exact years of completing

[11] Due to missing data and imbalanced sample sizes, we excluded the 2012 and 2015 surveys from the MLFS dataset, as mentioned previously.

secondary education or dropping out in response to demographic changes in urban areas could provide a more comprehensive understanding of the complex relationship involved.

Migration measure

The primary independent variable in this study is the urban migration rate in 22 urban areas. We determined the migrant population through a question from the Mongolia Labor Force Survey (MLFS). Respondents were asked if they had lived elsewhere for more than six continuous months in the past five years before moving to their current urban areas, either in one of the capital cities of 21 provinces or in Ulaanbaatar. Additionally, they were asked to identify the type of location (urban or rural) of their previous residence, allowing for the determination of migration direction, namely from urban-to-urban or rural-to-urban areas. The participants' current location in the MLFS was categorized into four groups: the capital city (Ulaanbaatar), Aimag center (provincial capital city), soum center, and rural areas.

In this study, the host communities of interest, residing in urban areas, were defined as individuals who have continuously lived in Ulaanbaatar or an Aimag center (provincial capital city) for a minimum of five years. The independent variable was the immigration rate in the urban areas. This rate was calculated by dividing the number of immigrant participants who had resided in other provinces for at least six months within the past five years by the total number of respondents in each 22 urban areas in Mongolia.

Control Variables

In this study, we incorporated multilevel control variables, including individual, household, and province-level characteristics, to address for potential confounding factors and control for heterogeneity across different regions in Mongolia. The individual-level covariates included in the analysis are education level, age, gender, marital status, and disability status. These covariates provide information on the characteristics of the respondents that may influence their educational attainment. Gender, marital status, and disability status were included as binary dummy variables, taking a value of 1 if the respondent is male, married, or has a disability, and 0 otherwise. The household-level control variables included in the analysis are household size, the educational level of the household head, and the weekly hours devoted to agricultural work. At the province level, we included control variables such as Regional Gross Domestic Product (RGDP), which is calculated in real terms using the Consumer Price Index (CPI). This variable provides information on the economic conditions of each province and may have an impact on educational outcomes. Additionally, we included province-level general education infrastructure indicators, specifically the number of schools and the number of teachers. These variables reflect the availability and quality of educational resources in each province. The data for these covariates and control variables were obtained from publicly available Mongolia Statistical Information Service, ensuring the reliability and accessibility of the information.

Table 3.1 represents the descriptive statistics for the observations, and the dependent variable, Secondary Education, represents the educational outcome of individuals aged 21 to 30 in terms of the completion of general secondary education. The mean value for this variable is 0.62, indicating that, on average, 62% of urban youth in the sample have completed general secondary education. Another educational outcome variable, Secondary Education (Technical/vocational), captures the completion of secondary education, including both general education and technical or vocational certificates or diplomas. The mean value for this variable is 0.78, indicating that, on average, 78% of urban youth in the sample have completed secondary education in any form.

The independent variable, Migration Rate, measures the overall immigration rate at the province level specifically for urban areas. The average value of 4.34% suggests that, on average, 4.34% of the population in urban areas across provinces and years consists of immigrants who have resided in another province, whether urban or rural, for more than 6 months within the past 5 years. In addition, the variable Rural Migration Rate captures the rate of rural-to-urban immigration, specifically indicating the migration from rural areas to the current urban regions. The mean value of 1.01% indicates that, on average, 1.01% of the total population in urban areas have lived in rural areas for more than 6 months over the past 5 years.

The individual-level covariates in our analysis offer insights into the characteristics of the individuals included in the sample. The variable "Age" has a mean value of 25.54 years, indicating that, on average, the individuals in the sample are around 25 and a half years old. Approximately 50% of the individuals in the sample are male, while around 4% of the individuals have a disability. The variable "Never Married" is a binary variable, serving as a dummy variable indicating whether an individual has never been married. It has a mean value of 0.48, suggesting that approximately 48% of the individuals in the sample have never been married.

The household-level covariates provide information about the characteristics of the households to which the individuals belong. HH Head No Education is a dummy variable indicating whether the household head has never completed primary education, with a mean value of 0.04, suggesting that 4% of household heads have not completed primary education. HH Size represents the size of households, with a mean value of 4.23, indicating that, on average, households in the sample consist of 4.23 individuals. Agricultural Hours represents the weekly hours of agricultural work for a household, with a mean value of 11.85 hours per week.

The province-level covariates provide information about the characteristics of the provinces in which the individuals reside. RGDP represents real term Regional Gross Domestic Product, measured in million Mongolian Tugrik, with an average value of 129.20. # Schools indicates the number of general education schools in each province, with a mean value of 95.96 and a standard deviation of 92.81. # Teachers represents the number of teachers in each province, with an average value of 4,232.14 and a standard deviation of 4,526.55.

The sample consists of 43,352 observations, providing a robust dataset for the analysis of the relationship between migration rates and secondary education outcomes among individuals aged 21 to 30 in urban areas of Mongolia.

Empirical Strategies and Results

Fixed Effect Logistic Regression Model

This research aims to investigate how individuals residing in urban areas (one of the urban areas among Ulaanbaatar or 21 Aimag centers) in a given year t are influenced by the presence of immigrants in their current location (urban area j) when it comes to their likelihood of attaining complete secondary education or higher degrees.

When analyzing a variable that takes on a finite number of discrete values, such as the dependent variable in this study, it is not appropriate to treat as a continuous variable. However, the discreteness of the variable does not necessarily mean that a linear model for the expected value of the variable given a set of predictors, $E(y|\mathbf{x})$, is unsuitable. Alternative methods, such as logistic regression, can be utilized to effectively model the relationship between dependent variables in studies with discrete outcomes. Hence, to analyze this relationship, we employed a logit model.

$$P(EDU_{itj} = 1) = \Phi(\beta_0 + \beta_1 Mig_{jt} + \beta_2 X_{it} + \beta_3 X_{ht} + \beta_4 X_{pt} + \beta_5 Urban_j + \beta_6 Year_t + \varepsilon_{it}) \qquad (1)$$

The probability $P(EDU_{itj} = 1)$ represents whether respondent i, located in urban area j at time t, has completed secondary education or holds a higher degree, which is the primary outcome of interest in this study. Mig_{jt} denotes the percentage of immigration rate at the urban area j at time t. X_{it}, X_{ht}, X_{pt} represents multi-level control variables, individual, households, and province level characteristics potentially correlated with immigration rate and educational attainment of young population, respectively. The study also takes into account the fixed effects of location and time (year), denoted by $Urban_j$ and $Year_t$, respectively, on the likelihood of achieving secondary or higher levels of education.

Results

The study aimed to examine how the concentration of immigrants in urban areas affects the educational investment decisions of urban youth in developing countries, with a theoretical basis that considers the interaction between school and labor market dynamics. The empirical analysis used a logistic regression model with the dependent variable being the completion of secondary education among individuals in urban areas, represented as a binary variable, among individuals residing in urban areas. On the other hand, the independent variable was the immigration rate at the province-level urban areas at the time of the survey, which was a continuous variable ranging from 0 to 1. However, for consistency and ease of interpretation, the immigration rate was converted into a percentage value by multiplying it by 100.

In logistic regression, the coefficients do not possess a direct interpretation like in linear regression models. Instead, the focus is on calculating partial derivatives, also known as marginal effects, to determine the expected change in the probability of the dependent variable (educational attainment in this study) for a one-unit change in the independent variable (1 percentage point increase in urban immigration rate in this study). These marginal effects provide a more meaningful interpretation of the model coefficients and allow for a better understanding of the relationship between the immigration rate and educational attainment among urban youth.

Therefore, our results in Tables 3.2 to 3.5 display the marginal effect of the urban immigration rate on secondary education completion for urban youth aged 21 to 30. The results can be interpreted as the impact of a 1% increase in the urban immigration rate on the probability of achieving secondary education.

Results presented in Tables 3.2 to 3.5 display the outcomes of logistic regression analysis, highlighting the marginal effect of immigrant rates in province-level urban areas on the secondary education completion of native youth. The independent variable of Table 3.2 and 3.4 is total immigration rate, measured as people who has lived more than 6 months over the past 5 years at province level urban areas including capital city, Ulaanbaatar, which is 21 urban areas in this study, while the independent variable of Table 3.3 and 3.5 is only consider immigrants from rural areas. In addition, the dependent variable in Table 3.2 and 3.3 is whether individual-level observation completed secondary education, while specification of Table 3.4 and 3.5 not only consider general secondary education, but also technical or vocational certificates or diplomas.

The tables presented in this study, specifically Tables 3.2 to 3.5, contain various specifications based on the Equation (1) organized into Column (1) to (5). Column (1) shows a result of pooled OLS estimate, only considering individual characteristics, namely age, gender, disability status, and marital status. In contrast, Column (2) includes both location and year dummy variables to control location and time heterogeneity among observations, in addition to individual-level covariates. Furthermore, Column (3) accounts for individual and household characteristics, such as the educational attainment of the household head and size of household, while still hold constant for location and year specific factors by adding dummy variables. Lastly, Column (4) includes individual, household, and regional characteristics, such as the real-term province level gross domestic product (GDP), as well as the province level educational infrastructure such as, the number of schools and teachers, while adjusting for location and year fixed effects.

The results presented in Table 3.2 demonstrate a statistically significant and positive relationship between the rate of immigration in urban areas and the educational achievement of urban youth, specifically in terms of completing secondary education. In column (1), which does not account for location and year fixed effects, a 1% increase in the district-level immigration rate is associated with a 2.5% increase in the probability of an individual completing secondary education, holding individual heterogeneous characteristics. After controlling for individual, household, and regional characteristics, as well as province and year fixed effects in columns (2) to (4), the coefficient magnitude, representing the marginal effect of the immigration rate, remains similar and statistically significant. Specifically, an increase of 1% in the migration rate is associated with a 1% increase in the likelihood of completing secondary education among urban youth in columns (2) and (3), and a 0.9% increase in column (4). These effects are statistically significant at the 1% level in all cases.

The findings presented in Table 3.3, which examined rural-to-urban migration rate as the independent variable, exhibited similar results of the calculated partial derivative from the logistic regression estimate with Table 3.2 for total migration rate. For instance, when only individual characteristics were controlled for without province and year fixed effects, a 1% increase in rural-to-urban migration rate was associated with a 2.9% increase in probability of complete secondary education among urban young population. When we control for individual, level in Column (2), it shows that 1% increase in rural-to-urban population rate raise the 0.9% probability of complete secondary education in urban areas at 5% statistically significant level. Additionally, as we hold constant of individual, household level heterogeneity and location and year fixed effect, 1% increase in rural-to-urban immigrant rate is associated with 1.2% increase in likelihood of secondary education completion at 1% statistically significant level. Finally, when we control for individual, households and regional level characteristics along with location and year fixed effect, the results represent that when 1% increase in rural-to-urban immigrant rate in urban areas, it increases 0.7% of probability of secondary education completeness at 10% statistically significant level.

Table 3.4 and 3.5 exhibit a subtle negative correlation and less significant correlation between immigrant rates and educational outcomes when the dependent variable for educational attainment encompasses not only general education but also technical/vocational certificates or diplomas. Table 3.4 displays the marginal effects of the total immigrant rate, while Table 3.5 illustrates the marginal effects of the rural-to-urban immigrant rate. For instance, in Table 3.4, a 1% increase in the total immigrant rate is associated with a 0.3%, 0.2%, and 0.4% decrease in the likelihood of completing secondary education or obtaining job-related certificates or diplomas in Column (2) to (4), respectively. These outcomes, derived while controlling for various groups of covariates, are statistically significant at the 10%, 10%, and 5% levels. When the independent variable shifts to the rural-to-urban immigrant rate, the results remain consistent with those in

Table 3.4, which considers the entire immigrant population as the independent variable. However, these findings are statistically insignificant.

The logistic regression outcomes in Tables 3.2 to 3.5 reveal an essential observation: increased immigrant rates in urban areas, regardless of the origin of the immigrant population, display a strong association with educational investment, specifically in general secondary education. This implies that a higher influx of immigrants in urban areas is positively correlated with a higher likelihood of urban youth completing general secondary education.

However, it is crucial to consider that a rise in the immigrant population within urban areas could potentially discourage urban youth from pursuing job-related training, such as technical or vocational certificates or diplomas. This is reflected in the results, where the relationship between immigrant rates and the attainment of technical/vocational certificates or diplomas is not statistically significant.

The regression results also highlight the importance of individual characteristics in educational outcomes. Being disabled, married, or male is negatively associated with the likelihood of completing secondary education, compared to the reference groups of non-disabled youth, never-married individuals, and females. In terms of household characteristics, larger household size is negatively associated with the probability of secondary education completion among urban youth. However, the educational level of the household head exhibits a positive and significant relationship with the educational attainment of urban youth. Among the regional-level control variables, the real-term regional GDP and the number of teachers show negligible relationships with the likelihood of secondary education completion among native youth in urban areas. However, the number of schools at the province level is positively associated with the probability of secondary education attainment. This indicates that access to educational infrastructure, as measured by the number of schools in a province, plays a more crucial role in deciding to pursue additional education than the overall quality of education, as simply measured by the number of teachers.

It is important to note that these positive effects of the number of schools are significant only for the first type of educational outcome, which considers general education without technical or vocational certificates or diplomas.

Robustness Check

Instrumental Variable Identification

In this section, we provide a robustness analysis of our primary empirical results, utilizing the Two-Stage Residual Inclusion (2SRI) estimation method. This methodology aims to scrutinize the causal relationship between the completion of secondary education at the individual level (binary dependent variable) and the immigration rate across 22 urban areas in Mongolia (endogenous variables). This approach addresses potential endogeneity concerns using instrumental variables.

We incorporated weather shocks as an instrumental variable, particularly lagged rainfall variables, according to the ideas of prior studies (Boustan et al., 2010; Kleemans & Magruder, 2018; Munshi, 2003; Strobl & Valfort, 2015). These variables demonstrate a strong correlation with the migration rate in countries where rain-fed agriculture plays a critical economic role. The significant negative impact of rainfall on migration from areas reliant on rain-fed agriculture provides an appropriate instrument for capturing exogenous variation in immigration rates.

In a similar vein, in Mongolia, mobility has historically served as a key risk management strategy for herders, who have traditionally migrated from regions experiencing resource scarcity to those with abundance, taking advantage of the spatial and temporal structure of weather patterns. The seasonality of precipitation and the replenishment of water sources have thus been critical factors in maintaining the herding lifestyle, as noted by Chatty and Sternberg (2015) and Sattler et al. (2021). However, climate change in Mongolia has disrupted the seasonality of precipitation patterns, leading to a reduction in water resources (Chatty & Sternberg, 2015). Munkhtsetseg et al. (2007) have reported a significant positive correlation between precipitation and pasture yield, supporting the finding of Xu et al. (2021) that migration from rural to urban areas is negatively associated with the previous year's precipitation levels. Therefore, to introduce exogenous variation in the level of migration across Mongolia's 22 urban areas, we utilize the count of rainy days at the regional level for the three years preceding the reference year as instrumental variables.

Regarding instrumental variable methodologies, a two-stage approach is typically employed in IV methods, with the first stage entailing the estimation of endogenous variable expectations conditional on observed covariates and one or more instrumental variables. The second stage subsequently predicts outcomes as a function of the estimated treatment values obtained in the first stage, measured confounders, and possibly additional control variables. However, it is important to note that the widely employed 2SLS method, originally conceived within a linear framework for continuous endogenous treatment and outcome variables, might not be apt for the context of this study, which involves binary educational attainment outcome[12] variables (Basu et al., 2017; Terza et al., 2008).

Instead, in this study, we attempted to investigate the causal impact of province-level urban immigration rates on the probability of an individual completing secondary education by utilizing the Two-Stage Residual Inclusion (2SRI) method, with the number of rainy days in the previous

[12] In this study, the outcome variable is a binary indicator of whether the observation includes to a native youth aged between 21 and 30 in urban areas who completed secondary education or not.

years at the regional level serving as the instrumental variable. This method has gained acceptance in addressing endogeneity concerns through the application of IVs in nonlinear models. Instead of using predicted endogenous variables, in 2SLS, we calculate residuals on the first-stage linear regression, and in the second stage, we regress binary outcome variables on the endogenous variable, calculated residual from the first stage, and other covariates using logistic models such as logit or probit.

In the first stage of the 2SRI method, we estimate a linear regression model with the incoming migration rate as the dependent variable, using instrumental variables such as regional-level rainfall in the preceding years. The equation for the first stage is as follows:

$$Mig\ Rate_{jt} = \delta_0 + \delta_1 rainfall_{rt-1} + \delta_2 rainfall_{rt-2} + \delta_3 rainfall_{rt-3} + \delta_4 X_{it} + \delta_5 X_{ht} + \delta_6 X_{jt} + \delta_7 Aimag_j + \delta_8 Year_t + \varepsilon_{jt} \qquad (2)$$

In this equation, $Mig\ Rate_{jt}$ refers to the immigration rate of the urban areas j at time t. Additionally, $rainfall_{rt-1}$, $rainfall_{rt-2}$, and $rainfall_{rt-3}$ signify the number of rainy days in the r region at time $t-1$, $t-2$ and $t-3$ from the reference year. Furthermore, X_{it}, X_{ht}, and X_{jt} are incorporated as a control variable to account for individual, household, and province-level covariates, while ε_{jt} is included as an error term.

After estimating the first stage equation, we obtain the residuals (ε_{jt}), which represent the unexplained variation in the immigration rate after accounting for the instrumental variables and other covariates.

Subsequently, for the second-stage equation, logistic regression is employed to represent the individual-level secondary education attainment, binary variables, as a function of the province-level urban areas immigration rate, other explanatory variables, and residuals derived from the first stage equation. When dealing with a binary outcome variable, incorporating these residuals into a second-stage logistic regression model suggests that the IV estimate is contingent upon these residuals. This, in turn, brings the IV estimate closer to the subject-specific log odds ratio (Burgess et al., 2017). The estimation of the second-stage binary educational attainment outcome for individual and the continuous immigrant rate equations is conducted as follow equation:

$$P(EDU_{ijt} = 1) = \Phi(\pi_0 + \pi_1 \widehat{Mig\ Rate}_{jt} + \pi_2 X_{it} + \pi_3 X_{ht} + \pi_4 X_{jt} + \pi_5 Residual_{jt} + \pi_6 Urban_j + \pi_7 Year_t + \varepsilon_{ijt}) \qquad (3)$$

The variable denoted as $P(EDU_{ijt} = 1)$ corresponds to the educational outcome variable. The endogenous independent variable $Mig\ Rate_{jt}$ and control variables, including X_{it}, remain the same as in the first stage. $Residual_i$ indicates the residual value derived from the first stage estimate, and ε_{it} is included as an error term.

2SRI Results

In this section, we present the key findings of our study, which were derived from the application of the two-stage residual inclusion (2SRI) estimation method. Table 3.6 to 3.11 show the results of the 2SRI estimate, including the two first-stage linear regression[13] results and four second-stage logistic regression models[14].

13 In the 2SRI estimation, we conducted separate first-stage linear regressions for the total migration rate and the rural-to-urban migration rate, resulting in Table 3.6 and Table 3.7 displaying the respective first-stage results.

14 We considered two different outcome variables for education attainment. Therefore, the second-stage logistic regression models for the 2SRI method yield four sets of results, corresponding to the two independent variables (total migration rate and rural-to-urban migration rate) and the two outcome variables (education attainment measured only by general education and education attainment including technical/vocational certificates).

Table 3.6 and 3.7 show the results of the first step of linear regression within the two-stage process, exploring the relationship between two types of immigration, total immigration rate and rural-to-urban immigration rate, and the regional-level precipitation-affected days, in the preceding three years relative to time t. The results are broken into three different groups, shown in column (1), column (2), and column (3). Column (1) has the results when we only consider factors related to individuals. Column (2) shows the results when we also think about factors related to the individual's household characteristics. Finally, column (3) gives the results when we include factors about individuals, their households, and their province level control variables. Furthermore, in every specification, location- and year- specific heterogeneity is controlled by dummy variables.

The results in Table 3.6 indicate that when solely considering individual and household-level characteristics in column (1) and column (2), the current rate of urban immigration rate positively correlated with the precipitation levels of the recent past periods, which contradicts hour hypothesis. Nevertheless, after taking into account the socio-economic characteristics at the provincial level, the findings reveal a negative relationship between the number of rainy days in the previous three years ($t-1$, $t-2$, and $t-3$) and the rate of urban immigration. Notably, the coefficients for the number of rainy days in $t-1$ and $t-3$ is statistically significant, in line with our hypothesis, whereas the coefficient for $t-2$ is not. Table 3.7, which present the results of the first stage linear regression in the 2SRI process for rural-to-urban immigration rates, demonstrates findings similar to those of the first-stage results shown in Table 3.6 for entire immigrants. As per hour hypothesis, while the coefficient for the number of rainy days in $t-2$ indicates a positive correlation with rural-to-urban immigration rate, those for $t-1$ and $t-3$ show a negative correlation.

According to the first stage linear regression results from various specification with two different urban immigration rate variables, when we account for all related control variables – individual, household, and province characteristics – recent past precipitation significantly influences the migration of people to urban areas in the preceding year. In the next step, we incorporated the endogenous variable, immigration rate, and the residual obtained from the first stage into a second stage logistic regression. This allowed us to estimate the causal effect of the urban immigration rate on individual educational investment, considering two distinct educational attainment outcomes: general education along, and secondary education combined with technical/vocational certificates or diplomas.

Table 3.8 displays the calculated marginal effect of regressing individual-level secondary education outcomes, focusing solely on general education, on the total province-level urban immigration rate. The results indicate that with a 1% increase in the immigration rate, the probability of individuals completing secondary education rises by 0.016%, with a 5% level of statistically significant, taking into account all other covariates in column (3). In Table 3.9, we modify the independent variable to account for only the province-level urban immigration rate of individuals from rural areas. The findings indicate a positive correlation between rural-to-urban immigration rate and the probability of urban youth obtaining a secondary education. However, the marginal effect is not statistically significant.

The results in Tables 3.10 and 3.11 features a change in the dependent variable to determine whether urban youth aged between 21 and 30 have completed secondary education or achieved a technical/vocational certificate or diploma at time t, coded as 1 if yes and 0 if no. Table 3.10 presents the logistic regression results where the binary dependent variable is regressed on the province-level total urban immigration rate. Despite the fixed effect specification shows a negative coefficient, the instrumental variable estimates propose that the total immigration rate exhibits a positive correlation with the probability to complete secondary education or technical/vocational training among urban individuals. The coefficient value is 0.04, although the result is not statistically significant. Similarly, Table 3.11 demonstrates a positive correlation between educational attainment and rural-to-urban immigration rates when considering the independent variable of immigration rates for individuals relocating from rural domains, even though the result is not statistically significant.

Conclusion and Discussion

In conclusion, this study has provided valuable insights into the relationship between internal migration and the decision to invest in education among urban native youth. The previous literature has not reached a consensus on whether there is a crowd-in or crowd-out effect because two intercorrelated factors can influence local youth's decision to pursue more education: the effects on the education market and the labor market.

On one hand, the growing demand for education can reduce available school resources, leading to a crowding-out effect. On the other hand, a decrease in the relative wages of low-skilled workers compared to high-skilled workers encourages native students to pursue higher degrees, resulting in a crowding-in effect.

Therefore, these two interactions between the education and labor markets significantly influence the decision of native students to invest in education. If the marginal benefit of education outweighs the marginal cost after an increase in immigration, they are more likely to invest more in education. Conversely, if the marginal cost exceeds the marginal benefit, they are more likely to invest less. To assess the impact of urban immigration on the likelihood of completing secondary education among young individuals aged 21 to 30 in 22 urban areas, including 21 province capitals and Ulaanbaatar, Mongolia, we utilized data from the repeated cross-sectional Mongolia Labor Force Survey (MLFS). Our goal was to investigate whether immigration contributes to increased educational investment among local youth. Using logistic regression analyses and various specifications, we examined the marginal effect of urban immigration rates on secondary education outcomes.

The findings indicate that, after accounting for multilevel control variables such as individual, household, and province-level characteristics, as well as adjusting for location and time-specific differences, urban immigration rates significantly influence the secondary education completion of urban native youth. In particular, an increase in total immigration rates positively affects the likelihood of completing secondary education among the urban young population. Additionally, rural-to-urban immigration rates demonstrate a positive correlation with educational attainment outcomes in fixed-effect specifications. However, when examining a broader range of educational achievements, including technical or vocational certificates and diplomas, a negative relationship is observed, while rural-to-urban immigration is not statistically significant.

To ensure the robustness of our primary empirical findings, we employed instrumental variables using the Two-Stage Residual Inclusion (2SRI) specification, and the results largely correspond with our main findings from the fixed effect model. As initially hypothesized, the number of rainy days at that regional level during the $t-1$, $t-2$ and $t-3$ time periods before the reference point are significantly and negatively associated with the immigration rate. This indicates that lower levels of precipitation encourage people to move to urban areas. Then, we incorporated the calculated residual from the first stage into the second stage logistic regression, which included the endogenous variable (urban immigration rate) and other control variables. The marginal effect results from the second stage indicate that the total immigration rate has a statistically significant positive impact on the probability of secondary education completion among urban youth. Although the rural-to-urban immigration rate also showed a positive association, it was not statistically significant. Moreover, in contrast to the primary results from the fixed-effect model, the 2SRI findings revealed that a higher total immigration rate increases the probability of completing general secondary education or obtaining technical/vocational certificates or diplomas. However, this was not consistent when we altered the independent variable to rural-urban migration.

The findings of this research indicate that a higher proportion of immigrants, who are generally less educated or possess lower skills compared to native population, in urban areas could be a significant factor that encourages human capital accumulation in developing countries. The empirical evidence primarily supports this for general education, as it increases the likelihood of native youth pursuing higher degrees in this domain. However, the empirical results regarding the relationship between increased immigration, urban schools, and labor market effects on encouraging urban populations to pursue job-related training programs remain inconclusive and debated.

Nonetheless, these findings have significant implications for policymakers and educators seeking to understand the socio-economic consequences of urban immigration in developing countries, particularly in terms of human capital accumulation. Identifying the factors associated with educational attainment among urban populations can inform the design of targeted interventions and polices aimed at improving education outcomes in the context of increasing urbanization and migration for future economic development policies. Future research could investigate the underlying mechanisms driving these relationships, as well as focusing on vocational training in developing countries as primary outcome variables.

<Table 3.1. Summary of Statistics of Age between 21 and 30>

	Mean	SD	Min	Max
Dependent Variables				
Secondary Education	0.62	0.49	0.00	1.00
Secondary Education (Technical/Vocational)	0.78	0.41	0.00	1.00
Independent Variables				
Migration Rate	4.34	3.73	0.00	19.55
Rural Migration Rate	1.01	0.98	0.00	8.60
Individual Covariates				
Age	25.54	2.86	21.00	30.00
Male (=1)	0.50	0.50	0.00	1.00
Disability (=1)	0.04	0.19	0.00	1.00
Never married (=1)	0.48	0.50	0.00	1.00
Households Covariates				
HH Head No Education (=1)	0.04	0.20	0.00	1.00
HH Size	4.23	1.66	1.00	17.00
Farming Hours (week)	11.85	23.53	0.00	111.00
Province Covariates				
RGDP	129.29	164.97	1.14	457.97
# Schools	95.96	92.81	4.00	245.00
# Teachers	4232.14	4526.55	165.00	12117.00
Observations	43352			

Note: The mean, standard deviation, minimum and maximum values are reported for the dependent and independent variables, as well as individual, household, and province-level covariates.

<Table 3.2. Logistic Regression Analysis: Impact of Total Migration Rate on General Secondary Education Completion *(N=43352)*>

	(1)	(2)	(3)	(4)
Migration Rate	0.028***	0.010***	0.010***	0.009***
	(0.001)	(0.001)	(0.001)	(0.001)
Age	-0.008***	-0.008***	-0.009***	-0.009***
	(0.001)	(0.001)	(0.001)	(0.001)
Male	-0.145***	-0.141***	-0.140***	-0.141***
	(0.004)	(0.004)	(0.004)	(0.004)
Disability	-0.364***	-0.353***	-0.346***	-0.346***
	(0.011)	(0.011)	(0.011)	(0.011)
Never Married	0.059***	0.046***	0.031***	0.033***
	(0.005)	(0.005)	(0.005)	(0.005)
HH Head Education			-0.335***	-0.332***
			(0.013)	(0.013)
HH Size			-0.014***	-0.015***
			(0.001)	(0.001)
Farming Hours (week)			-0.003***	-0.003***
			(0.000)	(0.000)
RGDP				-0.001***
				(0.000)
# Schools				0.005***
				(0.001)
# Teachers				-0.000***
				(0.000)
YEAR FE	NO	YES	YES	YES
REGIONAL FE	NO	YES	YES	YES

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

< Table 3.3. Logistic Regression Analysis: Impact of Rural-to-Urban Migration Rate on General Secondary Education Completion *(N=43352)*>

	(1)	(2)	(3)	(4)
Rural-Urban Migration Rate	0.029^{***}	0.009^{**}	0.012^{***}	0.007^{*}
	(0.003)	(0.003)	(0.003)	(0.003)
Age	-0.009^{***}	-0.008^{***}	-0.009^{***}	-0.009^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
Male	-0.150^{***}	-0.141^{***}	-0.140^{***}	-0.141^{***}
	(0.005)	(0.004)	(0.004)	(0.004)
Disability	-0.374^{***}	-0.353^{***}	-0.346^{***}	-0.346^{***}
	(0.011)	(0.011)	(0.011)	(0.011)
Never Married	0.070^{***}	0.046^{***}	0.031^{***}	0.033^{***}
	(0.005)	(0.005)	(0.005)	(0.005)
HH Head Education			-0.336^{***}	-0.333^{***}
			(0.013)	(0.013)
HH Size			-0.014^{***}	-0.015^{***}
			(0.001)	(0.001)
Farming Hours (week)			-0.003^{***}	-0.003^{***}
			(0.000)	(0.000)
RGDP				-0.001^{***}
				(0.000)
# Schools				0.004^{***}
				(0.001)
# Teachers				-0.000^{***}
				(0.000)
YEAR FE	NO	YES	YES	YES
REGIONAL FE	NO	YES	YES	YES

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

< Table 3.4. Logistic Regression Analysis: Impact of Total Migration Rate on Secondary Education Completion, including Technical/Vocational Certificates or Diplomas *(N=43352)*>

	(1)	(2)	(3)	(4)
Migration Rate	0.027***	-0.003*	-0.002*	-0.004**
	(0.001)	(0.001)	(0.001)	(0.001)
Age	-0.009***	-0.011***	-0.012***	-0.012***
	(0.001)	(0.001)	(0.001)	(0.001)
Male	-0.080***	-0.077***	-0.075***	-0.075***
	(0.004)	(0.004)	(0.003)	(0.003)
Disability	-0.413***	-0.396***	-0.387***	-0.387***
	(0.013)	(0.013)	(0.012)	(0.012)
Never Married	0.020***	0.011**	-0.008*	-0.008*
	(0.004)	(0.004)	(0.004)	(0.004)
Education			-0.306***	-0.306***
			(0.013)	(0.013)
HH Size			-0.009***	-0.009***
			(0.001)	(0.001)
Agricultural Hours			-0.003***	-0.003***
			(0.000)	(0.000)
RGDP				0.000
				(0.000)
# Schools				-0.001
				(0.001)
# Teachers				-0.000
				(0.000)
YEAR FE	NO	YES	YES	YES
REGIONAL FE	NO	YES	YES	YES

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

< Table 3.5. Logistic Regression Analysis: Impact of Rural-to-Urban Migration Rate on Secondary Education Completion, including Technical/Vocational Certificates or Diplomas *(N=43352)*>

	(1)	(2)	(3)	(4)
Rural-Urban Migration Rate	-0.001	-0.004	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)
Age	-0.010***	-0.011***	-0.012***	-0.012***
	(0.001)	(0.001)	(0.001)	(0.001)
Male	-0.084***	-0.077***	-0.075***	-0.075***
	(0.004)	(0.004)	(0.003)	(0.003)
Disability	-0.427***	-0.396***	-0.387***	-0.387***
	(0.013)	(0.013)	(0.012)	(0.012)
Never Married	0.029***	0.011**	-0.008*	-0.008*
	(0.004)	(0.004)	(0.004)	(0.004)
HH Head Education			-0.306***	-0.306***
			(0.013)	(0.013)
HH Size			-0.009***	-0.009***
			(0.001)	(0.001)
Farming Hours (week)			-0.003***	-0.003***
			(0.000)	(0.000)
RGDP				0.000
				(0.000)
# Schools				-0.000
				(0.001)
# Teachers				-0.000
				(0.000)
YEAR FE	NO	YES	YES	YES
REGIONAL FE	NO	YES	YES	YES

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

<Table 3.6. First Stage Linear Regression Analysis for Total Migration in 2SRI>

	(1)	(2)	(3)
# $Rainy\_Days_{t-1}$	0.032***	0.032***	-0.009***
	(40.89)	(40.85)	(-6.50)
# $Rainy\_Days_{t-2}$	0.053***	0.053***	-0.002
	(52.89)	(52.88)	(-1.51)
# $Rainy\_Days_{t-3}$	-0.047***	-0.047***	-0.071***
	(-33.67)	(-33.66)	(-50.67)
R^2	0.819	0.819	0.833

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

< Table 3.7. First Stage Linear Regression Analysis for Rural-to-Urban Migration in 2SRI >

	(1)	(2)	(3)
# $Rainy\_Days_{t-1}$	0.001	0.001	-0.005***
	(1.23)	(1.19)	(-6.46)
# $Rainy\_Days_{t-2}$	0.005***	0.005***	0.003***
	(14.76)	(14.78)	(4.88)
# $Rainy\_Days_{t-3}$	-0.016***	-0.016***	-0.015***
	(-24.80)	(-24.79)	(-19.56)
R^2	0.522	0.523	0.528

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

<Table 3.8. Second Stage Logistic Regression Analysis: Total Migration Rate on General Secondary Education Completion *(N=43352)*>

	(1)	(2)	(3)
Migration Rate	0.036***	0.035***	0.016**
	(8.90)	(8.89)	(2.69)
Age	-0.008***	-0.009***	-0.009***
	(-9.64)	(-11.20)	(-11.23)
Male	-0.141***	-0.141***	-0.141***
	(-32.18)	(-32.81)	(-32.98)
Disability	-0.353***	-0.346***	-0.346***
	(-31.07)	(-30.30)	(-30.29)
Never Married	0.044***	0.029***	0.033***
	(8.81)	(5.96)	(6.62)
HH Head Education		-0.332***	-0.332***
		(-25.71)	(-25.92)
HH Size		-0.015***	-0.015***
		(-11.27)	(-11.49)
Farming Hours (week)		-0.003***	-0.003***
		(-34.93)	(-34.89)
RGDP			-0.001***
			(-4.84)
# Schools			0.006***
			(4.68)
# Teachers			-0.000***
			(-3.84)
Residuals	0.007***	0.008***	0.009***
	(5.35)	(5.84)	(6.51)
Individual	YES	YES	(-1.18)
Household	NO	YES	YES
Regional	NO	NO	YES

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

< Table 3.9. Second Stage Logistic Regression Analysis: Rural-to-Urban Migration Rate on General Secondary Education Completion *(N=43352)*>

	(1)	(2)	(3)
Rural-Urban Migration Rate	0.169***	0.157***	0.038
	(6.90)	(6.59)	(1.53)
Age	-0.008***	-0.010***	-0.009***
	(-9.69)	(-11.30)	(-11.27)
Male	-0.142***	-0.141***	-0.141***
	(-32.24)	(-32.87)	(-32.94)
Disability	-0.351***	-0.345***	-0.345***
	(-30.82)	(-30.10)	(-30.23)
Never Married	0.043***	0.028***	0.032***
	(8.52)	(5.68)	(6.51)
HH Head Education		-0.331***	-0.332***
		(-25.48)	(-25.80)
HH Size		-0.015***	-0.015***
		(-11.74)	(-11.42)
Farming Hours (week)		-0.003***	-0.003***
		(-35.19)	(-34.49)
RGDP			-0.001**
			(-3.19)
# Schools			0.005***
			(4.10)
# Teachers			-0.000***
			(-5.34)
Residuals	0.007*	0.009**	0.007*
	(2.12)	(3.02)	(2.24)
Individual	YES	YES	YES
Household	NO	YES	YES
Regional	NO	NO	YES

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

<Table 3.10. Second Stage Logistic Regression Analysis: Total Migration Rate on General Secondary Education Completion, including Technical/Vocational Certificates or Diplomas *(N=43352)*>

	(1)	(2)	(3)
Migration Rate	0.007	0.009**	0.004
	(1.88)	(2.70)	(0.75)
Age	-0.011***	-0.012***	-0.012***
	(-15.70)	(-17.60)	(-17.61)
Male	-0.077***	-0.075***	-0.075***
	(-21.67)	(-22.40)	(-22.40)
Disability	-0.396***	-0.387***	-0.387***
	(-31.58)	(-31.25)	(-31.25)
Never Married	0.011*	-0.009*	-0.008*
	(2.56)	(-2.20)	(-2.08)
HH Head Education		-0.304***	-0.305***
		(-24.12)	(-24.08)
HH Size		-0.009***	-0.009***
		(-8.23)	(-8.17)
Farming Hours (week)		-0.003***	-0.003***
		(-51.57)	(-51.36)
RGDP			-0.000
			(-0.07)
# Schools			0.000
			(0.09)
# Teachers			-0.000
			(-0.71)
Residuals	-0.004**	-0.004***	-0.004***
	(-3.14)	(-3.39)	(-3.46)
Individual	YES	YES	YES
Household	NO	YES	YES
Regional	NO	NO	YES

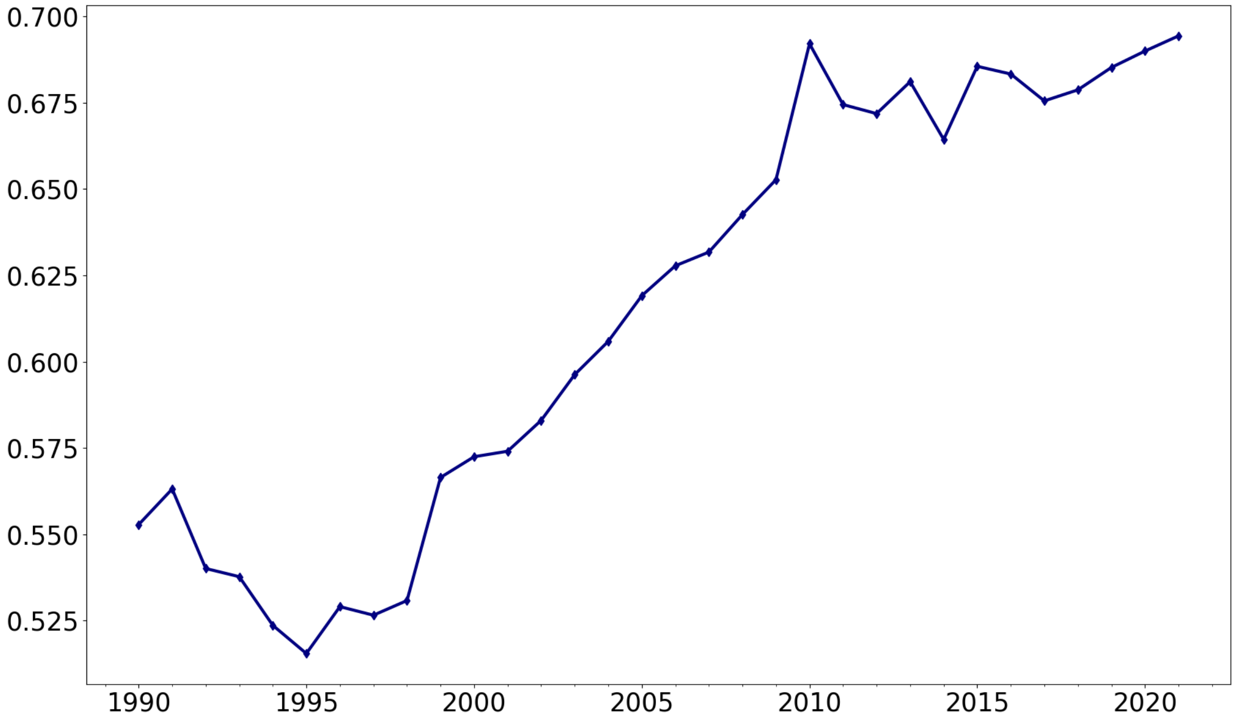
t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

<Table 3.11. Second Stage Logistic Regression Analysis: Rural-to-Urban Migration Rate on General Secondary Education Completion, including Technical/Vocational Certificates or Diplomas *(N=43352)*>

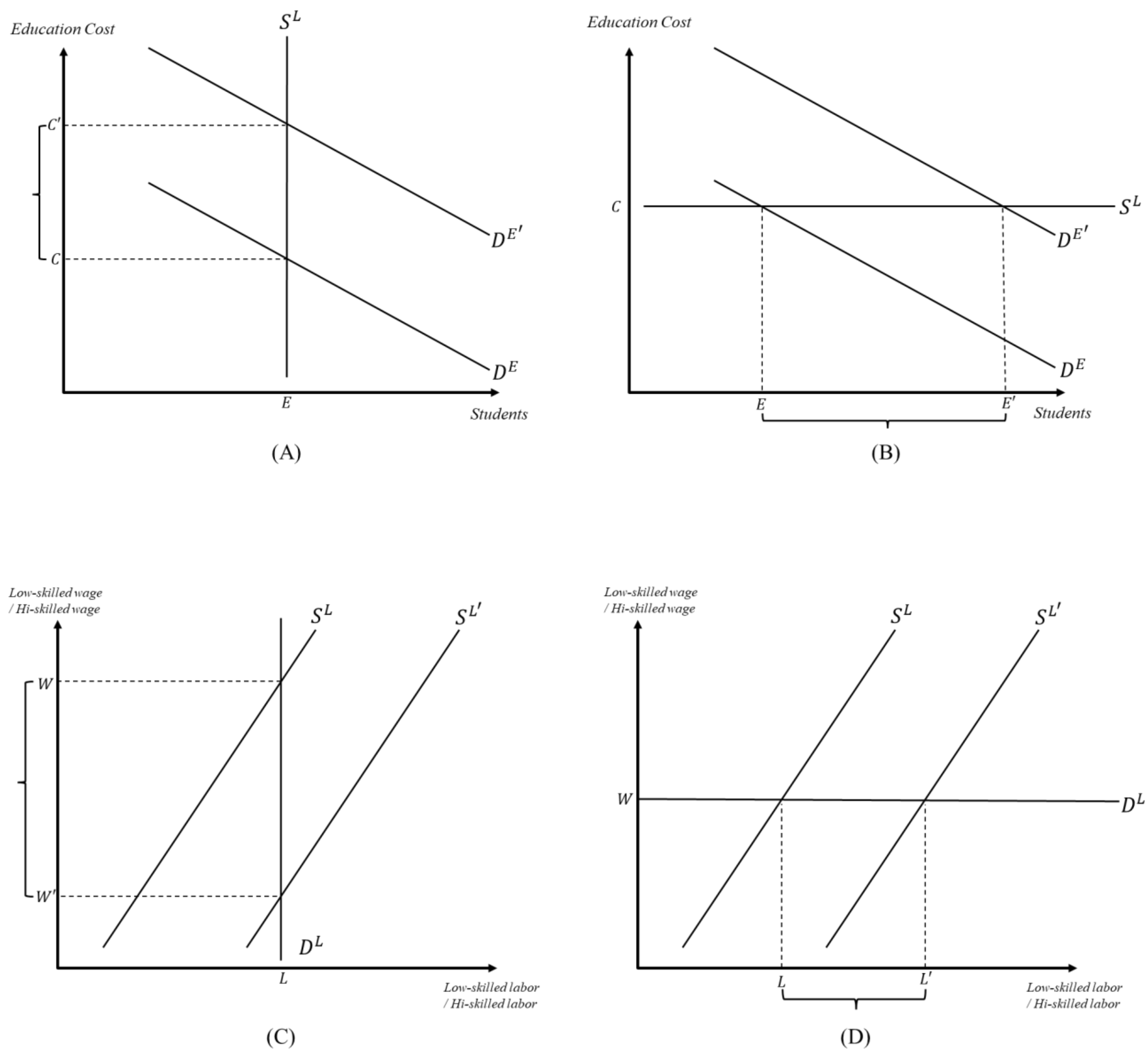
	(1)	(2)	(3)
Rural-Urban Migration Rate	0.029	0.032	0.015
	(1.39)	(1.65)	(0.74)
Age	-0.011***	-0.012***	-0.012***
	(-15.73)	(-17.64)	(-17.63)
Male	-0.078***	-0.075***	-0.075***
	(-21.67)	(-22.41)	(-22.40)
Disability	-0.395***	-0.387***	-0.387***
	(-31.54)	(-31.22)	(-31.23)
Never Married	0.010*	-0.009*	-0.008*
	(2.52)	(-2.20)	(-2.09)
HH Head Education		-0.304***	-0.305***
		(-24.06)	(-24.00)
HH Size		-0.009***	-0.009***
		(-8.32)	(-8.19)
Farming Hours (week)		-0.003***	-0.003***
		(-50.55)	(-50.13)
RGDP			0.000
			(0.95)
# Schools			0.000
			(0.01)
# Teachers			-0.000
			(-1.19)
Residuals	-0.005*	-0.003	-0.003
	(-2.02)	(-1.24)	(-1.25)
Individual	YES	YES	YES
Household	NO	YES	YES
Regional	NO	NO	YES

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

<Figure 3.1. Total Urbanization Rate in Mongolia between 1990 and 2021>



<Figure 3.2. Supply and Demand Curves in the Education and Labor Markets>



CHAPTER 4

THE IMPACT OF HOUSEHOLD MEMBER MIGRATION ON CHILDREN'S NUTRITION IN AGRICULTURAL HOUSEHOLDS: EVIDENCE FROM MEXICAN FAMILIES[15]

[15] Yoon, T., and Filipski, M. To be submitted to *Journal of Development Studies*

Abstract

The final chapter of this dissertation explores the correlation between international migration of family members and the welfare of the children who remain in their original locations within agricultural households in the developing world.

The study employs longitudinal data from the Mexican Family Life Survey (MxFLS) to scrutinize the relationship between the migration of household members and the nutritional status of children in agricultural households in Mexico. We assess the net impact of migration on children's anthropometric indicators, accounting for the effect of cash transfers through programs like PROCAMPO and PROGRESA, using a Triple-Differences specification.

Our empirical findings indicate a positive correlation between the migration status of a household and children's nutritional health, specifically when interacting with cash transfers, specifically from PROCAMPO. The results demonstrate that when we account for the cash transfer, PROCAMPO, in comparison to non-cash transfer recipients and children from migrant households, the migration of household members enhances the long-term nutritional status, as indicated by height-for-age z-scores. By focusing on the well-being of agricultural families left behind, the study strives to clarify the intricate interplay between international migration, cash transfers, remittances, labor allocation, and child welfare in developing regions.

Introduction

It is widely believed that international migration of family members significantly impacts the welfare of children remaining in their countries of origin, especially within the context of developing nations. The interplay between the absence of family members and the increase in remittances due to international migration, and its subsequent effect on children's nutritional status in developing countries, has drawn increasing attention from both researchers and policymakers. Despite the emergence of a substantial body of literature dedicated to investigating this complex relationship, the findings remain controversial, and consensus is yet to be reached (Antén, 2010; Hildebrandt et al., 2005; Lu, 2015).

In order to address the knowledge gap, this study aims to investigate the impact of international migration of one or more household members on the nutritional status of children in developing countries, specifically on households engaged in agricultural activities. We have concentrated on the agricultural labor channel of the migration impact for several reasons. Firstly, agriculture remains a predominant source of income and employment in the rural regions of many developing countries. Moreover, the environmental and socio-economic conditions in rural areas have become increasingly challenging, leading to rural people opting for migration as a means to alleviate these difficulties. Consequently, the rural residents, primarily engaging in agriculture, have emerged as a major contributor to international migration and remittances in developing countries. Hence, a comprehensive understanding of the welfare of families left behind in rural areas, specifically agricultural households, in developing countries is important for both domestic policymakers and rural populations (Obi et al., 2020). Secondly, the loss of agricultural labor due to migration can have a significant impact on the welfare of family members, particularly the children's health status. Although previous studies have found that financial and social remittances can improve the economic conditions of households left behind in origin countries, the loss of labor can have a negative impact on children's health, even if economic conditions are relaxed through remittances. For instance, the spouses or children of migrants may bear a greater workload in the agricultural household to compensate for labor loss. This could result in increased stress and fatigue for women, which could have a detrimental effect on children's health, or children's increased workload could worsen their health directly. Several previous studies presented empirical evidence that the out-migration of household members leads to a reduction in the working hours of urban and rural women in off-farm income-generating activities, despite the increase in income due to remittances. Women from agricultural households are more likely to spend time on non-wage activities and agricultural tasks when there is out-migration in the households (Amuedo-Dorantes and Pozo 2006; Hanson 2007; Chen 2006; Mu and van de Walle 2011; Binzel and Assassd 2011; Lokshin, M., & Glinskaya 2009). Moreover, Chang et al. (2011) find that the household members' migration significantly increases the farming hours of children and elderly individuals who remain in rural China.

In our empirical analysis, a straightforward linear regression approach to examine the causal relationship between international migration from developing countries and the health status of children left behind would yield biased estimates due to the presence of unobservable factors that simultaneously influence household members' migration decisions and the nutritional well-being of their children. To mitigate this endogeneity concern, we employ the difference-in-differences (DD) estimation method. Additionally, we enhance our analysis by incorporating the triple-differences (DDD) framework, which allows us to isolate the correlation of receiving cash transfers and children's health while accounting for both observed and unobserved factors that may confound the relationship between migration decisions and children's health.

Furthermore, to strengthen the case for causality, we propose an additional identification strategy by combining our primary empirical approaches, namely difference-in-differences and triple-differences, with propensity score matching (PSM) as a robustness check. We observe significant differences in the distribution of propensity scores between migrant and non-migrant households. To address this issue, we narrow our sample to the common support regions, which helps minimize the discrepancy between the curves. Subsequently, we estimate the impact of international migration by household members on the health status of children in agricultural households.

Our findings indicate that the difference-in-differences results reveal a positive but statistically insignificant association between U.S. migration of household members and children's health status. However, when employing the triple-differences approach, which includes an interaction term for migration status and cash transfer recipients, we observe a significant positive relationship between migration, cash transfer programs, and children' nutritional status. Specifically, the results demonstrate statistical significance at the 1% level for PROCAMPO beneficiaries.

We also conduct analyses using a matched sample, but the results do not yield statistical significance in the identification of standard difference-in-differences. However, when applying the triple-difference approach to the matched sample, we find consistent results showing a positive association between children from PROCAMPO beneficiary and migrant households and only improved height-for-age z-scores over time at a 1% statistically significant level. Similar patterns are observed for the PROGRESA program, with only the height-for-age z-scores exhibiting a positive relationship at the 10 % level of statistical significance.

This study offers several significant contributions. Firstly, we concentrate specifically on children from households involved in agriculture. Agricultural work is a significant source of migration and remittances from developing countries and demonstrates that rural people make the decision to migrate to developed nations to enhance their family's welfare. Additionally, previous studies have demonstrated a positive correlation between international migration and remittances and the financial well-being of households left behind in source countries. However, it remains unclear whether non-economic aspects of household welfare, such as children's growth, are also positively associated with international migration of household members. Furthermore, if households participate in labor-intensive household production, such as farming, the loss of labor due to migration could harm family well-being despite the alleviation of overall poverty from remittances. Secondly, we estimate the heterogenous treatment effect on cash transfer recipients among farming households, PROGRESA, and PROCAMPO. Cash transfers in developing countries have direct links to both migration decisions and children's health conditions. However, to the best of our knowledge, previous studies have not considered cash transfer recipient status when studying the impact of migration on children's health. Therefore, to better understand the impact of international migration on children's nutritional status, we take into account the heterogenous migration impact on children's health outcomes based on the respondents' cash transfer recipient status, PROGRESA or PROCAMPO, at the first round of the survey.

The structure of this paper is organized as follows: Section 1 provides an introduction of our study. In Section 2, we present the conceptual framework that explores the interplay between the absence of family members due to migration and the income effect resulting from remittances. Section 3 offers a description of the background and study area, while Section 4 reviews the relevant literature. In Section 5, we provide an overview of the Mexican Family Life Survey

(MxFLS), including a description of the observational statistics, comparing characteristics between migrant and non-migrant households. Section 6 presents our empirical strategies and the main results obtained from our primary identification strategy. Section 7 introduces our PSM-DD identification strategies and the results as a robustness check. Finally, in Section 8, we conclude the paper and provide further discussion points.

Conceptual Framework

There are two primary pathways through which the migration of household members can impact the nutritional status of children who are left behind. The first channel is the income effect, and according to the Hildebrandt et al. (2005), the most obvious channel for positively affecting child health is through increased income from remittances. A higher income allows households to improve their nutritional status by consuming better quality and quantity of foods. For instance, consuming more foods rich in micronutrients and proteins such as fruits, vegetables, and meats, can improve children's nutritional condition (Nguyen & Winters, 2011; Subramanian & Deaton, 1996). In addition, improving the household's hygienic environment and accessing better medical services can have a positive impact on children's health (De Brauw & Mu, 2011; Hildebrandt et al., 2005).

The other channel that affects children's nutritional status is through changes in the time and labor allocation of remaining household members when there is a labor loss due to international migration. The impact of changes in labor allocation on children's health status is not clear cut. For example, when the father migrates, the spouse left behind may have less time and energy for breastfeeding and child-rearing due to the increased workload, which could negatively affect child growth. On the other hand, the spouse may have more time to care for their children due to remittances, which can reduce their need for wage work in the origin country (Azzarri & Zezza, 2011; Carletto et al., 2011; De Brauw & Mu, 2011; Hildebrandt et al., 2005; Lu, 2015). Therefore, the allocation of time and labor due to household migration can both benefit and harm children, depending on the direction of the changes in their well-being. Two key factors, rising income levels and alterations in household productivity, shape the way in which household members who are left behind allocate their time and labor (Lokshin & Glinskaya, 2009). On the one hand, when remittances raise the income of the remaining household members, they are likely to spend more time which children and reduce their participation in wage-paying jobs. For instance, a higher reservation wage due to remittances means that these individuals are less likely to trade leisure time for labor force participation (Cox-Edwards & Rodríguez-Oreggia, 2009). As a result, remaining adults will have more time for childcare, leading to improved health outcomes for children. On the other hand, the migration of household members leads to a loss of labor and impacts the productivity of the remaining household members. If the missing labor is a complementary input regarding household production, the productivity per unit of labor in household production decreases. This implies that the remaining household labor must exert more effect and time to produce the same amount of output, particularly in labor-intensive household production. Conversely, if the labor of migrant household members was a substitute input, the marginal productivity of the remaining household labor for household production increases. In this scenario, migration makes household production more valuable and incentivizes the remaining members to increase their participation in household production (Kniesner, 1976; Leeds & Allmen, 2004; Lokshin & Glinskaya, 2009).

Overall, it appears likely that increased income in agricultural households can have a positive impact on children's health, while the implications of time and labor reallocation due to the migration of a household member are yet to be definitively assessed. Theoretically, while migration could enhance the overall economic condition of households, the impact on children's health – whether positive or negative- remains inconclusive. It could be subject to the complex interplay between increased remittances and adjustments in time or labor allocation.

Background and Study Area

Mexico Migration

In this study, Mexico is the focus country as it presents a unique opportunity to address research questions due to the distinctive characteristics of Mexican migration. Firstly, Mexican migration to the United States has been dominated by agricultural labor, which has been the oldest and largest migration flow from Mexico to the U.S. in the 20th century (Durand et al., 1999; Durand et al., 2001; Fussell & Massey, 2004; Massey & Espinosa, 1997; Massey et al., 1994; Phillips & Massey, 2000; Winters et al., 2001). Secondly, rural Mexico is characterized by a predominant pattern of circular migration. Migrant workers primarily from farming households, tend to migrate seasonally, thus not reducing the overall agricultural production in their homeland (Fussell & Massey, 2004). Thirdly, while not exclusive to Mexico, agricultural labor is often viewed as a substitute labor input for household farming. Additionally, the rural labor market in many developing countries is not extensive and flexible, which means that agricultural households in rural Mexico are likely to offset the labor loss from migrant household members to maintain the production volume (Chen, 2006; Mu & Van de Walle, 2011; Sadoulet et al., 2001).

Considering the context of international migration from rural Mexico, the substantial flow of seasonal migration, the labor input characteristics of the agricultural sector, and the rigid labor market in developing countries could play a significant role. Remaining household members may find less time for activities such as childcare, which includes monitoring dietary habits, conducting regular health check-ups, or administering vaccinations. Consequently, if the positive impact of increased income does not outweigh the negative effects of time and labor reallocation among agricultural household members, children's nutritional status may deteriorate.

Cash Transfer in Mexico

We also extend our discussion to how conditional rural cash transfers are linked with migration and children's nutritional status. In Mexico, there are two significant conditional cash transfer programs for agricultural households, namely PROCAMPO and PROGRESA, and it is crucial to control the beneficiaries of these programs in order to reduce bias in our estimation. Firstly, cash transfers constitute a significant portion of the average total household income in rural Mexico. Secondly, the recipients of these transfers possess distinct household characteristics as both programs have different eligibility criteria. For instance, PROGRESA primarily determines eligibility based on household well-being and income levels, while the eligibility of PROCAMPO recipients is dependent on the households' total hectares of specific crops (corn, beans, rice, wheat, sorghum, barley, soybeans, cotton, and cardamom). Finally, the conditions required to maintain eligibility can affect household behavior, which in turn can directly or indirectly influence migration decisions and children's health (Ruiz-Arranz et al., 2006). For example, PROGRESA requires recipients to regularly visit health clinics, attend public health lectures, and send their children to school. Furthermore, PROGRESA payments are directly made to women as they are more likely to use the income to improve their children's nutrition and education. Thus, it is plausible that the health of children receiving PROGRESA benefits will improve through enhanced health knowledge and increased access to health services. Conversely, the additional income from PROCAMPO is more likely to be associated with agricultural production. The household members are required to invest cash transfers to agricultural production, and the multiplier effects of productivity from investment are expected to increase long-term household welfare (Sadoulet et al., 2001). Also, since the PROCAMPO payment can be used as collateral, it may relax the credit constraints of households. Therefore, PROCAMPO will have a significant impact on quantity of nutritional intake of household members through agricultural investment and improved agricultural productivity (Ruiz-Arranz et al., 2006).

Several previous literatures suggest that cash transfer can either foster or deter international migration. Stecklov et al. (2005) find that the PROGRESA decrease overall migration to the U.S, but Angelucci (2012) argues that the program increased labor-induced migration to the US by reliving the credit constraints of eligible households. Scott-Andretta and Cuecuecha (2010) discover that the PROCAMPO decrease the migration flows from Mexico to the US. Hence, the relationship between cash transfers, migration, and children's nutritional status is complex and can be influenced by various factors such as household characteristics and program requirements. The empirical evidence from the literature suggests that both PROGRESA and PROCAMPO can have different impacts on migration and children's health. While PROGRESA has been associated with improved health knowledge and access to health services, PROCAMPO can have a significant impact on agricultural productivity and household welfare through investment in agriculture and alleviation of credit constraints. Therefore, it is important to consider the specific context and program design in order to understand the impact of cash transfers on migration and children's nutritional status.

Literatures Review

There exists a growing body of literature that examines the effects of migration and remittances on the health outcomes of children. However, the empirical evidence regarding these effects remains mixed and controversial. Kanaiaupuni and Donato (1999) estimated the effect of migration on infant mortality in five states of Mexico and discovered that migration and remittances are positively associated with the long-term survival rate of children in Mexico. Similarly, studies by Hildebrandt et al. (2005) and López Córdova (2018) found that parental migration has a positive impact on the health outcomes of their children in Mexico, leading to a decrease in infant mortality and a lower risk of low birth weight. However, Hildebrandt et al. (2005) revealed that children in migrant households are less likely to receive preventive health services, be fully vaccinated, and breastfed. Another study, (Nobles, 2011), also focused on Mexico and explores the impact of international migration on child health outcomes. This study uses anthropometric indicators, particularly Height-for-age Z-scores (HAZ scores) and finds that migration negatively affects HAZ scores among children in Mexico. In addition to studies focused on Mexico, several studies have examined the relationship between migration and children's health outcomes in other Central American countries (Antén, 2010; Carletto et al., 2011; De Brauw, 2011). Antén (2010) reported that remittances have a positive impact on the short-term nutritional status, as measured by WHZ and WAZ, of children under the age of 5 in Ecuador, but not on the long -or middle-term nutritional status, as measured by HAZ. De Brauw (2011) estimated whether migration and remittances could mitigate the negative impact of a food crisis on children's dietary status in El Salvador and finds that the HAZ of children from migrant households did not decrease compared to that of children from non-migrant households during the food crisis in 2008. Carletto et al. (2011) also observed a positive relationship between the international migration of household members and children's anthropometric statistics, including Height-for-Age Z-scores (HAZ) and the prevalence of stunting. Studies conducted by Azzarri and Zezza (2011) and De Brauw and Mu (2011) utilized data from Asian countries, specifically Tajikistan and China, respectively. Azzarri and Zezza (2011) found a positive relationship between migration and the remaining children's food security and nutrition, as measured by migration data and HAZ statistics. Furthermore, De Brauw and Mu (2011) revealed that the migration of household members has heterogenous effects on children's nutritional status outcomes for left-behind children in rural areas. The results show that older children between the ages of 7 and 12 are more likely to be underweight when their parents migrate, while this is not the case for younger children between the ages of 2 and 6. The authors suggest that when parents migrate, older children take on more household chores, which leads to be underweighted, while remaining adults spend less time on housework.

In summary, research investigating the correlation between migration and children's health outcomes indicates a multifaceted and context-dependent relationship. Some studies report positive effects on infant mortality and anthropometric statistics, while others highlight negative effects on preventive health services and breastfeeding practices. Previous research emphasizes the importance of understanding the effects of migration and remittances on children's health to inform policymaking. However, it is crucial to note the need for additional research to address the existing knowledge gap.

Data and Descriptive Statistics

Mexican Family Life Survey (MxFLS) Data

The analysis of the connection between international migration of agricultural household members and the health of children who are left behind is conducted using the first two waves of the Mexican Faily Life Survey (MxFLS). The MxFLS is a nationally representative longitudinal household survey carried out in Mexico, designed by researchers from both Mexico and the United States. The first wave of the survey, conducted in 2002, involved 8,440 households from 150 communities. The second wave followed up on over 90% of the respondents from the first wave, regardless of their decision to reside in either Mexico or the U.S. (Rubalcava & Teruel, 2008). The respondents were also asked detailed questions about individual, household, and community attributes. The survey also includes detailed questions regarding individual, household, and community characteristics.

The MxFLS is a highly suitable data source for this study due to several reasons. Firstly, the MxFLS questionnaires provide extensive information regarding short- and long-term migration histories and maintain a high re-contact rate of over 90% of respondents in the second wave (Rubalcava & Teruel, 2008). Secondly, it includes detailed anthropometric information about children, which is utilized as the primary outcome variable of interest. Lastly, the dataset contains information on the receipt of cash transfers during the first wave. This detail enables the control of heterogenous impact of international migration on children's nutritional status, relying on the standing of beneficiaries of cash transfers within rural Mexico, PROGRESA and PROCAMPO.

Children's Anthropometric (Nutritional) Indicators

This study employs two anthropometric indicators to assess the health of children as the main outcome variables of interest. These indicators are Height-for-age Z-scores (HAZ) and Weight-for-height Z-scores (WHZ) or BMI Z-scores for children over five years of age. The Z-scores are determined by comparing each child to an international (Antén, 2010) standard population of children of the same monthly age and sex as defined by the World Health

Organization (WHO) (De Onis et al., 2007; WHO, 2006). Different measures of children's anthropometric indicators reflect different dimensions of nutritional status. HAZ is a long-term indicator that indicates chronic malnutrition or diseases and is less likely to measure the body mass relative to height and weight for z-scores and is useful to represent short-term changes in nutritional status. WHZ and BMIZ are used for children under five years and over five years of age, respectively. To evaluate the long- and short-term effects of international migration on children's nutritional status, we utilize HAZ and WHZ or BMIZ as dependent variables (Acosta et al., 2007; Alves & Belluzzo, 2004; Antén, 2010; Attanasio et al., 2004; Duflo, 2003; Gibson et al., 2011; Langworthy, 2011; Morales et al., 2004; Valdivia, 2004)[16].

Migration Identification

In terms of migration measures, a key concept that we are interested in identifying in children from migrant households is whether the children are influenced by the family member who has migrated to another country (Azzarri & Zezza, 2011). To construct an appropriate sample for this study, I have selected children who were interviewed in the first survey wave and remain in the same households in the second wave. Specifically, the treatment group comprises children from agricultural households that have one or more migrant members who are older than 15 years and have resided in the United States for more than a month between the first and second survey waves[17]. The control group, on the other hand, comprises children from households with working-

[16] As per the World Health Organization (WHO) guidelines, the parameters for exclusion include Height-for-Age Z (HAZ) scores that fall below -6 or above +6, and Weight-for-Height Z (WHZ) scores that are below -5 or greater than +5. The excluded samples is highly unlikely that any sample selection bias will be introduced Antén, J.-I. (2010). The impact of remittances on nutritional status of children in Ecuador. *International migration review*, *44*(2), 269-299.

[17] Currently, a household member who is a migrated between the first and second wave may be located in either Mexico or the United States.

age member who have not visited the United States in the past five years, as of the second wave. This implies that we can safely assume that the control group children have not been impacted by the loss of labor and remittance due to U.S. migration for at least five years. However, it is unfortunate that data on temporary migration, which refers to household members who stayed in the U.S. for less than a year, is only available for the two years preceding the second survey wave. Although there is possibility of attrition among control group children who were influenced by a migrant member and who had traveled to the United States three years before the second survey wave for a period of less than one year, it is unlikely to have significantly changed the nutritional status of the children. In addition, in order to assess the effect of current labor loss on the health of children, we have defined two treatment groups to estimate the heterogenous impact of current migration on the short-term nutritional status of children left behind in Mexico. The first treatment group includes households with migrant members, irrespective of their current location. In contrast, the second treatment group is limited to families with migrant members who are currently residing in the United States.

This method has enabled us to identify 1610 individual-period observations. Among these observations, 258 individual-period samples belong to the treatment group and reside in agricultural households with a migrant member. The remaining 1352 individual-period observations belong to the control group. Among the total observations, 42% are beneficiaries of PROGRESA, and 33% are beneficiaries of PROCAMPO. While households receiving PROGRESA are ineligible for other anti-poverty or educational subsidies, they are allowed to apply for PROCAMPO. Moreover, a significant portion of farming households, 84% to be precise, are located in rural communities with a population of less than 2,500.

Empirical strategies and Results

Theoretical motivation

The identification of the impact of international migration on children's nutritional status is a complex issue because migrant households are not randomly assigned. Additionally, since migration is not an exogenous variable, it may be associated with several variables that also influence the nutritional status of children. As a result, comparing children in migrant households with those in non-migrant households could result in either "upward" or "downward" biased estimate of the impact of migration. This poses a significant practical problem in accurately measuring the effect of migration on children's nutritional status.

Three major concerns make a migrant decision is not assumed random:

1. Selection bias may arise due to differences in the characteristics of migrant and non-migrant households, where parents who are more educated and have better health knowledge for children are more likely to migrate (positive selection), and those who are less likely to migration (negative selection) may have worse nutritional outcomes.
2. An omitted variable could also create and endogeneity problem for instance, if a household experience economic shocks such as crop failure or natural disasters that affect children's nutritional status, this variable may be correlated with the household's decision to migrate for additional income to compensate agricultural loss.
3. Reverse causality may lead to an endogeneity problem where parents migrate due to their children's malnutrition or illness.

This section aims to tackle the aforementioned challenges through the explanation of two identification strategies: the standard difference-in-differences (DID) and triple-differences (DDD) approaches. These methodologies will be employed to estimate the treatment effect of U.S. migration from farming households in Mexico on the nutritional status of children.

Difference-in-Difference method

Supposed we have child i who belongs to group $M_i = m \in \{0,1\}$, where $M_i = 1$ represents households with migrant members, and the child is observed in period $T_i = t \in \{0,1\}$, where $T_i = 1$ is the second wave. Let $Y_i(0)$ and $Y_i(1)$ be the potential outcomes for the child before and after migration, respectively, such as Height-for-Age Z-scores (HAZ). Then, the average treatment effect on the treated (ATT), the DID estimand, δ_{DID}, is given by

$$
\begin{aligned}
ATT = \delta_{DID} = E[Y_i(1) - Y_i(0)] \\
&= (E[Y_i|M_i = 1, T_i = 1] - E[Y_i|M_i = 1, T_i = 0]) \\
&- (E[Y_i|M_i = 0, T_i = 1] - E[Y_i|M_i = 0, T_i = 0])
\end{aligned}
$$

where $(E[Y_i|M_i = 1, T_i = 1] - E[Y_i|M_i = 1, T_i = 0])$ and $(E[Y_i|M_i = 0, T_i = 1] - E[Y_i|M_i = 0, T_i = 0])$ show the difference between the treatment and control groups in the post-program period and pre-program period respectively. We can remove the bias caused by pre-existing differences between the treatment and control group through a double differencing procedure represented in the equation above.

In practice, the DID estimation approach can be implemented using a linear regression framework. Let y_{it} denote the height-for-age (HAZ) or weight-for-height (WHZ or BMIZ) of observations on date t. The primary explanatory variable is $USMIG \cdot Year$, which is an indicator variable equal to 1 if children live in an agricultural household which has migrant household members who have been to the U.S. between the first and second wave in time of the second wave and 0 otherwise. Let $USMIG_i$ denotes indicator variables for migration status, $Year_i$ time indicators, and X_i is a vector of control variables. Additionally, let ε_i represent an unobserved individual error term.

$$HAZ_i = \alpha + \delta_{DID} \cdot USMIG \cdot Year_i + \gamma \cdot Year_i + \lambda \cdot USMIG_i + X_i \cdot \beta + \varepsilon_i \quad (1)$$

Here, λ represent the difference in the mean outcome variable (i.e. HAZ, WHZ, or BMI) between children living in households with migrant members and children living in households without migrant members before migration, and δ_{DID} represents the DD estimator, which captures the difference in the change in the outcome variable between the two groups after migration, while controlling for any pre-existing differences in the outcome variable.

Difference-in-Difference-in-Differences (DDD) method

Difference-in-differences estimation methodology, which offers an opportunity to account for variables that potentially influence the nutritional status of children in migrant households. This is achieved by comparing the nutritional status of children in migrant households with that of children in non-migrant households and examining variations in the outcome variable before and after migration. By employing this method, we can effectively exclude the influence of invariant temporal trends and other confounding variables that may equally affect both the treatment and control groups under study.

However, it is important to acknowledge that there may still be unobserved factors that influence migration decisions and correlate with the outcome variable. To address this concern, we utilized the Triple-difference (DDD) method, incorporating and interaction term that represents both the indicators of cash transfer recipients and migrant households. This inclusion allows us effectively to control for and estimate the heterogenous treatment effect specific to the two prominent cash transfer initiatives in Mexico, namely PROCAMPO and PROGRESA. The Triple-differences approach offers a notable advantage by effectively controlling for time-varying unobservable factors that may differentially impact individuals based on their migration status and receipt of cash transfers. In our analysis, we consider three essential dimensions: migration status, cash transfer beneficiary status, and time.

The first difference in our analysis pertains to migration status, enabling a comparison of outcomes between individuals who have migrated and those who have not. This comparison allows us to identify divergences in the effects experienced by these two groups. The second difference is based on whether individuals receive cash transfers or not, allowing us to differentiate between beneficiaries and non-beneficiaries. This differentiation provides valuable insights into the varying impacts of cash transfers. Lastly, by incorporating time as our third dimension, we can observe how these effects evolve over time, providing a comprehensive understanding of the dynamics involved.

$$y_{it} = \beta \cdot USMIG \cdot Time \cdot Cash_i + \gamma \cdot X_i + \zeta \cdot USMIG_i \cdot Time_i + \lambda \cdot USMIG_i \cdot Cash_i + \mu \cdot Time_i \cdot Cash_i + \nu \cdot USMIG_i + \xi \cdot Time_i + \sigma \cdot Cach_i + \varepsilon_i \quad (2)$$

The triple-difference specification incorporates an interaction term among migration, cash transfer beneficiary status, $USMIG \cdot Time \cdot Cash_i$, and time variables. These interaction terms enable the triple-difference analysis to account for group-specific unobserved temporal shifts or consumption patterns pertinent to children's nutritional status and migration.

Results

Table 4.2 presents the results of the standard difference-in-differences (DID) estimation on the long-term and short-term nutritional status of children in agricultural households, and the relationship with U.S. migration of household members. The Height-for-Age Z-scores (HAZ) is an indicator of children's long-term nutritional status, while the Weight-for-Height (WHZ) or BMI-for-Age Z-scores (BMIZ) is an indicator of children's short-term nutritional status. The first and second columns report the DID regression results for HAZ, and the third and fourth columns show the DID regression estimates on WHZ (BMIZ). Furthermore, in this study, we identified children from migration households in two different ways. In columns (1) and (3) of Table 4.2, the regression results with treatment groups include children from agricultural families with migrant members who have been to the U.S. between the first and second waves, regardless of their current residential locations. In contrast, in columns (2) and (4), the treatment groups include children from agricultural households with migrant members who currently reside in the U.S., implying the absence of current labor loss. The standard difference-in-differences results in Table 4.2 from columns (1) to (4) show that the household members' U.S. migration is positively associated with the long-term and short-term health status of children from agricultural households, but the results are not statistically significant.

The results derived from the application of the triple-differences specification, including an interaction term for migration status and cash transfer recipients, are presented in Tables 4.3 and 4.4. Table 4.3 provides an analysis with a focus on PROCAMPO beneficiary status. The interaction term for PROCAMPO beneficiary status, migration status, and year yield a significant outcome as shown in column (1). The coefficient was found to be 0.729, indicating statistical significance at the 1% level, with the height-for-age z-score employed as the dependent variable. This suggests a strong and statistically significant relationship between the three factors and children's nutritional status in the context of international migration. It reveals that, over time, children from households with a migrant member and receiving PROCAMPO cash transfers exhibited an improvement in nutritional status by 0.864 and 1.03 z-scores, in contrast to children from non-migrant households or those not receiving PROCAMPO cash transfers.

Column (2) of Table 4.3 expands upon this analysis by focusing on migrant households in which the migrant member currently resides in the U.S., which we assume to be the treatment group. The results of this regression demonstrate that, over time, children from agricultural households with a current U.S.-based migrant member and receiving PROCAMPO cash transfers show improved height-for-age z-scores by 0.986 and 1.152 z-scores, compared to those in non-migrant households or those not receiving PROCAMPO cash transfers. These results were statistically significant at the 1% level. Nevertheless, the triple difference regression results for WHZ or BMIZ scores for the two treatment groups presented in columns (3) and (4) yielded both negative and positive coefficients but were not statistically significant.

Table 4.4 offers an analysis involving an interaction term between PROGRESA recipients and household migration status. The data suggest that children from PROGRESA beneficiary and migrant households exhibited improved HAZ scores of 0.106 and 0.2756 z scores compared to children from non-migrant households and non-PROGRESA beneficiary households, irrespective of the current migrant status. Additionally, column (2) takes into account households with current U.S. migrant members as the treatment group, indicating that children from households receiving PROGRESA cash transfers and with migrant members saw improvements of 0.062 and 0.191 z scores compared to non-migrant and non-beneficiary households. However, these differences were not statistically significant. Columns (3) and (4) of Table 4.4 present the triple differences estimation for WHZ and BMIZ scores of children, but no statistically significant outcomes were observed.

Robustness Check

Propensity Score Matching with Difference-in-Differences (PSM-DD)

The conventional approach to Difference-in-Differences analysis assumes that, in absence of the intervention, the average outcomes for both the treated and control groups would have followed parallel paths over time. However, this assumption may not hold if the decision to migrate is correlated with pre-period characteristics that are themselves correlated with the dynamics of the outcome variable, which in our study is children's nutritional status, and such correlation may lead to biased estimate if they are not addressed in the analysis (Abadie, 2005).

In non-experimental studies, the propensity score methods are commonly used to reduce selection bias by improving baseline balance between treatment and control groups. In the context of difference-in-differences analysis, the propensity score is employed as the probability of receiving the treatment, which in this study is international migration from Mexico to the U.S. This involves the use of a matching method to ensure that treatment and control groups are as similar as possible in terms of observed baseline characteristics (Rosenbaum and Rubin, 1983;

Stuart, 2010; Stuart et al., 2014). Specifically, we identify a matched subset from the treatment and control groups based on measures of baseline characteristics.

Table 4.5 presents a comparative evaluation of individual, household, and regional level characteristics between treatment and control groups. It illustrates these characteristics for the entire sample before the application of propensity score matching (first column), and the refined sample post-propensity score matching (second column). Figure 4.1, meanwhile, visually demonstrates the propensity scores of observations, both before and after the application of the propensity score matching technique. As mentioned above, the purpose of using propensity score matching was to ensure the selection of treatment and control groups that exhibit similar baseline characteristics, propensity score. This process is clearly depicted in Figure 4.1, where the propensity score shows a significant alignment in characteristics across both groups following the matching process. By employing this methodology, we were able to construct a robust sample of 386 individual-period observations. This sample, in turn, was balanced in its representation and allowed for the effective migration of potential confounding variables.

PSM-DD Results

Table 4.6 represents the results of a standard difference-in-differences estimation performed on a propensity-score matched sample. Although the long-term and short-term nutritional indicators for children from migrant households appeared to increase relative to those from non-migrant households (as evidenced by columns (1) to (4)), these results lacked statistical significance.

Table 4.7 and 4.8 display the outcomes of a triple difference estimation, employing interaction terms of migration and cash transfer beneficiary status among households, applied to the matched sample. These tables reveal findings that are statistically significant and align with the triple difference estimation conducted on the sample in our primary specification results.

In Table 4.7, Columns (1) and (2) report a statistically significant interaction term for PROCAMPO beneficiary households, migrant households, and time, showing a coefficient of 1.072 and 1.045, respectively significant at 1% level. This suggests that over time, height-for-age z-scores for children from PROCAMPO beneficiary and migrant households have improved 1.551 and 1.514 z-scores in Column (1), and 1.366 and 1.378 z-scores in Column (2) compared to those from non-migrant and non-cash transfer beneficiary households. However, when considering weight-for-height z-scores (WHZ) and BMI-for-age z-scores (BMIZ) as dependent variable (Column (3) and (4)), these results do not attain statistical significance.

Table 4.8 focuses on the triple difference estimation for height-for-age, adding an interaction term for PROGRESA beneficiary status, migrant households, and time. In Column (1), which considers migrant households regardless of the current location of residence, we observe statistically significant results at the 10% level. The interaction term for cash transfer beneficiaries, migrant households, and time was 0.740, which implies that over time, the height-for-age z-scores of children from PROGRESA beneficiary and migrant households improved by 1.15 and 0.976 z-scores compared to children from non-migrant and non-cash transfer beneficiary households. However, from Column (2) onwards, when considering the treatment group with only current migrant household members or when the dependent variable is either WHZ or BMIZ scores of children from agricultural households (Column (3) and (4)), the results from the triple difference regression analysis with matched samples based on baseline characteristics did not achieve statistical significance.

Conclusion and Discussion

In conclusion, this research conducts a comprehensive exploration of the influence of household migration status and enrollment in cash transfer programs (PROCAMPO and PROGRESA) on children's nutritional status in rural Mexico, focusing specifically on the agricultural labor channel. Methodologically, we utilized both standard difference-in-differences and triple-differences models to account for potential endogeneity arising from the non-random selection of migrant households for the treatment group. To further enhance our analysis, we applied a propensity score matching technique to create a balanced sample based on baseline individual, household, and regional characteristics, and employed difference-in-differences and triple differences estimations on this matched sample as a robustness check.

Our primary models use children's nutritional status indicators (height-for-age z-scores (HAZ), weight-for-height z-scores (WHZ), or BMI-for-age z-scores (BMIZ) by age) as dependent variables, with the migration status of households as the independent variable. Our results suggest a positive correlation between household migration in agricultural areas and children's nutritional outcomes. However, these associations did not achieve statistical significance.

Incorporating triple difference estimations allowed us to control for the heterogenous impacts of migration on children's health arising from observable and unobservable confounders linked to cash transfer eligibility. This provided insights into potential interaction effects between migration status and cash transfer participation. Our findings indicate that the long-term nutritional indicator, HAZ, showed consistent improvement over time for children in migrant households benefiting from cash transfers, compared to those in non-migrant households. However, only PROCAMPO showed statistically significant results. For short-term nutritional indicators, WHZ and BMIZ, the results were mixed and statistically insignificant.

The parallel trends assumption, important for implementing a difference-in-differences identification strategy to explore causal links, suggests that absent the treatment – in our case, migration of household members – the average outcomes (health status, in this case) for both the treatment and control groups would exhibit parallel trends over time. In the context of our study, the treatment group consists of children from households that experienced migration to the U.S. between the first and second waves, whereas the control group consists of households with no U.S. migration within the same period. Given the intricate and long-standing history of agricultural migration from Mexico to the U.S., defining a cutoff for treatment using the survey data was challenging. Testing for parallel trends requires several years of panel data. This, in effect, made it difficult to definitively establish the proven parallel trends in the difference-in-differences study.

To address this constraint, in the robustness check section, we aligned our sample based on characteristics prior to the treatment period. We opted to employ propensity score matching, which allowed us to make a more balanced comparison between our study groups. This strategy effectively reduces potential bias to confounding variables and strengthens overall findings. The outcomes aligned consistently with the main specification, showing a positive and significant association between migration and long-term nutritional status when an interaction term encompassing PROCAMPO beneficiary status, household migration status, and time was introduced in the triple difference specification.

The present study provides notable contributions to both the academic discourse and policy design. By focusing on children from agricultural households, it addresses es a significant gap in the literature of international migration on household welfare, specifically children's nutritional status. This focus is particularly relevant considering that agricultural labor is a major source of migration and remittances in developing countries, a situation which can produce contrasting impact on family well-being due to potential adverse effect due to labor loss.

The finding highlights the interconnectedness of economic factors like migration and cash transfer programs on children's health outcomes in rural communities. This underscores the need for more integrated approaches in economic and health research to unravel the complex causal pathways that influence child nutrition. The study also calls for greater consideration of context-specific factors, such as agricultural productivity, in understanding the effects of economic policies on health. For policy design, the findings suggest the potential benefits of coupling cash transfer programs with strategies that enhance household income and food security. It also underscores the need for more targeted interventions that take into account the specific circumstances of migrant households in rural areas. Future research could build on these findings by exploring other socio-economic factors that could affect children's nutritional status and by conducting more rigorous evaluation of cash transfer programs in different contexts.

<Table 4.1. Summary Statistics of Agricultural Households: A Comparison between Migrant Households and Non-Migrant Households *(N=1610)* >

	(1) Whole Sample	(2) Migration HH	(3) Non-Migration HH	(4) Diff.
Health Status				
HAZ	-0.9725	-0.8200	-1.0016	-0.1816**
BMI	0.3970	0.3886	0.3986	0.0101
Individual Characteristics				
Female	0.4994	0.5194	0.4956	-0.0238
Age Under 2	0.0335	0.0078	0.0385	0.0307**
Age b/w 2&5	0.2335	0.2481	0.2308	-0.0173
Age b/w 5&10	0.5404	0.5194	0.5444	0.0250
Household Characteristics				
Household Edu	0.0882	0.0388	0.0976	0.0589***
% Under 15	0.5282	0.5539	0.5233	-0.0306***
Own Telephone	0.1963	0.2558	0.1849	-0.0709***
Own wash machine	0.7491	0.8760	0.7249	-0.1511***
Own Toilet	0.4634	0.6202	0.4334	-0.1867***
Water source	0.2733	0.2636	0.2751	0.0116
Wal material	0.6261	0.6047	0.6302	0.0255
Room per Capita	0.3583	0.3523	0.3594	0.0071
Regional Characteristics				
Regional Safety	0.4373	0.2713	0.4689	0.1976***
pop1	0.0435	0.0233	0.0473	0.0241*
pop2	0.0236	0.0155	0.0251	0.0096
pop3	0.1019	0.0620	0.1095	0.0475**
pop4	0.8311	0.8992	0.8180	-0.0812***
Cash Transfers				
PROGRESA	0.4236	0.3953	0.4290	0.0336
PROCAMPO	0.3317	0.4806	0.3033	-0.1774***

< Table 4.2. Difference-in-Differences Analysis: Impact of US Migration on Children's Nutritional Status among Agricultural Households *(N=1610)* >

	(1) HAZ	(2) HAZ (Current Absent)	(3) BMI	(4) BMI (Current Absent)
Migration ***** Year	0.0249	0.0847	0.134	0.164
	(0.0926)	(0.101)	(0.111)	(0.121)
Migration	0.000785	-0.0125	-0.129	-0.134
	(0.110)	(0.119)	(0.116)	(0.126)
Year	0.178^{***}	0.171^{***}	0.00723	0.00746
	(0.0371)	(0.0364)	(0.0445)	(0.0436)
Constant	-1.129^{**}	-1.123^{**}	0.513	0.510
	(0.413)	(0.413)	(0.425)	(0.425)
adj. R^2	0.321	0.321	0.0693	0.0692
Control	YES	YES	YES	YES

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

< Table 4.3. Triple Differences Analysis: Impact of US Migration on Children's Nutritional Status among PROCAMPO and NON-PROCAMPO Households *(N=1610)* >

	(1) HAZ	(2) HAZ (Current Absent)	(3) BMI	(4) BMI (Current Absent)
PROCAMPO * Migration * Year	0.729^{***}	0.855^{***}	-0.128	0.0241
	(0.187)	(0.203)	(0.226)	(0.246)
PROCAMPO * Migration	-0.623^{**}	-0.695^{**}	0.0945	0.145
	(0.217)	(0.239)	(0.231)	(0.255)
PROCAMPO * Year	-0.135	-0.131	0.0298	0.00375
	(0.0800)	(0.0779)	(0.0968)	(0.0944)
Migration * Year	-0.301^{*}	-0.297^{*}	0.190	0.152
	(0.125)	(0.136)	(0.151)	(0.165)
PROCAMPO	0.00897	-0.00287	-0.136	-0.141
	(0.0961)	(0.0936)	(0.102)	(0.0996)
Migration	0.290^{*}	0.307	-0.159	-0.191
	(0.146)	(0.161)	(0.156)	(0.171)
Year	0.219^{***}	0.212^{***}	-0.00180	0.00629
	(0.0441)	(0.0434)	(0.0533)	(0.0526)
Constant	-1.113^{**}	-1.096^{**}	0.531	0.509
	(0.413)	(0.413)	(0.426)	(0.426)
adj. R^2	0.327	0.327	0.0711	0.0715
Control	YES	YES	YES	YES

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

< Table 4.4. Triple Differences Analysis: Impact of US Migration on Children's Nutritional Status among PROGRESA and NON-PROGRESA Households *(N=1610)* >

	(1) HAZ	(2) HAZ (Current Absent)	(3) BMI	(4) BMI (Current Absent)
PROGRESA * Migration * Year	0.218	0.181	0.0819	-0.0195
	(0.189)	(0.203)	(0.226)	(0.243)
PROGRESA * Migration	0.239	0.288	0.244	0.304
	(0.213)	(0.228)	(0.227)	(0.242)
PROGRESA * Year	0.112	0.119	-0.216*	-0.208*
	(0.0748)	(0.0737)	(0.0896)	(0.0883)
Migration * Year	-0.0576	-0.00990	0.0943	0.187
	(0.119)	(0.139)	(0.143)	(0.167)
PROGRESA	-0.182	-0.183	0.172	0.177
	(0.0948)	(0.0938)	(0.100)	(0.0991)
Migration	-0.0959	-0.137	-0.227	-0.298
	(0.139)	(0.160)	(0.148)	(0.171)
Year	0.130**	0.122*	0.0997	0.0937
	(0.0490)	(0.0475)	(0.0587)	(0.0569)
Constant	-1.049*	-1.052*	0.512	0.502
	(0.414)	(0.413)	(0.426)	(0.426)
adj. R^2	0.325	0.325	0.0747	0.0748
Control	YES	YES	YES	YES

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

<Table 4.5. Differences in Characteristics between Migrant and Non-Migrant Households: Before and After Propensity Matching>

	(1) Before Matching *Diff.*	(2) After Matching *Diff.*
Household Characteristics		
Household Head Education	0.0589***	0.0081
Female Household Head	-0.1290***	-0.0451
Own Telephone	-0.0709***	-0.0214
Source of water to drink	0.0116	0.0333
Own wash machine	-0.1511***	-0.0635*
Do you have toilet?	-0.1867***	-0.0420
Number of Rooms per capita	0.0071	-0.0135
Regional Characteristics		
Area with high crime	0.1976***	0.0881*
Rural area (Population below 2,500)	-0.0812***	-0.0086
States Indicators	-1.2193***	-0.8266*
Cash Transfers		
PROGRESA beneficiaries HH	0.0336	0.0265
PROCAMPO beneficiaries HH	-0.1774***	-0.0119
N	**1610**	**386**

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

<Table 4.6. Difference-in-Differences Analysis: Impact of US Migration on Children's Nutritional Status among Matched Observations of Agricultural Households *(N=386)* >

	(1) HAZ	(2) HAZ (Current Absent)	(3) BMI	(4) BMI (Current Absent)
Migration ***** Year	0.0672	0.163	0.229	0.231
	(0.154)	(0.145)	(0.177)	(0.167)
Migration	0.00444	-0.0229	-0.320	-0.302
	(0.188)	(0.180)	(0.176)	(0.168)
Year	0.136	0.0930	-0.0883	-0.0597
	(0.126)	(0.107)	(0.145)	(0.123)
Constant	-2.161	-2.123	-2.057	-2.105
	(1.217)	(1.218)	(1.075)	(1.076)
adj. R^2	0.338	0.339	0.196	0.196
Control	YES	YES	YES	YES

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

< Table 4.7. Triple Differences Analysis: Impact of US Migration on Children's Nutritional Status among Matched Observations of PROCAMPO and NON-PROCAMPO Households *(N=386)*>

	(1) HAZ	(2) HAZ (Current Absent)	(3) BMI	(4) BMI (Current Absent)
PROCAMPO * Migration * Year	1.072***	1.045***	-0.426	-0.0440
	(0.298)	(0.280)	(0.355)	(0.337)
PROCAMPO * Migration	-0.801*	-0.772*	0.282	-0.0200
	(0.373)	(0.360)	(0.353)	(0.340)
PROCAMPO * Year	-0.479*	-0.321	0.328	0.0718
	(0.244)	(0.205)	(0.291)	(0.247)
Migration * Year	-0.442*	-0.333	0.430	0.253
	(0.205)	(0.193)	(0.244)	(0.233)
PROCAMPO	0.0000716	-0.131	-0.463	-0.278
	(0.301)	(0.251)	(0.285)	(0.238)
Migration	0.364	0.324	-0.459	-0.306
	(0.250)	(0.247)	(0.237)	(0.233)
Year	0.360*	0.248	-0.242	-0.0943
	(0.167)	(0.143)	(0.199)	(0.172)
Constant	-2.512*	-2.395*	-2.156*	-2.250*
	(1.210)	(1.215)	(1.080)	(1.084)
adj. R^2	0.368	0.369	0.207	0.205
Control	YES	YES	YES	YES

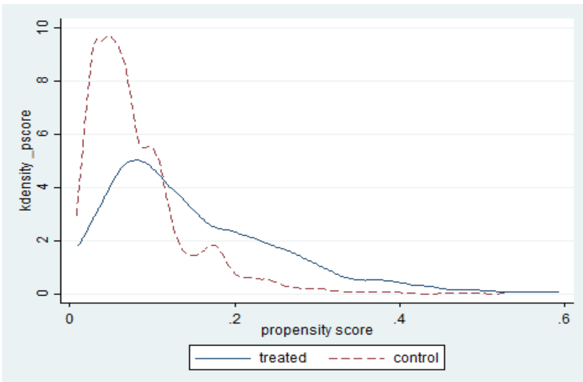
Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

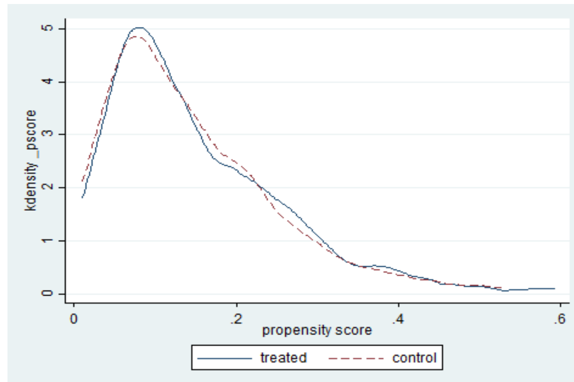
< Table 4.8. Triple Differences Analysis: Impact of US Migration on Children's Nutritional Status among Matched Observations of PROGRESA and NON-PROGRESA Households *(N=386)*>

	(1) HAZ	(2) HAZ (Current Absent)	(3) BMI	(4) BMI (Current Absent)
PROGRESA * Migration * Year	0.740*	0.580	0.0468	-0.0637
	(0.310)	(0.301)	(0.361)	(0.349)
PROGRESA * Migration	-0.0469	0.0564	-0.141	-0.144
	(0.375)	(0.362)	(0.352)	(0.339)
PROGRESA * Year	-0.410	-0.281	-0.180	-0.164
	(0.253)	(0.228)	(0.294)	(0.265)
Migration * Year	-0.236	-0.0695	0.206	0.289
	(0.199)	(0.187)	(0.232)	(0.217)
PROGRESA	0.185	0.128	0.317	0.373
	(0.313)	(0.284)	(0.293)	(0.266)
Migration	0.0305	-0.0691	-0.262	-0.296
	(0.239)	(0.228)	(0.225)	(0.214)
Year	0.309	0.181	-0.0122	-0.00820
	(0.164)	(0.128)	(0.191)	(0.149)
Constant	-2.099	-2.062	-2.036	-2.058
	(1.222)	(1.220)	(1.083)	(1.079)
adj. R^2	0.350	0.349	0.201	0.204
Control	YES	YES	YES	YES

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

<Figure 4.1. Distribution of Propensity Score: Before and After Matching >

<table>
<tr><td>Before Matching</td><td>After Matching</td></tr>
<tr><td>


</td><td>


</td></tr>
</table>

CHAPTER 5

CONCLUSION

In conclusion, this dissertation conducts an extensive analysis of the socio-economic consequences of migration within developing nations, focusing particularly on Mongolia and Mexico. The first chapter navigates the intricate link between informal settlements and air pollution within Ulaanbaatar, Mongolia's capital city. Our study, utilizing pioneering deep learning techniques and satellite imagery, provides an innovative viewpoint on how the population of informal settlements affects air pollution levels. A significant and positive correlation was found between the number of '*gers*' and several air pollutant levels, underlining the environmental implications of urbanization and migration in developing nations. The second chapter explores the nuanced relationship between internal migration and educational investment among Mongolia's urban native youth. Through the lens of secondary education completion rates among young urban individuals, our research offers crucial insights into the crowd-in and crowd-out effects of heightened immigration on education and labor markets. The findings suggest that increased immigration rates generally foster human capital development, resulting in a higher propensity for native youth to pursue higher educational attainment. The third and concluding chapter delves into the impact of household migration and enrollment in cash transfer programs on children's nutritional status in rural Mexico. Our research demonstrates that while migration in agricultural households typically has a positive correlation with children's nutritional outcomes, it only holds statistical significance when intertwined with cash transfer programs. The results emphasize the complicated interaction of migration, economic policies, and health outcomes, underscoring the need for more integrated and context-specific policy interventions. Overall, this dissertation highlights the multifaceted environmental and socio-economic impacts of migration and demographic transition in developing countries. The research delivers a comprehensive exploration of the multidimensional aspects of migration impact, offering valuable insights for future research and policy formulation. The innovative use of unconventional data sets and computational methodologies lays a strong foundation for further exploration of this intricate topic, catalyzing future research that can continue to broaden the boundaries of demographic studies. We anticipate that the methodologies and insights presented in this dissertation will lay a firm foundation for future studies, deepening our understanding of the complex phenomena of migration in the developing world.

REFERENCES

Abadie, A. (2005). Semiparametric difference-in-differences estimators. *The Review of Economic Studies*, *72*(1), 1-19.

Acosta, P., Fajnzylber, P., & Lopez, J. H. (2007). *The impact of remittances on poverty and human capital: evidence from Latin American household surveys* (Vol. 4247). World Bank Publications.

ADB. (2018). Mongolia: Urban Sector Fact Sheet. *Asian Development Bank*. https://www.adb.org/publications/mongolia-urban-sector-fact-sheet

Aliero, M. M., Mukhtar, N., & Al-Doksi, J. (2014). The usefulness of unmanned airborne vehicle (UAV) imagery for automated palm oil tree counting. *Journal of Forestry. Researchjournali*, *1*(1).

Alola, A. A., & Kirikkaleli, D. (2019). The nexus of environmental quality with renewable consumption, immigration, and healthcare in the US: wavelet and gradual-shift causality approaches. *Environmental Science and Pollution Research*, *26*(34), 35208-35217.

Altonji, J. G., & Card, D. (1991). The effects of immigration on the labor market outcomes of less-skilled natives. In *Immigration, trade, and the labor market* (pp. 201-234). University of Chicago Press.

Alves, D., & Belluzzo, W. (2004). Infant mortality and child health in Brazil. *Economics & Human Biology*, *2*(3), 391-410.

Angelucci, M. (2012). Conditional cash transfer programs, credit constraints, and migration. *Labour*, *26*(1), 124-136.

Angrist, J. D., & Kugler, A. D. (2003). Protective or counter-productive? Labour market institutions and the effect of immigration on EU natives. *The Economic Journal*, *113*(488), F302-F331.

Antén, J.-I. (2010). The impact of remittances on nutritional status of children in Ecuador. *International migration review*, *44*(2), 269-299.

Assaad, R., Ginn, T., & Saleh, M. (2018). Impact of Syrian refugees in Jordan on education outcomes for Jordanian youth. Economic Research Forum Working Paper Series No,

Attanasio, O., Gomez, L. C., Rojas, A. G., & Vera-Hernandez, M. (2004). Child health in rural Colombia: determinants and policy interventions. *Economics & Human Biology*, *2*(3), 411-438.

Azzarri, C., & Zezza, A. (2011). International migration and nutritional outcomes in Tajikistan. *Food Policy*, *36*(1), 54-70.

Bartlett, A. A., & Lytwak, E. P. (1995). Zero growth of the population of the United States. *Population and Environment*, *16*(5), 415-428.

Basu, A., Coe, N., & Chapman, C. G. (2017). *Comparing 2SLS VS 2SRI for binary outcomes and binary exposures*.

Bauner, S., & Richter, B. (2006). Real estate market, mortgage market and cadastre in Ulaanbaatar and Darkhan-City, Mongolia. In: Report by Deutsche Gesellshaft für Technische Zusammenarbeit (GTZ). Eschborn ….

Betts, J. (1998). Educational crowding out: do immigrants affect the educational attainment of American minorities?

Betts, J. R., & Fairlie, R. W. (2003). Does immigration induce ‘native flight’from public schools into private schools? *Journal of Public Economics*, *87*(5-6), 987-1012.

Betts, J. R., & Lofstrom, M. (2000). The educational attainment of immigrants: trends and implications. In *Issues in the Economics of Immigration* (pp. 51-116). University of Chicago Press.

Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The quarterly journal of economics*, *118*(4), 1335-1374.

Borjas, G. J. (2004). Do foreign students crowd out native students from graduate programs? In: National Bureau of Economic Research Cambridge, Mass., USA.

Borjas, G. J. (2017). The Wage Impact of the Marielitos: A Reappraisal. *Ilr Review*, *70*(5), 1077-1110. https://doi.org/10.1177/0019793917692945

Boustan, L. P., Fishback, P. V., & Kantor, S. (2010). The effect of internal migration on local labor markets: American cities during the Great Depression. *Journal of Labor Economics*, *28*(4), 719-746.

Briggs, D. J., de Hoogh, C., Gulliver, J., Wills, J., Elliott, P., Kingham, S., & Smallbone, K. (2000). A regression-based method for mapping traffic-related air pollution: application and testing in four contrasting urban environments. *Science of the Total Environment*, *253*(1-3), 151-167.

Brunello, G., Lodigiani, E., & Rocco, L. (2020). Does low skilled immigration increase the education of natives? Evidence from Italian provinces. *Labour economics*, *63*, 101794.

Brunello, G., & Rocco, L. (2013). The effect of immigration on the school performance of natives: Cross country evidence using PISA test scores. *Economics of education review*, *32*, 234-246.

Bryant, J., & Rukumnuaykit, P. (2013). The labour market impacts of immigration to developing countries: evidence from a registration campaign in Thailand. *The Journal of Development Studies*, *49*(6), 785-800.

Buchholz, S., Junk, J., Krein, A., Heinemann, G., & Hoffmann, L. (2010). Air pollution characteristics associated with mesoscale atmospheric patterns in northwest continental Europe. *Atmospheric Environment*, *44*(39), 5183-5190. https://doi.org/https://doi.org/10.1016/j.atmosenv.2010.08.053

Burgess, S., Small, D. S., & Thompson, S. G. (2017). A review of instrumental variable estimators for Mendelian randomization. *Statistical methods in medical research*, *26*(5), 2333-2355.

Burke, M., Driscoll, A., Lobell, D. B., & Ermon, S. (2021). Using satellite imagery to understand and promote sustainable development. *Science*, *371*(6535), eabe8628.

Byambadorj, T., Amati, M., & Ruming, K. J. (2011). Twenty-first century nomadic city: Ger districts and barriers to the implementation of the Ulaanbaatar City Master Plan. *Asia Pacific Viewpoint*, *52*(2), 165-177.

Calderón-Mejía, V., & Ibáñez, A. M. (2016). Labour market effects of migration-related supply shocks: evidence from internal refugees in Colombia. *Journal of Economic Geography*, *16*(3), 695-713.

Card, D. (1990). The impact of the Mariel boatlift on the Miami labor market. *Ilr Review*, *43*(2), 245-257.

Card, D. (2005). Is the new immigration really so bad? *The Economic Journal*, *115*(507), F300-F323.

Carletto, C., Covarrubias, K., & Maluccio, J. A. (2011). Migration and child growth in rural Guatemala. *Food Policy*, *36*(1), 16-27.

Carrington, W. J., & De Lima, P. J. (1996). The impact of 1970s repatriates from Africa on the Portuguese labor market. *Ilr Review*, *49*(2), 330-347.

Chatty, D., & Sternberg, T. (2015). Climate effects on nomadic pastoralist societies. *Forced Migration Review*(49), 25.

Chen, J. J. (2006). Migration and imperfect monitoring: Implications for intra-household allocation. *American Economic Review*, *96*(2), 227-231.

Chen, Z. H., Cheng, S. Y., Li, J. B., Guo, X. R., Wang, W. H., & Chen, D. S. (2008). Relationship between atmospheric pollution processes and synoptic pressure patterns in northern China. *Atmospheric Environment*, *42*(24), 6078-6087. https://doi.org/https://doi.org/10.1016/j.atmosenv.2008.03.043

Cole, M. A., Elliott, R. J., & Liu, B. (2020). The impact of the Wuhan Covid-19 lockdown on air pollution and health: a machine learning and augmented synthetic control approach. *Environmental and Resource Economics*, *76*(4), 553-580.

Contreras, D., & Gallardo, S. (2022). The effects of mass migration on the academic performance of native students. Evidence from Chile. *Economics of education review*, *91*, 102314.

Cox-Edwards, A., & Rodríguez-Oreggia, E. (2009). Remittances and labor force participation in Mexico: an analysis using propensity score matching. *World development*, *37*(5), 1004-1014.

Cui, D., Wu, D., Liu, J., Xiao, Y., Yembuu, B., & Adiya, Z. (2019). Understanding urbanization and its impact on the livelihood levels of urban residents in Ulaanbaatar, Mongolia. *Growth and Change*, *50*(2), 745-774.

Daily, G. C., Ehrlich, A. H., & Ehrlich, P. R. (1995). Response to Bartlett and Lytwak (1995): Population and immigration policy in the United States. *Population and Environment*, 521-526.

De Brauw, A. (2011). Migration and child development during the food price crisis in El Salvador. *Food Policy*, *36*(1), 28-40.

De Brauw, A., & Mu, R. (2011). Migration and the overweight and underweight status of children in rural China. *Food Policy, 36*(1), 88-100.

De Onis, M., Garza, C., Onyango, A. W., & Borghi, E. (2007). Comparison of the WHO child growth standards and the CDC 2000 growth charts. *The Journal of nutrition*, *137*(1), 144-148.

Dedeoğlu, M., Koçak, E., & Uucak, Z. Ş. (2021). The impact of immigration on human capital and carbon dioxide emissions in the USA: an empirical investigation. *Air Quality, Atmosphere & Health*, *14*(5), 705-714.

Dickinson-Craig, E., Bartington, S. E., Watts, R., Mandakhbayar, O., Khurelbaatar, E.-O., Ochir, C., Boldbaatar, D., Warburton, D., Thomas, G. N., & Pope, F. D. (2022). Carbon monoxide levels in households using coal-briquette fuelled stoves exceed WHO air quality guidelines in Ulaanbaatar, Mongolia. *International journal of environmental health research*, 1-12.

Dietz, T., & Rosa, E. A. (1997). Effects of population and affluence on CO2 emissions. *Proceedings of the National Academy of Sciences*, *94*(1), 175-179.

Dore, G., & Nagpal, T. (2006). Urban transition in Mongolia: pursuing sustainability in a unique environment. *Environment: Science and Policy for Sustainable Development*, *48*(6), 10-24.

Duflo, E. (2003). Grandmothers and granddaughters: old-age pensions and intrahousehold allocation in South Africa. *The World Bank Economic Review*, *17*(1), 1-25.

Durand, J., Massey, D. S., & Parrado, E. A. (1999). The new era of Mexican migration to the United States. *The Journal of American History*, *86*(2), 518-536.

Durand, J., Massey, D. S., & Zenteno, R. M. (2001). Mexican immigration to the United States: Continuities and changes. *Latin American research review*, 107-127.

Dustmann, C., Glitz, A., & Frattini, T. (2008). The labour market impact of immigration. *Oxford Review of Economic Policy*, *24*(3), 477-494.

Dustmann, C., Schönberg, U., & Stuhler, J. (2016). The impact of immigration: Why do studies reach such different results? *Journal of Economic Perspectives*, *30*(4), 31-56.

Edo, A. (2019). The impact of immigration on the labor market. *Journal of Economic Surveys*, *33*(3), 922-948.

El Badaoui, E., Strobl, E., & Walsh, F. (2017). Impact of internal migration on labor market outcomes of native males in Thailand. *Economic Development and Cultural Change*, *66*(1), 147-177.

Enkh-Undraa, D., Kanda, S., Shima, M., Shimono, T., Miyake, M., Yoda, Y., Nagnii, S., & Nishiyama, T. (2019). Coal burning-derived SO2 and traffic-derived NO2 are associated with persistent cough and current wheezing symptoms among schoolchildren in Ulaanbaatar, Mongolia. *Environmental health and preventive medicine*, *24*(1), 1-11.

Environmental Migration. (2021). MIGRATION DATA PORTAL. https://www.migrationdataportal.org/themes/environmental_migration_and_statistics

Farre, L., Ortega, F., & Tanaka, R. (2018). Immigration and the public–private school choice. *Labour economics*, *51*, 184-201.

Fekade, W. (2000). Deficits of formal urban land management and informal responses under rapid urban growth, an international perspective. *Habitat international*, *24*(2), 127-150.

Friedberg, R. M. (2001). The impact of mass migration on the Israeli labor market. *The quarterly journal of economics*, *116*(4), 1373-1408.

Friedberg, R. M., & Hunt, J. (2018). The impact of immigrants on host country wages, employment and growth. In *Interdisciplinary Perspectives on the New Immigration* (pp. 89-110). Routledge.

Fussell, E., & Massey, D. S. (2004). The limits to cumulative causation: International migration from Mexican urban areas. *Demography*, *41*(1), 151-171.

Gibson, J., McKenzie, D., & Stillman, S. (2011). What happens to diet and child health when migration splits households? Evidence from a migration lottery program. *Food Policy*, *36*(1), 7-15.

Gould, E. D., Lavy, V., & Daniele Paserman, M. (2009). Does immigration affect the long-term educational outcomes of natives? Quasi-experimental evidence. *The Economic Journal*, *119*(540), 1243-1269.

GSG. (2022). *INFORMAL SETTLEMENTS: NO LONGER INVISIBLE*.

Hanson, G. H. (2009). The economic consequences of the international migration of labor. *Annu. Rev. Econ.*, *1*(1), 179-208.

Hildebrandt, N., McKenzie, D. J., Esquivel, G., & Schargrodsky, E. (2005). The effects of migration on child health in Mexico [with comments]. *Economia*, *6*(1), 257-289.

Hoxby, C. M. (1998). Do immigrants crowd disadvantaged. *Help or hindrance*, 282.

Huang, Y. K., Luvsan, M. E., Gombojav, E., Ochir, C., Bulgan, J., & Chan, C. C. (2013). Land use patterns and SO2 and NO2 pollution in Ulaanbaatar, Mongolia. *Environmental Research*, *124*, 1-6. https://doi.org/10.1016/j.envres.2013.02.006

Hunt, J. (1992). The impact of the 1962 repatriates from Algeria on the French labor market. *Ilr Review*, *45*(3), 556-572.

Hunt, J. (2017). The impact of immigration on the educational attainment of natives. *Journal of Human Resources*, *52*(4), 1060-1118.

Jackson, O. (2015). Does Immigration Crowd Natives into or out of Higher Education? Working Papers. No. 15-18. *Federal Reserve Bank of Boston*.

Jensen, P., & Rasmussen, A. W. (2011). The effect of immigrant concentration in schools on native and immigrant children's reading and math skills. *Economics of education review*, *30*(6), 1503-1515.

Jun, S. (2021). Is the Raw Coal Ban a Silver Bullet to Solving Air Pollution in Mongolia?: A Study of the Mongolian Government's Air Pollution Reduction Policies and Recommendations in the Context of COVID-19. *Journal of Public & International Affairs*.

Kanaiaupuni, S. M., & Donato, K. M. (1999). Migradollars and mortality: The effects of migration on infant survival in Mexico. *Demography*, *36*(3), 339-353.

Kattenborn, T., Leitloff, J., Schiefer, F., & Hinz, S. (2021). Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, *173*, 24-49.

Kleemans, M., & Magruder, J. (2018). Labour market responses to immigration: Evidence from internal migration driven by weather shocks. *The Economic Journal*, *128*(613), 2032-2065.

Kniesner, T. J. (1976). An indirect test of complementarity in a family labor supply model. *Econometrica: Journal of the Econometric Society*, 651-669.

Komatsu, S., Ha, H. D., & Kaneko, S. (2013). The effects of internal migration on residential energy consumption and CO2 emissions: A case study in Hanoi. *Energy for Sustainable Development*, *17*(6), 572-580.

Kugler, A., & Kugler, M. (2009). Labor market effects of payroll taxes in developing countries: Evidence from Colombia. *Economic Development and Cultural Change*, *57*(2), 335-358.

Kuo, L. (2018). Pollution pushes Mongolia's herders to reconsider city life. *The Guardian*. https://www.theguardian.com/world/2018/sep/24/pollution-pushes-mongolias-herders-to-reconsider-city-life

LaLonde, R. J., & Topel, R. H. (1991). Labor market adjustments to increased immigration. In *Immigration, trade, and the labor market* (pp. 167-199). University of Chicago Press.

Langworthy, B. (2011). The effects of parental migration on child nutrition.

Lazear, E. P. (2001). Educational production. *The quarterly journal of economics*, *116*(3), 777-803.

Leeds, M. A., & Allmen, P. v. (2004). Spousal complementarity in home production. *American Journal of Economics and Sociology*, *63*(4), 795-811.

Lemos, S., & Portes, J. (2014). New labour? The effects of migration from Central and Eastern Europe on unemployment and wages in the UK. *The BE Journal of Economic Analysis & Policy*, *14*(1), 299-338.

Liu, W., Yamazaki, F., & Vu, T. T. (2010). Automated vehicle extraction and speed determination from QuickBird satellite images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *4*(1), 75-82.

Lokshin, M., & Glinskaya, E. (2009). The effect of male migration on employment patterns of women in Nepal. *The World Bank Economic Review*, *23*(3), 481-507.

Long, H., Li, J., & Liu, H. (2022). Internal migration and associated carbon emission changes: Evidence from cities in China. *Energy Economics*, *110*, 106010.

Longhi, S., Nijkamp, P., & Poot, J. (2005). A meta-analytic assessment of the effect of immigration on wages. *Journal of Economic Surveys*, *19*(3), 451-477.

López Córdova, J. E. (2018). *Globalization, Migration and Development: The Role of Mexican Migrant Remittances*.

López, O. S., Bartolomei, R. S., & Lamba-Nieves, D. (2019). Urban Informality: International Trends and Policies to Address Land Tenure and Informal Settlements. *Center for a New Economy (CNE)*.

Lu, Y. (2015). Internal migration, international migration, and physical growth of left-behind children: a study of two settings. *Health & place*, *36*, 118-126.

Lucas, R. E. (2015). Internal migration in developing economies: An overview. *KNOMAD's working paper*, *6*.

Massey, D. S., & Espinosa, K. E. (1997). What's driving Mexico-US migration? A theoretical, empirical, and policy analysis. *American journal of Sociology*, *102*(4), 939-999.

Massey, D. S., Goldring, L., & Durand, J. (1994). Continuities in transnational migration: An analysis of nineteen Mexican communities. *American journal of Sociology*, *99*(6), 1492-1533.

Mastrorillo, M., Licker, R., Bohra-Mishra, P., Fagiolo, G., Estes, L. D., & Oppenheimer, M. (2016). The influence of climate variability on internal migration flows in South Africa. *Global Environmental Change-Human and Policy Dimensions*, *39*, 155-169. https://doi.org/10.1016/j.gloenvcha.2016.04.014

Mayer, B. (2016). Climate migration and the politics of causal attribution: a case study in Mongolia. *Migration and Development*, *5*(2), 234-253.

McHenry, P. (2015). Immigration and the human capital of natives. *Journal of Human Resources*, *50*(1), 34-71.

McIntosh, M. F. (2008). Measuring the labor market impacts of Hurricane Katrina migration: Evidence from Houston, Texas. *American Economic Review*, *98*(2), 54-57.

McLeman, R., & Smit, B. (2006). Migration as an adaptation to climate change. *Climatic Change*, *76*(1-2), 31-53. https://doi.org/10.1007/s10584-005-9000-7

Meng, Z., Ding, G., Xu, X., Xu, X., Yu, H., & Wang, S. (2008). Vertical distributions of SO2 and NO2 in the lower atmosphere in Beijing urban areas, China. *Science of the Total Environment*, *390*(2-3), 456-465.

Miller, R. (2017). Settling between Legitimacy and the Law: At the Edge of Ulaanbaatar's Legal Landscape. *Traditional Dwellings and Settlements Review*, 7-20.

Morales, R., Aguilar, A. M., & Calzadilla, A. (2004). Geography and culture matter for malnutrition in Bolivia. *Economics & Human Biology*, *2*(3), 373-389.

Mu, R., & Van de Walle, D. (2011). Left behind to farm? Women's labor re-allocation in rural China. *Labour economics*, *18*, S83-S97.

Mubin, N. A., Nadarajoo, E., Shafri, H. Z. M., & Hamedianfar, A. (2019). Young and mature oil palm tree detection and counting using convolutional neural network deep learning method. *International Journal of Remote Sensing*, *40*(19), 7500-7515.

Munkhtsetseg, E., Kimura, R., Wang, J., & Shinoda, M. (2007). Pasture yield response to precipitation and high temperature in Mongolia. *Journal of Arid Environments*, *70*(1), 94-110.

Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the US labor market. *The quarterly journal of economics*, *118*(2), 549-599.

Murray, T. J. (2016). Public or private? The influence of immigration on native schooling choices in the United States. *Economics of education review*, *53*, 268-283.

Nguyen, M. C., & Winters, P. (2011). The impact of migration on food consumption patterns: The case of Vietnam. *Food Policy*, *36*(1), 71-87.

Nobles, J. (2011). Parenting from abroad: Migration, nonresident father involvement, and children's education in Mexico. *Journal of Marriage and Family*, *73*(4), 729-746.

Ottaviano, G. I., & Peri, G. (2012). Rethinking the effect of immigration on wages. *Journal of the European economic association*, *10*(1), 152-197.

Padilla-Romo, M., & Peluffo, C. (2023). Violence-induced migration and peer effects in academic performance. *Journal of Public Economics*, *217*, 104778.

Paganini, M., Petiteville, I., Ward, S., Dyke, G., Steventon, M., Harry, J., & Kerblat, F. (2018). Satellite earth observations in support of the sustainable development goals. *The CEOS Earth Observation Handbook*.

Pandey, R., Alatalo, J. M., Thapliyal, K., Chauhan, S., Archie, K. M., Gupta, A. K., Jha, S. K., & Kumar, M. (2018). Climate change vulnerability in urban slum communities: investigating household adaptation and decision-making capacity in the Indian Himalaya. *Ecological Indicators*, *90*, 379-391.

Park, H., Fan, P., John, R., Ouyang, Z., & Chen, J. (2019). Spatiotemporal changes of informal settlements: Ger districts in Ulaanbaatar, Mongolia. *Landscape and urban planning*, *191*, 103630.

Peri, G. (2016). Immigrants, productivity, and labor markets. *Journal of Economic Perspectives*, *30*(4), 3-30.

Persello, C., Wegner, J. D., Hansch, R., Tuia, D., Ghamisi, P., Koeva, M., & Camps-Valls, G. (2022). Deep Learning and Earth Observation to Support the Sustainable Development Goals: Current Approaches, Open Challenges, and Future Opportunities. *IEEE Geoscience and Remote Sensing Magazine*.

Phillips, J. A., & Massey, D. S. (2000). Engines of immigration: Stocks of human and social capital in Mexico. *Social Science Quarterly*, 33-48.

Pires, J., Sousa, S., Pereira, M., Alvim-Ferraz, M., & Martins, F. (2008a). Management of air quality monitoring using principal component and cluster analysis—Part I: SO2 and PM10. *Atmospheric Environment*, *42*(6), 1249-1260.

Pires, J., Sousa, S., Pereira, M., Alvim-Ferraz, M., & Martins, F. (2008b). Management of air quality monitoring using principal component and cluster analysis—Part II: CO, NO2 and O3. *Atmospheric Environment*, *42*(6), 1261-1274.

Pischke, J.-S., & Velling, J. (1997). Employment effects of immigration to Germany: an analysis based on local labor markets. *Review of Economics and Statistics*, *79*(4), 594-604.

Porio, E. (2011). Vulnerability, Adaptation, and Resilience to Floods and Climate Change-Related Risks among Marginal, Riverine Communities in Metro Manila. *Asian Journal of Social Science*, *39*(4), 425-445. https://doi.org/10.1163/156853111x597260

Price, C. E., & Feldmeyer, B. (2012). The Environmental Impact of Immigration: An Analysis of the Effects of Immigrant Concentration on Air Pollution Levels. *Population Research and Policy Review*, *31*(1), 119-140. https://doi.org/10.1007/s11113-011-9216-3

Qin, H., & Liao, T. F. (2016). The association between rural-urban migration flows and urban air quality in China. *Regional Environmental Change*, *16*(5), 1375-1387. https://doi.org/10.1007/s10113-015-0865-3

Rafiq, S., Nielsen, I., & Smyth, R. (2017). Effect of internal migration on the environment in China. *Energy Economics*, *64*, 31-44.

Rahman, A., Luo, C., Khan, M. H. R., Ke, J., Thilakanayaka, V., & Kumar, S. (2019). Influence of atmospheric PM2. 5, PM10, O3, CO, NO2, SO2, and meteorological factors on the concentration of airborne pollen in Guangzhou, China. *Atmospheric Environment*, *212*, 290-304.

Rubalcava, L., & Teruel, G. (2008). User's guide for the Mexican family life survey second wave. *Retrieved July*, *1*, 2010.

Ruiz-Arranz, M., Davis, B., Handa, S., Stampini, M., & Winters, P. (2006). Program conditionality and food security: the impact of Progresa and Procampo transfers in rural Mexico. *EconomiA, Brasılia (DF)*, *7*(2), 249-278.

Sadoulet, E., De Janvry, A., & Davis, B. (2001). Cash transfer programs with income multipliers: PROCAMPO in Mexico. *World development*, *29*(6), 1043-1056.

Sanbata, H., Asfaw, A., & Kumie, A. (2014). Indoor air pollution in slum neighbourhoods of Addis Ababa, Ethiopia. *Atmospheric Environment*, *89*, 230-234.

Sang, J., Wu, Z., Guo, P., Hu, H., Xiang, H., Zhang, Q., & Cai, B. (2018). An improved YOLOv2 for vehicle detection. *Sensors*, *18*(12), 4272.

Sarzin, Z. (2021). *The impact of forced migration on the labor market outcomes and welfare of host communities* 70th Anniversary of the 1951 Refugee Convention, Sattler, D. N., Bishkhorloo, B., & Graham, J. M. (2021). Climate change threatens nomadic herding in Mongolia: A model of climate change risk perception and behavioral adaptation. *Journal of Environmental Psychology*, *75*, 101620.

Schoeni, R. F. (1997). The Effect of Immigrants on the Employment and Wages of Native Workers: Evidence from the 1970s and 1980s. *Unpublished paper RAND, Santa Monica, CA*, *17*.

Scott-Andretta, J., & Cuecuecha, A. (2010). *The Effect of Agricultural Subsidies on Migration and Agricultural Employment*.

Selod, H., & Shilpi, F. (2021). Rural-urban migration in developing countries: Lessons from the literature. *Regional Science and Urban Economics*, *91*, 103713.

Smith, C. L. (2012). The impact of low-skilled immigration on the youth labor market. *Journal of Labor Economics*, *30*(1), 55-89.

Soyol-Erdene, T.-O., Ganbat, G., & Baldorj, B. (2021). Urban air quality studies in Mongolia: pollution characteristics and future research needs. *Aerosol and Air Quality Research*, *21*(12), 210163.

Squalli, J. (2010). An empirical assessment of US state-level immigration and environmental emissions. *Ecological Economics*, *69*(5), 1170-1175.

Srestasathiern, P., & Rakwatin, P. (2014). Oil palm tree detection with high resolution multi-spectral satellite imagery. *Remote Sensing*, *6*(10), 9749-9774.

Stecklov, G., Winters, P., Stampini, M., & Davis, B. (2005). Do conditional cash transfers influence migration? A study using experimental data from the Mexican PROGRESA program. *Demography*, *42*(4), 769-790.

Strobl, E., & Valfort, M.-A. (2015). The effect of weather-induced internal migration on local labor markets. Evidence from Uganda. *The World Bank Economic Review*, *29*(2), 385-412.

Subramanian, S., & Deaton, A. (1996). The demand for food and calories. *Journal of political economy*, *104*(1), 133-162.

Suen, W. (2000). Estimating the effects of immigration in one city. *Journal of population economics*, *13*(1), 99-112.

Tayara, H., Soo, K. G., & Chong, K. T. (2017). Vehicle detection and counting in high-resolution aerial images using convolutional regression neural network. *IEEE Access*, *6*, 2220-2230.

Terza, J. V., Basu, A., & Rathouz, P. J. (2008). Two-stage residual inclusion estimation: addressing endogeneity in health econometric modeling. *Journal of health economics*, *27*(3), 531-543.

Tumen, S. (2018). The impact of low-skill refugees on youth education. *Available at SSRN 3273708*.

Tumen, S. (2021). The effect of refugees on native adolescents' test scores: Quasi-experimental evidence from PISA. *Journal of Development Economics*, *150*, 102633.

UN-HABITAT. (2016). *World Cities Report 2016: Urbanization and Development - Emerging Futures*.

UN-HABITAT. (2018). *SDG Indicator 11.1.1 Training Module: Adequate Housing and Slum Upgrading.* .

UNEP. (2019). Air pollution hurts the poorest most.

UNSD. (2019). *Make cities and human settlements inclusive, safe, resilient and sustainable*. https://unstats.un.org/sdgs/report/2019/goal-11/

Valdivia, M. (2004). Poverty, health infrastructure and the nutrition of Peruvian children. *Economics & Human Biology*, *2*(3), 489-510.

Verme, P., & Schuettler, K. (2021). The impact of forced displacement on host communities: A review of the empirical literature in economics. *Journal of Development Economics*, *150*, 102606.

Wang, H., Cheng, Z., & Smyth, R. (2018). Do migrant students affect local students' academic achievements in urban China? *Economics of education review*, *63*, 64-77.

Wang, L., Li, M., Yu, S., Chen, X., Li, Z., Zhang, Y., Jiang, L., Xia, Y., Li, J., & Liu, W. (2020). Unexpected rise of ozone in urban and rural areas, and sulfur dioxide in rural areas during the coronavirus city lockdown in Hangzhou, China: implications for air quality. *Environmental Chemistry Letters*, *18*(5), 1713-1723.

West, S. E., Bowyer, C. J., Apondo, W., Büker, P., Cinderby, S., Gray, C. M., Hahn, M., Lambe, F., Loh, M., & Medcalf, A. (2021). Using a co-created transdisciplinary approach to explore the complexity of air pollution in informal settlements. *Humanities and Social Sciences Communications*, *8*(1), 1-13.

WHO. (2006). *WHO child growth standards: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: methods and development*. World Health Organization.

WHO. (2018). World Health Organization issues recommendations to tackle health impacts of air pollution in Mongolia.

WHO. (2021a). *Ambient (outdoor) air pollution*. https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health

WHO. (2021b). *WHO global air quality guidelines: particulate matter (PM2. 5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide*. World Health Organization.

Winters, P., De Janvry, A., & Sadoulet, E. (2001). Family and community networks in Mexico-US migration. *Journal of Human Resources*, 159-184.

Xu, Y., Zhang, Y., & Chen, J. (2021). Migration under economic transition and changing climate in Mongolia. *Journal of Arid Environments*, *185*, 104333.