2021

Assessing the Variations of Educational Attainment at National and Subnational Levels Using Hierarchical Linear Models

Bingxin Qi

Assessing the Variations of Educational Attainment at National and Subnational Levels Using Hierarchical Linear Models

---

A Dissertation

Presented to

the Faculty of the Morgridge College of Education

University of Denver

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

Bingxin Qi

November 2021

Advisor: Dr. Bruce Uhrmacher

Author: Bingxin Qi
Title: Assessing the Variations of Educational Attainment at National and Subnational Levels Using Hierarchical Linear Models
Advisor: Dr. Bruce Uhrmacher
Degree Date: November 2021

Abstract

Education is a human right, and equal access to education is not only crucial for an individual's well-being, but also essential for eradicating poverty, ensuring long-term prosperity for all, transforming the society, and achieving sustainable development. Measuring education development, especially the variations of educational attainment, in a timely and accurate manner can help educators, practitioners, scientists, and policymakers compare and evaluate various education indicators at both subnational and national levels. This research presents an approach that combines multi-source and multi-dimensional data including population distribution, human settlement, and education data to assess and explore educational attainment trajectories at both national and subnational levels across multiple years. In addition, this study contributes to the power discussions by validating the robustness of models using replication datasets with missing values.

Acknowledgements

First of all, I would like to express my sincere gratitude to all the members of the dissertation committee for their invaluable support and guidance, and their time spent on the completion of this dissertation.

Special thanks to Dr. Bruce Uhrmacher, the dissertation chair. Without his help, advice, and encouragement, this research and dissertation would not have happened. I would also like to thank Dr. Paul Sutton for his knowledge and Dr. Antonio Olmos for his thoughtful feedback, which always aimed at moving me forward.

In addition to my committee members, I would like to thank Dr. Nicholas Cutforth and Dr. Jing Li for their assistance.

Last but not least, I would also like to give special thanks to my family: my parents Mr. Shiying Qi and Mrs. Guizhen Yu for their continuous support and understanding, and Xuantong Wang for his constant support in my life.

Table of Contents

List of Figures

Chapter One: Introduction

1.1 Education and United Nations' Sustainable Development Goals

Education is constantly influencing the development of our social progress. Sustainable Development Index #4 'Quality Education' acknowledges this reality as does the Human Development Index. Measuring education outcomes is a recognized mechanism for measuring the quality of life and assessing levels of socioeconomic development. In addition, education is also the foundation for improving the standard of living and achieving sustainable development (UNDP, 2015). According to the United Nations' human development report, the education index is measured with major dimensions such as education quality and accessibility (UNDP, 2015). Numerous studies have specifically focused on investigating the interrelationships between these factors. For instance, these factors include (1) students' socioeconomic status (e.g., ethnicity, free/reduced lunch rate, and family income), (2) education accessibility (e.g., school leadership and education policies), and (3) educational attainment, which includes higher institution drop-out rate, high school retention rate, and academic performances (Alspaugh, 1998 ; Berkowitz et al., 2017; Crawford, 2015; Okpala et al., 2000; Reason, 2003).

In past decades, researchers have proposed various methods to investigate the interrelationships among the above-mentioned variables that are related to education accessibility, quality, and student achievements (Cortez & Silva, 2008; Kabra & Bichkar, 2011; Saa, 2016). For example, traditional statistical models (including multiple linear regression, logistics regression, and canonical correlation) have been widely used to examine the mechanism of the interrelationships among key variables of interests (Bielefeldt, 2005; Griffith, 1996; Topor et al., 2010). Many researchers have also proposed the use of model-based approaches in validating education indicators as measures of educational attainment. Specifically, various advanced modeling techniques including structural equation modeling and longitudinal growth curve modeling (Muthén, 1994) have been implemented to improve the accuracy of students' attainment prediction by identifying significant attributes (Huitt et al., 2009), classifying predictors (Johnson & Hull, 2014), and differentiating unique contributions of characteristics associated with educational attainment (Kaplan & Elliott, 1997).

Therefore, this study will adopt one of the multilevel modeling techniques, which is hierarchical linear modeling (HLM), to investigate associations between development trajectories of educational attainment and other socioeconomic factors. This research will also incorporate multi-source data to analyze the different growth patterns of these factors at national and subnational levels.

1.2 Research Questions

The overarching goal of this proposed research is to estimate and evaluate the variations of educational attainment on various scales. The key questions are:

1) In what ways can multi-source data and multilevel growth models be incorporated to explore and assess the associations between developmental trajectories and patterns of education and socioeconomic development?

2) Can multilevel models be used to measure and assess national and subnational data to better capture the spatial heterogeneity of variations of educational attainment? If so,

3) What are the initial status and growth rates of educational attainment for countries with various economic and human development status?

   (a) Are there any variations in initial status and growth rates within countries across different regions? Are there any differences in initial status and growth rates among countries?

   (b) For both regional and national levels, are there variations in initial status and growth rates related to factors that are addressed in this study? If so, what are the magnitudes of the relations?

4) How might one utilize the results from multilevel growth models related to educational attainment as empirical evidence to support assumptions regarding power estimations of 3-level HLM growth models without randomized trials?

This research will develop and produce growth multilevel models that can be applied to evaluate heterogeneity of educational attainment at various national and subnational levels across multiple years. The main objectives of this study are to (1) explore and evaluate educational attainment and socioeconomic development using multi-source data, and (2) investigate general educational attainment growth patterns on various national and subnational scales.

1.3 Rationale

Educational attainment is an important aspect of socioeconomic development. Nevertheless, unlike various economic activities, educational attainment has been greatly overlooked by many countries, especially by countries in the less developed regions, as it cannot generate economic profits in the short run and requires lots of financial and human resources as input. For example, according to the report from the United Nations (UN, 2019), it is estimated that more than 50% of children that are not enrolled in school are living in sub-Saharan Africa. In addition to that, about 617 million youth around the world are not equipped with basic mathematics and literacy skills. Although education is not able to generate a large amount of profit in the short run, it can facilitate upward socioeconomic mobility and help people escape poverty (Haskins, 2009; Ladd, 2012). Therefore, there is an urgent need for us to evaluate and track development of educational attainment on various scales in order to support better education policies and practices.

Given the significance of education factors on identifying problems of socioeconomic development on a global scale, in recent years, scientists have incorporated multi-source data to further enhance the model performances that evaluate various socioeconomic indicators associated with human development. For instance, many researchers have been using geospatial data to study human activities at subnational levels (Bundervoet et al., 2015; Ma et al., 2014; Pesaresi et al., 2016; Smits & Permanyer, 2019). Moreover, based on the remotely sensed nighttime light data, Sutton and Costanza (2002) were able to estimate global marketed and non-marketed economic value from two classified satellite images with global coverage. They discovered that the Gross Domestic Product (GDP) was correlated with the amount of light energy emitted by that nation. Therefore, there is also a potential that spatiotemporal data can be used in combination with education data and models to help us better evaluate education development at subnational levels for countries around the world.

1.4 Research Design

In order to better assess the variations of educational attainment, this research contains two different components that are described as follows:

1) Explore and develop growth HLMs for measuring and predicting educational attainment trajectories: I will combine and explore the applications of traditional statistical methods and multi-source data for assessing the educational attainment growth patterns.

2) Use multi-source data to estimate the variations of educational attainment at subnational levels: it is important to estimate and analyze the variations of educational attainment on various subnational scales for countries around the world to support quality education and sustainable development. Nevertheless, it is insufficient to use education data alone to analyze the impacts. Hence, there is a potential for us to combine education data with other forms and sources of data (e.g., geospatial data) to evaluate educational attainment on various scales.

1.5 Research Significance

Many education researchers have only utilized traditional statistical models and advanced modeling techniques for evaluating and predicting educational attainment at individual, school and district levels. Only a few empirical studies were conducted at regional levels, and these studies tend to focus on a specific region or country. Therefore, there is a lack of objective, consistent, and comparable evaluation of educational attainment on a global scale to illustrate the variations of education development trajectories and patterns. The results of this study can potentially affect other aspects of policies regarding socioeconomic development including resource allocation, aid allocation, poverty reduction, urban planning, government spending, and even healthcare. Although many studies have demonstrated the associations between education and socioeconomic variables, they tend to rely on models and statistical tests to establish the interrelationships at the individual level. Thus, there is a lack of research emphasizing modeling the interactions of these variables with multi-source data. In addition, this research also contributes to the research methods and statistics fields by exploring the feasibility of incorporating multi-source and multi-dimensional data for supporting education development evaluations to overcome the limitations of single-source and single-method research design (Holmbeck et al., 2002). Moreover, the results of the HLM model fit indices and parameter estimates with different subsamples will also contribute to the power estimation discussions, especially for HLM growth models without randomized trials at either subnational (i.e., state or province) level and national (i.e., country) level as outlined in many HLM power analysis manuals.

1.6 Broader Impacts

In 2016, world leaders adopted the 17 Sustainable Development Goals (SDGs) of the 2030 Agenda for Sustainable Development, which was built on the success of the Millennium Development Goals (MDGs). The United Nations' SDG education goal includes a specific target that aims to promote "equal access for all women and men to affordable and quality technical, vocational and tertiary education, including university." Moreover, these SDGs also specify that we should "eliminate gender disparities in education and ensure equal access to all levels of education and vocational training for the vulnerable." In order to achieve sustainable development, it is crucial to address the current education issues in a timely, effective, and efficient manner. The results generated by this project can provide substantive knowledge of the current education status for countries around the world on the regional and national scales for future planning and policy recommendations that promote better education opportunity, improve the quality of education, mitigate the adverse impacts of development, improve the overall standards and diversification of learning and education, and seek to attain sustainable development in the long run.

Additionally, this project not only seeks to support various educational and development goals, but also aims to produce data, models, and concepts that can benefit other educational, scientific, and political projects to support education development on various scales.

The rest of this study proceeds as follows. In Chapter 2, I review the existing work related to the relationships between education and socioeconomic development, as well as the development of multilevel growth models. In Chapter 3, I describe (1) the overall research design, (2) data collection and data pre-processing procedures, and (3) the configurations of HLMs for educational attainment analyses. In Chapter 4, I present the details of building various HLMs with time series analyses and interpret the results, and this study ends at Chapter 5, where I make conclusions based on the findings from HLM results and provide suggestions for the future work.

Chapter Two: Literature Review

Assessing socioeconomic development in a frequent, rapid, and accurate manner is important for achieving the SDGs on various national and subnational scales. The United Nations' 17 SDGs and 168 associated targets, are developed to transform the world by urging countries around the world to solve current development challenges related to education, poverty, inequality, climate change, etc. (Griggs et al., 2013; Robert et al., 2005; Sachs, 2012; UNDP, 2015). In recent years, many countries and regional organizations have made significant progress towards the achievement of these goals. Nevertheless, due to the complexity of socioeconomic development, many countries are still suffering from these problems, and some of the actions and policies are not implemented effectively and efficiently.

In order to support the 2030 Agenda for Sustainable Development, it is important to monitor and evaluate the current socioeconomic development status to provide scientific evidence for facilitating the policy and decision-making processes. Measuring socioeconomic development, especially the variations of educational attainment, in a timely and accurate manner can help us better evaluate the effectiveness of the educational systems and processes of education development (Thomas et al., 1999). In the long-run, since education is the foundation of development and growth, measuring socioeconomic data related to the variations in educational attainment and achievement will also help countries achieve many of the SDGs including stable economic growth, eradication of poverty, reduction of inequality and exclusion (Yakunina & Bychkov, 2015).

2.1 Urbanization and Development

To achieve the United Nations' MDGs and SDGs, many governments have invested significantly to improve the basic infrastructure and provide better access to social services. Currently, about half of the world population lives in urban areas. The rate of urbanization is especially significant in developing countries (Cohen, 2006). This rapid industrialization and urbanization process can improve our material lives by providing higher standards of living and better access to services and resources. It is projected that urban populations are expected to reach 5 billion in 2030 (DESA, 2010). Therefore, there is an urgent need to evaluate the impacts of urbanization on sustainable development on various spatial scales in order to mitigate its adverse impacts.

Urbanization generally refers to the process of population transitions from rural to urban areas (Grimm et al., 2000). Moreover, with the expanded extent of urbanization over time, it predominantly involves the procedures by which cities and towns are formed and enlarged as an increased number of people choose to work and live in central urban areas (Vlahov & Galea, 2002).

Urbanization is an indication of social and economic development. Urbanization has been a global phenomenon which demonstrates the transition of development patterns since it does not only refer to increase in the number of urban residents (Nguyen, 2018; Satterthwaite, 2009; Wang & Su, 2019), but also reflects the serial changes regarding key social aspects such as the industry structure (Parikh & Shukla, 1995), employment (Sato & Zenou, 2015), and living conditions (Lin & Liu, 2015). Thus, an extensive amount of literature has focused on investigating the interactions between urbanization and development, especially economic growth (Turok & McGranahan, 2013), and many scholars have established the causal relationship between the two in the short term. For instance, a unidirectional panel causality was established from urbanization to GDP in European Union countries (Kasman & Duman, 2015). Moreover, in the context of emerging-market countries, Ghosh and Kanjilal (2014) used threshold cointegration tests to establish the unidirectional causality from economic development to urbanization in India from 1971 to 2008. Similarly, Zhao and Wang (2015) obtained a unidirectional causal relationship running from economic growth to urbanization in China from 1980 to 2012.

In the long run, the causal relationship between economic growth and urbanization has also been validated in various research contexts. For example, the Granger test has been applied to demonstrate the causality from urbanization to GDP in Saudi Arabia from 1971 to 2012 (Belloumi & Alshehry, 2016). Some researchers have also found the interaction effects (not unidirectional) of regional variability in terms of the relationship. For example, Sadorsky (2013) investigated 76 developing countries and regions, and a mixed effect was found for the relationship between the increased levels of income and levels of urbanization. Meanwhile, the interaction effects were also found at provincial levels (Elliott et al., 2017).

However, some researchers have found little or negative effects of urbanization on economic development (Bertinelli & Black, 2004; Njoh, 2003). For instance, Chen et al. (2014) studied the interrelations between the accelerated urbanization and the expected income growth on a global scale with the panel estimation method. A nonsignificant relationship was found over the last 30 years. In addition, the relationship was re-examined by a number of researchers with a changing of understanding and definitions of urbanization. For instance, with controlled demographic and socioeconomic variations, the economic development (measured by the increased income) varies significantly for both industrialized and developing countries (Lenzen et al., 2006).

Therefore, there is a need to further investigate the associations between urbanization and development on various scales and at multiple time points.

2.2 Educational Attainment and Development

Education indices are key performance indicators for assessing development as they can reveal and explain some of the socioeconomic phenomena from both participation and success in the labor market (Jenkins & Sabates, 2007). Particularly, educational attainment is one of the most important variables that can help us understand socioeconomic status and characteristics (Aghion et al., 2009; Hanushek & Woessmann, 2010; Klasen, 2002; Psacharopoulos, 1994). Educational attainment also plays an important role in evaluating socioeconomic well-being (Hanushek & Woessmann, 2010).

The theoretical framework has also supported the mechanisms of effects of education on economic growth. According to Mankiw et al. (1992), education contributes to increased human capital and thus promotes labor productivity. Moreover, education also significantly contributes to increased innovation capacity in the economy through the distribution of knowledge, skill, and technology (Aghion et al., 1998).

The empirical evidence of the impact of education on economic development has also been widely investigated. Researchers have mixed findings in terms of the relationship. A positive contribution of education to development can be found from extensive studies with classical regression models (Benos & Zotou, 2014; Hill & King, 1995; Knowles et al., 2002). However, there is a substantial controversy on the interpretation of this association. For example, through research reviews and syntheses, researchers have found that the strength of the association between years of schooling and levels of economic growth varies across empirical research contexts (Krueger & Lindahl, 2001). Nevertheless, some scholars have also challenged the plausibility of simple regression models with years of schooling as independent variables on economic growth predictions (Pritchett, 2006).

Although there are disputes about the interpretation of the interrelationship between education and development, many scholars have raised questions on the measurement of education since it is affected not only by educational attainment but also other influential factors (e.g., cognitive skills and health status). Therefore, it is of key significance to re-evaluate the associations among educational attainment, economic growth, and urbanization with (1) various geospatial scales (i.e., provincial and national levels); (2)

various temporal scales (i.e., over a relatively longer period); and (3) appropriate aggregation and disaggregation with regard to measurements of variables.

2.3 Measurement of Human Development with Nighttime Lights

In recent years, many scientists have incorporated multi-source data to enhance model performances for evaluating various socioeconomic indicators that are related to human development. There are many difficulties associated with collecting traditional socioeconomic data for measuring human well-being. For example, accurate information about the human population distribution and human settlements are not available for many regions of the world. Remote sensing technology and geospatial data can be an alternative way for scientists to assess and monitor human activities and presence in a timely and consistent way. For instance, the nighttime light data is widely used for estimating and evaluating socioeconomic activities since it captures the artificial lights on Earth's surface (Baugh et al., 2013; Elvidge et al., 1999; Zhang & Seto, 2011). Elvidge et al. (2009) produced a global poverty map at 30 arcsec resolution based on population and nighttime light data. Therefore, the subnational data generated from nighttime lights can greatly help scientists measure human activities on various spatial scales.

Many scientists have also adopted income Gini concepts for calculating other socioeconomic indexes based on the Lorenz curve. For example, Elvidge et al. (2012) produced the Nighttime Light Development Index (NLDI) based on the Defense Meteorological Satellite Program (DMSP) nighttime light data and LandScan population density data to measure human development. The NLDI can be used to measure the distribution of income and wealth on national and subnational scales. NLDI for each country is calculated based on the Lorenz curve produced from the cumulative proportion of nighttime lights and the cumulative proportion of population (Figure 1). Their results showed that NLDI has a strong correlation with other indicators like the Human Development Index (HDI), poverty rate, and the proportion of urban population. Therefore, NLDI can be an alternative way for measuring human development using spatial data. Song et al. (2010) have also used the Spatial Lorenz Curve (SLC) and Gini coefficients to measure land use changes based on an unsupervised land use classification method with cloud-free Landsat TM images. Similar to NLDI, the SLC is calculated based on the cumulative proportion of land use and the cumulative proportion of total land. Therefore, these studies show that there is a great potential for scientists to utilize geospatial data to monitor the allocation of resources, distribution of population, and different levels of development on various spatiotemporal scales.

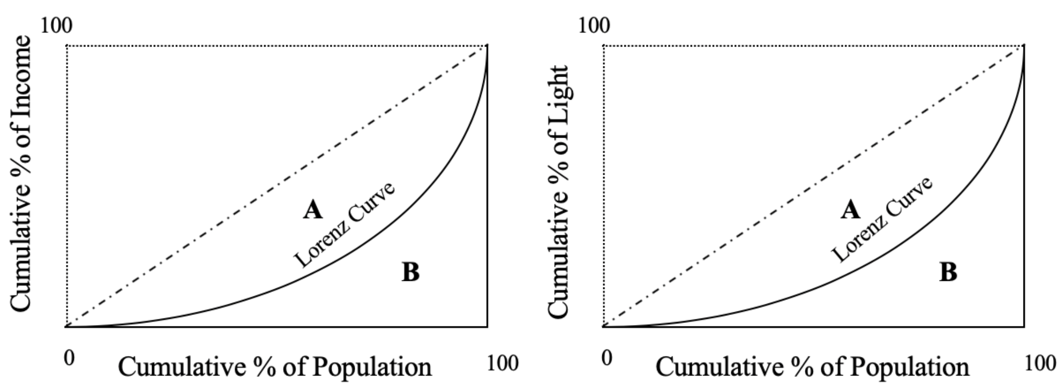


Figure 1. Calculation of (a) Income Gini Using Cumulative Proportion of Income and Population and (b) NLDI Using Spatial Lorenz Curve Based on Cumulative Proportion of Population and Light

2.4 Measurements of Educational Attainment

Although education indicators such as educational attainment are important since they allow researchers to explore their associations with other socioeconomic factors, they are relatively difficult to measure. For instance, researchers have utilized various educational attainment measurements in their studies for the following reasons: (1) there is a lack of a universal standard for measuring educational attainment because many countries and regions have developed and adopted very complex educational systems that are ever changing (Connelly et al., 2016); (2) it is difficult to scale various school grades with complex psychometric structures (Buis, 2010); and (3) there is no consensus on standards to measure the various details and aspects of individuals' educational backgrounds in order to develop standardized procedures for assessment and evaluation.

Currently, many international institutions and organizations are producing education datasets that usually incorporate academically-driven public opinion surveys (Jowell et al., 2007; Smith et al., 2011) as measurements of educational attainment. Moreover, there is an increasing number of international achievement studies (Chiu, 2007; Mayer, 2008; Samdal et al., 1999) that lead to the increased popularity of cross-national and sub-regional comparative studies (Hoffmeyer-Zlotnik et al., 2003) to explore the differences of education measurements.

Traditionally, researchers in the fields of educational sociology and social satisfaction tend to include educational measures in their studies more frequently (Paterson & Iannelli, 2007). Nevertheless, more researchers from other fields are incorporating educational indexes as secondary data (e.g., as outcome or explanatory variables) to reduce the bias and improve the relevancy and accuracy of their research.

Generally speaking, the following are common broad categories and approaches used to measure educational attainment: (1) time spent in education (Eikemo et al., 2008; Schneider, 2013) because it fits the statistical models such as regression models as continuous covariates (Treiman, 2014); (2) qualification-based measures such as students' educational backgrounds and previous subjects undertaken because these can provide researchers with additional information; and (3) scaling education measures that include educational instruments developed and based on certain relevant criteria. For example, some researchers advocate the use of scaling education measures such as the development of qualifications by ranking average income of workers with certain degrees of education (Treiman, 2005). This approach has been advocated by many other scholars (Lambert, 2012) because a large number of attributes can be represented in a single scale. In a statistical modelling framework, scoring offers a parsimonious way of summarizing detailed education data. In conclusion, schooling measures show advantages over other educational attainment variables and instruments such as school-enrollment ratios or adult literacy rates because: (1) it is aligned with concepts of human capital which influence current decision factors like fertility and health (Barro & Lee, 1996); (2) numerical values are preferred rather than categorical ones, such as highest degrees of education received, since there is a lack of a universal standard for primary, secondary, and tertiary degrees for many countries and regions, and the classifications of these degrees from educational systems are changing over time (Bolton, 2012); (3) it is effective in avoiding the repeated measures of accumulated years of schooling as a measure of attainment. For example, the measure of years of schooling could be inaccurate since people may dropout and re-enter schools (Dearden et al., 2002); (4) compared to the accumulated schooling measures, the mean years of schooling will include more people in general rather than people within certain age ranges (e.g., years of schooling for people aged 25 and over); and (5) they have been included in many empirical cross-country studies since there is an increasing amount of national census data available (Barro & Lee, 1993).

2.5 Hierarchical Linear Modeling (HLM) Design

2.5.1 Limitations of Current Models

In the field of geography, conventional statistical models, such as multiple linear regression and logistic regression are commonly used techniques to explore associations between variables of interest. For example, at regional levels, Ma et al. (2014) confirmed the statistically strong connections between the Visible Infrared Imaging Radiometer Suite (VIIRS)-derived nighttime radiances and multiple urbanization variables across cities in China. The results from linear regressions show that increases in satellite-observed night light signals of cities are generally responsive to linear growth in urban population, GDP, electric power consumption and road area. VIIRS nighttime light therefore can be indicative of demographic and economic dynamics during the urbanization processes. Xu et al. (2014) used a piecewise linear model to examine the spatiotemporal trends in urban development. Yu et al. (2015) used linear regression analysis on relationships between the average light index and the integrated poverty index to evaluate regional poverty in China.

At the global level, Elvidge et al. (2009) derived a global poverty map from DMSP nighttime lights based on the assumption that nighttime lights can be used as a proxy for wealth. Therefore, a linear relationship between the national poverty index and the proportion of population living below the poverty line was established to measure the poverty per grid cell using LandScan population data. Shi et al. (2016) adopted the linear regression model to quantify the correlation between the electric power consumption and inter-calibrated nighttime lights from 1992 to 2012. Kummu et al. (2018) produced a global subnational GDP dataset by disaggregating the national GDP based on the population per grid cell.

In addition, spatial autocorrelation has been widely applied to detect and quantify the correlation between a value of some variables at one location in space and nearby values of the same variable (Griffith, 1987). These neighboring values can be identified by an n-by-n binary geographic weighted matrix. As a variant and extension of conventional correlation, Moran's I spatial autocorrelation (Cliff & Ord, 1981) examines the heterogeneity of variables at the same location.

Since most spatial analyses employ model-based analytical techniques, and the underlying assumptions of homoscedasticity (i.e., the random disturbance in the relationship between independent and dependent variables) and independence observation (i.e., the probability of a value taken on by a model's error terms does not affect the probability of a value taken on from the remaining error terms in the model) are violated, the index or measures of spatial autocorrelation are needed to solve the problems mentioned above. Moreover, the Moran's I index is needed since (1) it measures the degrees of violation of assumptions; and (2) it describes an overall pattern and detects deviations across geographic locations in terms of the patterns.

Though Moran's I index has been predominantly used in empirical studies to explore clustering effects of locations like disease distribution and emergence (Hoen et al., 2009), traffic congestion and crashes (Moons et al., 2009), and spatiotemporal events (Prasannakumar et al., 2011), there is still space for improvement of Moran's I index since: (1) it is still unclear for researchers to determine the spatial contiguity matrix (Chen, 2013); and (2) there are problems of scaling consistencies (Chen, 2011) for variables.

Similar to conventional statistical models, spatial autocorrelation has limitations since it fails to better estimate nested structures, and this leads to the following concerns: (1) the reduced accuracy caused by aggregation bias, misestimated parameters, and unit of analysis problems (Hopkins, 1982); (2) impoverished conceptualization that discourages the formulation of explicit multilevel models with hypotheses about effects occurring at each level and across levels (Aguinis et al., 2013); and (3) growth models without levels (Goldstein et al., 1994).

In the field of economy, the following empirical studies have explored interrelationships between various factors of socioeconomic development. Yang et al. (2014) analyzed the relationship between average years of schooling (AYS) and educational Gini coefficient for each subgroup by decomposing the datasets based on regions, income, gender, and age. They measured the within-group contribution and between-group contribution to measure different factors' contribution to education inequality. Mesa (2007) utilized the Gini coefficient measures proposed by Thomas et al.

(2001) to assess the relationship between education Gini and income Gini using trend analysis and linear regression models.

Chen and Nordhaus (2011) examined the structural relationship between nighttime lights and GDP by using 1° by 1° arc grid cells to aggregate spatial data. They concluded that nighttime lights can be used to predict population and economic statistics of Sub-Saharan Africa since these regions have very low population density and economic activities. Gregorio and Lee (2002) explored the relationship between educational attainment and income inequality using the technique of seemingly unrelated regressions (SURE) at the national level. Castelló-Climent and Doménech (2014) used a fixed effects estimation model and identified positive relationships between education and income inequality.

Although HLMs also rely on assumptions such as homoscedasticity with other ordinary least-squares (OLS) regression models, such as spatial autocorrelation models, the advantages of using HLMs over other regression methods are obvious. One of the primary advantages of modeling growth with HLMs is that it is comparatively flexible. For instance, Moran's I index cannot indicate multivariate associations (Lee, 2001). For example, if a researcher plans to examine the shared characteristics between adjacent regions, only one variable of interest can be explored at a time. Moreover, though Moran's I index, or other spatial autocorrelation models are able to evaluate degrees of violations of assumptions shared with HLMs, they fail to analyze time-varying covariates because they are univariate in nature (Bian-Ling, 2014). More importantly, they are not as flexible as HLMs because HLMs allow researchers to model fixed or random growth parameters such as linear slope, higher order polynomial functions, and they allow researchers to determine and constrain the fixed and random effects at higher levels (Rasbash & Goldstein, 1994).

Another primary advantage of HLMs is that it assumes little about the data structure. For example, repeated measures often require equal time intervals, and HLMs are robust with missing data, and even missing data at different measurement occasions. However, if missing values occur at different measurement occasions, general linear regression models, such as spatial autocorrelation, may not be able to generate unbiased estimates. Therefore, HLM is the appropriate approach for handling missing values (Rogosa & Saner, 1995).

HLM is also an appropriate approach for handling longitudinal data in that various growth change patterns can be modeled. For example, except for time scores, time-varying covariates can be included at level-1 of HLMs (McCoach & Kaniskan, 2010). However, other OLS regression methods for studying geographical clusters do not have the capabilities to examine time-varying covariates together with variables of interest. Therefore, the assumptions of HLMs are shared with other OLS regressions, such as linearity, normality, homogeneity of variance, and homoscedasticity, and Moran's I index demonstrates its advantages of modeling spatial similarities. However, HLM is selected for this study as the analytic technique due to its flexibility and better capabilities of detecting clustering effects in different units of samples collected at different time points.

In sum, these conventional models suffer from several main deficiencies: (1) there is no objective model for comparing various factors at global and regional levels; (2) there is a need to take spatiotemporal heterogeneity into consideration; and (3) it is important to use multilevel models to further improve the estimations of lower-, higher- and cross-level variations. Therefore, this study incorporates HLM technique to overcome the above-mentioned limitations.

2.5.2 Development of HLMs

HLMs are often referred to as multilevel linear models (Goldstein, 2011), mixed-effects models (Elston & Grizzle, 1962), random coefficient regression models (Longford, 1993), and covariance components models (Dempster et al., 1981) in various research domains. It is developed to analyze the nested structural data, and to capture the shared characteristics, patterns, and growth trajectories of individuals within the hierarchical groupings (Heck & Thomas, 2015). By modeling the variability within and between clusters, random coefficient variables at micro and macro levels, and interactions between levels without biased aggregation and disaggregation methods, HLMs tend to provide researchers with more accurate parameter and relationship estimates. In addition, HLMs have the flexibility and capability of handling longitudinal data with hierarchical structures. Therefore, this study purposes to use HLMs over other traditional regression models for longitudinal data for the following reasons: (1) improved estimates of effects within individual units; (2) the formulation and testing of hypotheses about cross-level effects (e.g., how varying income levels might affect the relationship between urbanization and educational attainment within and across countries); and (3) the partitioning of variance and covariance components among levels (e.g., decomposing the covariation among sets of subnational-level variables into within and between country components).

In situations where data are grouped or nested, the effects of variables on the outcome is conditional to that nesting. If data are dependent upon the effects of higher-level units, then residuals of individuals within the higher-level unit will be correlated. Thus, it is of great importance for researchers to represent the nesting effects in the model.

Similar to OLS regression models, the base level HLM model (the developmental trajectory model in this study) is referred to as the level-1 model. The analysis of HLMs is also similar to OLS regression: the outcome variable at level-1, educational attainment is predicted as a function of a linear combination of level-1 variables (the time scores and growth parameters) and an intercept.

At subsequent levels, the level-1 slopes and intercept become the dependent variables being predicted by level-2 variables. At the level-3 model, the intercepts and slopes at level-2 are further explained by level-3 variables. Through this process, the effects of level-1 variables on the outcome and the effects of level-2 variables on the outcome are more accurately modeled.

Moreover, the slopes and intercepts are predicted by models within and across levels, so the differences in the relationship between variables at both levels and the outcome can be better understood.

The application of multilevel models to analyze longitudinal data is prevalent in the field of educational research. Specifically, educational researchers tend to study the growth of the individual student learner within the organizational context of classrooms and schools. For instance, Bryk and Raudenbush (1988) formulated a three-level model that enabled a decomposition of the variation in individual growth trajectories into within and between school components. They found that 83% of the variance in growth rates was due to school-level clusters.

Other than univariate and multivariate analysis of variance (ANOVA) with repeated measures, HLMs with time series are applied in this study to: (1) model the growth patterns of educational attainment, especially the variations of educational attainment within each individual region; (2) investigate the between cluster (i.e., between regions) variations on the developmental trajectories of the outcome with estimated growth parameters at a within-cluster (i.e., within regions) level; (3) analyze the interrelationship between intra- (differences on growth within each regional cluster) and inter-individual (differences between regional clusters) changes; and (4) uniquely identify the key and influential determinants or predictors of intra-individual and inter-individual changes.

Theoretically, rationales of using HLMs (Bryk & Raudenbush, 1992) with multiple time points over general linear models are as follows: (1) compared to multivariate analyses and ANOVA with repeated measures, HLMs display greater flexibility on estimating random variations of individual growth (Van der Leeden et al., 1996); (2) HLMs provide better control for time-invariant between-group variation (Galla et al., 2014); and (3) HLMs provide the covariance structure analysis that is capable of demonstrating and estimating level-2 relationships between slope and a single predictor of change (McArdle & Epstein, 1987).

Conceptually, HLMs can be viewed as a modeling process where separate OLS regression models are examined at different levels. However, HLMs use a full maximum likelihood (MLF) or a restricted maximum likelihood estimation method to evaluate the fixed (e.g., the regression coefficients and intercepts) as well as random effects.

Therefore, it is the capability of handling the nesting effect (determined by the intra-class correlation) that distinguishes HLMs from OLS models. HLMs take the nesting effects into account by effectively estimating the random effects in the model. Compared to conventional OLS simple regression models, these effects are explicitly specified in the model, so the biases caused by nesting effects can be appropriately addressed.

The essential task of constructing HLMs is to correctly specify the random components associated with the growth parameters and variables at every level. However, OLS models assume equal variances of growth parameters and intercepts at every level. The Chi-square statistics are used to indicate whether a random effect should be included in the models.

Moreover, the likelihood ratio test (also known as the deviance statistic) is used to compare HLMs with different specifications on fixed and random components. The deviance statistic is also called -2 log likelihood (-2LL), and a statistically smaller value of deviance statistic indicates a better model fit for the data (Garson, 2013). To conclude, HLMs demonstrate greater flexibility of modeling the fixed and random effects to OLS models by using Chi-square statistics to suggest the appropriateness of model configurations.

Empirically, there is an increasing amount of literature showing the applications of HLMs for analyzing the development of the variations of educational attainment over time. For example, Kunovich and Hodson (2002) developed multilevel models with individual-level and county-level data to study the county-level variation on academic achievement. Similarly, since the educational attainment and achievement gap has been one of the central concerns of education policy in the U.S., Xiang (2009) developed multilevel models to indicate school-level and district-level differences of mathematical achievement by including time-varying covariates. Furthermore, longitudinal large scale data was increasingly included in HLM empirical studies to reveal educational achievement gaps (Ichou & Vallet, 2011).

2.6 Summary

The purpose of this literature review is to explore the trends and associations between socioeconomic and education development. Therefore, developing advanced statistical models using multi-source data, especially geospatial data, also provides a new opportunity for improving current model performances. In addition to that, this can help us generate more accurate and meaningful results of education data at various levels to monitor the status of education development using HLMs.

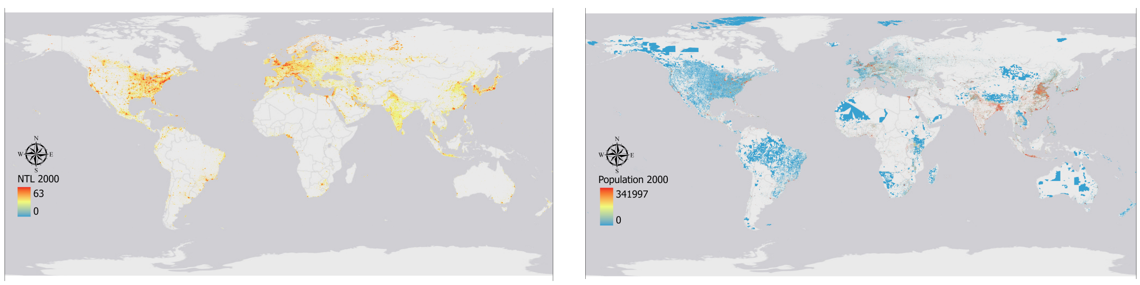
Chapter Three: Datasets and Methods

This chapter presents an approach that combines multi-source data (including population distribution, human settlement, and nighttime light data) to assess change patterns of human development and educational attainment at both national and subnational levels across multiple years. This research utilizes nighttime light imagery collected by the various satellites, including the DMSP and VIIRS and human settlement data from the Global Human Settlement Layer (GHSL) framework to assess human development and evaluate its association with the education indicator.

3.1 Data Preparation

In order to combine multi-source data to assess change patterns of human development and educational attainment at both national and subnational levels across multiple years, this study incorporates: (1) geospatial data including nighttime lights and human settlements and (2) education data related to educational attainment. The geospatial data can support the visualization and analysis of human activities and demographic transitions on Earth's surface over time. Human development can be monitored through the intensity of nighttime lights collected by satellites. The global human settlement information is obtained from the GHSL, which was mapped based on

Landsat imagery to show the global built-up areas from 1975 to 2014 (Pesaresi et al., 2015). The GHSL represents global spatial information in the form of built-up, population density, and settlement maps characterizing human presence on Earth's surface over time. I propose to use the GHSL to extract the population in urban and rural areas within various national and subnational entities. This information will be used to measure the proportion of urban population and estimate the human development levels. Figure 2 shows an example of nighttime light intensity and population distribution on a global scale.



(a) (b)

Figure 2. Examples of (a) Nighttime Light Intensity and (b) Population Distribution

The study has three stages. The first stage is to pre-process the geospatial data to calculate NLDI and urban population ratios across different years. The second stage is to assess and explore human development and educational attainment trajectories at the national level across multiple time periods to better analyze the developmental trajectories of educational attainment, human development, and urbanization for each country. The third stage is to construct 3-level HLMs to assess the variations of educational attainment in different subnational entities.

3.1.1 Data Collection and Pre-processing

This research incorporates multisource data (including urbanization, population, human development, economic, and education datasets) to analyze educational attainment on various spatiotemporal scales. Although this study utilizes multiple variables as inputs for model configurations and statistical analyses, I will aggregate them based on common attributes (including ISO 3166-1 alpha-3 standard for countries and areas, source: https://www.iso.org/iso-3166-country-codes.html) so that these datasets with multi-dimensionalities can be incorporated into the models. The geospatial administrative boundaries provided by the United Nations and the Database of Global Administrative Areas (GADM) can facilitate the aggregation of spatial data in order to join results with tabular datasets. The datasets included in this study are summarized in table 1.

Table 1. Summary of Datasets

Dataset	**Description**	**Sources**	**Data Type**
Educational Attainment Data	National educational attainment data	The United Nations' Development Program (http://hdr.undp.org/en/data)	Tabular (csv)
Subnational Human Development Index (4.0)	Subnational educational attainment data	Global Data Lab (https://globaldatalab.org/shdi/shdi/)	Tabular (csv)
World Inequality Database on Education	National and subnational education indicators	The United Nations Educational, Scientific and Cultural Organization (https://www.education-inequalities.org/)	Tabular (csv)

Global Human Settlement Layers	Global geospatial dataset for human settlement and population distribution on earth for 1990, 2000, and 2015	European Commission (https://ghsl.jrc.ec.europa.eu/)	Raster (Geotiff)
DMSP/VIIRS nighttime light	DMSP nighttime light product from 1992-2013 and VIIRS nighttime light products from 2015-2017	NOAA/NASA (https://ngdc.noaa.gov/)	Raster (Geotiff)
Administrative Boundaries	National and subnational administrative boundaries from Database of Global Administrative Areas (v3.6)	GADM (https://gadm.org/)	Vector (Shapefile)

This study incorporates education, population, and development data for countries and their subnational entities to build models. Based on the United Nations' educational attainment database, a total number of 187 countries with 1689 subnational entities are included in this study. The education data are available from 1980 to 2018 for most of the subnational entities. The summary of subnational entities by regions and income groups are included in table 2.

Table 2. Summary of subnational entities by region and income group

National Education Indicators	**Count**
Total Number of Countries	187
Total Number of Subnational Entities	1689
By Regions	
Sub-Saharan Africa	512
Europe and Northern America	450

Latin America and the Caribbean	305
Eastern and South-eastern Asia	203
Northern Africa and Western Asia	195
Central and Southern Asia	181
Oceania	34
By Income Groups	
Lower middle-income countries	526
Upper middle-income countries	513
High-income countries	489
Low-income countries	352

NLDI at national and subnational levels were constructed using level 0, 1, and 2 administrative units obtained from GADM. Level 0 represents national-level administrative boundaries, level 1 represents state and provincial-level boundaries, and level 2 represents county and district-level boundaries. In order to construct the Lorenz curve for each country based on the cumulative proportion of nighttime light and population, this study used the level 1 subdivisions' administrative boundary layer (state or province) to calculate the sum of population and nighttime light within each subdivision. Based on the cumulative percentage of nighttime light and population data, this study calculates the NLDI value for each country for that corresponding year. The subnational NLDI at level 1 subdivisions is calculated based on level 2 subdivisions' data using the same procedures.

This study utilizes urban population ratio in order to measure the demographic transition and population concentration patterns caused by urbanization and development. The urban regions are defined by the human settlement data obtained from GHSL. This study uses the urban center and urban cluster grid cells as masks to extract urban population. The extracted urban population pixels will be aggregated based on national and subnational entities' boundaries. Therefore, the urban population ratio is defined as follows:

$$\text{urban population ratio} = \frac{total\ urban\ population}{total\ population} \tag{1}$$

Specifically, the table below summarizes how the variables are accessed, calculated, and processed:

Table 3. Summary of Variables in HLMs

Variable	**Dataset**	**Format/Calculation Formula**	**Data Descriptions**
Mean Years of schooling	National educational attainment data	raw data; converted from attainment levels using official durations of each level	Average number of years of education received by people ages 25 and older
National NLDI	National development index data	details of combining three dimensions into the calculation http://hdr.undp.org/sites/default/files/hdr2020_technical_notes.pdf	A composite index measuring average achievement in basic dimensions of human development: health, education, and standard of living
Subnational HDI (4.0)	Regional HDI data	The composite score combining three dimensions at regional levels	Average of the subnational values of three dimensions: education, health, and standard of living
Urbanization Indicator	Human settlement and population distribution on a global scale	urban population ratio = (total urban population)/ (total population)	These data contain a multitemporal information layer on built-up presence as derived from Landsat image collections
Income	World Bank national accounts data, and OECD National Accounts data	raw data; GDP per capita	In current US currency

GDP	World Bank national accounts data, and OECD National Accounts data	raw data; annual GDP growth	Annual growth rate (%)

3.2 Configurations of HLM Growth Models

Following the procedures of configuring the HLM growth models, the unconditional three-level growth model is constructed first, and Intercepts- and Slopes-as-Outcomes Models with covariates will be constructed until unconditional models yield acceptable model fits (Raudenbush & Bryk, 2002).

The developmental trajectories for economic and education growth have been explored and studied by many researchers. For example, Cai et al. (2002) have applied the Neoclassical Growth Theory (i.e., the initial status is negatively related to rates of change in following years) to validate quadratic associations between economic and education growth patterns in the context of developing countries and regions. Moreover, the Solow Growth Model, a variant of Neoclassical Growth Model, has been applied to investigate the association between education and economic growth (Vinod & Kaushik, 2007). It is also found that education as human capital input has demonstrated to be an influential indicator of economic growth using modified neoclassical models in the long-run for 78 countries from 1960 to 1995, and the growth shows a nonlinear manner (Bassetti, 2009). Thus, the theoretic framework of education development has been examined by many empirical cross-national studies. For instance, using the neoclassical models, researchers have demonstrated that the average years of schooling can be used to explain a significant proportion of the cross-country variations in economic growth rate (De la Fuente & Doménech, 2002; Fleisher & Chen, 1997) over time.

3.2.1 Research Hypotheses

Based on my research questions, theoretical and empirical support, the following research hypotheses are developed to answer key questions for this study:

H1: Level-1 Unconditional Model – Within-Country and Within-State Level

$$Y_{ti} = \pi_{oi} + \pi_{1i}\, a_{ti} + \pi_{2i}\, a_{ti}{}^{2} + e_{ti}, \quad e_{ti} \sim N\,(0, \sigma^2) \tag{2}$$

where t represents the coded time scores, i denotes the regions, a_{ti} is the time score at time t for region i, π_{1i} is the linear growth parameter for region i, π_{2i} is the growth trajectory parameter for region i associated with quadratic change polynomial, and e_{ti} is the error variance with repeated measures. It is commonly assumed that e_{ti} is independently and normally distributed with a mean of 0 and constant variance σ^2. There is a quadratic (non-linear) growth trajectory pattern at the regional level (provincial level) within countries.

H2: Level-2 Within-Country and Between-State Level

The less developed regions (states) will demonstrate higher linear and quadratic rates of change. The values of instantaneous growth parameters will be statistically significantly smaller than the acceleration parameters.

$$\pi_{oi} = \beta_{00} + r_{0i} \tag{3}$$

$$\pi_{1i} = \beta_{10} + \beta_{11}(\text{Urbanization})_{1i} + r_{1i} \tag{4}$$

$$\pi_{2i} = \beta_{20} + \beta_{21}\,(\text{Urbanization})_{1i} + \beta_{22}\,(\text{Human Development})_{2i} + r_{2i} \tag{5}$$

where $(\text{Urbanization})_{1i}$ and $(\text{Human Development})_{2i}$ indicate that they are measured characteristics of regions to predict level-2 growth parameters π_{1i} and π_{2i}, β_{00} is the cluster mean of intercept parameter for level-2 unit, β_{10} is the cluster mean of linear slope parameter for level-2 unit, and β_{20} is the cluster mean of quadratic growth parameter for each level-2 unit, r_{pi} is the random effects associated with regions with means of 0 and a normally distributed full covariance matrix (T).

H3: Level-3 Between-Country and Between-State Level

The growth patterns across countries show different development trajectories, but the values of economic growth and human development variables are statistically significant predictors of educational attainment.

$$\beta_{00} = \gamma_{00} + \gamma_{01}\,(\text{Income})_k + \gamma_{02}\,(\text{Human Development})_k\ \textit{(non-randomly varying)} \tag{6}$$

$$\beta_{10} = \gamma_{10} + \gamma_{11}\,(\text{Urbanization})_k + \gamma_{12}\,(\text{Human Development})_k\ \textit{(non-randomly varying)} \tag{7}$$

$$\beta_{20} = \gamma_{20} + \gamma_{21}\,(\text{GDP})_k + \gamma_{12}(\text{Human Development})_k + \mu_{2k} \tag{8}$$

where γ_{p0} is the grand mean for the corresponding polynomial order of change, γ_{01} is the main effect of variable (i.e., income), γ_{02} is the main effect of variable (i.e., human development) to predict the intercept parameter β_{00} for each level-3 unit (i.e., country), and μ_{2k} is the only random effect specified at level-3, which is associated with the quadratic change parameter β_{20} at the country-level. Table 4 and figure 3 below offer a better visual representation of the HLMs in terms of its structures, hierarchies, and configurations:

Table 4. Detailed Descriptions of HLMs

Level	**Model**	**Geographical Level**	**Descriptions**
Level 1	$Y_{ti} = \pi_{oi} + \pi_{1i}\, a_{ti} + \pi_{2i}\, a_{ti}^2 + e_{ti}$	state or province	unconditional
Level 2	$\pi_{oi} = \beta_{00} + r_{0i}$ $\pi_{1i} = \beta_{10} + \beta_{11}(\text{Urbanization})_{1i} + r_{1i}$ $\pi_{2i} = \beta_{20} + \beta_{21}\,(\text{Urbanization})_{1i} + \beta_{22}\,(\text{Health})_{2i} + r_{2is}$	subnational	predictors at the subnational level are included
Level 3	$\beta_{00} = \gamma_{00} + \gamma_{01}\,(\text{GDPPC})_k + \gamma_{02}\,(\text{HD})_k$ $\beta_{10} = \gamma_{10} + \gamma_{11}\,(\text{GDPPC})_k + \gamma_{12}\,(\text{HD})_k$ $\beta_{20} = \gamma_{20} + \gamma_{21}\,(\text{GDPPC})_k + \gamma_{12}(\text{NLDI})_k + \mu_{2k}$	national	more predictors at the national level are included

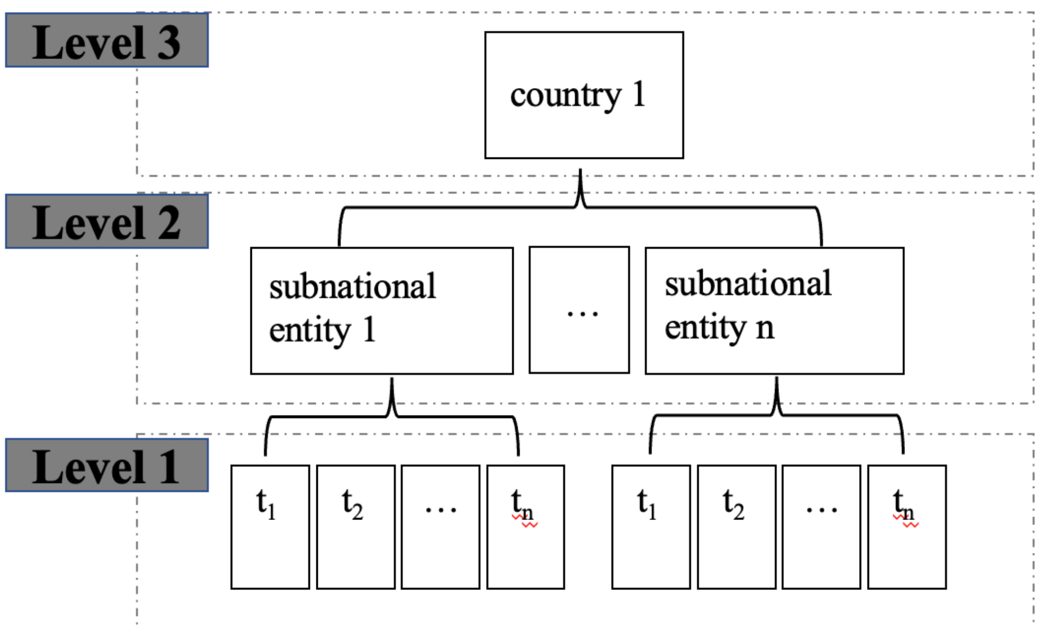


Figure 3. Visual Representations of the HLMs

3.3 Power Analysis for the Proposed HLM models

Many researchers, methodologists, and scholars have been contributing to address the sample size issue for HLMs, and many publications are simulation studies that are mainly focused on estimating the impacts of number of micro- and macro-level units on the precisions of parameter estimates, variance components, and cross-level interactions.

Specifically, there are many scholars who have adopted a fixed total sample size to ensure statistical power. For instance, Kreft (1996) has proposed a 30/30 rule of thumb for 2-level designs, which requires a minimum of 900 of total samples for any type of effects to be studied. Similarly, Hox (1995) suggests another rule of thumb that a total number of 50 clusters with 20 individuals per cluster is appropriate for multilevel modeling. Empirically, Maas and Hox (2005) have conducted a simulation study with varying numbers of cluster sizes, clusters, and intra-class correlation coefficients (ICCs). They found that when the clusters were substantially lower than 100, the sampling error estimates for macro-level variances tended to be underestimated.

Their findings represent the central topic to the sample size issues for HLMs, and lead to discussions about the significant role of numbers of macro-level units played on precisions of higher-level variance estimates. Particularly, Snijders and Bosker (1993) have argued that in the context of a 2-level model with fixed effects, as the number of clusters decreases, the sampling errors increase with total sample size being kept constant. Moreover, other researchers also pointed out the allocation of sample sizes depends on practical aspects such as treatment conditions (Moerbeek et al., 2000).

Nevertheless, there is an increasing number of researchers who stress to readdress the sample size allocation issues with careful considerations of research objectives (Snijders & Bosker, 1993). Some researchers put forward the role of covariates in explaining (Raudenbush, 1997) the variations of dependent variables. In addition, some scholars also mention that the sample size requirements depend on the magnitude of clustering effects (i.e., the values of ICCs), and it matters when variables of interest are included at different levels. For instance, small cluster sizes are found to be unproblematic when testing regression coefficients, but it has a negative impact on testing power when constraining random slope variances at the macro-level (Snijders, 2005). Therefore, the power analysis procedures are outlined in the following section.

3.4 Checking for Assumptions

Based on the assumption checking procedures proposed by Raudenbush & Bryk (2002), the following assumptions will be checked before conducting statistical analyses for this study:

1) at level-1 of the model, each e_{ti} is independent and normally distributed with a mean of 0 and variance of σ^2 for every measurement occasion within each level-2 unit (i.e., regional levels such as states or provinces). Therefore, this assumption will not be affected by the data characteristics of this study if residuals at different time occasions were not associated with variables selected in the model and residual variances show a normal distribution over different measurement occasions.

2) at level-2, the predictors are independent with level-2 residual variance. In other words, the predictors included at the regional level are not affected by regional error variance r_{ij}.

3) the vectors of Q + 1 random error at level-2 are multivariate normal, each with a mean of 0, some variance, τ_{qq}, and covariance among the random elements, q and q', of τ_{qq}'. The random-error vectors are independent among the J level-2 units [i.e., $r_j = (r_{0j}, \dots, r_{Qj})' \sim$ iid $\mathrm{N}(0, \mathrm{T})$].

4) the set of level-3 predictors (i.e., all the unique elements in W_j across the Q +1 equations) are independent with μ_{2k}.

5) the errors across levels are independent with one another.

6) the predictors at each level are not correlated with cross-level random effects.

Thus, these assumptions will be examined during data pre-processing and data analyses procedures. For example, if significant numbers of outliers are detected, the influential cases will be dropped to make sure that assumption of normality is met and will not be affected by data characteristics. In addition, for the fifth assumption the residual plots within and between levels will be examined during the data analysis procedures to ensure that the assumption will not be affected by the data.

3.5 Power Analysis Procedures

According to Raudenbush and Bryk, (2002), in longitudinal studies, the sample size is T, the number of time points per region (i.e., state/province) for the duration of the study is D, the number of regions is i, and J is the number of clusters (i.e., countries). Some researchers argue that by adding more time points T, it would be helpful to increase the power when the within cluster variance σ^2 is large (Usami, 2014). Some researchers argue that by adding sample size, n, it could significantly increase the power when between cluster variations and cross-level interaction effects in terms of their developmental trajectories are large (Mathieu et al., 2012).

Due to the lack of literature in the fields of educational and social studies that focuses on studying power issues in multilevel models without randomized trials (Heo & Leon, 2008), this study assumes that there are two groups of countries (e.g., by separating countries with into low- and high-income groups (World Bank, 2021) or by grouping countries into North-South divide (Arrighi et al., 2003; McFarlane, 2006). Therefore, a dichotomous grouping variable X_g will be introduced at level-2 models only for the purpose of power examinations. I further assume that $X_g = 1$ represents the treatment group and $X_g = 0$ represents the control group, following the power calculation procedures that are put forward by Spybrook et al. (2011). Therefore, the power of detecting treatment effects depend on the following noncentrality parameter:

$$\varphi = n\lambda\delta^2/4 \tag{9}$$

$$\delta = \frac{\gamma_{p01}}{\sqrt{\tau_{\beta p} + \tau_{\pi p}}} \tag{10}$$

where δ^2 is the group difference on the polynomial of interest divided by the standard deviation (*SD*) for that polynomial, or the square root of the sum of the between cluster variance. γ_{p01} is the main effect of treatment for quadratic change, $\tau_{\beta p} + \tau_{\pi p}$ is the total between country and between region variance, and λ is the reliability parameter:

$$\lambda = \tau_{11}/(\tau_{11} + V_1) \tag{11}$$

Thus, the hypotheses to test the significance of the main effect (i.e., the treatment effect) for the quadratic change are:

$$H_0: \gamma_{p01} = 0$$
$$H_1: \gamma_{p01} \neq 0$$

When the H_0 is true, the test statistics F follows a central F (1, J-2) distribution:

$$F = \frac{\hat{\gamma}_{p01}}{\mathrm{Var}(\hat{\gamma}_{p01})} \tag{12}$$

When the H_1 is true, which means that the treatment effect is statistically significant, the test statistics remain the same but follow a noncentral F (1, J-2; λ) distribution. Thus, the noncentrality parameter above can be rewritten as:

$$\lambda = \frac{\gamma_{p01}^2}{\mathrm{Var}(\hat{\gamma}_{p01})} = \frac{J\gamma_{p01}^2}{4[\tau_{\beta p} + (\tau_{\pi p} + V_p)/n]} \tag{13}$$

Therefore, the larger the noncentrality parameter, the greater the power of the test. In addition, another parameter that influences the power is the value of ICC, which is:

$$\rho = \frac{\tau_{\beta p}}{\tau_{\beta P} + \tau_m} \tag{14}$$

The nominator is the between cluster (i.e., country level) variance and the denominator is the total variance. Thus, the treatment effect estimate is written as:

$$\mathrm{Var}(\hat{\gamma}_{p01}) = \frac{4\left[\rho + \frac{(1-\rho+V_p)}{n}\right]}{J} \tag{15}$$

Adding the ICC value as a parameter in calculating the main effect, the noncentrality parameter becomes:

$$\lambda = \frac{J\delta^2}{4[\rho + (1-\rho)/(\alpha_p n)]} \tag{16}$$

Based on the equation above, it can be concluded that the power is a function of number of higher-level J (i.e., country level), the cluster size n (i.e., number of regions in each country), the standardized effect size δ, the within region variance, σ^2, the between region variance $\tau_{\pi p}$, the study duration D, and the number of measurement occasions T.

Therefore, assuming that each country has an average of 11 regions (i.e., cluster size n = 11), and without knowing the actual values for within country and between country variance, this study conducts power analyses based on the following scenarios:

Variations at Level-2 > Variations at Level-3

In other words, $\sigma^2 > \tau_{11}$, the reliability value decreases. In this case, increasing the duration is more effective for increasing the power. Therefore, I use Optimal Design software Version 3.01 (Raudenbush et al., 2011) with data collected from 2013 to 2017 (duration D = 5 in figure 4), measurement occasions (i.e., number of time points) T = 5, and α_p= 0.05. The results in Figure 4 show the number of total clusters needed to maintain power at 0.8 with ICC values of 0.1 and 0.15:

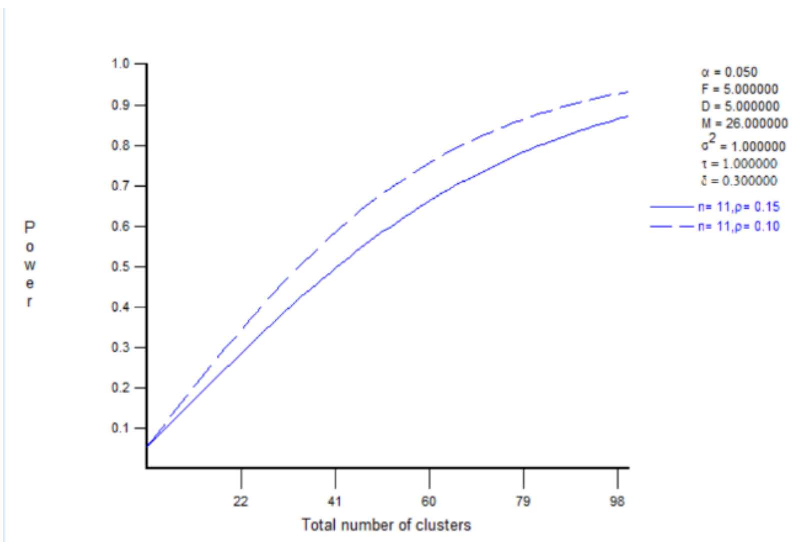


Figure 4. Power and Total Number of Clusters (with duration = 5, occasions = 5)

As the duration and measurement occasions increase, the number of clusters needed to maintain power at 0.8 is significantly reduced as shown below:

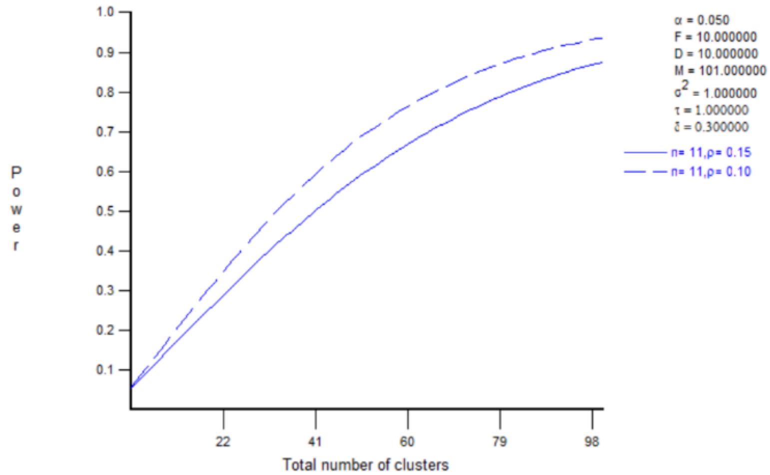


Figure 5. Power and Total Number of Clusters (with duration = 10, occasions = 10)

Thus, this study with a total number of clusters of 130 and duration of 10 and measurement occasions of 10 is sufficient to maintain power at 0.8.

Variations at Level-2 < Variations at Level-3:

In other words, $\sigma^2 < \tau_{11}$, and regions vary greatly compared to within region variations in terms of the growth trajectories. The reliability will converge toward 1.0 (Raudenbush & Bryk, 2002), and in this case the total sample size (J) and cluster size (n) are much more statistically influential than duration and measurement occasions for increasing power. Therefore, assuming the small to medium effect size of 0.3 (Cohen, 1992) of income effect on countries, the results show (Figure 6) the number of clusters (i.e., countries) needed to maintain power at 0.8:

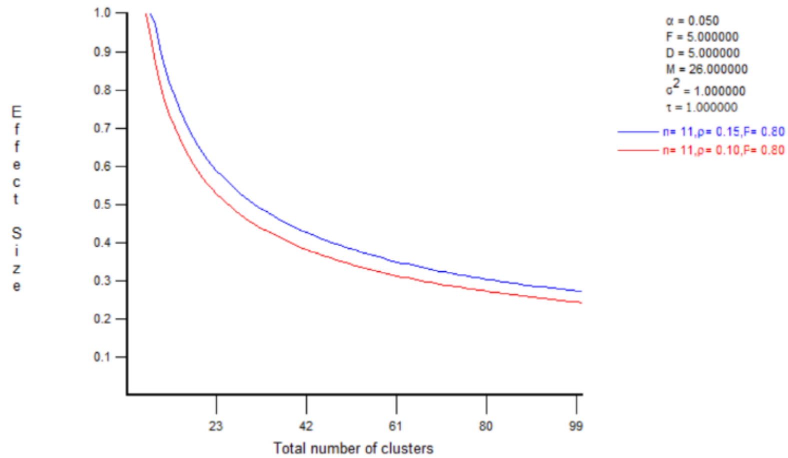


Figure 6. Effect Size and Total Number of Clusters

Thus, the total number of clusters (i.e., for 2 groups of countries) needed to maintain power at 0.8 with a small to medium income effect size, for example, is 80 to 99. Therefore, based on the data included in this study more than 130 clusters (i.e., countries) and a conservative estimation of income effects are sufficient to maintain power at 0.8.

To better check for assumptions of the power analyses, including the definite number of sample sizes required at each level of HLMs, this study will divide the data into subsamples to test if the model fits and if parameter estimates will be significantly different from one another. The results from different subsamples will be summarized into tables and charts and compared to the results yielded from the total sample.

Chapter Four: Results and Discussion

4.1 Data Pre-processing and Assumption Checking

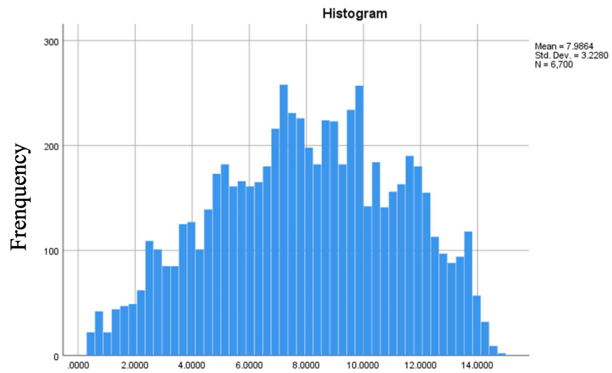
4.1.1 Checking Missing Values with Little's Test

The associations between missing data and variables included in the models are generally reported before model building processes. This is to ensure that problems of estimation bias caused by non-random missingness can be avoided. However, Little's test (Little, 1988) will not be used since data included for HLMs in this study do not have missing values.

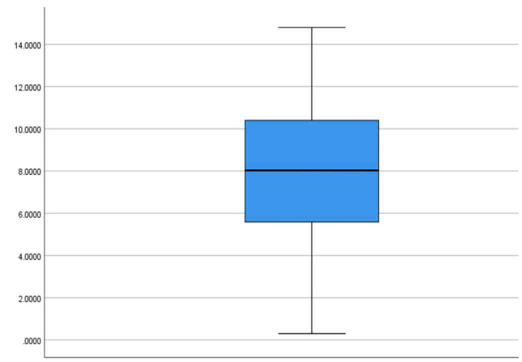
4.1.2 Detecting Univariate Outliers

To avoid the influence of outliers in distorting the statistical estimates including means, variances, correlation coefficients and so forth, univariate outliers will be checked. Following the guidance of Cohen et al. (2003), the outliers will not be addressed and removed if the total number of outliers is less than 1 to 2 percent of the total sample within a specific variable. Thus, box plots will be generated to ensure that there are no extreme outliers in each variable included in the HLMs.

Based on the box plots and histograms below, no extreme influential cases were identified for each of the variables. Therefore, no cases were removed from data at both subnational and national levels.

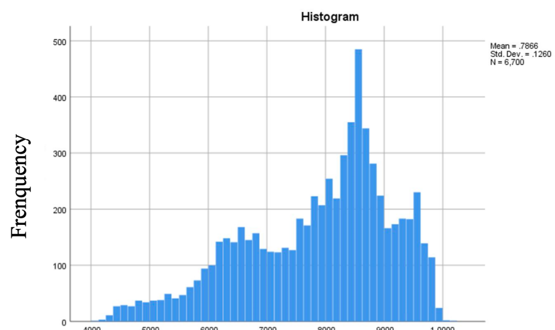


(a) Histogram of MYS

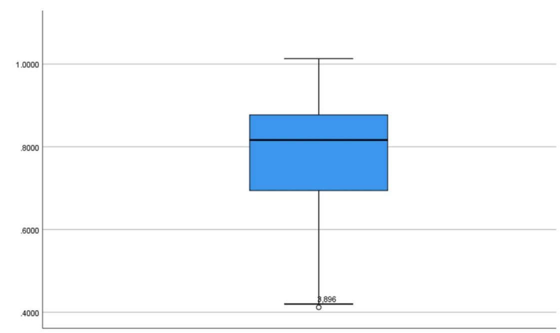


(b) Boxplot of MYS

Figure 7. The (a) Distribution and (b) Boxplot of the Outcome Variable: Mean Years of Schooling Indicator at Level-1

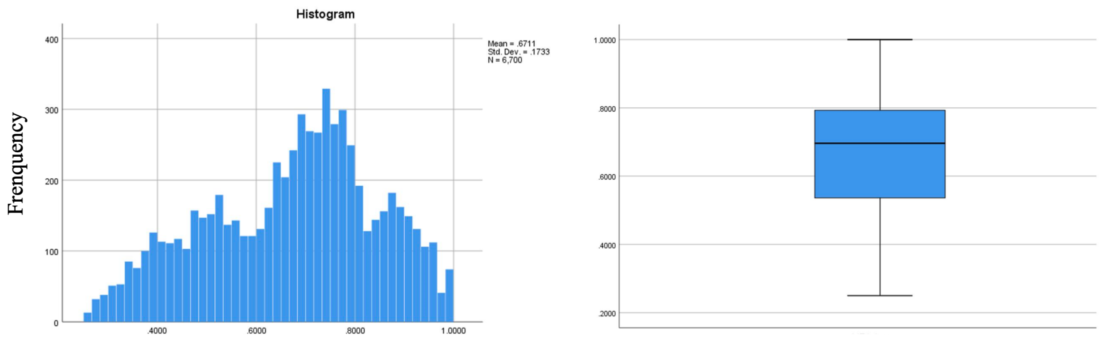


(a) Histogram of HDI Health



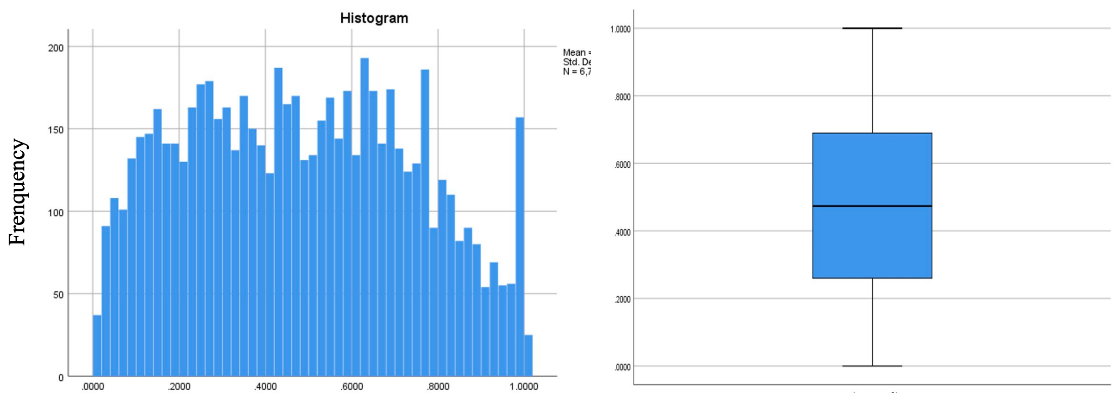
(b) Boxplot of HDI Health

Figure 8. The (a) Distribution and (b) Boxplot of Human Development Index Health Indicator at Level-2



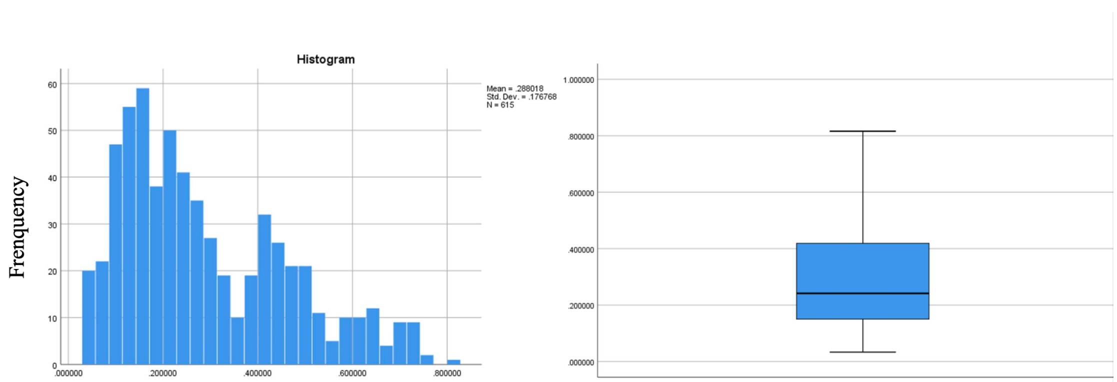
(a) Histogram of HDI Income (b) Boxplot of HDI Income

Figure 9. The (a) Distribution and (b) Boxplot of the Human Development Index Income Indicator at Level-2



(a) Histogram of Urban Population Ratio (b) Boxplot of Urban Population Ratio

Figure 10. The (a) Distribution and (b) Boxplot of Urban Population Ratio at Level-2



(a) Histogram of NLDI (b) Boxplot of NLDI

Figure 11. The (a) Distribution and (b) Boxplot of NLDI at Level-3

4.1.3 Identifying Multivariate Outliers

The Mahalanobis distance (De Maesschalck et al., 2000) is a widely applied metric for detecting multivariate outliers by capturing the extent to which cases differ from the centroid (i.e., the means of all variables) to the other cases and variables:

$$\text{Mahalanobis Distance} = D^2 = (N - 1) * (h_{ii} - \frac{1}{N}) \quad (17)$$

where h_{ii} is the leverage, which indicates the extent to which cases are far from the others, either in the same or off the trend. In addition, the Mahalanobis distance metric can be evaluated with the χ^2 distribution with p (i.e., number of variables) degrees of freedom.

Level-1 and Level-2 Multivariate Outliers

Setting up the p-value associated with significant values of Mahalanobis distance as less than 0.001, influential cases were flagged as "1" and tables below represent the frequencies of influential cases of variables at level-1 and level-2 models. To sum up, among the total of 6700 cases from the subnational level data, 14 multivariate outliers ($p < 0.001$) were identified which consists of 0.2% of the total sample.

Table 5. The Descriptive Statistics on Mahalanobis Distance Estimate

	N	Minimum	Maximum	Mean	S.D.
Mahalanobis Distance	6700	0.1097	25.3317	4.9993	3.3994
Valid N (listwise)	6700				

Table 6. Frequency Statistics of Mahalanobis Distance Flag

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	6686	99.8	99.8	99.8
	1	14	0.2	0.2	100
	Total	6700	100	100	

Level-3 Multivariate Outliers

The tables below summarize the frequencies of multivariate outliers in variables at the national level. The total number of multivariate outliers is 2, which corresponds to 0.3 percent of the total sample in the data.

Table 7. The Descriptive Statistics on Mahalanobis Distance Estimate

	N	Minimum	Maximum	Mean	S.D.
Mahalanobis Distance	615	0.0971	29.1045	3.9935	3.0709
Valid N (listwise)	615				

Table 8. Frequency Statistics of Mahalanobis Distance Flag

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	613	99.7	99.7	99.7
	1	2	0.3	0.3	100
	Total	615	100	100	

4.1.4 Checking Univariate Normality

For each variable included in the HLMs, regardless of the levels at which these variables are used as predictors or outcome variables, Q-Q plots will be generated as they are one of the most popular graphical techniques to visually examine the shape of the distribution (Oppong & Agbedra, 2016). In a Q-Q plot, the quantiles of the sample are plotted against the quantiles that would be expected if the sample came from a normal distribution. Therefore, the sample dots will be in a perfect straight diagonal line if the sample is normally distributed. In other words, the sampling distribution would be considered normal if the data points show a linear trend that is close to the perfect diagonal line in the Q-Q plots.

The formal statistical tests for normality are also used as complementary tools to ensure that the normality assumption is satisfied. The values of skewness, kurtosis, and Kolmogorov-Smirnov tests are reported in this study. From the table below, one can conclude that the values of skewness and kurtosis are in acceptable ranges to retain the assumption of normality (West et al., 1995).

Table 9. Skewness, Kurtosis, and Kolmogorov-Smirnov Test Statistics for Variables at Level-1 and Level-2 Models

	Kolmogorov-Smirnov			skewness	kurtosis
	Statistic	df	sig		
HDI_health	0.098	6700	<.001	-0.63	-0.328
HDI_income	0.062	6700	<.001	-0.306	-0.721
Urban Population Ratio	0.055	6700	<.001	0.117	-1.002
MYS_indicator	0.036	6700	<.001	-0.143	-0.736

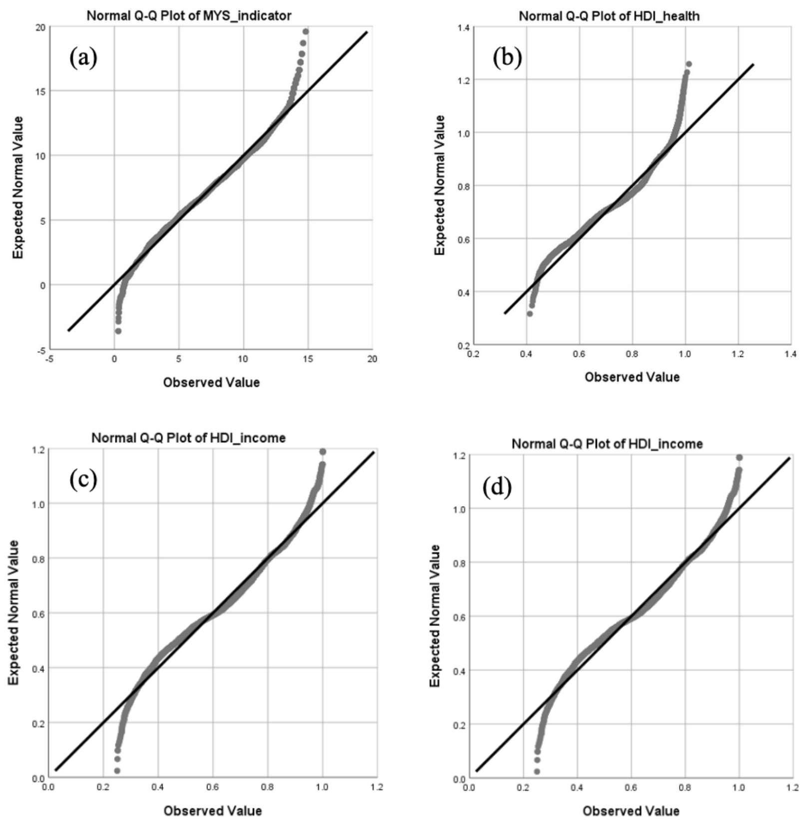


Figure 12. The Q-Q Plots for (a) Mean Years of Schooling Indicator (b) Human Development Index Heath Indicator (c) Human Development Index Income Indicator and (d) Human development Index Income Indicator

4.1.5 Confirming Multivariate Normality

Multivariate normality assumes that each variable in the datasets and all the possible linear combinations of these variables are normally distributed. Moreover, the normality of residuals is assumed if the multivariate normality assumption is met. Therefore, it is necessary to check this assumption since it has a direct impact on the robustness of statistical tests for estimates from HLMs (Micceri, 1989). From the figure below, one can conclude that the assumption of multivariate normality of variables and residuals are satisfied.

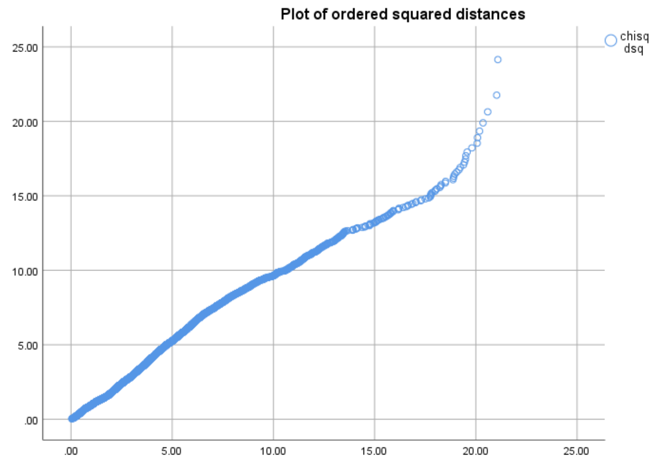


Figure 13. The Multivariate Q-Q Plot

4.2 Model Configurations

4.2.1 Two-Level Unconditional Growth HLMs

Level-1 Model: Individual Subnational Entity Growth Trajectories

As mentioned by Raudenbush and Bryk (2002), three-level model configuration procedures start with building and assessing two-level models. This procedure also applies to building unconditional models without any independent variables to explore and examine the distribution and variations of educational attainment (measured by mean years of schooling). Specifically, the null models without estimating the organizational and characteristic effects of subnational entities and countries are developed. The null models are important because they partition the variance in the outcome variable into within and between individual subnational entity components so that the clustering effects can be tested and confirmed. In addition, level-1 of the null model is constructed to explore and determine the shape of developmental trajectories of education development for individual subnational entities.

The level-1 model specification procedures are organized into three parts: first, the linear time scores will be included to examine the feasibility of linear growth function; second, if the linear growth trajectory is confirmed, the quadratic and other higher order time scores will be added to the level-1 model, for assessing the feasibility of higher functions of growth; and third, to evaluate the shape of developmental trajectories, models with different configurations will be compared and interpreted based on their parameter estimates and deviance statistics. HLM8 is used to build HLMs and analyses on model fits and parameter estimates.

The research question that can be answered by the unconditional two-level model is: what is the shape of developmental trajectories of educational attainment for all subnational entities over the five consecutive years (from 2013 to 2017)? Therefore, the status of educational attainment at time *t* of subnational entity *i* is represented as:

$$Y_{ti} = \pi_{0i} + (\pi_{1i} * TIME\_LIN_{1i}) + (\pi_{2i} * TIME\_QUA_{2i}) + \ldots + (\pi_{pi} * TIME\_F_{pi}) + e_{ti}, \; e_{ti} \sim N(0, \sigma^2) \tag{18}$$

where

Y_{ti} is the mean years of schooling at time *t* for subnational entity *i;*

($TIME\_LIN_{1i}$) represents linear time scores, which are coded as 0 for the starting year 2013, and 1, 2, 3, and 4 for years 2014 to 2017, respectively;

($TIME\_QUA_{2i}$) represents quadratic time scores, which are coded as 0 for the starting year 2013, and 1, 4, 9, and 16 for years 2014 to 2017, respectively;

($TIME\_F_{pi}$) represents time scores of higher order growth functions, and exponential operations are used to indicate the corresponding order of the growth function (e.g., cubic for cubic growth change, etc.);

π_{0i} is the initial status of subnational entity *i.* In other words, it represents the expected status of mean years of schooling for that entity in 2013 (when *TIME_LIN* = 0);

π_{1i} is the linear development rate for entity i, and π_{pi} is the p^{th} function of growth rate for entity *i* during the five consecutive years.

e_{ti} is the residual variance within each individual subnational entity after controlling for the time scores. It is assumed to have a mean of 0 and variance of σ^2 with independent distribution N.

Level-2 Model: Unconditional Subnational Entity Growth

As level-1 model specification mentioned above, at level-2, whether a specific subnational-level model parameter β_{pqk} is included in the model depends on the significance of first-level parameter π_{pjk}. For instance, if the quadratic growth parameter π_{2i} is justified in the level-1 model, then β_{20k} will be included in the level-2 model. However, if evidence of quadratic growth estimates does not suggest the inclusion of π_{2i}. In other words, if the model estimates do not demonstrate quadratic change trajectories for subnational entities, the corresponding level-2 estimate, β_{20k}, will not be included in the model. Therefore, the level-2 model is specified as follows to indicate meaningful random variations in π_{pi}:

$$\pi_{pi} = \beta_{p0} + \sum_{q=1}^{Qp} \beta_{q0} + r_{pi} \, , \tag{19}$$

where

β_{p0} is the average initial status (intercept);

β_{q0} is growth rate of the corresponding growth function;

r_{pi} is the random variation associated with the growth rate, and is assumed to have a normal distribution.

4.3 Results from Two-level Unconditional Models

4.3.1 Model 1: Two-level Unconditional Linear Growth Model

The preliminary results suggest significant random variation effects in π_{0i} and π_{1i} at level-2. As shown in table 10, the reliability estimates for random effects for both π_{0i} and π_{1i} are larger than 0.88. OLS Regression Coefficient Estimate Reliability is another estimate that justifies the applications of HLMs. According to Raudenbush et al. (2019), smaller values of reliability coefficients do not necessarily invalidate the HLM analysis. However, extremely low reliabilities (e.g., < 0.10), often suggest model misspecifications. For example, a random growth parameter might be considered as fixed in subsequent analyses. Thus, the null hypotheses for $\tau_{\pi 0i}$ and $\tau_{\pi 1i} = 0$ are rejected. Thus, by specifying random effects for level-2 outcomes (π_{0i} and π_{1i}), level-2 models are written as:

$$\pi_{0i} = \beta_{00} + r_{0i} \tag{20}$$

$$\pi_{1i} = \beta_{10} + r_{1i} \tag{21}$$

Table 10: Two-level Unconditional Linear Model of Growth in Mean Years of Schooling

Fixed Effect		Coefficient	*S.E.*	t Ratio
Average initial status, β_{00}		7.8254	0.0886	88.321***
Average developmental rate per year, β_{10}		0.0805	0.0024	32.895***
Random Effect	Variance Component	*df*	χ^2	p value
Level 1				
Temporal variation, e_{tij}	0.0090			
Level 2 (subnational entity)				
Initial status, r_{0i}	10.5220	1339	2617116.1957	<0.001
Development rate in mean years of schooling, r_{1i}	0.0071	1339	11980.3977	<0.001

OLS Regression Coefficient Estimate Reliability	
Initial Status, π_{0i}	0.999
Development rate, π_{1i}	0.888

Deviance = 1981.9135 with 4 estimated parameters

Model Parameter Estimates

Fixed Effects

For the fixed effect estimates, average initial status β_{00} and average growth rate β_{10} are 7.8254 and 0.0805, respectively. These indicate that the average mean years of schooling for all subnational entities at the starting year (i.e., 2013) is 7.8254, and the mean years of schooling, on average, increases by 0.0805 each year. The standard errors (*SE*) for these fixed effect estimates, 0.0886 and 0.0024 are relatively small. In addition, the significant *p*-values ($p < 0.001$) associated with these two fixed effect estimates suggest that the initial status of mean years of schooling across subnational entities is statistically significantly different from 0, and subnational entities also demonstrate significant linear growth per year.

Random Effects

Specific to this two-level unconditional linear growth model (i.e., model 1), the random effects estimates refer to variations of growth trajectories for subnational entities that are associated with initial status (i.e., the intercept π_{0i}) and linear growth rate (i.e., π_{1i}). The estimates for variance of intercept and linear growth rate for this model are 10.521 and 0.0071, respectively. In addition, the corresponding χ^2 statistics for two variance components are also significant to reject the null hypotheses. Therefore, the intercept and linear slope are justified to be included in the model.

From results of growth parameters, the values of outcome variable (i.e., mean years of schooling) scatter around the mean intercept of 7.8254 with standard deviation of $(10.521)^{1/2}$. Thus, with 95% of confidence intervals (*CI*s), 95% of the values of the outcome variable for the starting year fall within the range from 1.3383 to 14.3126. For the linear growth rate, values scatter around the mean linear slope of 0.0805 with standard deviation of $(0.0071)^{1/2.}$ Hence, 95% of the linear growth rate values scatter between -0.0881 and 0.2491. To sum up, for all subnational entities included in the model, they vary significantly in terms of their initial status and linear growth rate. For instance, a subnational entity with one standard deviation above the average initial status is expected to have 11.0690 mean years of schooling, and this subnational entity is expected to increase 0.1648 in mean years of schooling per year.

Correlations between Mean Intercept and Mean Linear Slope

The correlation between initial status (i.e., the intercept π_{0i}) and linear growth rate (i.e., π_{1i}) can be obtained through the following formula:

$$\hat{\rho}(\pi_{0i}, \pi_{1i}) = \frac{\hat{\tau}_{01}}{(\hat{\tau}_{00}\hat{\tau}_{11})^{1/2}} = -0.0402/(10.521*0.0071)^{1/2} = -0.1468 \quad (22)$$

where

$\hat{\tau}_{00}$ is the variance associated with mean initial status r_{0i};

$\hat{\tau}_{11}$ is the variance associated with mean linear growth rate r_{1i};

$\hat{\tau}_{01}$ is the covariance between the r_{0i} and r_{1i}.

The estimated correlation coefficient ($\hat{\rho}$ = -0.1468) shows that there is a small to moderate negative correlation (Cohen, 1988) between intercept and linear growth rate. For example, if the individual subnational entity has a higher initial value in mean years of schooling, it is expected to show a slower rate of linear growth.

4.3.2 Model 2: Two-Level Unconditional Quadratic Growth Model

Since the results from model 1 show that there are significant random effects on variations among initial status and linear growth in mean years of schooling for subnational entities, a quadratic growth model will be constructed and evaluated to further explore the shape of developmental trajectories. Therefore, for the quadratic growth model, the mean years of schooling of an individual subnational entity *i* at time *t* at level-1 is:

$$Y_{ti} = \pi_{0i} + \pi_{1i} *(\text{TIME\_LINEAR}_{1i}) + \pi_{2i} *(\text{TIME\_QUADRATIC}_{2i}) + e_{ti.} \quad (23)$$

Compared to Model 1, a new function of growth π_{2i}, the quadratic growth parameter is added to explore the feasibility of quadratic change. Noticeably, since this study does not impose any specialized coding scheme, there are no issues related to centering for time scores. Thus, the quadratic time scores are coded with squared linear time scores for that specific year.

The level-1 model parameter estimates become outcome variables at higher levels, which lead to the following model configurations at level 2:

$$\pi_{0i} = \beta_{00} + r_{0i} \tag{24}$$

$$\pi_{1i} = \beta_{10} + r_{1i} \tag{25}$$

$$\pi_{2i} = \beta_{20} + r_{2i} \tag{26}$$

where

β_{20} represents the average quadratic rate of change, and

r_{2i} is the random variation associated with the mean quadratic slope.

Table 11: Two-level Unconditional Quadratic Model of Growth in Mean Years of Schooling

Fixed Effect		Coefficient	*S.E.*	t Ratio
Average initial status, β_{00}		7.8120	0.0885	88.265***
Average linear growth rate per year, β_{10}		0.1073	0.0049	22.067***
Average quadratic growth rate per year, β_{20}		-0.0067	0.0008	-8.082***
Random Effect	Variance Component	*df*	χ^2	*p* value
Level 1				
Temporal variation, e_{tij}	0.0067			
Level 2 (subnational entity)				
Initial status, r_{0i}	10.4986	1339	2373870.3484	<0.001
Linear growth rate in mean years of schooling, r_{1i}	0.0234	1339	5109.6068	<0.001

Quadratic growth rate in mean years of schooling, r_{2i}	0.0005	1339	2588.2494	<0.001

OLS Regression Coefficient Estimate Reliability	
Initial Status, π_{0i}	0.999
Linear growth rate, π_{1i}	0.738
Quadratic growth rate, π_{2i}	0.483

Deviance = 1420.4714 with 7 estimated parameters

Model Parameter Estimates

Fixed Effects

The average initial status β_{00} and average linear growth rate β_{10} are 7.8120 and 0.1073, respectively. Compared to model 1, estimate reliability coefficients for model 2 are also high (i.e., above 0.738). This means that the mean initial status and linear slope coefficients (i.e., π_{0i} and π_{1i}) do not vary significantly from those of model 1. Thus, on average, mean years of schooling is 7.8120 for subnational entities at the starting year, and these entities increase at a linear growth rate of 0.1073 for the following years. Moreover, the quadratic slope coefficient is -0.0067 ($p < 0.001$), which means that subnational entities also demonstrate an instantaneous negative rate of change. In other words, for a subnational entity that shows a greater positive linear growth, it is also expected to have a slower quadratic rate of change in mean years of schooling.

Random Effects

The standard deviation of individual trajectories within subnational entities is 0.0819 in mean years of schooling $(\sigma^2 = 0.0067)^{1/2}$. Moreover, the χ^2 statistics associated with variance in π_{0i}, π_{1i}, and π_{2i} are all statistically significant ($p < 0.001$). These indicate that there are significant differences among subnational entities with respect to mean initial status, linear and quadratic rates of change. Specifically, the values of outcome variable (i.e., mean years of schooling) scatter around the average intercept ($\beta_{00} = 7.8120$) with standard deviation of $(r_{0i} = 10.4986)^{1/2}$. Thus, approximately 95% of the values of the outcome variable for the starting year fall within the range from 1.3316 to 14.2924. For the linear growth rate parameter, values scatter around the mean linear slope with standard deviation of $(r_{1i} = 0.02341)^{1/2.}$ Hence, 95% of the linear growth rates scatter between -0.1987 and 0.4133. As for instantaneous quadratic growth with standard deviation of $(r_{2i} = 0.0005)^{1/2}$, 95% of the quadratic slopes are in the range between -0.0515 and 0.0381.

Model Comparison

To determine whether model 1 or model 2 is a closer fit for the data, the model deviance statistics with associated degrees of freedom (*df*) are evaluated. Model 1 with only linear growth (1981.9135 with 4 *df*) is compared with model 2 where a quadratic growth rate is included (1420.4714 with 7 *df*). As a result, the difference between these two deviance statistics is 561.4422, with an approximate χ^2 distribution with 3 *df*. The difference in two deviance statistics is statistically significant ($p < 0.001$). Therefore, the result implies that model 2 with a smaller value of deviance statistic is preferred, and the quadratic growth parameters should remain in the models.

Moreover, after a closer examination on estimates of variance components from model 2, Chi-square statistics for both mean intercept and mean linear slope are reduced compared to those of model 1. Nevertheless, the values remain significant. These indicate that there are still unexplained variances in growth parameters. Thus, to further explore the developmental trajectories, an additional growth function (i.e., cubic growth) will be added to examine whether the added growth function can further explain the random variations in these growth parameters.

4.3.3 Model 3: Two-Level Unconditional Cubic Growth Model

Assuming that educational attainment of subnational entities is progressing in a cubic growth pattern, the mean years of schooling of an individual subnational entity *i* at time *t* can be written as:

$$Y_{ti} = \pi_{0i} + (\pi_{1i} * TIME\_LIN_{1i}) + (\pi_{2i} * TIME\_QUA_{2i}) + (\pi_{3i} * TIME\_CUB_{3i}) + e_{ti} \tag{27}$$

where

$TIME\_CUB_{3i}$ stands for the cubic growth function for subnational entity *i;*

π_{3i} is the cubic rate of change (i.e., slope).

The level-1 model configurations further lead to the following model specification details at level-2:

$$\pi_{0i} = \beta_{00} + r_{0i} \tag{28}$$

$$\pi_{1i} = \beta_{10} + r_{1i} \tag{29}$$

$$\pi_{2i} = \beta_{20} + r_{2i} \tag{30}$$

$$\pi_{3i} = \beta_{30} \tag{31}$$

where

β_{30} is the mean cubic rate of change, and for model building and comparison purposes, π_{3i} is specified as non-randomly varying.

Table 12: Two-level Unconditional Cubic Model of Growth in Mean Years of Schooling (with non-randomly varying cubic slope)

Fixed Effect		Coefficient	*S.E.*	t Ratio
Average initial status, β_{00}		7.8139	0.0885	88.326***
Average linear growth rate per year, β_{10}		0.0939	0.0079	11.836***
Average quadratic growth rate per year, β_{20}		0.0027	0.0043	0.6170
Average cubic growth rate per year, β_{30}		-0.0016	0.0007	-2.382*
Random Effect	Variance Component	*df*	χ^2	*p* value
Level 1				
Temporal variation, e_{tij}	0.0067			
Level 2 (subnational entity)				
Initial status, r_{0i}	10.4986	1339	2379277.7081	<0.001
Linear growth rate in mean years of schooling, r_{1i}	0.0234	1339	5121.2458	<0.001
Quadratic growth rate in mean years of schooling, r_{2i}	0.0005	1339	2594.1451	<0.001

OLS Regression Coefficient Estimate Reliability	
Initial Status, π_{0i}	0.999
Linear growth rate, π_{1i}	0.739
Quadratic growth rate, π_{2i}	0.484

Deviance = 1428.3170 with 7 estimated parameters

Model Parameter Estimates

Fixed Effects

From the table 12 above, the mean intercept and linear rate of change do not vary from those of model 2, and the reliability coefficients for both the two fixed effects estimates are above 0.739. However, after including the cubic growth parameter at level-1 model, the quadratic slope varies from -0.0067 (model 2) to 0.0027, which is not statistically significant. This implies certain degrees of model misspecifications for growth parameters. In addition, the cubic slope is -0.0016 ($p < 0.05$), which means that on average, a subnational entity with a higher initial status, and linear and quadratic rates of change also demonstrates a slower cubic change in mean years of schooling over time.

To examine and compare models, deviance statistics of model 3 (Deviance = 1428.3170 with 7 *df*) and model 2 (Deviance = 1420.4714 with 7 *df*) are evaluated. A non-significant result of difference in deviance statistics ($\chi^2 = 7.8456$, $p > 0.50$) shows that the inclusions and specifications of the cubic growth parameters $TIME\_CUB_{3i}$, π_{3i}, and β_{30} can be inappropriate. Hence, instead of constraining the cubic slope as non-randomly varying, the slope will be constructed as a random effect, which makes the configurations of level-2 models as follows:

$$\pi_{0i} = \beta_{00} + r_{0i} \tag{32}$$

$$\pi_{1i} = \beta_{10} + r_{1i} \tag{33}$$

$$\pi_{2i} = \beta_{20} + r_{2i} \tag{34}$$

$$\pi_{3i} = \beta_{30} + r_{3i} \tag{35}$$

Table 12a: Two-level Unconditional Cubic Model of Growth in Mean Years of Schooling (with random cubic slope)

Fixed Effect		Coefficient	*S.E.*	t Ratio
Average initial status, β_{00}		7.8139	0.0885	88.326***
Average linear growth rate per year, β_{10}		0.0939	0.0079	11.836***
Average quadratic growth rate per year, β_{20}		0.0027	0.0043	0.6170
Average cubic growth rate per year, β_{30}		-0.0016	0.0007	-2.382*
Random Effect	Variance Component	*df*	χ^2	*p* value
Level 1				
Temporal variation, e_{tij}	0.0047			
Level 2 (subnational entity)				
Initial status, r_{0i}	10.4904	1339	3056535.0640	<0.001
Linear growth rate in mean years of schooling, r_{1i}	0.0553	1339	3799.8838	<0.001
Quadratic growth rate in mean years of schooling, r_{2i}	0.0133	1339	2770.9923	<0.001
Cubic growth rate in mean years of schooling, r_{3i}	0.0003	1339	2378.8944	<0.001

OLS Regression Coefficient Estimate Reliability	
Initial Status, π_{0i}	0.999
Linear growth rate, π_{1i}	0.650
Quadratic growth rate, π_{2i}	0.526
Cubic growth rate, π_{3i}	0.453

Deviance = 888.0970 with 11 estimated parameters

Model Parameter Estimates

Fixed Effects

From the table shown above, the fixed effects estimates are identical to model-3 where the cubic growth slope is specified as non-randomly varying. Moreover, the quadratic slope coefficient remains non-significant ($\beta_{20} = 0.0043$, $p > 0.50$). This can be caused by the large amount of unexplained variance components introduced to the model with the inclusions of cubic growth parameters. Another hypothesis for this is that subnational entities are showing non-traditional growth patterns, such as piecewise growth trajectories where the first two years can be viewed as the initial stage, and last two to three years are demonstrating higher order functions of growth.

Random Effects

The random variation associated with the average intercept remains significant and does not vary significantly from that of the model with fixed effects cubic growth parameters. Thus, the results reject the null hypothesis that subnational entities do not vary in terms of their initial status. Moreover, the random variance associated with linear slope is 0.0234 ($p < 0.001$), and this suggests that there are also significant variations in linear growth at the subnational level. The presence of significant variations also applies to quadratic ($r_{2i} = 0.0133$, $p < 0.001$) and cubic growth parameters ($r_{3i} = 0.0003$, $p < 0.001$). Hence, these significant variance components suggest the inclusions of both quadratic and cubic growth parameters. More importantly, the non-significant fixed coefficient for quadratic slope indicates that other coding schemes such as piecewise schemes can be used to indicate alternative forms of developmental trajectories.

Model Comparison

To examine and compare models, deviance statistics of model 3 with random effect specified to cubic growth parameters (Deviance = 888. 0970 with 11 *df*) and model 2 (Deviance = 1420.4714 with 7 *df*) are evaluated. The result of difference in deviance statistics is statistically significant ($\chi^2 = 532.3744$, $p < 0.001$), which shows the necessity of the inclusion of higher order growth parameters such as $TIME\_CUB_{3i}$, π_{3i}, and β_{30}. Thus, the two-level unconditional piecewise growth model will be specified and evaluated.

4.3.4 Model 4: Two-Level Piecewise Unconditional Quadratic Growth Model

According to Raudenbush and Bryk (2002), piecewise growth modeling can be an alternative and option to analyze curvilinear trajectories when the exploratory analyses show evidence of non-linearity of growth trajectories. In particular, this approach explores different forms of developmental patterns by separating curvilinear trajectories into discrete linear components. Therefore, piecewise coded growth models assume distinctive forms of growth for different time periods.

There are empirical research findings that support the distinct forms of growth for different periods in terms of education development across countries and cultures (Van Deursen et al., 2015; Avendano et al., 2009). Based on the empirical evidence and literature, this specific piecewise model will separate forms of growth into the following two time periods: (1) from 2013 to 2014, and (2) from 2015 to 2017. The summarized codes for piecewise time scores are presented in the table below:

Table 13. Coding Schemes for the Unconditional Two-Piece Quadratic Growth Model

Two-Rate Model						
			Years			
	2013	2014	2015	2016	2017	Interpretation of πs:
a_{1ti}	0	0	1	2	3	π_{1i} growth rate period 1
a_{2ti}	0	0	1	4	9	π_{2i} growth rate period 2

This model is proposing a faster growth rate in later years than earlier ones. For period 1, the linear growth rate is hypothesized, and a second growth rate (i.e., quadratic) is hypothesized for period 2. These two periods consist of the level-1 individual subnational entity growth model expressed in the following equation:

$$Y_{ti} = \pi_{0i} + (\pi_{1i} * a_{1ti}) + (\pi_{2i} * a_{2ti}) + e_{ti} \tag{36}$$

where

a_{1ti} and a_{2ti} are coded time scores to indicate the piecewise regression.

Thus, the level-2 models become:

$$\pi_{0i} = \beta_{00} + r_{0i} \tag{37}$$

$$\pi_{1i} = \beta_{10} + r_{1i} \tag{38}$$

$$\pi_{2i} = \beta_{20} \tag{39}$$

Table 13a. Two-Level Piecewise Unconditional Quadratic Growth Model in Mean Years of Schooling

Fixed Effect	Coefficient	*S.E.*	t Ratio
Average initial status, β_{00}	7.8620	0.0884	88.905***
Average linear growth rate per year, β_{10}	0.1542	0.0060	25.870***
Average quadratic growth rate per year, β_{20}	-0.0216	0.0013	-16.367***

Random Effect	Variance Component	*df*	χ^2	*p* value
Level 1				
Temporal variation, e_{tij}	0.0129			
Level 2 (subnational entity)				
Growth rate for period 1, r_{0i}	10.4825	1339	265228.6220	<0.001
Growth rate for period 2, r_{1i}	0.0088	1339	7539.1511	<0.001

OLS Regression Coefficient Estimate Reliability	
Growth rate 1, π_{0i}	0.999
Growth rate 2, π_{1i}	0.822

Deviance = 3292.7607 with 4 estimated parameters

Model Parameter Estimates

Fixed Effects

The average initial status for subnational entities is 7.8620, and for period 1 (2013 to 2014), the mean growth rate is 0.1542. For period 2 (2015 to 2017), the mean developmental rate is -0.0216. The growth rate coefficients for both periods are statistically significant, which indicates that both growth rates at two periods are significantly different from 0.

Random Effects

The random variation associated with growth rate for period 1 is 10.4825 ($p < 0.001$). Therefore, the subnational entities vary significantly in terms of the growth rate for period 1. Moreover, the random variation associated with developmental rate for period 2 is 0.0088 ($p < 0.001$), and this indicates that for period 2, the growth rates of subnational entities also vary significantly from one another. However, compared to previous models (e.g., model 2) with traditional coding schemes, this model yields greater values of Chi-square statistics for random components. Thus, compared to the traditional quadratic growth model, piecewise model configuration does not demonstrate a better fit for characterizing developmental trajectories in educational attainment at the subnational level.

The piecewise quadratic growth model also yields high reliability for parameter estimates and demonstrates an alternative form of trajectories. Thus, in building three-level unconditional models, this study will continue to explore this type of trajectories and have it compared with traditional development models.

Summary

Table 14. Model Comparison with Deviance Statistics

Model	*Model 1*		*Model 2*		*Model 3*		*Model 4*	
Fit Index	DS	*df*	DS	*df*	DS	*df*	DS	*df*
	1981.9135	4	1420.4714	7	888.0970	11	3292.7607	4

Note. DS (Deviance Statistic); df (Number of Parameters Estimated)

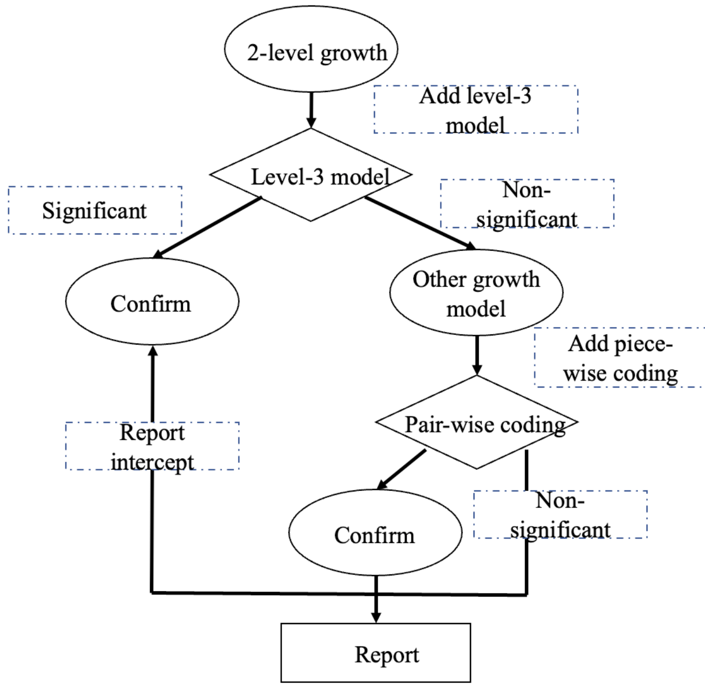


Figure 14. Decision Tree for Building Unconditional Growth Models

Key findings from two-level unconditional models are of great importance because they determine the shape of growth trajectories for subnational entities, and the findings are: (1) model 4 has the worst model performance with the largest value of deviance statistic; (2) with the same number of parameters being estimated, model 1 demonstrates a better fit for the data. Therefore, we can conclude that the two-piece growth trajectory does not fit the data adequately; (3) model 3 yields the best model fit results. However, the quadratic growth becomes nonsignificant with the inclusion of higher order growth function. Thus, cubic growth cannot be established; and (4) model 2 is selected among these models with a significantly smaller value of deviance statistic compared with model 1 ($\chi^2_{(3)} = 561.4422$, $p < 0.001$). To conclude, for subnational entities, a quadratic growth trajectory is evaluated and confirmed.

4.4 Three-Level Unconditional Growth Models

4.4.1 Model 5: Unconditional Three-Level Quadratic Growth Model

Since the quadratic growth pattern is validated at the subnational level, this three-level unconditional growth model will be constructed with identical model specifications to model 2. Meanwhile, a level-3 model is added upon the validated two-level quadratic growth model (model 2). These level-3 model equations are specified to answer the following question: do countries progress in a quadratic pattern in terms of the development of educational attainment?

The model equations from different levels include:

(1) the level-1 model regarding developmental trajectories of educational attainment for individual subnational entity *i* in country *j* at different time occasions *t*:

$$Y_{tij} = \pi_{0ij} + \pi_{1ij} *(\text{TIME\_LINEAR}_{1ij}) + \pi_{2ij} *(\text{TIME\_QUADRATIC}_{2ij}) + e_{tij}, \quad e_{tij} \sim N(0, \sigma^2) \tag{40}$$

(2) the level-2 model equations with respect to variability in growth rates for subnational entities:

$$\pi_{0ij} = \beta_{00j} + r_{0ij} \tag{41}$$

$$\pi_{1ij} = \beta_{10j} + r_{1ij} \tag{42}$$

$$\pi_{2ij} = \beta_{20j} + r_{2ij} \tag{43}$$

and,

(3) the level-3 model equations to investigate country-level variations in growth trends:

$$\beta_{00j} = \gamma_{000} + u_{00j} \tag{44}$$

$$\beta_{10j} = \gamma_{100} + u_{10j} \tag{45}$$

$$\beta_{20j} = \gamma_{200} \tag{46}$$

where

β_{00j} is the mean initial status within country j,

γ_{000} is the average initial status across countries,

β_{10j} represents average linear growth rate of five consecutive years within country j,

γ_{100} represents average linear growth rate of five consecutive years across countries j,

β_{20j} represents average quadratic growth rate of five consecutive years within country j,

γ_{200} represents average quadratic growth rate of five consecutive years across countries j.

Model Parameter Estimates

Table 15 displays results of fixed and random effects estimates of this three-level unconditional quadratic growth model with the deviance statistic. Specifically, β_{20j} (i.e., the average quadratic growth rate within a country) is initially configured as a random effect estimate, and the non-significant random variation component (u_{20j} = 0.0001, p > 0.50) suggests that β_{20j} should be configured as a fixed effect.

Fixed Effects

The fixed effects estimates summarized in table 15 indicate that the average mean years of schooling across all countries (γ_{000}) starts at 7.7744 in 2013 and increases at a linear rate of 0.1063 (γ_{100}) and a quadratic rate of -0.0067 (γ_{200}) per year. These coefficients are close to those mean initial status (7.8120), linear growth rate (0.1073), and a quadratic rate (-0.0067) for individual subnational entities (see Table 11), indicating that subnational entities are approximately evenly distributed across countries. In addition, the small standard errors for these three between country fixed parameters (0.2794, 0.0118, and 0.0023) also suggest that the true estimates fall into relatively narrow ranges (i.e., narrow CIs).

Table 15. Three-Level Unconditional Quadratic Growth Model in Mean Years of Schooling

Fixed Effect		Coefficient	*S.E.*	t Ratio
Average initial status across countries, γ_{000}		7.7744	0.2794	27.822***
Average linear growth rate per year across countries, γ_{100}		0.1063	0.0118	8.998***
Average quadratic growth rate per year across countries, γ_{200}		-0.0067	0.0023	-2.897**
Random Effect	Variance Component	*df*	χ^2	*p* value
Level 1				
Temporal variation, e_{tij}	0.0065			
Level 2 (subnational entity)				
Initial status, r_{0ij}	1.4730	1220	316192.5857	<0.001
Linear growth rate in mean years of schooling, r_{1ij}	0.0295	1220	6280.4920	<0.001
Quadratic growth rate in mean years of schooling, r_{2ij}	0.0005	1339	2675.9412	<0.001
Level 3 (country)				
Initial status, u_{00j}	9.3312	119	8807.3988	<0.001
Linear growth rate, u_{10j}	0.0036	119	1196.9674	<0.001

OLS Regression Coefficient Estimate Reliability	
Initial Status, π_{0i}	0.996

Linear growth rate, π_{1i}	0.786
Quadratic growth rate, π_{2i}	0.514

Deviance = -1203.5596 with 13 estimated parameters

Random Effects

The second panel of table 15 shows the partitioned variance of initial status, linear, and quadratic growth parameters into within and between country components. Significant variations are found within countries (among subnational entities) for initial status, growth rates (r_{0ij}, r_{1ij}, and r_{2ij}) as well as between country average initial status and growth rates (u_{00j} and u_{10j}). Moreover, by comparing χ^2 statistics for corresponding parameter estimates, one can conclude that the variations in initial status and linear growth rate between countries are smaller than the variations within countries. Therefore, countries demonstrate less variability in developmental trajectories in educational attainment than subnational entities.

Based on the estimates of variance components, the proportion of variation that lies between countries to the total variation of both initial status and linear growth rate can be calculated to examine the magnitudes of clustering effects (i.e., ICCs). In other words, values of unconditional models' ICCs will be evaluated to examine the magnitudes of clustering effects. In particular, below is the percentage of variance in initial status explained at the country level:

$$\tau_{\beta 00} / \tau_{\beta 00} + \tau_{\pi 00} = 9.3312/ 9.3312 + 1.4730 = 0.8637 \tag{47}$$

and percentage of variance in linear growth rate accounted by countries is:

$$\tau_{\beta 11} / \tau_{\beta 11} + \tau_{\pi 11} = 0.0036/\ 0.0036 + 0.0295 = 0.1088 \tag{48}$$

Thus, approximately 86% of the total variance in initial status lies between countries. The result of the percentage of variance in linear growth rate is, however, significantly lower. Only about 11% of the variance is explained by countries. These results indicate that countries differ less in terms of their initial status of educational attainment, and they tend to vary significantly in their linear growth rates. In order to explore country level effects on developmental patterns, some country-level characteristics will be included in the three-level conditional model to explain the variability in educational attainment development.

Variance-Covariance Components

Another approach to examining the within and between country effects is to decompose the correlations between initial status and growth rates into within and between country components. Results show that within a country, the estimated correlation between initial status and linear growth rate is -0.199, and this correlation is stronger at the country level (-0.270).

The variance-covariance and correlation matrices at level-2 and level-3 are presented as:

$$\text{Level 2} \begin{pmatrix} 1.4730 & -0.199 & 0.242 \\ -0.199 & 0.0295 & -0.999 \\ 0.242 & -0.999 & 0.0005 \end{pmatrix} = \widehat{\mathrm{T}}_{\pi} = \begin{pmatrix} \hat{\tau}_{\pi 00} & \tau_{\pi 01} & \tau_{\pi 02} \\ \tau_{\pi 10} & \hat{\tau}_{\pi 11} & \tau_{\pi 12} \\ \tau_{\pi 20} & \tau_{\pi 21} & \hat{\tau}_{\pi 22} \end{pmatrix}$$

$$\text{Level 3} \begin{pmatrix} 9.3312 & -0.27 \\ -0.27 & 0.0036 \end{pmatrix} = \widehat{\mathrm{T}}_{\pi} = \begin{pmatrix} \hat{\tau}_{\beta 00} & \tau_{\beta 01} \\ \tau_{\beta 10} & \hat{\tau}_{\beta 11} \end{pmatrix} \tag{49}$$

To sum up, the three-level unconditional model provides parameter estimate values of great importance to illustrate shapes of developmental trajectories in educational attainment across subnational entities and countries. More importantly, this model provides decomposed variability and variance components in growth parameters at each level. The correlations between growth parameters and partitioned variability indicate an important characteristic of the data: there is a high percentage of explained variation accounted for at the country level (i.e., level-3).

4.4.2 Model 6: Unconditional Three-level Piecewise Growth Model

From previous comparisons among two-level growth models, the model fit deviance statistics for different models show a possibility that growth rates may appear faster and more variable in early years (2013-2014) than the latter period (2015-2017). In addition, the traditional cubic growth model fails to fit the data adequately, which means that the traditional growth model can be insufficient to explain the variability of growth parameters at each level. Thus, an alternative solution is to construct a three-level piecewise growth model to evaluate the assumption of distinct growth patterns for two different periods.

Since the previous two-level piecewise growth model (model 4) yields high reliability coefficients for parameter estimates, this three-level piecewise quadratic growth model is used to investigate the following questions: (1) For countries, is there more variability in education development at early years than later years? and (2) At the country level, do the correlations between growth parameters differ in these two periods? To address the questions, the level-1 model is specified in the following form:

$$Y_{tij} = \pi_{0ij} + \pi_{1ij} * a_{1tij} + \pi_{2ij} * a_{2tij} + e_{tij}, \quad e_{tij} \sim \mathrm{N}(0, \sigma^2) \tag{50}$$

where

a_{1tij} and a_{2tij} are coded time scores.

As defined in Table 13 regarding the piecewise coding scheme, the level-1 model specification is as identical as the level-1 equation of model 4. This specification is to reflect the hypothesis assuming a linear growth pattern for period 1, and a quadratic growth for period 2.

Different from model 4, the level-2 equations of this model assume random effect components associated with (1) initial status π_{0ij}; and (2) growth parameter for period 1 (π_{1ij}). As for the growth parameter for period 2 (π_{2ij}), a fixed effect estimate is configured since the random variance component is tested as nonsignificant ($\chi^2_{(1339)} = 1330.5218$, $p > 0.5$). Therefore, the level-2 equations are constructed as follows:

$$\pi_{0ij} = \beta_{00j} + r_{0ij} \tag{51}$$

$$\pi_{1ij} = \beta_{10j} + r_{1ij} \tag{52}$$

$$\pi_{2ij} = \beta_{20j} \tag{53}$$

and for the country-level model at level-3, random effects are assumed to be associated with each growth parameters: (1) the average initial status within country j (u_{00j}); (2) the average growth rate for period 1 within country j (u_{10j}); and (3) average growth rate for period 2 within country j (u_{20j}).

$$\beta_{00j} = \gamma_{000} + u_{00j} \tag{54}$$

$$\beta_{10j} = \gamma_{100} + u_{10j} \tag{55}$$

$$\beta_{20j} = \gamma_{200} + u_{20j} \tag{56}$$

Table 16. Three-Level Piecewise Quadratic Growth Model in Mean Years of Schooling

Fixed Effect		Coefficient	*S.E.*	t Ratio
Average initial status across countries, γ_{000}		7.8234	0.2797	27.969***
Average linear growth rate per year across countries, γ_{100}		0.1497	0.0145	10.294***
Average quadratic growth rate per year across countries, γ_{200}		-0.0204	0.0036	-5.652***
Random Effect	Variance Component	*df*	χ^2	*p* value
Level 1				
Temporal variation, e_{tij}	0.0113			
Level 2 (subnational entity)				
Initial status, r_{0ij}	1.4207	1220	372403.7602	<0.001
Linear growth rate, r_{1ij}	0.0058	1220	5481.9715	<0.001
Level 3 (country)				
Initial status, u_{00j}	9.1958	119	8625.4875	<0.001
Linear growth rate, u_{10j}	0.0219	119	967.4030	<0.001
Quadratic growth rate, u_{20j}	0.0013	119	659.0907	<0.001
OLS Regression Coefficient Estimate Reliability				
Initial Status, π_{0i}			0.997	
Linear growth rate, π_{1i}			0.777	
Deviance = 506.1294 with 13 estimated parameters				

Model Parameter Estimates

Fixed Effects

The mean initial status across countries, γ_{000}, is 7.8234, and the linear growth rate for period 1 across countries is 0.1497. These two fixed effect parameters are not significantly different than those from the conventional three-level quadratic growth model (model 5). However, the significant difference is found in estimation of the quadratic growth rate for period 2, γ_{200}, which equals -0.0204. This indicates that on average, countries tend to increase by 0.1497 in mean years of schooling for period 1and then to decelerate for period 2.

Random Effects

The random effect estimates at both level-2 and level-3 indicate that subnational entities and countries vary significantly in terms of initial status and growth rate for period 1. In addition, from the results decomposing the variance components into within and between country levels, one can see that approximately 87 percent of total variance in initial status and 79 percent of total variance in growth rate for period 1 are explained by level 3 units (countries). Thus, this two-piece quadratic growth model indicates higher proportions of explained variances accounted for by level 3 units:

$$\tau_{\beta 00} / \tau_{\beta 00} + \tau_{\pi 00} = 9.1958 / 9.1958 + 1.4207 = 0.8662 \tag{57}$$

$$\tau_{\beta 11} / \tau_{\beta 11} + \tau_{\pi 11} = 0.0219 / 0.0219 + 0.0058 = 0.7906 \tag{58}$$

Moreover, the significant random parameter estimates at both level-2 and level-3 suggest that for period 2, countries also vary significantly from one another.

Model Comparison

Table 17. Model Comparison with Deviance Statistics

Model Fit Index	*Model 5*		*Model 6*	
	DS	*df*	DS	*df*
	-1203.5596	13	506.1294	13

Note. DS (Deviance Statistic); df (Number of Parameters Estimated)

Since the conventional three-level quadratic growth model and three-level piecewise quadratic growth model have the same number of estimated parameters (n = 13), the difference in deviance statistics for these two models is nonsignificant ($\chi^2_{(0)}$ = 1709.6890, p > .50). Nevertheless, the conventional quadratic growth model is selected due to: (1) its simpler coding scheme for time scores at level-1 and (2) smaller value of the deviance statistic. Thus, model 5 is selected for conditional model building, in other words, predictors that highlight characteristics of different analysis units' levels will be included to identify significant features that can be used to explain the variability of developmental trajectories.

Summary

Results from the three-level unconditional model are crucial because: (1) the growth patterns for countries are evaluated and confirmed; (2) variances are decomposed into level-2 and level-3 components; and (3) correlations between growth parameters are examined.

To conclude, three-level unconditional model results include: (1) the variability across countries are smaller than that of subnational entities, which means that countries are much more similar in development trajectories; (2) large proportions of variances in growth parameters are explained by countries rather than subnational entities. Therefore, strong clustering effects at the country level are examined; and (3) quadratic growth trajectories are also validated for countries. Given the key findings, strong emphasis will be put to explore three-level conditional models to identify significant country level characteristics that affect the growth trajectories of educational attainment.

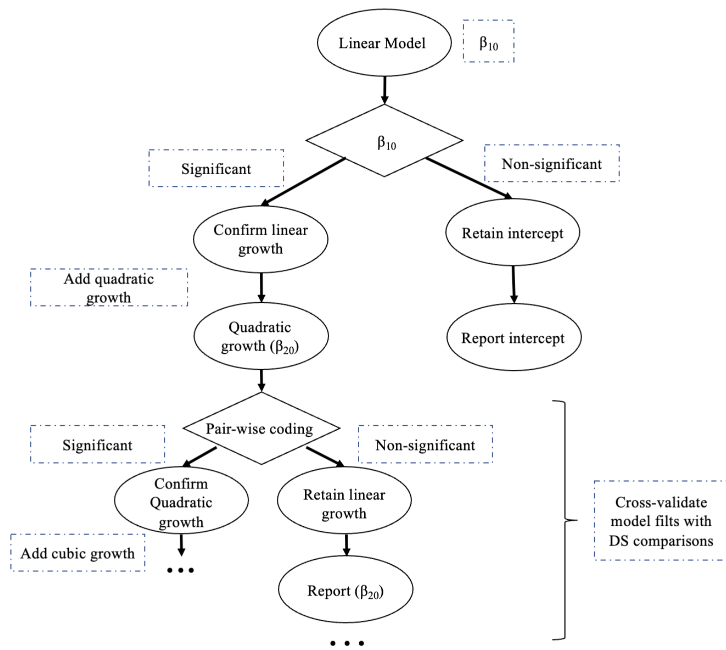


Figure 15. Three-level Unconditional Model Building Flow Chart

4.5 Characteristics of Subnational Entities that Explain Differences in Education Development

4.5.1 Configurations of Two-Level Conditional Models

The two-level unconditional models constructed in the previous section mainly focus on estimating growth parameters to determine the shape of growth with only time scores included in models. Thus, in this section, conditional models will be discussed and compared especially with various predictors being included at level-2 model equations.

By adding other variables at level-2 models, especially the covariates of the income, health indicator, and urban population ratio, the two-level conditional models are built to investigate: (1) whether educational attainment of a subnational entity with a higher income level develops at a faster rate than a subnational entity at a lower income level; (2) whether a subnational entity with a greater urban population ratio or higher health level increases at a faster rate in educational attainment than that of a subnational entity with a lower urban population ratio or health level; (3) whether growth rates and development parameters are significantly correlated; and (4) to which extent the variability in growth parameters can be explained by characteristics of subnational entities. Thus, to address different questions, the following two-level conditional growth HLMs will be specified.

4.5.2 Model 7: Two-Level Conditional Quadratic Growth Model with the Effect of Income

Model 7 is constructed by including a continuous predictor at level-2, HDI income, to answer the question: whether income levels indicating the economic development levels of subnational entities can be used to explain the variability in growth (initial status, linear and quadratic growth rates) for educational attainment.

In particular, the level-1 model equations of model 7 remain the same as those of the two-level unconditional quadratic growth model (i.e., model 2). To model the variability of growth in educational attainment for subnational entities with different income levels, the level-2 model equations are formulated as follows:

$$\pi_{0i} = \beta_{00} + \beta_{01}(HDI\_INCOME) + r_{0i} \quad (59)$$

$$\pi_{1i} = \beta_{10} + \beta_{11}(HDI\_INCOME) + r_{1i} \quad (60)$$

$$\pi_{2i} = \beta_{20} + \beta_{21}(HDI\_INCOME) + r_{2i} \quad (61)$$

Compared with the two-level unconditional growth model (i.e., model 2), the additional parameters in this level 2 model are β_{01}, β_{11}, and β_{21}. β_{01} is the difference in initial status between subnational entities with various economic developmental status.

β_{11} is the difference in linear growth rate between subnational entities with various income levels, and β_{21} represents the variability in quadratic growth between entities with various economic status.

The table below summarizes the estimated fixed and random effect parameters of this two-level conditional model with the effect of income levels.

Table 18. Two-Level Conditional Model of Quadratic Growth in Educational Attainment with the Effect of Income

Fixed Effect		Coefficient	*S.E.*	t Ratio	Approx. *d.f.*
For initial status, π_{0i}					
Average intercept, β_{00}		-2.5961	0.2036	-12.749***	1338
Average intercept with effect of the income indicator, β_{01}		15.6340	0.2844	54.970***	1338
For linear growth rate, π_{1i}					
Average intercept, β_{10}		0.0777	0.0147	5.288***	1338
Average intercept with the effect of the income indicator, β_{11}		0.0446	0.0205	2.175*	1338
For quadratic growth rate, π_{2i}					
Average intercept, β_{20}		0.0082	0.0026	3.203**	1338
Average intercept with the effect of the income indicator, β_{21}		-0.0225	0.0037	-6.083***	1338
Random Effect	Variance Component	*df*	χ^2	*p* value	

Level 1				
Temporal variation, e_{ti}	0.0067			
Level 2 (subnational entity)				
Initial status, r_{0i}	3.1283	1338	707755.4180	<0.001
Linear growth, r_{1i}	0.0233	1338	5099.9629	<0.001
Quadratic growth, r_{2i}	0.0004	1338	2545.5909	<0.001

OLS Regression Coefficient Estimate Reliability	
Initial Status, π_{0i}	0.998
Linear growth rate, π_{1i}	0.738
Quadratic growth rate, π_{2i}	0.474

Deviance = -216.8009 with 7 estimated parameters

p*-value < 0.05, ** *p*-value < 0.01, * *p*-value < 0.001

Model Parameter Estimates

Fixed Effects

The coefficients of income levels (β_{01}, β_{11}, & β_{21}) are tested to be significantly related to educational attainment development ($p < 0.05$). On average, with a unit of increase in economic level, there is an expected 15.6340 increase on initial status in educational attainment. Therefore, with average value of initial status in educational attainment across subnational entities (β_{00} = -2.5961) in the starting year of 2013, the average initial status of a subnational entity with 1 unit increase in income level is 13.0379 (= -2.5961 + 15.6340) for mean years of schooling.

In addition, the β_{11} represents the effect of income levels on differences in linear growth rates across subnational entities. Specifically, on average, a subnational entity with 1 unit increase in income level develops at a higher linear rate compared with entities with 1 unit less in income level (β_{11} = 0.0446). Thus, when the income level increases by 1 unit for a subnational entity, on average, the linear rate of change increases to 0.1223 (= 0.0777 + 0.0446) in educational attainment development. Nevertheless, the effect of income level on quadratic growth rate is negative (β_{21} = -0.0225). Thus, when the income level increases by 1 unit for a subnational entity, there is an associated decrease at a quadratic rate of -0.0143 (= 0.0082 - 0.0225) in mean years of schooling.

Random Effects

The estimates for the variances of initial status (r_{0i}), linear (r_{1i}), and quadratic growth rates (r_{2i}) are 3.1283, 0.0234, and 0.0004, respectively. Both variance components are examined to be significant, indicating that after including the income predictor, individual subnational entities still vary significantly in terms of their initial status and growth rates in educational attainment development.

Except for the random effect estimates, the correlation between initial status and linear growth rate, $\hat{\rho}_1$, (-0.152) indicates that there was a negative correlation between initial status and linear growth parameter, which means that if a subnational entity has a higher initial status in educational attainment, it will develop at a slower linear rate of change in the following years. In addition, the linear growth and quadratic rate of change also correlate significantly ($\hat{\rho}_2 = -0.904$):

$$\hat{\rho}_1(\pi_{0i}, \pi_{1i}) = \frac{\hat{\tau}_{01}}{(\hat{\tau}_{00}\hat{\tau}_{11})^{1/2}} = \frac{-0.041}{(3.1283*0.0234)^{1/2}} = -0.152 \tag{62}$$

$$\hat{\rho}_2(\pi_{1i}, \pi_{2i}) = \frac{\hat{\tau}_{12}}{(\hat{\tau}_{11}\hat{\tau}_{22})^{1/2}} = \frac{-0.0028}{(0.0234*0.0004)^{1/2}} = -0.904 \tag{63}$$

Compared to the 2-level unconditional quadratic growth model (i.e., model 2), the magnitude of the correlation between intercept and linear growth rate ($\hat{\rho}_1$ = -0.152) as well as that of correlation between linear and quadratic growth parameter ($\hat{\rho}_2$= -0.904) in this model do not change significantly. ($\hat{\rho}_{1null}$ = -0.04; $\hat{\rho}_{2null}$ = -0.897). These results indicate that after taking the income indicator as the key characteristic of subnational entities, there are still unexplained variances regarding the relationships between initial status and rates of change with regard to their education development.

Variance-Explained Statistics

The notion of proportion reduction in variance is used to examine the model configuration improvements by examining level-2 variances (τ_{00}, τ_{11}, & τ_{22}) from these two models (model 7 and model 2). This is to understand how much more variances can be accounted for by this conditional growth model. Therefore, the model performance with the inclusion of the income indicator can be evaluated.

Proportion of variance explained in initial status β_{0i}:

$$= \frac{\hat{\tau}_{00}\text{ (Null Model)} - \hat{\tau}_{00}\text{(Model 7)}}{\hat{\tau}_{00}\text{ (Null Model)}} = \frac{10.4986 - 3.1283}{10.4986} = 0.7020$$

Proportion of variance explained in linear growth parameter β_{1i}:

$$= \frac{\hat{\tau}_{11}\text{ (Null Model)} - \hat{\tau}_{11}\text{(Model 7)}}{\hat{\tau}_{11}\text{ (Null Model)}} = \frac{0.0234 - 0.0233}{0.0234} = 0.0043$$

Proportion of variance explained in quadratic growth parameter β_{1i}:

$$= \frac{\hat{\tau}_{22}\text{ (Null Model)} - \hat{\tau}_{22}\text{(Model 7)}}{\hat{\tau}_{22}\text{ (Null Model)}} = \frac{0.0005 - 0.0004}{0.0005} = 0.2$$

Results from the equations above show variance explained in intercept and rates of change by adding the income indicator to the conditional model. The income indicator can account for an additional approximately 70% of variance in initial status, and 20% in quadratic growth rate. However, it can only explain 0.4% of variance in linear rate of change in educational attainment status for subnational entities. Thus, more covariates and indicators will be included and evaluated in the following conditional growth models.

4.5.3 Model 8: Two-Level Conditional Quadratic Growth Model with Effect of Urban Population (Holding the Effect of Income Indicator Constant)

The large proportion of unexplained variances in linear growth parameter at the level-2 model provides the rationale to include additional and alternative predictors to investigate the joint effects of predictors on explaining the variability of growth, especially for the linear growth parameter. Therefore, with the identical model specifications at level-1 from the previous two-level conditional quadratic growth model (i.e., model 7), the level two model is constructed as:

$$\pi_{0i} = \beta_{00} + \beta_{01}(HDI\_INCOME) + \beta_{02}(UrbanPopulation) + r_{0i} \tag{64}$$

$$\pi_{1i} = \beta_{10} + \beta_{11}(HDI\_INCOME) + \beta_{12}(UrbanPopulation) + r_{1i} \tag{65}$$

$$\pi_{2i} = \beta_{20} + \beta_{21}(HDI\_INCOME) + r_{2i} \tag{66}$$

The predictor, urban population, is added to the intercept and linear growth parameters to examine whether the joint effects of these predictors (e.g., the cross-level interaction term) are significantly explaining the variances in linear growth. The preliminary results show that after adding the urban population predictor, the coefficient of income indicator, β_{11}, becomes non-significant (β_{11} = -0.0289, p = 0.227). Nevertheless, the coefficient of urban population predictor is instead statistically significant (β_{12} = 0.0744, p < 0.001). Thus, the urban population predictor will remain in the model to replace the effect of the income indicator to better reveal the patterns of linear growth trajectories. Moreover, holding the effect of income indicator in intercept and quadratic growth parameter constant, and adding the effect of urban population in linear growth, this conditional quadratic growth model at level-2 is reformulated as follows:

$$\pi_{0i} = \beta_{00} + \beta_{01}(HDI\_INCOME) + \beta_{02}(UrbanPopulation) + r_{0i} \quad (67)$$

$$\pi_{1i} = \beta_{10} + \beta_{11}(UrbanPopulation) + r_{1i} \quad (68)$$

$$\pi_{2i} = \beta_{20} + \beta_{21}(HDI\_INCOME) + r_{2i} \quad (69)$$

Table 19. Two-Level Conditional Model of Quadratic Growth in Educational Attainment with the Effect of Urban Population (Holding the Income Indicator Constant)

Fixed Effect	Coefficient	*S.E.*	t Ratio	Approx. *d.f.*
For initial status, π_{0i}				
Average intercept, β_{00}	-2.3558	0.2084	-11.304***	1337
Average intercept with effect of the income indicator, β_{01}	14.3796	0.3830	37.548***	1337
Average intercept with effect of urban population, β_{02}	1.2428	0.2463	5.045***	1337
For linear growth rate, π_{1i}				
Average intercept, β_{10}	0.0767	0.0071	10.730***	1338
Average intercept with the effect of urban population, β_{11}	0.0640	0.0121	5.280***	1338

For quadratic growth rate, π_{2i}					
Average intercept, β_{20}		0.0101	0.0020	5.123***	1338
Average intercept with the effect of income indicator, β_{21}		-0.0253	0.0027	-9.308***	1338
Random Effect	Variance Component	*df*	χ^2	*p* value	
Level 1					
Temporal variation, e_{ti}	0.0067				
Level 2 (subnational entity)					
Initial status, r_{0i}	3.0599	1337	691830.8899	<0.001	
Linear growth, r_{1i}	0.0234	1338	5111.7760	<0.001	
Quadratic growth, r_{2i}	0.0004	1338	2546.3358	<0.001	

OLS Regression Coefficient Estimate Reliability	
Initial Status, π_{0i}	0.998
Linear growth rate, π_{1i}	0.738
Quadratic growth rate, π_{2i}	0.474

Deviance = -292.0421 with 7 estimated parameters

p*-value < 0.05, ** *p*-value < 0.01, * *p*-value < 0.001

Model Parameter Estimates

Fixed Effects

Compared to the fixed effect estimates from the previous conditional model (model 7), the coefficient of the income indicator remains stable and significant (β_{01} = 14.380) after the inclusion of urban population as the predicting variable in initial status for subnational entities. Meanwhile, the urban population is tested to be significant, after controlling the effect of the income indicator. This illustrates that the variability in intercept within each subnational entity can be further explained by urban population (β_{02} = 1.243). Therefore, with one unit increase in income level, there is an associated (14.3796) increase in mean years of schooling, holding the urban population indicator constant. Likewise, when urban population ratio increases by 1 percent, there is an expected 1.2428 increase in mean years of schooling, holding the income indicator constant.

As mentioned above, the urban population has replaced the income indicator as the explanatory variable for linear rate of change. Thus, except for the initial status, the coefficient of the urban population on linear rate of change is significant (β_{11} = 0.0640). In addition, the effect of the income indicator on quadratic rate of change remains stable and significant (β_{12} = -0.0253).

Random Effects

The variance associated with initial status slightly decreases (r_{0i} = 3.0599) compared to that from the previous model (model 7). This indicates that adding the urban population indicator further explains the variability of initial status for subnational entities. In addition, the variance associated with quadratic growth rate remains identical (r_{2i} = 0.0004). Therefore, the effect of the income indicator is tested to be stable on explaining the quadratic growth. Nevertheless, the variance of linear growth parameter increases (r_{1i} = 0.0234) even though the urban population is demonstrated to be a stronger predictor (β_{11} = 0.0640) than income indicator. Thus, other alternative predictors will be included and evaluated to explore the patterns of linear growth trajectories.

4.5.4 Model 9 Two-Level Conditional Quadratic Growth Model (Holding Other Subnational-Level Characteristics Constant)

To better model and characterize the developmental trajectories in educational attainment, other subnational-level characteristics, such as the health indicator of human development and educational quality indicator will also be included in this model. Specifically, the correlations between the initial status and growth rates will be examined to determine the differences in educational attainment development across subnational entities with various levels of health and educational quality.

Thus, the level-2 equations for model 9 are specified as follows:

$$\pi_{0i} = \beta_{00} + \beta_{01}(HDI\_EDU) + \beta_{02}(HDI\_Health) + \beta_{03}(HDI\_Income) + \beta_{04}(UrbanPopulation) + r_{0i} \tag{70}$$

$$\pi_{1i} = \beta_{10} + \beta_{11}(HDI\_EDU) + \beta_{12}(HDI\_Health) + \beta_{13}(UrbanPopulation) + r_{1i} \tag{71}$$

$$\pi_{2i} = \beta_{20} + \beta_{21}(HDI\_EDU) + \beta_{22}(HDI\_Health) + \beta_{23}(HDI\_Income) + \beta_{24}(UrbanPopulation) + r_{2i} \tag{72}$$

Table 20. Two-Level Conditional Quadratic Growth Model (Holding Other Subnational-Level Characteristics Constant)

Fixed Effect	Coefficient	*S.E.*	t Ratio	Approx. *d.f.*
For initial status, π_{0i}				
Average intercept, β_{00}	-2.0488	0.1432	-14.307***	1335
Average intercept with effect of education quality, β_{01}	17.5077	0.2856	61.311***	1335
Average intercept with effect of the health indicator, β_{02}	-1.9425	0.2633	-7.378***	1335

Average intercept with effect of income indicator, β_{03}	0.6873	0.3265	2.105*	1335
Average intercept with effect of urban population, β_{04}	0.2725	0.1154	2.361*	1335
For linear growth rate, π_{1i}				
Average intercept, β_{10}	-0.0313	0.0280	-1.118	1336
Average intercept with the effect of education quality, β_{11}	-0.2494	0.0601	-4.152***	1336
Average intercept with the effect of the health indicator, β_{12}	0.3505	0.0666	5.263***	1336
Average intercept with the effect of urban population, β_{13}	0.0418	0.0305	1.370	1336
For quadratic growth rate, π_{2i}				
Average intercept, β_{20}	0.0313	0.0042	7.374***	1335
Average intercept with the effect of education quality, β_{21}	0.0355	0.0110	3.225**	1335
Average intercept with the effect of health indicator, β_{22}	-0.0646	0.0107	-6.020***	1335
Average intercept with the effect of income indicator, β_{23}	-0.0188	0.0053	-3.562***	1335
Average intercept with the effect of urban population, β_{24}	0.0059	0.0052	1.122	1335

Random Effect	Variance Component	*df*	χ^2	*p* value
Level 1				
Temporal variation, e_{ti}	0.0067			
Level 2 (subnational entity)				

Initial status, r_{0i}	0.7011	1335	159290.77 62	<0.001
Linear growth, r_{1i}	0.0225	1336	4945.7041	<0.001
Quadratic growth, r_{2i}	0.0004	1335	2462.6825	<0.001

OLS Regression Coefficient Estimate Reliability	
Initial Status, π_{0i}	0.992
Linear growth rate, π_{1i}	0.730
Quadratic growth rate, π_{2i}	0.458

Deviance = -2256.1742 with 7 estimated parameters

Model Parameter Estimates

Fixed Effects

Based on table 20, all the alternative predictors are found to have significant main effects on initial status, linear and/or quadratic growth parameters. Nevertheless, after the inclusion of other predictors, urban population becomes nonsignificant in explaining the variability in linear and quadratic growth (β_{13}= 0.0418; β_{24}= 0.0059, $p > 0.05$).

Except for main effects of urban population and the income indicator on initial status, which have been discussed in the previous conditional model, other variables are also significant predictors to explain the initial status. Specifically, on average, the education indicator measuring educational quality has the strongest effect. With one unit increase, there is an associated additional 17.5077 (β_{01}) growth in initial status, holding other variables constant. Moreover, the health indicator has a negative main effect on initial status. In other words, a unit increase in the health index leads to a corresponding 1.9425 decrease in intercept of educational attainment with other variables being held constant.

The alternative variables have also been examined to have significant main effects on linear and quadratic growth. For instance, (1) the health indicator is identified as the strongest predictor for linear change, and for a subnational entity with a higher level of health, it is more likely to progress at 0.3505 faster rate in linear growth, and (2) educational quality is found to be the strongest indicator for quadratic growth. Specifically, a unit increase in education quality is associated with a 0.0355 faster rate in quadratic growth.

Random Effects

The variance components remain significant, which means that there are still unexplained variances for subnational entities development. However, the variances are significantly reduced, and this indicates that subnational characteristics significantly explain some variability of growth.

Variance-Explained Statistics

Proportion of variance explained by alternative variables in initial status β_{0i}:

$$= \frac{\hat{\tau}_{00}\text{ (Null Model)} - \hat{\tau}_{00}\text{(Model 9)}}{\hat{\tau}_{00}\text{ (Null Model)}} = \frac{10.4986 - 0.7011}{10.4986} = 0.9332$$

Proportion of variance explained in linear growth parameter β_{1i}:

$$= \frac{\hat{\tau}_{11}\text{ (Null Model)} - \hat{\tau}_{11}\text{(Model 9)}}{\hat{\tau}_{11}\text{ (Null Model)}} = \frac{0.0234 - 0.0225}{0.0234} = 0.0385$$

Proportion of variance explained in quadratic growth parameter β_{2i}:

$$= \frac{\hat{\tau}_{22}\text{ (Null Model)} - \hat{\tau}_{22}\text{(Model 9)}}{\hat{\tau}_{22}\text{ (Null Model)}} = \frac{0.0005 - 0.0004}{0.0005} = 0.2$$

and the correlations between growth parameters are:

$$\hat{\rho}_1(\pi_{0i}, \pi_{1i}) = \frac{\hat{\tau}_{01}}{(\hat{\tau}_{00}\hat{\tau}_{11})^{1/2}} = \frac{-0.008}{(0.711 * 0.0225)^{1/2}} = -0.0632$$

$$\hat{\rho}_2(\pi_{1i}, \pi_{2i}) = \frac{\hat{\tau}_{12}}{(\hat{\tau}_{11}\hat{\tau}_{22})^{1/2}} = \frac{-0.0028}{(0.0225 * 0.0004)^{1/2}} = -0.9303$$

Summary

Table 21. Model Comparison with Deviance Statistics

Model Fit Index	*Model 2*		*Model 7*		*Model 8*		*Model 9*	
	DS	*df*	DS	*df*	DS	*df*	DS	*df*
	1420.4714	7	-216.8009	7	-292.0421	7	-2256.1742	7

Note. DS (Deviance Statistic); df (Number of Parameters Estimated)

Table 22. Variance Explained in Initial Status and Growth Parameters as a Result of the Main Effect of Alternative Predictors

Model	*Initial Status Var* (π_{0i})	*Linear Growth Var* (π_{1i})	*Quadratic Growth Var* (π_{2i})
Model 2 (Unconditional)	10.4986	0.0234	0.0005
Model 9	0.7011	0.0225	0.0004
Proportion of Variance Explained	0.9332	0.0385	0.2

The findings of conditional models are worth noticing because: (1) with the same number of parameters estimated in model 2, and model 7 to model 9, results of deviance statistics show that conditional models do not demonstrate significant improvements on model fit ($p > 0.5$), however, (2) as shown in table 22, the main effects of predictors show significant proportions of variance explained compared to the unconditional model (model 2).

To explain the contradictory findings, one needs to closely examine model fit comparisons by using deviance statistics: To compare models, the differences in deviance statistics and numbers of parameters estimated are used. Nevertheless, the numbers of parameters estimated in model 2, model 7, model 8, and model 9 are identical, which means that the differences in numbers of parameters estimated are 0. Thus, the model comparison results always show as non-significant ($p > 0.5$).

To further test the model fit of model 9, in other words, whether it is appropriate to construct main effects of alternative predictors in the model, the random effect associated with quadratic growth is removed (r_{2i}). With four parameters estimated in model 9, one can conclude that model 9 shows a significant improvement of model fit compared to model 2 ($\chi^2_{(3)} = 3217.0365$, $p < .001$). Therefore, model 9 outperforms model 2 with inclusions of alternative predictors, and the specifications of main effects are justified. Moreover, r_{2i} is also retained since it is statistically significant in the model (r_{2i} = 0.0004).

This model further shows the effects of predictors on developmental trajectories at the subnational level. Due to the following reasons, only main effects are discussed: (1) there is negligible significance of discussing and interpreting interaction effects since covariates included in the model are continuous rather than discrete, and interactions between continuous variables can have infinite numbers of effects; (2) cross-level interaction effects are the primary investigations of HLMs, so that the main effects are the primary concerns for models within the corresponding levels; and (3) there is negligible practical significance to interpreting interaction effects between covariates in the model from a specific level because it conveys less information to interpret mathematical operations of multiplying effects of covariates (the majority of which are composite indexes). Nevertheless, the effects of (1) income level and (2) urban population on explaining the educational attainment development are evaluated and tested as significant.

4.5.5. Model 10: Three-Level Conditional Model with Effects of GDPPC and NLDI

A level-3 model will be constructed to address variability in growth parameters at the country level. The explanatory variables included at level-3 model equations are: (1) gross domestic product per capita (GDPPC), a widely used measure of economic development; (2) NLDI, a measurement of human activity levels (Elvidge et al., 2012); (3) urban population ratio, indicating the proportion of urban population, and (4) urban population Gini, an indicator calculated to further suggest the level of urbanization (Qi et al., 2021). These covariates will be included and examined with covariates included at the level-2 model so that country-level features and their effects on subnational development can be identified.

With the inclusion of level-3 covariates, some explanatory variables describing the development of subnational entities become nonsignificant. In addition, since the quadratic growth is validated for countries, the level-1 model equation remains identical with model 5.

After excluding non-significant covariates explaining growth patterns for subnational entities, the level-2 equations are specified as:

$$\pi_{0i} = \beta_{00} + \beta_{01}(HDI\_EDU) + \beta_{02}(HDI\_Income) + \beta_{03}(UrbanPopulation) + \mathrm{r}_{0i} \tag{73}$$

$$\pi_{1i} = \beta_{10} + \beta_{11}(UrbanPopulation) + r_{1i} \tag{74}$$

$$\pi_{2i} = \beta_{20} + \beta_{21}(HDI\_Health) + r_{2i} \tag{75}$$

The GDPPC and NLDI, over the other two level-3 covariates, are tested significantly explaining country-level variations in terms of the growth. Thus, equations of the level-3 model are:

$$\beta_{00j} = \gamma_{000} + \gamma_{001}\,(GDPPC) + u_{00j} \tag{76}$$

$$\beta_{01j} = \gamma_{010} + u_{01j} \tag{77}$$

$$\beta_{02j} = \gamma_{020} + \gamma_{021}\,(GDPPC) + u_{02j} \tag{78}$$

$$\beta_{03j} = \gamma_{030} + u_{03j} \tag{79}$$

$$\beta_{10j} = \gamma_{100} \tag{80}$$

$$\beta_{11j} = \gamma_{110} + \gamma_{111}\,(GDPPC) + u_{11j} \tag{81}$$

$$\beta_{20j} = \gamma_{200} \tag{82}$$

$$\beta_{21j} = \gamma_{210} + \gamma_{211}\,(NLDI) \tag{83}$$

The table below displays the effects of GDPPC and the fixed effect of NLDI on educational attainment development for countries.

Table 23. Three-level Conditional Growth Model with Effects of GDPPC and NLDI at Level-3

Fixed Effect	Coefficient	*S.E.*	t Ratio	Approx. *d.f.*
For initial status, π_{0i}				
For average intercept, β_{00}				
Average intercept, γ_{000}	-3.8757	0.3708	-10.453***	118
Average intercept with effect of GDPPC, γ_{001}	0.0001	< 0.0001	4.142***	118
For HDI_EDU, β_{01}				
Average intercept, γ_{010}	13.7551	0.5091	27.018***	119

For HDI_INCOME, β_{02}	4.6111	0.7449	6.190***	118
Average intercept, γ_{020}	-0.0001	< 0.0001	-4.095***	118
Average intercept with effect of GDPPC, γ_{021}				
For urban population, β_{03}				
Average intercept, γ_{030}	0.6832	0.0953	7.168***	119
For linear growth rate, π_{1i}				
For average intercept, β_{10}				
Average intercept, γ_{100}	0.0857	0.0121	7.058***	736
For urban population, β_{11}				
Average intercept, γ_{110}	0.0737	0.0186	3.960***	118
Average intercept with the effect of GDPPC, γ_{111}	<-0.0001	<0.0001	-3.112**	118
For quadratic growth rate, π_{2i}				
Average intercept, β_{20}				
Average intercept, γ_{200}	0.0075	0.0074	1.009	736
For HDI_HEALTH, β_{21}				
Average intercept, γ_{210}	-0.0226	0.0093	-2.440*	736
Average intercept with the effect NLDI, γ_{211}	0.0135	0.0057	2.363*	736

Random Effect	Variance Component	*df*	χ^2	*p* value
Level 1				

Temporal variation, e_{ti}	0.0066			
Level 2 (subnational entity)				
Initial status, r_{0i}	0.0897	869	14841.4254	<0.001
Linear growth, r_{1i}	0.0271	1204	5777.9427	<0.001
Quadratic growth, r_{2i}	0.0004	1315	2520.0896	<0.001
Level 3				
Initial status r_{0i}/Initial status, u_{00j}	5.3503	110	302.1782	<0.001
Initial status r_{0i}/HDI_EDU, u_{01j}	13.4917	111	320.6974	<0.001
Initial status r_{0i}/HDI_INCOME, u_{02j}	8.8189	110	182.6676	<0.001
Initial status r_{0i}/UrbanPop, u_{03j}	0.5510	111	225.4461	<0.001
Linear growth r_{1i}/ UrbanPop, u_{11j}	0.0134	110	1040.9512	<0.001
Deviance = -4631.7706 with 34 estimated parameters				

Model Parameter Estimates

Fixed Effects

From the first panel of table 23, GDPPC, an economic development index, is demonstrated to be a significant predictor to further illustrate the variability of countries' developmental trajectories. In particular, GDPPC is an effective covariate of the following growth parameters. As for initial status: (1) GDPPC significantly predicts the averaged intercept (β_{00}) of educational attainment development (γ_{001} = 0.0001, p <0.001). Therefore, the grand mean for intercept in 2013 is -3.8757 across countries, and with 1 unit increase in GDPPC, the country is expected to have an increase of 0.0001 for the grand mean in the starting year; and (2) GDPPC is also found to be a significant explanatory variable of the income indicator (β_{02}), which is a subnational characteristic that significantly describes the initial status. For instance, with 1 unit increase in GDPPC, there is an associated 0.0001 decrease in the intercept. To sum up, with 1 unit increase in GDPPC, the country is expected to increase 0.7354 (= -3.8757 + 0.0001 + 4.6111 – 0.0001) in initial status, holding the effect of urban population constant.

Except for the initial status, GDPPC is also significantly predicting the linear growth. As shown in the table, the averaged linear growth across countries (γ_{100}) is 0.0857. With the effect of urban population being held constant for subnational entities, with 1 unit increase in GDPPC at the country level, there will be an associated 0.00002 decrease in linear growth rate.

The nighttime light development index is the only covariate that is tested as a significant predictor to illustrate the variability in quadratic rate of change. Specifically, for the subnational entities with the same urban population level, with 1 unit increase in NLDI for the country that the region belongs to, there will be an associated 0.0135 increase in the quadratic growth rate.

Random Effects

The second panel from table 23 further decomposes the residual variances into different levels. For instance, after accounting for the effects of country-level covariates, the random components remain significant. This means that there is unexplained variance (u_{00j} = 5.3503; u_{02j} = 8.8189; u_{11j} =0.0134, $p < 0.001$) to indicate the patterns of developmental trajectories. Therefore, more data are needed to include country-level covariates to reveal the growth. More importantly, after including country-level predictors, the level-2 random components are still significant, indicating that the subnational entities are significantly different from one another. Thus, more covariates at both level-2 and level-3 are needed to reduce the residual variances.

Model Comparisons and The Clustering Effects

Table 24. Model Comparison with Deviance Statistics

Model Fit Index	*Model 5*		*Model 9*		*Model 10*	
	DS	*df*	DS	*df*	DS	*df*
	-1203.5596	13	-2256.1742	7	-4631.7706	34

Note. DS (Deviance Statistic); df (Number of Parameters Estimated)

Model 10 shows a significant improvement of model fit (χ^2 = 3428.2110, p < 0.001) for the data compared with the three-level unconditional model (model 5). Meanwhile, the three-level conditional model also shows a better fit (χ^2 = 2375.5963, p < 0.001) for the data compared to the two-level conditional model (model 9). Therefore, the country-level covariates are demonstrated to be effective in explaining the growth patterns of educational attainment.

As mentioned in previous sections, the clustering effects of level-3 are tested significant in model 5. Thus, the following equation is used to show the effectiveness of level-3 covariates on explaining the variances in initial status:

$$= \frac{\hat{\tau}_{00}\ (\text{Model9}) - \hat{\tau}_{00}(\text{Model 10 })}{\hat{\tau}_{00}\ (\text{Model9})} = \frac{0.7011 - 0.0897}{0.7011} = 0.8721$$

Thus, adding the GDPPC, an additional 87.21% of the variance is explained in the initial status of the growth. However, GDPPC and NLDI do not show additional explained variances in linear and quadratic growth. In other words, more covariates are needed to be included so that how countries and subnational entities are varying in terms of their linear and quadratic growth can be explained.

Summary

Throughout the model building procedures, fewer country-level covariates are found to be significant characteristics to reveal the linear and quadratic growth. Nevertheless, country-level models are demonstrated to have significant clustering effects to indicate the differences between regions. To sum up, the current results show that it is insufficient to use economic indices in models measuring educational development. A more comprehensive conditional model can be built to indicate the whole picture of growth with more comprehensive datasets collecting various aspects of development (e.g., educational resource allocation index, equity of development, etc.).

4.6 Power Analysis

4.6.1 Significance of the Power Analysis with Missing Values

Missing values are the common challenges for longitudinal studies. Previously, a lot of researchers have encountered missingness since participants are likely to drop out the studies over the repeated measures occasions (Bryant et al., 2003; Goldstein et al., 2014; Maas & Snijders, 1997). Moreover, most studies have addressed the issues of missingness for time-invariant covariates such as participants' demographic information. However, there are still research gaps because: (1) fewer studies have investigated the impacts of missing values for time-varying covariates on the accuracy of model parameter estimates; (2) fewer empirical research has studied the influences of missing values from non-linear growth model on the reliability of parameter estimates; and (3) fewer studies have the complete dataset so that the results from replicated datasets with missing values can be compared. Thus, in this section, the level-3 time-varying covariate, GDPPC, will be selected as the covariate with missing values so that the following questions can be answered.

4.6.2 The Questions

The previous power analysis section depicts the specific number of higher level (i.e., level-2 and level-3) clusters needed to maintain statistical power. Therefore, in this section, a cross-validation analysis will be conducted using identical model specifications to model 10. Two specific questions will be answered by results in this section: (1) whether the HLMs are robust to handle missing values for a covariate (GDPPC) from the higher-level model (i.e., level-3 model equations); and if so, (2) at which percentage of missingness can the three-level conditional HLMs generate acceptable model fits and reliable parameter estimates?

According to Rubin (1976), there are two general types of missing data – missing completely at random (MCAR), meaning that probability of the missing values on Y in a dataset is not related to the value of Y itself or not associated with values of any other variables in the dataset and missing at random (MAR). These two categories of missingness do not need special treatments during data pre-processing before the analyses. Compared to MAR, MCAR is more stringent in most missing scenarios and settings. Therefore, the replications of the datasets with various levels of missingness will reflect on the MAR setting, with the assumption that the probability of the missingness occurring on Y is not related to the value of Y after controlling for other variables in the dataset.

4.6.3 Methods of Handling Missingness

There are various ways and methods to handle missingness. The conventional methods include: (1) listwise deletion (LD), where only cases present in all variables are included for analyses; and (2) mean substitutions (MS), which is to replace the missing values with imputed means of corresponding variables. Both methods have advantages and drawbacks: first, LD is demonstrated to yield unbiased parameter estimates under MCAR condition (Wothke, 2000). Nevertheless, LD excludes cases from the incomplete data which causes substantial reduction of sample size. For instance, in this study, if there is only one missing value occurring among the five measurement occasions, the whole case (either the subnational entities or the countries) will be excluded. Therefore, the statistical power will be significantly reduced, and the standard errors of the parameter estimates will be inflated. MS can solve the problem of losing cases to include all data by imputing the means of variables. However, the imputed values will have significantly less variability, which can be a threat to the precisions of parameter estimates (Gibson & Olejnik, 2003).

To overcome the drawbacks of the conventional methods addressing missingness in datasets, a more recent approach, multiple imputation (MI), has been widely applied (Fichman & Cummings, 2003) by implementing the Markov Chain Monte Carlo methods (MCMC) to generate *m* imputations (Schafer, 1997) for missing values. The procedure of MI is basically conducting different imputations *m* times based on the same observed values. Therefore, this procedure yields multiple complete datasets rather than a single complete one. However, due to the complexity of the computation, the uncertainty of the imputation estimates is substantially increased (King et al., 2001). Other scholars have also found that MI does not perform well as expected for datasets, especially for datasets with larger numbers of clusters (Cheung, 2007). To conclude, MS is chosen as the method to handle missing values in this study, and the results will be compared with the parameter estimates generated from complete datasets.

4.6.4 Evaluation Criteria

According to the established criteria, there are two measures to evaluate the performance of MS. First, since the model is appropriately constructed and specified, the deviance statistic will be used to measure the overall fit of the proposed model (model 10). The deviance statistics from the replicated datasets are expected to be distributed with corresponding degrees of freedom if the MS method is appropriate. Second, to evaluate the reliability of parameter estimates, the measure of the relative percentage bias will be used:

$$\text{Bias}\ (\hat{\theta}) = \frac{\hat{\theta} - \theta}{\theta} \times 100\% \tag{84}$$

where

θ is the true population value, and $\hat{\theta}$ represents the average parameter estimate. The rule of thumb to determine the accuracy of an estimate is based on Hoogland and Boomsma (1998), that a parameter estimate can be considered accurate if the relative percentage bias measure is less than 5%.

4.6.5 Procedures

The replicated datasets with missing values are generated to test if the model is robust in generating unbiased model fit and parameter estimates. The procedure of testing the statistical power of HLMs with missing values are as follows: first, missing rates of 10%, 20%, and 30% are applied to GDPPC, a level-three covariate in HLMs. Second, 100 unique replicated datasets are produced under each missing rate setting. Third, to ensure the quadratic growth pattern, each country cannot have more than 2 missing values.

Using Python 3.7 random sampling function from numeric and mathematical modules, datasets are generated. Once replicated datasets are generated, they will be put into HLM8 for further analyses.

4.6.6 Results

The Model Fit Index

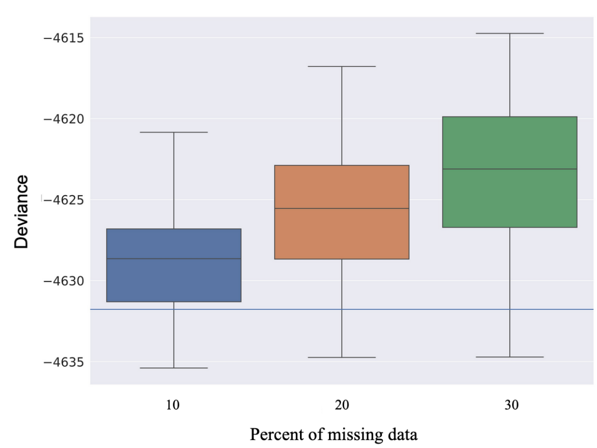


Figure 16. The Boxplots of Deviance Statistics for Percentage of Missingness at 10%, 20%, and 30%

The empirical means of deviance statistics of MS at three levels of missingness (10%, 20%, and 30%) are -4628.7976, -4625.5937, and -4623.4746 respectively. As shown in the figure above, the following observations are made: (1) the means of deviance statistic do not vary significantly from the actual value, which means that MS generally performs well across different levels of missingness; (2) the means of deviance statistics are larger compared to the actual value of -4631.7706, which means that MS generally underestimates model fits; (3) the differences between means of deviance statistics and the actual value increase as the percentage of missingness increases; and (4) the SDs of deviance statistics across three percentages of missingness are 3.1940, 4.1740, and 4.4373, and these indicate that as the number of missing values increases, the deviance statistics become unstable. Therefore, MS can be considered as an acceptable method that does not seriously affect the model fit statistics of three-level HLMs with univariate missingness occurring to only one level-3 time-varying covariate.

The Bias of Parameter Estimate

To examine the accuracy of fixed effect estimates, the grand mean of intercept coefficients (γ_{000}) and main effect of GDPPC (γ_{001}) on the average intercept are selected. From the results summarized in the tables below, one can conclude that: (1) the grand mean coefficient estimates can be considered unbiased, even though the relative percentage bias increases from 1.1% to 2.82% as the percentage of missingness increases from 10 to 30%, and (2) the main effects of GDPPC on mean intercepts are biased (the values of relative percentage bias are out of the acceptable range). Therefore, MS cannot be the most appropriate method to generate unbiased estimates that are regressed on the time-varying covariate with missing values.

Except for the fixed effect estimates, MS has been demonstrated to perform generally well to generate unbiased random components. For instance, the random variance estimates associated with intercept remain unbiased across the three missing scenarios and the relative percentage of bias values are all within the acceptable range (5%). Thus, MS is tested as a reliable method to generate accurate random effect estimates.

Table 25a. The relative percentage bias of parameter estimates with 10 percent of missingness at level-3

Relative Percentage Bias (%)	*Fixed-effect Estimate*		*Random Component*
	γ_{000}	γ_{001}	u_{02j}
	1.10	16.98	0.63

Table 25b. The relative percentage bias of parameter estimates with 20 percent of missingness at level-3

Relative Percentage Bias (%)	*Fixed-effect Estimate*		*Random Component*
	γ_{000}	γ_{001}	u_{02j}
	2.82	22.60	2.43

Table 25c. The relative percentage bias of parameter estimates with 30 percent of missingness at level-3

Relative Percentage Bias (%)	*Fixed-effect Estimate*		*Random Component*
	γ_{000}	γ_{001}	u_{02j}
	2.98	23.30	3.90

Chapter Five: Conclusion

5.1 Model Results

The key findings and inferences made from interpretations of model fits and parameter estimate results are summarized in this chapter. In addition, the contribution of this research in assessing the robustness of HLMs to handling missingness will be highlighted.

First, HLMs have the capability and flexibility to explore the shape of developmental trajectories of educational attainment for both subnational entities and countries. From unconditional model building procedures, the quadratic growth is tested and established for subnational entities and countries. In addition, the models fail to yield acceptable model fits for the cubic growth trajectory, indicating that cubic growth is not validated. Thus, to test the alternative shape of growth, piecewise growth models were constructed and built with a more complicated coding scheme of time scores, assuming distinct growth patterns and rates for two different time periods. The results showed that two-level and three-level piecewise growth HLMs did not outperform conventional quadratic growth models. To sum up, with current data available, the educational attainment for regions and countries are progressing in a quadratic manner.

Second, HLMs successfully overcome the challenges of aggregation bias compared to traditional statistical techniques to modeling hierarchical structures by assessing and partitioning the amount of variations at each level. The results show that: (1) 86 percent of variance in initial status is explained at the country level, in other words, the covariates characterizing countries are more effective to indicate the initial status of educational attainment; and (2) approximately 11 percent of the variance in linear slope is explained at the country level. Therefore, significant clustering effects at the country level are confirmed.

Third, the results from two-level conditional HLMs identified the following variables that can capture and explain the heterogeneity of education development for subnational entities: first, education development index, a subdimension to measure the overall human development level, is a composite scale considering the expected years of schooling and the actual mean years of schooling. A higher value of HDI education suggests a higher level of access to knowledge for the corresponding region. Education index has been tested to have a positive impact on the education development for subnational entities. Specifically, a subnational entity with a higher level of access to knowledge starts at a higher status, a slower linear instantaneous rate of change, and a higher quadratic change for the following years. In other words, the growth pattern for a subnational entity with higher education quality will show a clearer curvilinear trend than the subnational entity with lower education quality. Second, another subdimension index indicating human development, is the health index. It is also a composite considering life expectancy at birth, and higher values of health index suggest that individuals are living longer and healthier lives in the corresponding area. Nevertheless, the effect of health indicator is different. A subnational entity with higher health development level may start at a lower status in education at the beginning year and shows a faster linear but slower quadratic growth over time. In other words, the developmental trajectories of education will be flattened over time with the effect of health. The last subdimension for human development, the income index, is a composite considering gross national income (GNI) per capita to indicate the standard of living. A higher value of income index suggests a higher standard of living for the region. The income index is only affecting the initial status and the quadratic growth. It is within our assumption that income is positively affecting the initial status of educational attainment. However, income also has a strong negative effect on quadratic growth over time. Finally, urban population ratio, which is an indication of urbanization level, is only positively affecting the initial status for subnational entities.

Implications based on the developmental trajectories for subnational entities are: (1) more resources should be allocated to regions with lower education quality for promoting the education development and (2) less attention can be drawn to regions with higher human development levels since they tend to be stable in terms of education development (see figure 17 and 18).

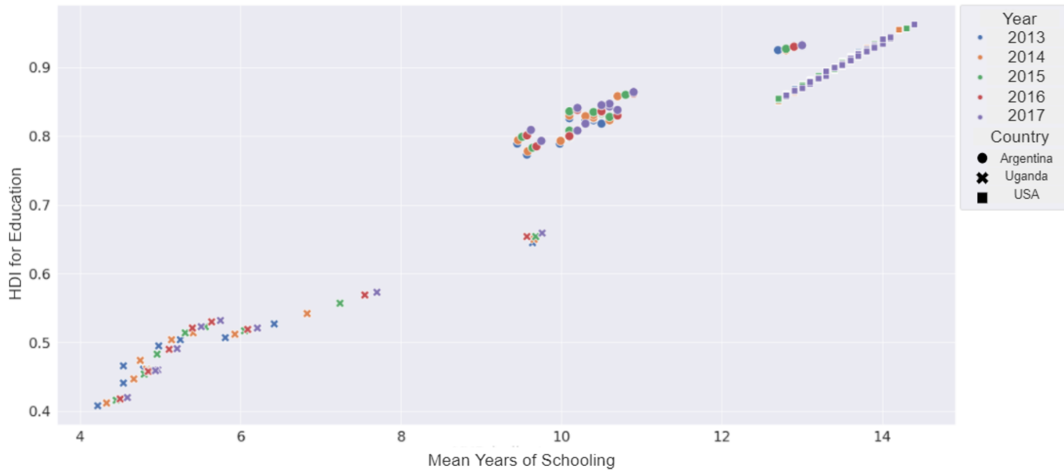


Figure 17. The Effect of Education Quality on Education Development at the Subnational Level

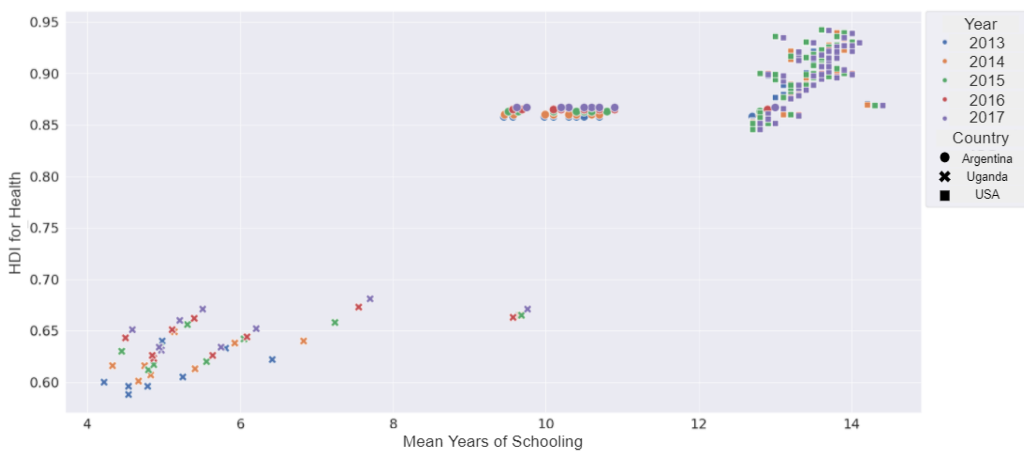


Figure 18. The Effect of Health Levels on Education Development at the Subnational Level

To sum up, even though there is a variety of variables found to significantly explain the variations of subnational entities in terms of the initial status, the instantaneous growth (i.e., the linear and quadratic growth) remain to be unexplained. In other words, the large values of variances associated with the growth parameters need further explorations with inclusions of other socio-economic variables. For instance, recently there is a growing number of literatures using inequality-adjusted indices (e.g., gender inequality index, income-consumption balance, inequality of life expectancy index, etc.) to better capture the patterns of human development for a specific region. Therefore, the model could have been improved and growth parameters could have been better explained and estimated with inequality-adjusted indices.

Fourth, as indicated in the values of ICCs, countries are explaining most of the heterogeneity in education development. Results from three-level conditional HLMs show similar findings as those of two-level conditional HLMs: (1) variables are effectively explaining variations in initial status for countries, but not explaining linear and quadratic growth; (2) GDPPC is the only significant economic measure at the country level to further reduce unexplained variances in initial status and linear slope, and countries with higher GDPPC, higher health status and higher income levels tend to have slower rates of quadratic change. In other words, the countries with higher human development levels tend to have slower changes over time; and (3) NLDI is tested to be an effective predictor at the country level to reduce the variance in the quadratic slope. Specifically, the countries with higher levels of human activities tend to have higher quadratic rates of change. Besides the coefficients of fixed effect estimates, the large values of variance components indicate that even though the clustering effects of countries are confirmed, there are fewer variables that can better capture the characteristics of countries to explain the heterogeneity of education development over the five consecutive years.

Based on the observations made from the 3-level conditional model, the implications are: (1) countries with lower initial status in terms of education development are the ones that need more resources; and (2) policies and resources can be put to countries that present greater human activities levels because they are demonstrating higher change rates and greater potential in education development (see figure 19).

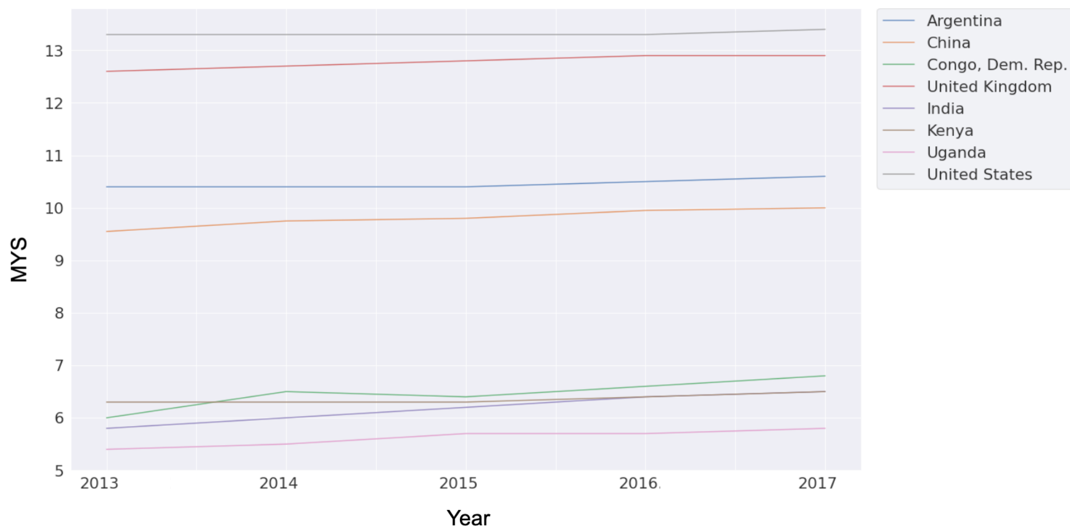


Figure 19. Examples of Developmental Trajectories of Educational Attainment at Country Level

5.2 Contributions

A specific contribution of this study is to demonstrate the robustness of HLMs to handling missing values, especially missingness occurred at the level-3 model. Previous studies have an extended discussion of power analyses within the scenario of randomized control trials, and little research has been done to demonstrate the robustness of longitudinal HLMs without randomized assigned groups. Moreover, there is a heated debate over the weights of missingness occurring at models from different levels, and no studies have been conducted to suggest the significant weight of level-3 covariate, particularly the time-varying covariate on model fit performances and accuracy of parameter estimates.

As indicated in the previous chapter, with sufficient measurement occasions and a total cluster of 130, the HLMs constructed in this study are confirmed to maintain power at 0.8. Thus, a cross-validation was conducted by generating 300 replicated datasets with missingness occurring for a level-3 covariate, GDPPC, under MAR assumption. MS method was used to address the missing values by overcoming the difficulties of data loss and computation complexity. The results show that: (1) MS is valid to generate unbiased model fit statistics at various percentages of missingness (10%, 20%, and 30%); (2) MS is robust to generate random effect estimates, no matter what percentage of missingness occurred at the level-3 time-varying covariate. The MS is found to be relatively unstable to generate unbiased fixed effect estimates: (1) MS can generate unbiased fixed effect coefficients on which the covariate with missing values are not regressed; and (2) MS cannot generate unbiased fixed effect coefficients on which the covariate with missingness is regressed, and even with a small percentage of missingness (10%), the relative percentage of bias is out of the acceptable range.

5.3 Future Work

There are a few problems that can be addressed in future research. First, the current HLMs only include conventional variables to characterize subnational entities and countries. As mentioned in the discussion section, more new indicators capturing nuanced differences across subnational entities and countries are needed. For instance, more patterns could have been found if HLMs would include more inequality-adjusted indices that can better reflect the conflicts and dynamics of education and human development. In addition, more reliable variables are needed for future studies to overcome the current limitations of variables included in HLMs for this study. For example, NLDI is a measurement regarding spatial heterogeneity of human development. However, errors are introduced when it is used to assess temporal variations.

Second, during the model building procedures, the cubic growth was once found to be significant, which means that the subnational entities and countries may demonstrate cubic developmental trajectories over time. Moreover, more alternative patterns of development could have been tested and established since the current HLMs only include data collected from 2013 to 2017. Therefore, a piecewise growth pattern may be established if more data from previous years can be combined with recent years of data.

Third, the power analysis can be extended by comparing multiple methods to address missingness and by replicating datasets with missingness occurring to multiple time-varying covariates. For example, MI can be used to address multivariate missingness because it can generate multiple parameter estimates and datasets instead of single value estimate and single imputed complete dataset. In addition, a comparison of methods to handling missing values can be helpful for future researchers who prefer to avoid using MS since it can substantially reduce the variability of variables with mean replacements. Thus, future research can be conducted to provide a more comprehensive picture and offer a guide for researchers to select the most appropriate method to address missing values so that challenges of biased estimates and model fits of three-level HLMs can be overcome.

Besides comparing methods to address missingness, other evaluation criteria can be used to better assess the robustness of HLMs. For example, *CI*s can be further constrained and specified so that the precisions of parameter estimates can be better evaluated. In addition, the *SE*s can also be used as another measure indicating the accuracy of parameter estimates. Thus, more future work can be done by including more criteria, and by adjusting the critical values (e.g., set wider *CI*s) to determine the bias.

Multivariate models and other spatial autocorrelation models can be compared with the performances of HLMs. For instance, the results of spatial autocorrelation models can be directly compared to the model parameter estimates from HLMs so that the capabilities of these models on handling hierarchical structures can be further evaluated. In addition, multivariate models such as loglinear regression models can be constructed so that researchers would have potential to observe other patterns that HLMs could not have generated.

HLMs have been widely applied to research from various disciplines, but more new models such as machine learning models can be integrated with HLMs to generate meaningful information and patterns when the variables and datasets become large. For instance, the long short-term memory (LSTM) models can be adopted for the purpose of prediction with accuracy so that the educators, practitioners and policy and decision-makers can better utilize the results to benefit individuals, organizations, regions, and countries with improved educational attainment. Nevertheless, future investigations can start by building upon the current HLMs constructed and proposed in this study. In addition, variables that are included, and conceptual framework of missingness established from this study can be further utilized.

References

Aghion, P., Askenazy, P., Bourlès, R., Cette, G., & Dromel, N. (2009). Education, market rigidities and growth. *Economics Letters*, *102*(1), 62–65.

Aghion, P., Howitt, P., Howitt, P. W., Brant-Collett, M., & García-Peñalosa, C. (1998). *Endogenous growth theory*. MIT press.

Aguinis, H., Gottfredson, R. K., & Culpepper, S. A. (2013). Best-practice recommendations for estimating cross-level interaction effects using multilevel modeling. *Journal of Management*, *39*(6), 1490–1528.

Alspaugh, J. W. (1998). The relationship of school-to-school transitions and school size to high school dropout rates. *The High School Journal*, *81*(3), 154–160.

Arnold, C. L. (1992). An introduction to hierarchical linear models. *Measurement and Evaluation in Counseling and Development, 25*(2), 58–90.

Arrighi, G., Silver, B. J., & Brewer, B. D. (2003). Industrial convergence, globalization, and the persistence of the North-South divide. *Studies in Comparative International Development*, *38*(1), 3.

Avendano, M., Jürges, H., & Mackenbach, J. P. (2009). Educational level and changes in health across Europe: longitudinal results from SHARE. *Journal of European Social Policy, 19*(4), 301-316.

Barro, R. J., & Lee, J. W. (1996). International measures of schooling years and schooling quality. *The American Economic Review*, *86*(2), 218–223.

Barro, R. J., & Lee, J.-W. (1993). *International comparisons of educational attainment*. National Bureau of Economic Research.

Bassetti, T. (2009). Education and Poverty in a Solow Growth Model. *Long-Run Growth and the Standard of Life. United Kingdom: Edward Elgar Publishing*.

Baugh, K., Hsu, F.-C., Elvidge, C. D., & Zhizhin, M. (2013). Nighttime lights compositing using the VIIRS day-night band: Preliminary results. *Proceedings of the Asia-Pacific Advanced Network*, *35*, 70–86.

Belloumi, M., & Alshehry, A. S. (2016). The impact of urbanization on energy intensity in Saudi Arabia. *Sustainability*, *8*(4), 375.

Benos, N., & Zotou, S. (2014). Education and economic growth: A meta-regression analysis. *World Development*, *64*, 669–689.

Berkowitz, R., Moore, H., Astor, R. A., & Benbenishty, R. (2017). A research synthesis of the associations between socioeconomic background, inequality, school climate, and academic achievement. *Review of Educational Research*, *87*(2), 425–469.

Bertinelli, L., & Black, D. (2004). Urbanization and growth. *Journal of Urban Economics*, *56*(1), 80–96.

Bian-Ling, O. (2014). Moran's I Tests for Spatial Dependence in Panel Data Models with Time Varying Spatial Weights Matrices. *2014 International Conference on Economic Management and Trade Cooperation (EMTC 2014)*.

Bielefeldt, T. (2005). Computers and student learnings: Interpreting the multivariate analysis of PISA 2000. *Journal of Research on Technology in Education*, *37*(4), 339–347.

Bolton, P. (2012). Education: Historical Statistics, House of Commons Library. *Social &*.

Bryant, A. L., Schulenberg, J. E., O'Malley, P. M., Bachman, J. G., & Johnston, L. D. (2003). How academic achievement, attitudes, and behaviors relate to the course of substance use during adolescence: A 6 ‐ year, multiwave national longitudinal study. *Journal of Research on Adolescence, 1*3(3), 361-397.

Bryk, A. S., & Raudenbush, S. W. (1988). Heterogeneity of variance in experimental studies: A challenge to conventional interpretations. *Psychological Bulletin*, *104*(3), 396.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods.* Sage Publications, Inc.

Buis, M. L. (2010). Scaling levels of education. *Amsterdam, Faculty of Social Sciences, VU-University Amsterdam*.

Bundervoet, T., Maiyo, L., & Sanghi, A. (2015). *Bright lights, big cities: Measuring national and subnational economic growth in Africa from outer space, with an application to Kenya and Rwanda*. The World Bank.

Cai, F., Wang, D., & Du, Y. (2002). Regional disparity and economic growth in China: The impact of labor market distortions. *China Economic Review*, *13*(2–3), 197–212.

Castelló-Climent, A., & Doménech, R. (2014). Human capital and income inequality: Some facts and some puzzles. *Retrieved from BBVA Research Https://Www. Bbvaresearch. Com/Wp-Content/Uploads/Migrados/WP_1228_tcm348-430101. Pdf.*

Chen, M., Zhang, H., Liu, W., & Zhang, W. (2014). The global pattern of urbanization and economic growth: Evidence from the last three decades. *PloS One*, *9*(8), e103799.

Chen, X., & Nordhaus, W. D. (2011). Using luminosity data as a proxy for economic statistics. *Proceedings of the National Academy of Sciences*, *108*(21), 8589–8594.

Chen, Y. (2011). Modeling fractal structure of city-size distributions using correlation functions. *PloS One*, *6*(9), e24791.

Chen, Y. (2013). New approaches for calculating Moran's index of spatial autocorrelation. *PloS One*, *8*(7), e68336.

Cheung, M. W. L. (2007). Comparison of methods of handling missing time-invariant covariates in latent growth models under the assumption of missing completely at random. *Organizational Research Methods, 10*(4), 609-634.

Chiu, M. M. (2007). Families, economies, cultures, and science achievement in 41 countries: Country-, school-, and student-level analyses. *Journal of Family Psychology*, *21*(3), 510.

Cliff, A. D., & Ord, J. K. (1981). *Spatial processes: Models & applications*. Taylor & Francis.

Cohen, B. (2006). Urbanization in developing countries: Current trends, future projections, and key challenges for sustainability. *Technology in Society*, *28*(1–2), 63–80.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. Routledge.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ: Erlbaum.

Connelly, R., Gayle, V., & Lambert, P. S. (2016). A review of educational attainment measures for social survey research. *Methodological Innovations*, *9*, 2059799116638001.

Cortez, P., & Silva, A. M. G. (2008). *Using data mining to predict secondary school student performance*.

Crawford, G. A. (2015). The academic library and student retention and graduation: An exploratory study. *Portal: Libraries and the Academy*, *15*(1), 41–57.

Dearden, L., McIntosh, S., Myck, M., & Vignoles, A. (2002). The returns to academic and vocational qualifications in Britain. *Bulletin of Economic Research*, *54*(3), 249–274.

De la Fuente, Á., & Doménech, R. (2002). *Educational Attainment in the OECD, 1960-95*.

De Maesschalck, R., Jouan-Rimbaud, D., & Massart, D. L. (2000). The mahalanobis distance. *Chemometrics and intelligent laboratory systems, 50*(1), 1-18.

Dempster, A. P., Rubin, D. B., & Tsutakawa, R. K. (1981). Estimation in covariance components models. *Journal of the American Statistical Association*, *76*(374), 341–353.

DESA, U. (2010). *United Nations, Department of Economic and Social Affairs, Population Division: World urbanization prospects, the 2009 revision: highlights*. UN publications, New York, h ttp://esa. un. org/unpd/wup/Documents/WUP2009 ….

Eikemo, T. A., Huisman, M., Bambra, C., & Kunst, A. E. (2008). Health inequalities according to educational level in different welfare regimes: A comparison of 23 European countries. *Sociology of Health & Illness*, *30*(4), 565–582.

Elliott, R. J., Sun, P., & Zhu, T. (2017). The direct and indirect effect of urbanization on energy intensity: A province-level study for China. *Energy*, *123*, 677–692.

Elston, R. C., & Grizzle, J. E. (1962). Estimation of time-response curves and their confidence bands. *Biometrics*, *18*(2), 148–159.

Elvidge, C. D., Baugh, K. E., Anderson, S. J., Sutton, P. C., & Ghosh, T. (2012). The Night Light Development Index (NLDI): a spatially explicit measure of human development from satellite data. *Social Geography, 7*(1), 23-35.

Elvidge, C. D., Baugh, K. E., Anderson, S. J., Sutton, P. C., & Ghosh, T. (2012). The Night Light Development Index (NLDI): A spatially explicit measure of human development from satellite data. *Social Geography*, *7*(1), 23–35.

Elvidge, C. D., Baugh, K. E., Dietz, J. B., Bland, T., Sutton, P. C., & Kroehl, H. W. (1999). Radiance calibration of DMSP-OLS low-light imaging data of human settlements. *Remote Sensing of Environment*, *68*(1), 77–88.

Elvidge, C. D., Sutton, P. C., Ghosh, T., Tuttle, B. T., Baugh, K. E., Bhaduri, B., & Bright, E. (2009). A global poverty map derived from satellite data. *Computers & Geosciences*, *35*(8), 1652–1660.

Fichman, M., & Cummings, J. N. (2003). Multiple imputation for missing data: Making the most of what you know. *Organizational Research Methods, 6*(3), 282-308.

Fleisher, B. M., & Chen, J. (1997). The coast–noncoast income gap, productivity, and regional economic policy in China. *Journal of Comparative Economics*, *25*(2), 220–236.

Galla, B. M., Wood, J. J., Tsukayama, E., Har, K., Chiu, A. W., & Langer, D. A. (2014). A longitudinal multilevel model analysis of the within-person and between-person effect of effortful engagement and academic self-efficacy on academic performance. *Journal of School Psychology*, *52*(3), 295–308.

Garson, G. D. (2013). Introductory guide to HLM with HLM 7 software. Hierarchical linear modeling: Guide and applications, 55-96.

Ghosh, S., & Kanjilal, K. (2014). Long-term equilibrium relationship between urbanization, energy consumption and economic activity: Empirical evidence from India. *Energy*, *66*, 324–331.

Gibson, N. M., & Olejnik, S. (2003). Treatment of missing data at the second level of hierarchical linear models. *Educational and Psychological Measurement, 63*(2), 204-238.

Goldstein, H. (2011). *Multilevel statistical models* (Vol. 922). John Wiley & Sons.

Goldstein, H., Carpenter, J. R., & Browne, W. J. (2014). Fitting multilevel multivariate models with missing data in responses and covariates that may include interactions and non-linear terms. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 553-564.

Goldstein, H., Healy, M. J., & Rasbash, J. (1994). Multilevel time series models with applications to repeated measures data. *Statistics in Medicine*, *13*(16), 1643–1655.

Gregorio, J. D., & Lee, J.-W. (2002). Education and income inequality: New evidence from cross-country data. *Review of Income and Wealth*, *48*(3), 395–416.

Griffith, D. A. (1987). Spatial autocorrelation. *A Primer. Washington DC: Association of American Geographers*.

Griffith, J. (1996). Relation of parental involvement, empowerment, and school traits to student academic performance. *The Journal of Educational Research*, *90*(1), 33–41.

Griggs, D., Stafford-Smith, M., Gaffney, O., Rockström, J., Öhman, M. C., Shyamsundar, P., Steffen, W., Glaser, G., Kanie, N., & Noble, I. (2013). Policy: Sustainable development goals for people and planet. *Nature*, *495*(7441), 305.

Grimm, N. B., Grove, J. G., Pickett, S. T., & Redman, C. L. (2000). Integrated approaches to long-term studies of urban ecological systems: Urban ecological systems present multiple challenges to ecologists—pervasive human impact and extreme heterogeneity of cities, and the need to integrate social and ecological approaches, concepts, and theory. *BioScience*, *50*(7), 571–584.

Hanushek, E. A., & Woessmann, L. (2010). Education and economic growth. *Economics of Education*, 60–67.

Haskins, R. (2009). What works is work: Welfare reform and poverty reduction. *Nw. JL & Soc. Pol'y*, *4*, 30.

Heck, R. H., & Thomas, S. L. (2015). *An introduction to multilevel modeling techniques: MLM and SEM approaches using Mplus*. Routledge.

Heo, M., & Leon, A. C. (2008). Statistical power and sample size requirements for three level hierarchical cluster randomized trials. *Biometrics*, *64*(4), 1256–1262.

Hill, M. A., & King, E. (1995). Women's education and economic well-being. *Feminist Economics*, *1*(2), 21–46.

Hoen, A. G., Margos, G., Bent, S. J., Diuk-Wasser, M. A., Barbour, A., Kurtenbach, K., & Fish, D. (2009). Phylogeography of Borrelia burgdorferi in the eastern United States reflects multiple independent Lyme disease emergence events. *Proceedings of the National Academy of Sciences*, *106*(35), 15013–15018.

Hoffmeyer-Zlotnik, J. H., Hoffmeyer-Zlotnik, J. H., & Wolf, C. (2003). *Advances in cross-national comparison: A European working book for demographic and socio-economic variables*. Springer Science & Business Media.

Holmbeck, G. N., Li, S. T., Schurman, J. V., Friedman, D., & Coakley, R. M. (2002). Collecting and managing multisource and multimethod data in studies of pediatric populations. *Journal of Pediatric Psychology*, *27*(1), 5–18.

Hoogland, J. J., & Boomsma, A. (1998). Robustness studies in covariance structure modeling: An overview and a meta-analysis. *Sociological Methods & Research, 26*(3), 329-367.

Hopkins, K. D. (1982). The unit of analysis: Group means versus individual observations. *American Educational Research Journal*, *19*(1), 5–18.

Hox, J. J. (1995). *Applied multilevel analysis*. TT-publikaties.

Huitt, W., Huitt, M., Monetti, D., & Hummel, J. (2009). A systems-based synthesis of research related to improving students' academic performance. *3rd International City Break Conference Sponsored by the Athens Institute for Education and Research (ATINER), October*, 16–19.

Ichou, M., & Vallet, L.-A. (2011). Do all roads lead to inequality? Trends in French upper secondary school analysed with four longitudinal surveys. *Oxford Review of Education*, *37*(2), 167–194.

Jenkins, A., & Sabates, R. (2007). *The classification of qualifications in social surveys*.

Johnson, U. Y., & Hull, D. M. (2014). Parent involvement and science achievement: A cross-classified multilevel latent growth curve analysis. *The Journal of Educational Research*, *107*(5), 399–409.

Jowell, R., Roberts, C., Fitzgerald, R., & Eva, G. (2007). *Measuring attitudes cross-nationally: Lessons from the European Social Survey*. Sage.

Kabra, R. R., & Bichkar, R. S. (2011). Performance prediction of engineering students using decision trees. *International Journal of Computer Applications*, *36*(11), 8–12.

Kaplan, D., & Elliott, P. R. (1997). A model-based approach to validating education indicators using multilevel structural equation modeling. *Journal of Educational and Behavioral Statistics*, *22*(3), 323–347.

Kasim, R. M., & Raudenbush, S. W. (1998). Application of Gibbs sampling to nested variance components models with heterogeneous within-group variance. *Journal of Educational and Behavioral Statistics, 23*(2), 93-116.

Kasman, A., & Duman, Y. S. (2015). CO2 emissions, economic growth, energy consumption, trade and urbanization in new EU member and candidate countries: A panel data analysis. *Economic Modelling*, *44*, 97–103.

King, G., Honaker, J., Joseph, A., & Scheve, K. (2001). Analyzing incomplete political science data: An alternative algorithm for multiple imputation. *American political science review, 95*(1), 49-69.

Klasen, S. (2002). Low schooling for girls, slower growth for all? Cross-country evidence on the effect of gender inequality in education on economic development. *The World Bank Economic Review*, *16*(3), 345–373.

Knowles, S., Lorgelly, P. K., & Owen, P. D. (2002). Are educational gender gaps a brake on economic development? Some cross-country empirical evidence. *Oxford Economic Papers*, *54*(1), 118–149.

Kreft, I. G. (1996). Are multilevel techniques necessary? An overview, including simulation studies. *Unpublished Manuscript, California State University, Los Angeles*.

Krueger, A. B., & Lindahl, M. (2001). Education for growth: Why and for whom? *Journal of Economic Literature*, *39*(4), 1101–1136.

Kummu, M., Taka, M., & Guillaume, J. H. (2018). Gridded global datasets for gross domestic product and Human Development Index over 1990–2015. *Scientific Data*, *5*, 180004.

Kunovich, R. M., & Hodson, R. (2002). Ethnic diversity, segregation, and inequality: A structural model of ethnic prejudice in Bosnia and Croatia. *The Sociological Quarterly*, *43*(2), 185–212.

Ladd, H. F. (2012). Education and poverty: Confronting the evidence. *Journal of Policy Analysis and Management*, *31*(2), 203–227.

Lambert, P. S. (2012). Comparative scaling of educational categories by homogamy–Analysis of UK data from the BHPS. *Technical Paper 2012-1 of the DAMES Node, Data Management through e-Social Science*.

Lee, S.-I. (2001). Developing a bivariate spatial association measure: An integration of Pearson's r and Moran's I. *Journal of Geographical Systems*, *3*(4), 369–385.

Lenzen, M., Wier, M., Cohen, C., Hayami, H., Pachauri, S., & Schaeffer, R. (2006). A comparative multivariate analysis of household energy requirements in Australia, Brazil, Denmark, India and Japan. *Energy*, *31*(2–3), 181–207.

Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American statistical Association, 83*(404), 1198-1202

Lin, B., & Liu, H. (2015). China's building energy efficiency and urbanization. *Energy and Buildings*, *86*, 356–365.

Longford, N. T. (1993). Regression analysis of multilevel data with measurement error. *British Journal of Mathematical and Statistical Psychology*, *46*(2), 301–311.

Ma, T., Zhou, C., Pei, T., Haynie, S., & Fan, J. (2014). Responses of Suomi-NPP VIIRS-derived nighttime lights to socioeconomic activity in China's cities. *Remote Sensing Letters*, *5*(2), 165–174.

Maas, C. J., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology*, *1*(3), 86–92.

Maas, C. J. M., & Snijders, T. A. B. (1997). The multilevel approach to repeated measures with missing data. Manuscript submitted for publication.

Mankiw, N. G., Romer, D., & Weil, D. N. (1992). A contribution to the empirics of economic growth. *The Quarterly Journal of Economics*, *107*(2), 407–437.

Mathieu, J. E., Aguinis, H., Culpepper, S. A., & Chen, G. (2012). Understanding and estimating the power to detect cross-level interaction effects in multilevel modeling. *Journal of Applied Psychology*, *97*(5), 951.

Mayer, A. P. (2008). Expanding Opportunities for High Academic Achievement: An International Baccalaureate Diploma Program in an Urban High Schoo. *Journal of Advanced Academics*, *19*(2), 202–235.

McArdle, J. J., & Epstein, D. (1987). Latent growth curves within developmental structural equation models. *Child Development*, 110–133.

McCoach, D. B., & Kaniskan, B. (2010). Using time-varying covariates in multilevel growth models. *Frontiers in Psychology*, *1*, 17.

McFarlane, C. (2006). Crossing borders: Development, learning and the North–South divide. *Third World Quarterly*, *27*(8), 1413–1437.

Mesa, E. P. (2007). *Measuring education inequality in the Philippines*. UPSE Discussion Paper.

Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological bulletin, 105*(1), 156.

Moerbeek, M., van Breukelen, G. J., & Berger, M. P. (2000). Design issues for experiments in multilevel populations. *Journal of Educational and Behavioral Statistics*, *25*(3), 271–284.

Moons, E., Brijs, T., & Wets, G. (2009). Improving Moran's index to identify hot spots in traffic safety. In *Geocomputation and urban planning* (pp. 117–132). Springer.

Muthén, B. O. (1994). Multilevel covariance structure analysis. *Sociological Methods & Research*, *22*(3), 376–398.

Nguyen, H. M. (2018). The relationship between urbanization and economic growth. *International Journal of Social Economics*.

Njoh, A. J. (2003). Urbanization and development in sub-Saharan Africa. *Cities*, *20*(3), 167–174.

Okpala, C. O., Smith, F., Jones, E., & Ellis, R. (2000). A clear link between school and teacher characteristics, student demographics, and student achievement. *Education*, *120*(3), 487–487.

Oppong, F. B., & Agbedra, S. Y. (2016). Assessing univariate and multivariate normality. a guide for non-statisticians. *Math Theory Modeling, 6*(2), 26-33.

Parikh, J., & Shukla, V. (1995). Urbanization, energy use and greenhouse effects in economic development: Results from a cross-national study of developing countries. *Global Environmental Change*, *5*(2), 87–103.

Paterson, L., & Iannelli, C. (2007). Social class and educational attainment: A comparative study of England, Wales, and Scotland. *Sociology of Education*, *80*(4), 330–358.

Pesaresi, M., Ehrilch, D., Florczyk, A. J., Freire, S., Julea, A., Kemper, T., Soille, P., & Syrris, V. (2015). GHS built-up grid, derived from Landsat, multitemporal (1975, 1990, 2000, 2014). *European Commission, Joint Research Centre, JRC Data Catalogue*.

Pesaresi, M., Melchiorri, M., Alice, S., & Kemper, T. (2016). *Atlas of the human planet 2016: Mapping human presence on earth with the global human settlement layer*. Publications Office.

Prasannakumar, V., Vijith, H., Charutha, R., & Geetha, N. (2011). Spatio-temporal clustering of road accidents: GIS based analysis and assessment. *Procedia-Social and Behavioral Sciences*, *21*, 317–325.

Pritchett, L. (2006). Does learning to add up add up? The returns to schooling in aggregate data. *Handbook of the Economics of Education*, *1*, 635–695.

Psacharopoulos, G. (1994). Returns to investment in education: A global update. *World Development*, *22*(9), 1325–1343.

Qi, B., Wang, X., & Sutton, P. (2021). Can Nighttime Satellite Imagery Inform Our Understanding of Education Inequality?. *Remote Sensing, 13*(5), 843.

Rasbash, J., & Goldstein, H. (1994). Efficient analysis of mixed hierarchical and cross-classified random structures using a multilevel model. *Journal of Educational and Behavioral Statistics*, 337–350.

Raudenbush, S. W. (1997). Statistical analysis and optimal design for cluster randomized trials. *Psychological Methods*, *2*(2), 173.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). sage.

Raudenbush, S., Bryk, A., Cheong, Y. F., & Congdon, R. (2019). *HLM 8 Hierarchical Linear and Nonlinear Modeling*. Scientific Software International.

Raudenbush, S. W., Spybrook, J., Congdon, R., Liu, X., Martinez, A., Bloom, H., & Hill, C. (2011). *Optimal design software for multi-level and longitudinal research (Version 3.01)[Software]*.

Reason, R. D. (2003). Student variables that predict retention: Recent research and new developments. *Naspa Journal*, *40*(4), 172–191.

Robert, K. W., Parris, T. M., & Leiserowitz, A. A. (2005). What is sustainable development? Goals, indicators, values, and practice. *Environment: Science and Policy for Sustainable Development*, *47*(3), 8–21.

Rogosa, D., & Saner, H. (1995). Longitudinal data analysis examples with random coefficient models. *Journal of Educational and Behavioral Statistics*, *20*(2), 149–170.

Rubin, D. B. (1976). Inference and missing data. *Biometrika, 63*(3), 581-592.

Saa, A. A. (2016). Educational data mining & students' performance prediction. *International Journal of Advanced Computer Science and Applications*, *7*(5), 212–220.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. CRC press.

Sachs, J. D. (2012). From millennium development goals to sustainable development goals. *The Lancet*, *379*(9832), 2206–2211.

Sadorsky, P. (2013). Do urbanization and industrialization affect energy intensity in developing countries? *Energy Economics*, *37*, 52–59.

Samdal, O., Wold, B., & Bronis, M. (1999). Relationship between students' perceptions of school environment, their satisfaction with school and perceived academic achievement: An international study. *School Effectiveness and School Improvement*, *10*(3), 296–320.

Sato, Y., & Zenou, Y. (2015). How urbanization affect employment and social interactions. *European Economic Review*, *75*, 131–155.

Satterthwaite, D. (2009). The implications of population growth and urbanization for climate change. *Environment and Urbanization*, *21*(2), 545–567.

Schneider, S. L. (2013). The international standard classification of education 2011. In *Class and stratification analysis*. Emerald Group Publishing Limited.

Shi, K., Chen, Y., Yu, B., Xu, T., Yang, C., Li, L., Huang, C., Chen, Z., Liu, R., & Wu, J. (2016). Detecting spatiotemporal dynamics of global electric power consumption using DMSP-OLS nighttime stable light data. *Applied Energy*, *184*, 450–463.

Smith, A. K., Ayanian, J. Z., Covinsky, K. E., Landon, B. E., McCarthy, E. P., Wee, C. C., & Steinman, M. A. (2011). Conducting high-value secondary dataset analysis: An introductory guide and resources. *Journal of General Internal Medicine*, *26*(8), 920–929.

Smits, J., & Permanyer, I. (2019). The Subnational Human Development Database. *Scientific Data*, *6*, 190038.

Snijders, T. A. (2005). Power and sample size in multilevel linear models. *Encyclopedia of Statistics in Behavioral Science*.

Snijders, T. A., & Bosker, R. J. (1993). Standard errors and sample sizes for two-level research. *Journal of Educational Statistics*, *18*(3), 237–259.

Song, Y., Qiu, Q., Guo, Q., Lin, J., Li, F., Yu, Y., Li, X., & Tang, L. (2010). The application of spatial Lorenz curve (SLC) and Gini coefficient in measuring land use structure change. *2010 18th International Conference on Geoinformatics*, 1–5.

Spybrook, J., Bloom, H., Congdon, R., Hill, C., Martinez, A., Raudenbush, S., & TO, A. (2011). Optimal design plus empirical evidence: Documentation for the “Optimal Design” software. *William T. Grant Foundation. Retrieved on November*, *5*, 2012.

Sutton, P. C., & Costanza, R. (2002). Global estimates of market and non-market values derived from nighttime satellite imagery, land cover, and ecosystem service valuation. *Ecological Economics*, *41*(3), 509–527.

Thomas, V., Wang, Y., & Fan, X. (1999). *Measuring education inequality: Gini coefficients of education*. The World Bank.

Thomas, V., Wang, Y., & Fan, X. (2001). *Measuring Education Inequality: Gini Coefficients of Education. Policy Research Working Paper.*

Topor, D. R., Keane, S. P., Shelton, T. L., & Calkins, S. D. (2010). Parent involvement and student academic performance: A multiple mediational analysis. *Journal of Prevention & Intervention in the Community*, *38*(3), 183–197.

Treiman, D. J. (2005). The legacy of apartheid: Racial inequalities in the new South Africa. *UCLA CCPR Population Working Papers*.

Treiman, D. J. (2014). *Quantitative data analysis: Doing social research to test ideas*. John Wiley & Sons.

Turok, I., & McGranahan, G. (2013). Urbanization and economic growth: The arguments and evidence for Africa and Asia. *Environment and Urbanization*, *25*(2), 465–482.

UN. (2019). *Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all*. https://unstats.un.org/sdgs/report/2019/goal-04/

UNDP, U. (2015). Human development report 2015: Work for human development. *United Nations Development Programme*.

Usami, S. (2014). A convenient method and numerical tables for sample size determination in longitudinal-experimental research using multilevel models. *Behavior Research Methods*, *46*(4), 1207–1219.

Van der Leeden, R., Vrijburg, K., & de Leeuw, J. (1996). A review of two different approaches for the analysis of growth data using longitudinal mixed linear models: Comparing hierarchical linear regression (ML3, HLM) and repeated measures designs with structured covariance matrices (BMDP5V). *Computational Statistics & Data Analysis*, *21*(5), 583–605.

Van Deursen, A. J., Van Dijk, J. A., & Peter, M. (2015). Increasing inequalities in what we do online: A longitudinal cross sectional analysis of Internet activities among the Dutch population (2010 to 2013) over gender, age, education, and income. *Telematics and informatics, 32*(2), 259-272.

Vinod, H. D., & Kaushik, S. K. (2007). Human capital and economic growth: Evidence from developing countries. *The American Economist*, *51*(1), 29–39.

Vlahov, D., & Galea, S. (2002). Urbanization, urbanicity, and health. *Journal of Urban Health*, *79*(1), S1–S12.

Wang, Q., & Su, M. (2019). The effects of urbanization and industrialization on decoupling economic growth from carbon emission–A case study of China. *Sustainable Cities and Society*, *51*, 101758.

West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies.

World Bank. (2021). *World Bank Country and Lending Groups*. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

Wothke, W. (2000). Longitudinal and multigroup modeling with missing data. In *Modeling longitudinal and multilevel data* (pp. 205-224). Psychology Press.

Xiang, Y. (2009). *Ethnic differences in achievement growth: Longitudinal data analysis of math achievement in a hierarchical linear modeling framework* [PhD Thesis]. Boston College.

Xu, T., Ma, T., Zhou, C., & Zhou, Y. (2014). Characterizing spatio-temporal dynamics of urbanization in China using time series of DMSP/OLS night light data. *Remote Sensing*, *6*(8), 7708–7731.

Yakunina, R. P., & Bychkov, G. A. (2015). Correlation analysis of the components of the human development index across countries. *Procedia Economics and Finance*, *24*, 766–771.

Yang, J., Huang, X., & Liu, X. (2014). An analysis of education inequality in China. *International Journal of Educational Development*, *37*, 2–10.

Yu, B., Shi, K., Hu, Y., Huang, C., Chen, Z., & Wu, J. (2015). Poverty evaluation using NPP-VIIRS nighttime light composite data at the county level in China. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *8*(3), 1217–1229.

Zhang, Q., & Seto, K. C. (2011). Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data. *Remote Sensing of Environment*, *115*(9), 2320–2329.

Zhao, Y., & Wang, S. (2015). The relationship between urbanization, economic growth and energy consumption in China: An econometric perspective analysis. *Sustainability*, *7*(5), 5609–5627.