Understanding the Dimensions of Education Inequality in China at Different Geographical Scales

Lili Xiang

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds
School of Geography

February 2019

The candidate confirms that the work submitted is his/her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

The work in Chapter 6 of the thesis has appeared in publication as follows:

Xiang, L.,Stillwell, J.,Burns, L.,Heppenstall, A. and Norman, P. 2018. A geodemographic classification of sub-districts to identify education inequality in Central Beijing. *Computers, Environment and Urban Systems.* **70**, pp.59-70.

Chapter 4 includes work from the joint publication:

Xiang, L.,Stillwell, J.,Burns, L., and Heppenstall, A. 2019. Measuring and assessing regional education inequalities in China under changing policy regimes. *Applied Spatial Analysis and Policy.*

Acknowledgements

The three years of PhD study has been a very memorable time for me, comprising stress, anxiety, excitement and happiness. I am extremely appreciative of the help from my supervisors, Professor Alison Heppenstall, Professor John Stillwell and Dr Luke Burns, for their invaluable advice and great support for my research. I am very grateful to John for his tirelessly reading and very detailed 'red marks' that have helped me greatly to improve my research and writing skills, to Alison for her encouragement and help during hard times, and to Luke for his support with the methods I have used and helpful reading materials. I would like to thank Dr Paul Norman for providing me the chance to study in Leeds and his help with my first published paper. I also want to thank the China Scholarship Council and the University of Leeds for their respective sponsorship.

I am very grateful to my research support groups members, Dr Paul Waley and Dr Gordon Mitchell. They have been very kind to provide me valuable advice and feedback on my research. Many thanks to Professor Philip Rees for his help on geodemographic classification, Professor Adrian Favell for providing me suggestions about capability approach, and Dr Myles Gould for kindly spending his time to help me conduct multilevel modelling.

I also want to thank Nicholas Hood, Nik Lomax, Graham Clarke, Andy Evans, Andy Turner, Alexis Comber, Helen Durham, Nicolas Malleson, Andy Newing, Rachel Oldroyd, Jacqui Manton, Nichola Wood and Alison Suckall, who kindly offered help on many occasions during my PhD. Many thanks to my colleagues and friends including Annie Connolly, Sarah Fell, Lawrence Eagle, Jiren Xu, Rui Yu, Yuanxuan Yang, Junfan Lin, Changjia Li, Yunxia Wang, Wanyun Ying, Zi Ye, Asib Ahmed, Antonio Sliver, Nikee Groot, and Arif Rohman, who provided me with a great deal of help and encouragement and made my life in Leeds bright and colourful.

Huge thanks to my parents, who continue to provide unconditional love and mental support. Finally, I want to give special thanks to my husband, Wen Zeng, who shares my sadness and joy.

Abstract

Although education equality has been valued and frequently discussed by scholars from different disciplines, theoretical discussions and empirical studies of education equality from a geographical perspective have been somewhat limited. Since the traditional two-dimensional non-spatial framework for measuring and analysing education inequality is inadequate, the research in this thesis is based on a more comprehensive and flexible three-dimensional framework, in which geography is included as an important dimension.

China is used as the case study country to examine education inequalities at different geographical scales. At a regional scale, a multidimensional Index of Regional Education Advantage (IREA), underpinned by Amartya Sen's capability approach, is introduced to evaluate the effectiveness of policies targeted at reducing regional/provincial educational inequalities in China since 2005. At a local scale, the thesis explores the use of geodemographics as a means of assessing potential inequality in access to compulsory education within urban areas. The thesis argues that applying an area classification, one of the first in China, allows consideration of multi-dimensional, socio-spatial influences which affect school choice within urban areas. The ideas are illustrated through a case study of Central Beijing. At the micro scale, multilevel modelling is used to reveal the influence of contextual factors and confounding individual level socio-economic characteristics on pupils' travel distance to school in Beijing.

The results at the regional scale revealed that education in north-eastern China is better than in the south-west of the country, a pattern which lacks conformity with the eastern, middle and western macro-divisions adopted by Central Government as the basis of policy implementation. Furthermore, the social and spatial disparities in terms of access to education facilities within urban areas were also identified. This research has, for the first time, revealed education inequality in China comprehensively from a geographical perspective, and provides some unique insights and crucial policy implications of education inequalities in China at different geographical scales.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AYS Average Years of Schooling

CHTS Comprehensive Household Travel Survey

CPH Capped-Price Housing

CRH Cheap Rent Housing

DME Deprivation and Marginalisation in Education database

DRES The Decision on the Reform of the Education System

EAH Economically Affordable Housing

ECH Economic and Comfortable Housing

EFA The Education for All Global Monitoring Report

GEM Global Education Monitoring Report

GER Gross Enrolment Ratio

GNP Gross National Product

ICC Intraclass Correlation Coefficients

IGLS Iterative Generalised Least Squares

IID Independent and Identical Distribution

ISCED International Standard Classification of Education

LEA Local Education Authority

LRT Likelihood Ratio Test

MAUP Modifiable Areal Unit Problem

MDGs Millennium Development Goals

ML Maximum Likelihood Method

NMLERD National Medium and Long-Term Educational Reform and Development Plan Outline

NR Notifications of the Reform of the Funding Guarantee System for Rural Compulsory Education

NU Notifications of the State Council on Exemption of Tuition and Miscellaneous Fees for Compulsory Education in Urban Areas

NUR Notifications of Further Improvement of the Funding Guarantee System for Urban and rural Compulsory Education

OAC Output Area Classification

OECD Organization for Economic Co-operation and Development

OM Hukou Outside A Municipality

ONS Office for National Statistics

PCA Principal Component Analysis

PCGDP Per capita Gross Domestic Product

PISA Programmed for International Student Assessment

PPEE Per-pupil Educational Expenditure

PTS Pupil Travel Survey

RA Rural and Rural-Urban Average PPEE Ratio(s)

PRC The People's Republic of China

RIGLS Restricted or Residual Iterative Generalised Least Squares

RUPR Rural-urban Teachers with Higher Professional Rank Ratio(s)

RUQT Rural-Urban Qualified Teacher Ratio(s)

RUTP Rural-urban Teacher-Pupil Ratio(s)

SDG4 Sustainable Development Goal 4

S.E. Standard Error

SoEs State-owned Enterprises

UNESCO The United Nations Educational, Scientific and Cultural Organisation

VPC The Variance Partition Coefficients

WD Hukou within A District

WIDE World Inequality Database on Education

WM Hukou outside A District but within A Municipality/ Hukou within A Municipality

Chapter 1
Introduction

1.1 Research context

Education is not only a fundamental human right, but it is also a catalyst to reduce poverty, promote social mobility and encourage economic growth by producing a skilled workforce (UNESCO, 2010; Wilkinson and Pickett, 2010). In 2015, the United Nations Sustainable Development Summit formally adopted the 2030 Agenda for Sustainable Development, the key theme of Goal 4 (SDG4) being to "*ensure inclusive and equitable quality education and promote lifelong learning opportunities for all*" (UNESCO, 2015, p1). Compared to previous agendas, it places more emphasis on inclusive and equitable quality education for all children. Education inclusion and equality at all levels can improve productivity, promote social cohesion and encourage economic competitiveness (OECD, 2012). For example, South Korea's rapid and remarkable social and economic success is partly related to its decreasing level of education inequality over the past 30 years (Burt and Park, 2009). Education inequality perpetuates existing income differentials (Holsinger and Jacob, 2009) and social disparities (Breen and Jonsson, 2005; Darling-Hammond, 2007). Lopez-Acevedo (2009), through a national case study of Mexico, asserts that in many cities, income disparities are the result of education inequality, as highly educated employees tend to earn more money than less educated employees.

Education equality has been discussed broadly by scholars from different perspectives and disciplines (Tyler, 1977; Baker *et al.*, 2016b). In terms of geography, there are various forms of education inequality among countries, regions, counties and neighbourhoods, such as unequal resource distribution and accessibility to schools (Soja, 2009; 2010; Holloway *et al.*, 2010; Hamnett and Butler, 2011; Hamnett and Butler, 2013; Waters, 2017). Spatial educational inequalities are not just derivatives of other forms of education inequality, like those caused by ethnicity and class, but may also be the cause of other education inequalities (Marcuse, 2009). Although education equality from a geographical perspective is attracting increasing attention, research in this field has been somewhat limited. Most theories or concepts about equality have invariably avoided the need to contemplate education equality from a geographical perspective. It is apparent that a discussion of education inequalities, which is devoid of spatial or geographical considerations, is inadequate. This history suggests that an opportunity exists for geographers to contribute to this field.

This thesis uses China as an example country to study education inequalities from a geographical perspective. Different geographical scales constitute and shape varied social practice and tend to yield different empirical outcomes (Howitt, 1998; Baden *et al.*, 2007). Thus, education equality at different geographical scales normally has to be theorised or measured respectively. Accordingly, education inequalities in China have been examined at different geographical scales to provide a view of spatial education inequalities across China and to propose more effective policy suggestions for different levels of government to promote such equality.

Providing quality and equal education is a high priority for every national government, including the world’s largest developing country, China. After 30 years of reform and the pursuit of an ‘opening up’ policy, remarkable economic and social achievements have been made by China. However, the goal of achieving education equality is facing unprecedented challenges.

Firstly, at the regional scale, education equality in China is challenged by decentralisation in the education sector. With the shift from a planned to a market oriented economy, education reform based on decentralisation has also been achieved, involving local governments taking the primary responsibility for providing education investment and administering their education systems. Although political and administrative decentralisation can enhance the allocative efficiency of governments and improve the capacity to mobilise underused resources (Kyriacou *et al.*, 2017), it is often tempered by the uneven geographical distribution of benefits and increased disparities across local governments (Rodríguez-Pose and Ezcurra, 2009). Educational decentralisation tightens the link between school resources and local economic circumstances (Hannum, 2003).

Due to this decentralisation of the education system, together with uneven regional economic development and an urban-oriented development mode prevailing in China, regional and rural-urban education disparities have increased. Since 2005, several policies aimed at reducing regional and rural-urban education inequalities have been implemented, including the Law of Compulsory Education (2006) and the National Medium and Long-term Educational Reform and Development Plan Outline (2010-20) (NMLERD) in 2010 (The State Council of China, 2010; Sun, 2012), which attach great importance to education equality. A key question is whether inequalities in education have decreased under these changes; the answer to this question is crucial for formulating and adjusting policies for the future. However, there remains a lack of research which can comprehensively measure spatial variations in education, monitor the impact of policies and evaluate the overall effects.

At a local scale, the uneven allocation of educational resources across space and the involuntary confinement of specific disadvantaged groups has received close attention in recent decades (Bradford, 1990; Breen and Jonsson, 2005; Butler *et al.*, 2013; Hamnett and Butler, 2013). China's continuing urban and socio-economic transformation has led to problems emerging in Chinese cities which are similar to those in megacities in western developed countries (Gu *et al*., 2005). In terms of education, there is a long history of criticism of class stratified school segregation in western cities (Allen, 2007), such as Chicago (Christmann, 2005; Lipman, 2005) and London (Hamnett and Butler, 2011; Harris, 2012; Butler *et al.*, 2013; Harris, 2013). This is largely based on residential segregation and leaves vulnerable children in inferior schools (Butler and Hamnett, 2007; Logan and Burdick-Will, 2016). In China, the gap between the urban rich and poor is widening, with the Gini coefficient for income inequality increasing from 0.30 in 1980 (Xie and Zhou, 2014) to 0.47 in 2016 (National Bureau of Statistics of the PRC, 2017). There is increasingly marked socio-spatial differentiation caused by income differentiation and increasing levels of residential isolation and social exclusion experienced within cities (Li, 2005; Gu *et al*., 2006; Fang *et al.*, 2015). In 1993, more than 40% of university students in China came from lower occupation families, e.g. farmers and workers; however, universities today are increasingly dominated by students from wealthy and more prosperous backgrounds (The Economist, 2016).

In addition, due to the aforementioned unbalanced regional economic development and an urban-oriented development mode in China, there are more job opportunities and better welfare systems in the developed areas and cities, resulting in an increasing number of rural migrants moving to urban areas (Li and Placier, 2015). In 2011, China's urban population exceeded its rural population for the first time and China’s rural population will decrease by one third in the next 20 years with an estimated 300 million more people from rural areas moving to cities and towns. This rapid urbanisation means there will be around 30 million more school-aged children requiring compulsory education in urban areas by 2040 (Xie, 2015). Combined with the recent termination of the ‘one child policy’ (Bi and Zhang, 2016), many cities, especially the megacities such as Beijing, are facing massive challenges in terms of the provision of educational services for the children of in-migrants (Zhang, 2011). Within urban areas in China, the migrants are often disadvantaged in terms of education because of their non-local hukou registration status (Fu and Ren, 2010; Zhao and Howden-Chapman, 2010; Zhang *et al.*, 2015). Local urban citizens are granted priority access to urban services and facilities, while migrants with a non-local hukou normally have limited access to the same services and facilities, including schools (Zhao and Howden-Chapman, 2010).

Combined with the inequalities between local and non-local residents, socio-spatial differentiation within urban areas of China is likely to occur due to multiple factors. So, how do the multi-dimensional socio-spatial structure and socio-spatial differentiation influence residents’ access to educational resources? How does where people live, including geographical location and social composition, affect their access to school? The answer to these questions is crucial for setting educational policies and promoting education equality in urban areas. However, there is a conspicuous lack of detailed studies of education inequality within urban areas and its spatial diversity in particular due to the absence of educational data relating to small areas in China.

It is apparent that geography is not only required to be thought of hierarchically but also has to be considered relationally (Jones, 2009). Different geographical scales are not fragmented but inherently related (MacKinnon, 2011). For example, education development at a lower level unit (e.g. district) will be influenced by the educational policies implemented at a higher level (e.g. province).

At the micro scale, individuals are nested within the neighbourhoods in which they live, so the socio-spatial context will greatly impact on the structure of available educational resources for them (Gordon and Monastiriotis, 2007; Hamnett and Butler, 2011). However, in reality, there is still considerable heterogeneity in social composition in each type of neighbourhood. The pupils living in the same neighbourhood may have different mobility and chances to get access to education facilities due to their varied social background (e.g. household income, hukou) (Talen, 2001; Wu *et al.*, 2008; Hamnett and Butler, 2011). Thus, the level of education service people can access is not only influenced by where they live (Hamnett and Butler, 2011), but also by who they are (Williams and Wang, 2014). In order to abolish inequalities and promote education equality by making sure all residents are treated equally in terms of access to education resources regardless of their geographical location and personal mobility potential (e.g. socioeconomic characteristics, disability), the complicated socio-spatial structure that creates these inequalities should be explored. However, there is still a scarcity of research on education equality within small areas (not only in China but also around the world), which links the meso scale context with individual level analysis.

Thus, there are still considerable challenges facing China when seeking to provide equal education in different regions and different neighbourhoods within urban areas as well as for different social groups. In order to meet the challenges outlined above and promote education equality, this research explores education inequalities in China at different geographical scales and gives corresponding suggestions for policy at each level.

The remainder of this chapter presents the overall research questions, aims and objectives in Section 1.2, and the main content/structure of the thesis in Section 1.3.

1.2 Research questions, overall project aims and specific research objectives

Four key research questions have been prompted, which can be identified as follows:

1. What are the spatial patterns of regional and rural-urban education inequalities in China?
2. Have these regional and rural-urban education inequalities decreased or increased under changing policy regimes?
3. Within urban areas, how do social and spatial factors affect pupils' access to education?
4. Is social and spatial disadvantage associated with poor access to education?

As the answers to these questions are crucial for setting policy for the future, the research questions form and align with the following two broad research aims of the thesis:

First, to accurately measure existing regional and rural-urban education inequalities and quantify how they have changed over time so as to assess the extent to which the policies have been successful in reducing regional education inequalities.

Second, to reveal the education inequalities existing within urban areas by exploring the influence of residents' socio-economic characteristics and areas' specific socio-economic composition on pupils' access to education.

In an attempt to achieve these research aims, the following more specific research objectives have been established, which are addressed in subsequent chapters. The research objectives and the corresponding chapters in which they are addressed are summarised in Table 1.1.

Table 1.1 Research objectives and corresponding chapters

	Research objective	Corresponding chapter(s)
I.	To review the concept and discussion of education equality; give the definition of education equality used in this research; explore and review key research literature on education inequality from different geographical scales; introduce the geographies, specific institutional background and education system in China; build a research framework for this research based on the given theoretical and empirical context.	Chapter 2 Chapter 4 Chapter 5 Chapter 6 Chapter 7
II.	To review the methodological approaches and critically introduce corresponding data used in analysing educational inequality in China.	Chapter 3
III.	To accurately measure regional education inequalities in China and evaluate the effectiveness of policies targeted at reducing these inequalities by introducing a new multidimensional index.	Chapter 4
IV.	To detect the education inequalities between rural and urban areas by comparing rural-urban education and their variations across space.	Chapter 5
V.	To recognise potential educational advantaged and disadvantaged areas in terms of access to compulsory education by creating a bespoke geodemographic classification which is associated with residents' demographic and socio-economic composition, including housing type, parental occupation, qualifications and, in particular, registration status in Central Beijing.	Chapter 6
VI.	To explore the influence of context factors (e.g. geodemographic classification) and confounding individual level predicators (e.g. hukou, household income, and parents' qualifications) on pupils' travel distance in Beijing by adopting multilevel modelling.	Chapter 7
VII.	To summarise the fulfilled aforementioned objectives to answer the main research questions; to review the success of the research project and then give advice on improvement and possible future research	Chapter 8

1.3 Thesis structure

As shown in Table 1.1, each chapter seeks to address one objective respectively aside from Chapter 2, in which the Objective I is met through a combination of reviews across multiple chapters. Chapter 2 seeks to define the concept of education inequality in this research and presents related empirical and theoretical discussions. The research gap of education inequalities from geographical perspectives in China has been summarised. Following this, China's specific institutional background, including its administrative divisions and geographical boundaries, the hukou system (i.e. household registration system in China that each household has to register with a community) and the education system, are explained in detail before the main subject matter of this research is introduced. Indeed, the later chapters (Chapters 4, 5, 6 and 7) provide more detailed and further reviews of the major policies and discussions relating to the relevant geographical scales (e.g. regional, rural-urban and within urban areas) of particular interest. Thus, Chapter 2 not only provides a brief empirical and theoretical context for this study but also contains signposting to the detailed theoretical discussions and policy background of relevance described in the later substantive chapters that follow. After an introduction of China's specific institutional background, an overarching research design is proposed, which serves as a framework and foundation for later analysis of education inequalities in China.

Chapter 3 is concerned with fulfilling Objective II and thus introduces and reviews the methodological approaches adopted in subsequent chapters. Multiple methods have been applied to address the proposed objectives, including composite indices, geodemographics and multilevel modelling. After that, a detailed review of the main datasets used in this research, like 2010 census data, data from Educational Statistics Yearbook of China, China Educational Finance Statistical Yearbook, the fifth Comprehensive Household Travel Survey (CHTS) and the Pupil Travel Survey (PTS) datasets, is presented. In order to meet the data requirements for further analysis, the data cleaning and preparation methods are also reported in this chapter. In addition, the strengths and weaknesses of each dataset are discussed.

Chapter 4 introduces a new multidimensional index, the Index of Regional Education Advantage (IREA), to measure regional education inequalities and to evaluate the effectiveness of policies targeted at reducing regional/provincial educational inequalities in China since 2005. The analysis of the distribution of IREA scores and the decomposition of the IREA reveals the patterns of regional education inequality and drawbacks within current employed education policy are also identified by this index (Objective III).

Chapter 5 seeks to compare the rural-urban education inequalities and their spatial variations. As there is no separate enrolment data for rural and urban areas, only variables related to educational attainment and education provision are compared between rural and urban areas. The spatial variations in rural-urban educational gaps are also discussed to identify which areas tend to have a larger rural-urban education gap (Objective IV). Corresponding policy implications and suggestions are proposed based on the descriptive analysis.

Chapter 6 involves a change of scale and focuses on Beijing as a case study area for reporting the application of a geodemographic classification presented in Chapter 3 to identify potentially education advantaged and disadvantaged areas in terms of access to compulsory education within the city (Objective V). According to current nearby enrolment policies, the variables with direct or indirect influence on pupils' access to education are employed to build the bespoke education related geodemographic classification. Correspondingly, the education-related implications and potential issues within each type of area identified by the classification process are also discussed.

Chapter 7 seeks to explore the influence of pupils' individual level demographic and socio-economic traits and confounding context effects, like the influence of the bespoke geodemographic classification and school density within each sub-district of Beijing, on pupils' travel distance to school (Objective VI) by using the fifth Comprehensive Household Travel Survey (CHTS) and multilevel modelling approach specified in Chapter 3. Before conducting the multilevel analysis, another larger dataset, the Pupil Travel Survey (PTS) dataset is used to validate the small CHTS dataset and complement the analysis results and patterns revealed by analysis of the CHTS. Consequently, not only the influence of micro-level covariates and the context effects of macro-level variables are identified by the detailed descriptive analysis and multilevel modelling, but some implications relating to the social and spatial disparities of pupils in terms of access to education are also revealed in this chapter.

The final chapter, Chapter 8, synthesises the findings of the research project, draws overall conclusions and presents their policy implications for the government (Objective VII). It also summarises the innovations and contributions of the thesis, comments on potential sources of errors and limitations of this research and suggests possible improvements and proposals for future research.

Chapter 2
Education Equality and Institutional Background

2.1 Introduction

As discussed in Chapter 1, education equality has been valued by numerous leading worldwide development agencies. Equality in education matters not only because of its intrinsic worth for every human being but also because it is essential for realising other human rights and offers the chance to counter existing inequalities in other systems (e.g. economic inequalities). However, what is education equality? What ideals of education equality should we believe in? How can we identify education inequalities? This chapter provides a three-dimensional framework for defining, thinking and analysing education equality. Section 2.2 reviews the discussions about the concepts of equality and education equality, in particular, stressing the importance of a geographical perspective in understanding, defining and assessing them (Section 2.3). The three-dimensional framework of looking at education equality has been proposed based on critiques of previous concepts.

There are only a few studies of education inequalities from a geographical perspective in developing countries (Qian and Smyth, 2008; Zhang *et al.*, 2015). This study focuses on exploring education inequalities in China. In order to form an analysis framework specific for the China case study, the institutional background, such as the administrative divisions and the hukou system of China, are introduced in Section 2.4. A review of the key literature of education inequalities from different geographical scales in China is also provided in this section, in which some gaps evident from existing studies also have been observed. The concluding remarks are summarised in Section 2.5. It worth noting that although education is not confined to formal education (Illich, 2002), the following discussions of this research area focus on this area.

2.2 Towards understanding what education equality means

Inequality is inevitable and pervasive in our world. In order to offer a fair basis for regulating education inequality, it is very crucial to define what education equality is (Baker *et al.*, 2016c).

2.2.1 The concept of equality

So, what is equality? Many different concepts of equality have been introduced by scholars (Hare, 1952; Rawls, 1971; Baker *et al.*, 2016a). In general, equality is considered to be a relationship between two or more individuals or groups concerning some aspect(s) of their lives (Baker *et al.*, 2016a). The main differences between different egalitarian theories exist in their answers to the core questions: 'equality of what?' (Sen, 1979) and 'equality for whom?' (Young, 2001).

Firstly, the relative advantages and disadvantages that individuals or groups have can be judged by comparing different variables, such as income, wealth, utility, resources and liberty. The selection of focal variables is vital for inequality analysis and measurement (Sen, 1992). Should researchers focus on whether the target groups enjoy equal overall well-being or good lives? Or should they be more interested in some tangible aims, such as equality of wealth, income and education attainment? Should scholars emphasise the equality of provision or access to resources? Or should they concentrate on the equality of people's capabilities? Different perspectives, applications and egalitarian theories normally have different answers to the question, 'equality of what?' (Sen, 1979). Thus, at the outset of defining and measuring equality, it is crucial to properly answer this question and select focal variables.

Secondly, the idea of equality sometimes refers to individuals, but it is more often applied to groups, such as equality between different social, gender or ethnic groups (Young, 2001). However, in reality, it can be more complicated when there are overlapping groups. In this instance, for example, equality between women and men is interwoven with equality between the working class and middle class or equality between non-disabled and disabled people (Young, 2001). So it is not always simple to define 'equality for whom?'.

2.2.2 Education equality from non-spatial perspectives

Education can be equal or unequal in a number of different ways (Walker and Unterhalter, 2007). Drawing on the framework of defining equality in the above section, the concept of education equality can be defined from two key aspects: ‘equality of what’ and ‘equality for whom’. The answer to the first question, equality of what, also refers to a specific theme or focal variable for educational equality measurement. Education equality related discussions have been conducted in different disciplines, like Philosophy and Sociology, which have offered us a broad menu of answers to this question and created numerous concepts of educational equality. Should researchers emphasise the processes that lead to education inequalities or the unequal outcomes of education? Or should they include both? (Israel and Frenkel, 2017). The choice of focal variables or the space for equality assessment will determine what education equality we prioritize; thus, the focal variables are always selected according to specific perspectives and applications.

In the current field of education, the key concerns for the widely accepted liberal egalitarians is about provision of equal or equivalent learning opportunities (Brighouse, 2000; Ball, 2003; Walker and Unterhalter, 2007). The majority emphasise equal opportunities in different stages of people’s learning process. For example, Húsen (1974) defined equality of educational opportunities in terms of the starting point of education, the process, and finally the outcome. Levin (1976) went further in developing the concept of education equality and summarised four standards for assessing the equality of educational opportunities, including equality of educational access, educational participation, education results and educational effects’ on one’s life chances. Similarly, Farrell (1999) concludes that education equality can be analysed from four aspects; equality of access; equality of survival (i.e. stay in the school system); equality of output (i.e. attainment) and equality of outcome (i.e. same probability of leading similar lives).

The above concepts of education equality only focus on answering the first question (equality of what?). However, the other crucial question (equality for whom?) should also be properly contemplated for defining education equality and equality measurement. It should be recognised by any education inequality study that education is a social good that is distributed under a system with competitive social relations (Baker *et al.*, 2016b). Given that education goods are distributed in a competitive context, different forms of education inequality are revealed as patterns by various social classifications, including education inequalities between individual or groups with different socioeconomic status (e.g. income, social class) (Satz, 2012; Wang, 2012b; Croxford and Raffe, 2014), demographic background (e.g. ethnicity, religion, gender, immigrant background, language, single parent) (Hulton and Furlong, 2001; Diette, 2012; Boterman, 2013) and/or health condition (e.g. disabilities) (Helander, 1993; Holt, 2007).

The existing research has provided us with ample empirical evidence that the economically or socially advantaged groups are best placed to achieve educational success in an economically or socially stratified society (Levin, 1976; Erikson *et al.*, 2005; Butler *et al.*, 2013; Harabet, 2014). Children's education outcomes are influenced by the 'habitus' of family origin, that is, the social origin. Advantaged groups are conferred educational advantages for their children with better access to high quality education resources, participation and therefore better education achievement (Breen and Jonsson, 2005; Allen, 2007; Hamnett and Butler, 2013). In contrast, children from low income or social backgrounds are normally restricted or unable to get access to good schools, to participate or obtain similar education achievement on equal terms compared to children of advantaged groups (Halsey *et al.*, 1980; Green *et al.*, 2003; Teese and Polesel, 2003).

In many societies, there is a strong correlation between the highest level of education one has received and social class background (Halsey *et al.*, 1980; Deer, 2009). A study by Hung and Cheng (2009) examined social economic backgrounds of higher education freshmen students in Taiwan, and the results showed that those students growing up in urban areas, with high incomes and a good educational background continued to enjoy better access to top universities. The difference in quality of and opportunities for education amongst the different social classes for children thus constrains social mobility and exacerbates social class differences (Butler and Hamnett, 2007). Therefore, equal education opportunity is shown to be a vital means of ensuring that everyone has the possibility to achieve success through education and reduce inter-class and intra-class inequality (Bilton, 2002; Law and Pan, 2009).

As well as the education inequality among different social classes, in most countries, it is frequently difficult for immigrants to achieve the necessary mathematical and reading skills. The OECD (2016) (Organization for Economic Co-operation and Development) assessments of science, mathematics and reading indicate that the children of first and second generation migrants tend to perform less well than their native counterparts. In some parts of the world, the education system is challenged to provide education to people of certain religions or who speak a different language such as Muslim migrants in Europe (Daun, 2009). Continuing gender inequality in education is also highlighted by scholars (Herz and Sperling, 2004; Holsinger and Jacob, 2009). In addition, informal and formal exclusion of disabled people from education is also a global phenomenon (Helander, 1993; Brighouse, 2000).

It is obvious that, in reality, education inequality occur due to multiple factors rather than a single factor since individuals or groups belong to overlapping groups (Holsinger and Jacob, 2009). In this instance, education inequality between different social classes may be interleaved with inequality between men and women or other forms of inequality. Figure 2.1 shows that a high probability of children (aged 15) in OECD countries performing poorly can be a result of a combination or accumulation of various disadvantages, such socio-economic background, gender or immigrant background.

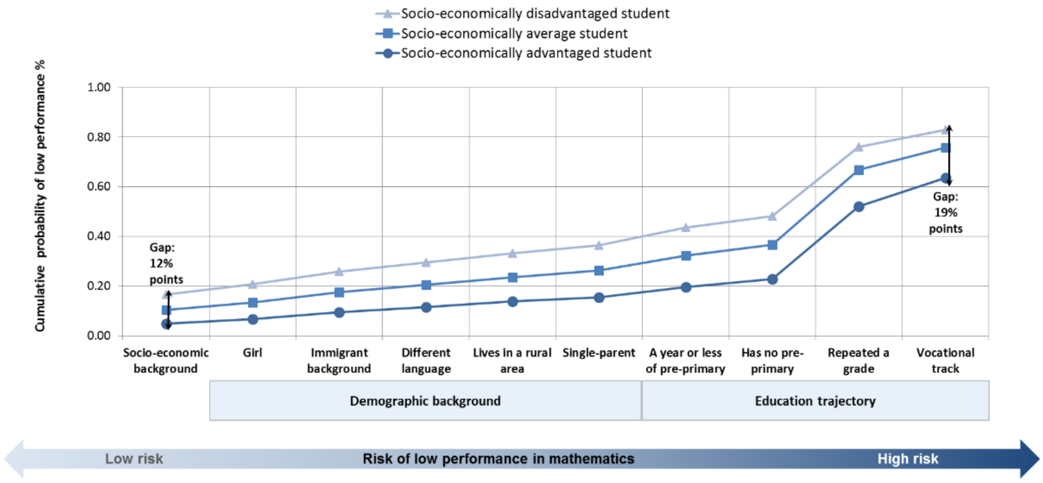


Figure 2.1 Cumulative average probability of low performance in mathematics across various socio-economic groups

Source: OECD (2016) *Education at a Glance* 2016

2.2.3 Education equality from a geographical perspective

In contrast to the above non-spatial perspectives, education equality also can be examined from the perspective of geography. There exist different forms of education inequality at various geographical scales (e.g. countries, regions, neighbourhoods), such as uneven educational resources allocation or differential accessibility to schools, which are attracting increasing attention from education researchers (Holloway *et al.*, 2010; Hamnett and Butler, 2011; Hamnett and Butler, 2013; Waters, 2017; Holt *et al.*, 2019). Spatial education inequalities are derived from or reflect other forms of education inequality. Unequal geographies of education are embedded in the asymmetry of power relations related to class cleavages, race and gender within certain social arrangements (Israel and Frenkel, 2017). Moreover, the spatial aspects of the arrangement of education can generate education inequalities or increase/decrease other forms of education inequality. Thus, considering education inequalities spatially cannot only enrich our theoretical understanding but also can help reveal new insights about education inequalities that extend our practical knowledge to take more effective actions to achieve education equality (Soja, 2009). However, due to the relative nascent studies in this field, the theories or concepts about education equality invariably elegantly avoid contemplating education equality from a geographical perspective.

Therefore, geography or spatiality has been incorporated into defining education equality in this research and extends previous non-spatial definitions and discussions about education equality. As spatial education inequality is not a substitute or alternative of other forms of education inequality, it is regarded as the third dimension of education equality, i.e. ‘equality of where?’. Geographical scale, as one of the fundamental concepts of geography, should be on the agenda for discussion (Howitt, 1998; Martin, 2008). Different spatial scales (e.g. national, regional and local) may produce different theories of education inequality, as well as having different practical implications (Israel and Frenkel, 2017). It is obvious that education equality can be investigated and defined from different perspectives and cannot be easily summarised. Rather than providing a specific definition of education equality, this research offers a comprehensive three-dimensional framework for defining and analysing education equality (Figure 2.2). This framework is designed to help users or researchers to define the concepts of education equality according to their specific research requirements and research areas.

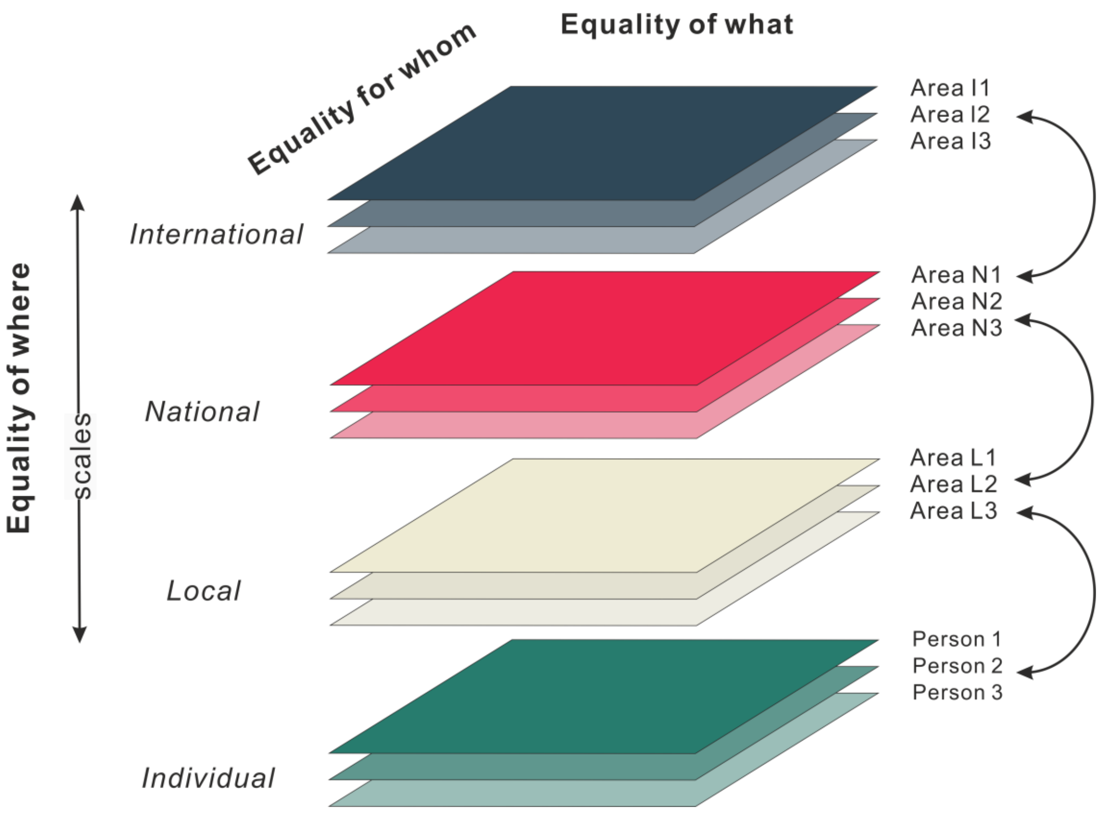


Figure 2.2 The three-dimensional framework of education equality

This three-dimensional framework is built according to the three main dimensions: 'equality of where?', 'equality for whom?' and 'equality of what?'. Previous frequently used non-spatial definitions of education equality are only one or two-dimensional. Given different geographical scales may require different theorisation of education inequality, each geographical scale thus could have different answers to the last two questions (whom? and what?). Accordingly, the first question should be used in the first place to confirm the target geographical scale and areas for defining, analysing and measuring education equality. For example, education equality can be considered at the international scale (e.g. all countries in the world, Asia and Europe), the national scale (e.g. different regions or provinces), the local scale (e.g. sub-districts or neighbourhoods). In this research, we focus on discussing education equality from the macro perspective, regional scale, and meso perspective, sub-district scale. For given geographical areas, 'who' can be the people in each area as a whole (e.g. inequality in high school enrolment rate between provinces) and one group (e.g. inequality in high school enrolment rates for girls between provinces). In addition, it also can be different socio-economic and demographic groups (e.g. socio-spatial inequalities in access to high performance school); in this context, multivariate explanations are required to understand the socio-spatial complexities associated with disparities in education.

After confirming the answers to the first two questions, egalitarians still vary substantially in their replies to the third question, 'equality of what'. As discussed in the early section, some researchers focus on the equal education opportunities in the learning process, like the four aspects stressed by Farrell (1999), including equality of access, survival, output and outcome. Some resources-based normative theories stress provision in education (Dworkin, 2002; 2018), that is, the equality in provision of education related resources. Sen (1979) contends that what we should prioritise and equalise is not education outcomes (e.g. pupils' educational attainment) or resources (e.g. per-pupil educational expenditure or teacher pupil ratio), but human capabilities; that is, the freedom people actually enjoy to do and to be what they value in and from education (e.g. the capability to access a good school or the capability to attain a certain level of qualification) (Walker and Unterhalter, 2007).

The answer to ‘equality of what?’ implies what ultimately matters and the space in which to assess education equality, and thus it also should be decided and confirmed according to specific research objectives. For example, Sen argues that evaluation of policies and educational arrangements should focus on accessing people’s capabilities (what people are able to do and to be) (Robeyns, 2005) rather than how many education resources they obtained or what qualifications they achieved. Because these policies normally aim at removing obstacles in people’s lives, they provide them with more freedom to live or achieve what they have reason to value (Walker and Unterhalter, 2007). Thus, the capabilities approach has been used to build the criteria and framework to evaluate regional education policies in China, which will be introduced in detail in Chapter 4.

The different geographical scales are inherently related; the most evident relationships are the ‘vertical’ links between various levels of geographies (Jones, 2009; MacKinnon, 2011). Recognising the relations among geographies can help reveal a more fluid and dynamic conception of space (Amin, 2002). Although there are some crucial horizontal linkages, the framework in this research only emphasises the vertical relationship in order to avoid making the definition and analysis of related concepts too complicated. Education development at the higher level geographical scale will influence education at the lower level geographies through context effects or other socio-spatial dynamics and *vice versa* (education at lower level geography will also influence its container, the higher level unit). For example, the educational or financial policies implemented at a higher level unit, like one province, tend to exert effects on all the lower level units, like counties, within it. In addition, one advantage of this framework is that it also links the meso scale analysis and individual level analysis and provides explanatory and descriptive tools to illustrate how context effects and individuals’ characteristics jointly constrain individual’s access to education.

Despite numerous attempts to define education equality and to classify different types of egalitarianism, the theories are invariably lacking considerations of geography. The three-dimensional framework developed in this research is one alternative that has extended the traditional two-dimensional framework for defining the concepts of education equality. This open-ended and networked framework offers a three-dimensional definition and evaluation space to enable researchers to define, analyse and measure education equality according to their specific research areas and targets. Moreover, the framework advocates relational thinking and acknowledges the inherent 'vertical' links and context effects (e.g. political and policy) between different scales and bounded territories, which is beneficial for education inequality analysis. For example, the macro perspectives of education inequality can be linked to and benefit from the micro perspectives. A specific research framework of education equality for this research is proposed (Section 2.4.6) based on the identified research gap in the later sections.

In the following section, a broad map of education inequalities within the world has been provided. The focus of this study is basic education, which includes primary school and secondary school education. Thus, the following review focuses on basic education.

2.3 Education inequality around the world

With various measures and initiatives in education, great strides have been achieved in reducing different forms of worldwide education inequalities. However, education opportunity variations between different countries and regions are still at the forefront of debates involving researchers and international organisations (Goesling, 2001; Holsinger and Jacob, 2009; UNESCO, 2010; UNESCO, 2016). According to the statistics of UNESCO (United Nations Educational, Scientific and Cultural Organization), most countries have very high Gross Enrolment Ratios (GER)[1] for primary and secondary education at present, but there are still some developing countries such as Sudan, Niger and Mali that have comparatively low GERs. On

[1] A country's Gross Enrolment Ratio (GER) at primary school level is defined as the total number of enrolled pupils regardless of age expressed as a percentage of the population at official primary education age. As it will include the over-age and under-age pupils, the GER may exceed 100%.

average, the developing countries' GER for primary education reached 91% by 2015, but there still remain 57 million children out of school around the world, more than half of whom are living in sub-Saharan Africa (United Nations, 2015). The GER for secondary education[2] is lower than that of primary school education worldwide, and there was a more significant differentiation between different countries with developing countries, like India (74.3%) and Malaysia (77.7%), having remarkably lower GERs than those of developed countries, Like UK (127.8%). Most African countries, such as Niger (18.8%) and Uganda (26.1%), have extremely low GERs for secondary education. It worth noting that the developing countries, like Chile (100.4%) and Argentina (106.8%), that have invested more in education, also have comparatively high GERs for secondary education. From the data on GER, we can conclude that the worldwide education development level is uneven and the development level of education seems more or less consistent with economic development in most countries; that is, the areas with lower GERs also tend to be the less developed areas, while developed countries usually have higher GERs.

Due to the uneven education development in the world, several worldwide development initiatives, like the previous Millennium Development Goals (MDGs) (2000-2015) and the current SDG4 (2015-2030) have included worldwide education equality and equity as their key theme. In order to monitor the progress of achieving MDGs and SDG4 and promote education equality worldwide, several authorities and institutions have put forward various indicators as well as monitoring reports, such as the *Education for All Global Monitoring Report* (EFA) and the *Education at a Glance* (OECD, 2016).

The *Education for All Global Monitoring Report* (EFA) is an authoritative report, which has been published on 12 occasions by UNESCO between 2002 and 2015. A Deprivation and Marginalisation in Education (DME) database was created to sustain the commitment towards EFA in 2010. It focused on 'marginalisation in education', which is persistent and unacceptable levels of

---

[2] Secondary education is defined as levels 5 to 8 of the International Standard Classification of Education (ISCED).

disadvantage. The DME dataset uses a number of measures to document the different dimensions of marginalisation; for example, education poverty, which is defined as the proportion of population with less than four years schooling, is the minimum education required to acquire basic literacy skills (UNESCO, 2010). It is apparent that the EFA report does not emphasise equal educational opportunities in every country, but is concerned with whether each country has achieved certain minimum educational standards, since receiving basic education is a universal right (Baker *et al.*, 2016c).

The EFA has been renamed and relaunched as the *Global Education Monitoring Report* (GEM) to monitor world education development and meet the aspirations of SDG4 in the next 15 years. Accordingly, the World Inequality Database on Education (WIDE) has been updated and extended from the DME dataset and draws attention to where education inequality is at unacceptable levels, aiming at informing policy design and public debate. WIDE brings together data from the Multiple Indicator Cluster Survey, the Demographic Health Survey, learning achievement surveys (e.g. Programmed for International Student Assessment or PISA) and national household surveys (e.g. China Family Panel Studies), which covers indicators from three domains: access; completion and learning resources.

Another international monitoring report on educational inequality and development, OECD's *Education at a Glance* report, is annually published by the OECD Directorate for Education and Skills to help member countries to develop policies to increase efficiency in schooling by comparing disparities in education among their member countries. OECD (2016) offers a rich set of quantitative, internationally comparable and up-to-date indicators, which reflect professional consensus on how to measure education internationally. The 2016 indicators of *Education at a Glance* cover the following four aspects: (1) the output of educational institutions and the impact of learning; (2) financial and human resources invested in education; (3) access to education, participation and progression; and (4) the learning environment and organisation of schools.

To conclude, although these international agendas on education inequality aspire to achieve different egalitarian objectives and these reports may have different sets of indicators, they measure education mainly from the following aspects: enrolment or transition rates, achievement or attainment and education provision. These domains provide references for other macro scale (e.g. regional) education inequality measurements, for example, provincial level education inequalities in China are measured from three aspects in this research: enrolment, attainment and provision.

2.4 Education inequality at different geographical scales and challenges facing China's education

Although the work on education inequalities from a geographical perspective is thriving, most of the theoretical frameworks and empirical case studies are focused on developed countries and there is a limited amount of studies in developing countries, like China. Therefore, this research has focused upon exploring the education inequalities from different geographical scales in China, which can provide a particular contribution and great reference value to discussions on developing countries as a whole. Therefore, the gaps in research on education inequalities at different geographical scales are summarised in this section. In addition, the Chinese institutional background and the specific analysis framework used for the research are introduced in the final part.

2.4.1 Regional education inequalities and decentralisation

In a national education context, since education resources are limited, the ability to accurately identify education inequalities is increasingly crucial for governments of each country to ensure that additional resources are directed to areas in greatest need. Over the last fifty years, political and administrative decentralisation has become increasingly commonplace in the world (Rodríguez-Pose and Ezcurra, 2009; Adam *et al.*, 2014). Decentralisation can also be observed in the field of education (Amelsvoort and Scheerens, 1997). However, the transfer of resources and powers from central to local government will be more beneficial to those regions which are most prosperous, with better institutions and socio-economic endowments (Cheshire and Gordon, 1998) and this is likely to perpetuate existing territorial disparities (Bardhan, 2002). In addition, decentralisation undermines the redistributive capability of the central government and its power to play an equalising role (Prud'homme, 1995).

As a consequence, the extensive trends of downwards shifting of power in some countries gradually lost some ground and re-centralisation occurred subsequently or simultaneously (De Vries, 2000). In terms of education, there is a tension between pursuing education equality on the one hand and striving for greater efficiency in the allocation and mobilisation of education resources on the other; the policies of the central government are at the heart of this dilemma. Thus, changes in the distribution of finance and power not only occur in one direction. Instead, the development of the social system, like the education system, is characterised by a continuous process of altering the responsibilities between local and central governments and involving complex negotiation among various stakeholders and actors (Minas *et al.*, 2012). The process of (re)-centralisation can occur simultaneously or after the processes of decentralisation, when the central government tries to follow regional autonomy in policy making and implementation to ensure that educational development takes place in accordance with national intentions.

Decentralisation and (re)-centralisation trends have been experienced not only in developed countries, like Sweden, the Netherlands and the UK (Minas *et al.*, 2012; Bardhan, 2002) but also in many developing countries, like China. Since the open-door policy was launched in 1978, China has achieved unprecedented economic growth and vigorous urbanisation (Li and Wei, 2010; Chen *et al.*, 2013). However, the rapid economic development has been accompanied by intensifying inequalities, such as increasing provincial economic inequalities and widening gap between rural and urban areas (Zhang and Kanbur, 2009; He *et al.*, 2018). Moreover, regional education disparities also have increased in China due to education reform based on decentralisation (Thomas and Peng, 2010). As a consequence, several policies have been implemented by China’s Central Government to reduce regional education inequalities since 2005 (The State Council of China, 2010; Sun, 2012).

Accordingly, evaluation of these policies is needed to assess their success and to set policy in the future. Firstly, it is crucial to accurately measure existing inequalities and quantify how they have changed over time. We can then attempt an assessment of the extent to which the policies have been successful and make suggestions about how policies can further reduce education inequalities in the future. However, there is a limited amount of published research on regional education inequalities in China and most of this work has used a Gini Index to measure inequality in terms of education attainment (Qian and Smyth, 2008; Wang, 2014). The Gini coefficient has been frequently used in worldwide studies to measure the educational attainment and distribution of educational resources, which provide an educational inequality index, giving a comparative and unique view of national education systems (Jacob and Holsinger, 2009). The Gini coefficient ranges from 1 which indicates perfect inequality to 0 which indicates perfect equality; that is, the smaller the Gini coefficient and the more equal the distribution is. In this context, the Gini is a score that reflects the extent of overall inequality within an area; however, a single score for a whole country (China) cannot distinguish the locations of disadvantaged and advantaged areas; different distributions of regional education inequalities may get the same score, so it becomes impossible to identify disadvantaged areas with the Gini Index.

The Theil index is also employed by scholars to compare the educational attainment distribution over time across countries (Lopez-Acevedo, 2009; Thomas and Wang, 2009). However, it also has the same limitation and cannot be used to identify disadvantaged areas. Moreover, education equality is a multidimensional issue, so the use of only one indicator (attainment) may lead to particularly narrow results. Implementing a policy based on improper evaluation may cause detrimental effects on education development (Vaughan, 2007). Thus, a composite index of education inequalities is required as a more robust measure to investigate regional education disparities and evaluate the effects of applied policy. Such a measure is developed and applied in Chapter 4.

2.4.2 Local level education inequalities and socio-spatial access to schools

In this section, education inequality has been defined at the local scale, e.g. spatial education inequality within urban areas. The discussion about education inequality in small areas from a geographical perspective normally focuses on two main forms: the uneven distribution of educational resources across space and the involuntary confinement of some groups with specific social characteristics (Israel and Frenkel, 2017). The former indicates supply constraints (Wu, 2009), which refer to unequal spatial separation or proximity to education facilities for residents (Chang and Liao, 2011), and the latter implies demand constraints. The assumption of education equality underpinning the first type of research is that the residents should be treated equally in terms of access to education resources, regardless of their location. It is usually assessed based on the distribution of educational resources, or the accessibility of pupils to educational resources (Talen, 2001; Dadashpoor *et al.*, 2016).

Accessibility is a frequently-used concept in geography, but it has no universal definition. In general, accessibility means the difficulty that a certain activity encounters when trying to reach a certain place or site in a given transportation system (Morris *et al.*, 1979; Johnston, 2000; Cowan, 2005; Gutiérrez *et al.*, 2009). As the use of public facilities is closely linked to accessibility, citizens' proximity to facilities and services can influence their wellbeing in a number of ways. If access to facilities is easier and more direct, this encourages opportunities by reducing the time and financial costs of access (Pearce *et al.*, 2007).

It is clear that different levels of accessibility will provide students with different amount of disposable time. Some of them need to spend large amounts of time commuting to school, while others can walk in no time (Talen, 2001). Research has concluded that the longer pupil travel time is correlated with lower qualifications, poorer health conditions, shorter sleep time, poor learning habits and a reduction of social activity amongst pupils (Fox, 1996). Considerable research in western countries had already shown that students' learning achievements and the probability of progressing on to post-compulsory education are correlated negatively with travel distance (from community to school) (Talen, 2001; Gordon and Monastiriotis, 2007; Dickerson and McIntosh, 2013).

A study of education in Tajikistan, a country which has undergone immense economic, political and social upheaval, found that the availability of secondary schools within each community is important. The school enrolment rate in a community without a secondary school is up to 10% lower amongst children living in communities with a secondary school (Baschieri and Falkingham, 2009). Moreover, children living far from school also have less opportunity to be involved in after-school activities and to get other additional user benefits (Pacione, 1989).

Accessibility research in western cities is largely associated with school consolidation and closure (Talen, 2001; Müller, 2011; Lin *et al.*, 2014). This occurs more frequently in rural areas of developed countries because of their scattered population (Lin *et al.*, 2014). Consolidation increases school size and creates longer travel distance for pupils from home to school, and greater separation between school and community (Talen, 2001). Some urban areas also struggle with the same problem. Many developed countries have seen decreasing fertility in the past few decades, which has led to school consolidation or closure because of the decline on the number of school-aged children, like the case study of Dresden in East Germany (Müller, 2011).

In developing countries, the variations in accessibility usually go hand in hand with the urbanisation process, and are normally due to insufficient facilities in newly urbanised areas (Tsou *et al.*, 2005; Baschieri and Falkingham, 2009). For example, a case study showed that Chile was currently facing a severe test in providing fair education and areas near urban centres had better accessibility, education quality and better school outcomes (De la Fuente *et al.*, 2013).

The above discussions emphasise the importance of spatial separation between schools and residential areas and only measure education inequality from two dimensions, 'equality of where?' and 'equality of what?' (e.g. accessibility to schools). However, its relation with some involuntary confinement should also be recognised. Given that spatial inequality is normally related to and/or derived from some social inequalities, the locational discrimination of education is also normally due to the biases of education imposed on certain social groups and linked to their geographical locations.

There are many social factors or forces shaping the spatial or locational discriminations in terms of access to education resources, such as social class and ethnicity (Soja, 2009). Different human clusters are formed by people (agency) and they tend to be gather into the spatial configurations following their habitus and socio-economic characteristics in certain economic and political processes. People belonging to similar neighbourhoods (social stratification) tend to share similar social status, have a similar habitus and potentially enjoy similar chances of accessing education resources (Israel and Frenkel, 2017).

Some researches argue that children's educational opportunities and aspirations of educational achievement are significantly influenced by the living environment (e.g. educational level) of the neighbourhood where they are raised (Brighouse *et al.*, 2010). In some segregated neighbourhoods, segregation denies certain social groups access to resources, services and employment opportunities, and thus the segregated groups are forced into marginalisation, deprivation, exclusion and isolation (Schnell *et al.*, 2015).

In terms of education, residential segregation usually leads to school segregation, which leaves vulnerable children in inferior schools (Butler and Hamnett, 2007). Therefore, residential segregation has been taken as the key to explain unequal education opportunities within urban areas, especially in the USA (Rickles *et al.*, 2004; Hallett and Venegas, 2011; Diette, 2012; Logan *et al.*, 2012; Satz, 2012; Rangel, 2013; Logan and Burdick-Will, 2016). There are quite a lot of single-race or segregated neighbourhoods in the cities across the USA. As a result, the children living there tend to attend racially isolated schools (Schofield, 1989). The significant racial segregation in the public education system in the USA is caused by the obvious social class, income and achievement differences between black and white families (Logan *et al.*, 2012; Satz, 2012). In addition to ethnic segregation in the USA, residential segregation in Europe tends to be caused by class division (Hamnett and Butler, 2011; Hallett and Venegas, 2011; Dickerson and McIntosh, 2013).

In the UK, distance from home to school has become the primary means or basis of allocating school places for schools; thus popular schools tend to have tighter catchment areas (Hamnett and Butler, 2011). How to gain access to popular schools with better education quality is often an overriding concern to parents, especially for the middle or higher classes, as school attainment has a significant influence on a student's future, like the entrance to university or employment. High income or social class families normally already live in educational privileged areas or are able to buy accommodation in privileged areas (Hamnett and Butler, 2013). On the contrary, low income households' schooling choice is very sensitive to school fees and distance to school (Alderman *et al.*, 2001). However, the socially marginalised people have to withstand education policies which largely disregard their vulnerabilities (Williams and Wang, 2014). Those living in privileged areas continue to have advantaged access to schools. Education inequalities (e.g. access to educational resources) for these groups further reinforce existing inequalities relating to social classification (e.g. ethnicity, social class) (Israel and Frenkel, 2017).

The above discussions of equality of access to education facilities not only stress 'equality of where?', but also are connected to 'equality for whom?', and revealed disparities of socio-spatial access to education resources. However, the research outlined above only focuses on one social variable, whereas, in reality, socio-spatial differentiation in urban areas can be complicated and associated with different social factors. Some scholars argued that socio-spatial analysis would be more reliable when based on combining different social variables rather than just employing one key explanatory variable (Tomintz *et al.*, 2009; Williams and Wang, 2014). Taking China as an example, rapid urbanisation and the specific institutional background in China (e.g. socialist market economic system and hukou system) generated new forms of socio-spatial differentiation and segregation in urban areas, which is based on multiple factors, like income, occupation and residential registration status ('hukou' will be introduced in the following section) (Gu *et al.*, 2006).

Due to the rapid urbanisation, China's rural population is set to decrease by half by 2050, and there are increasingly more rural migrants in urban areas (Li and Placier, 2015). These migrants have limited access to education resources, and thus they are disadvantaged in terms of education (Fu and Ren, 2010; Zhang *et al.*, 2015). As a consequence, except for the education inequalities between different income groups, there is also inequality between local residents and migrants in Chinese cities (Gu *et al.*, 2006; Li and Placier, 2015). However, due to lack of education data relating to small areas in China, there is limited research that considers education inequality under the country's special and complicated socio-spatial structure, and therefore most studies remain at the stage of description (Nielsen *et al.*, 2006; Wang, 2008; Lin *et al.*, 2010). So there is almost no research which comprehensively considers the multi-dimensional socio-spatial access to education resources in China. Accordingly, a multidimensional technique, geodemographic classification is applied in this thesis to assess potential educational inequalities in access to compulsory education in Beijing (Chapter 6).

2.4.3 Education inequalities at individual level and multilevel structure

The main hypothesis resulting from the discussion above assumes that people's living environment (educational resources distribution and social composition of neighbourhoods) influences or determines their opportunities of access to educational resources. That is, the pupils' residential location and social context ('where they live?') have a great influence on the structure of available educational opportunities through affecting available school places, quality of schools and bureaucratic rules defining the school catchment areas and enrolment criteria (Gordon and Monastiriotis, 2007; Hamnett and Butler, 2011). However, each pupil's social background, including demographic status (e.g. age, sex, race), and socioeconomic status (e.g. poverty level, home ownership, car ownership, income), i.e., 'Who they are', also impact on that pupil's mobility and access to education facilities (Talen, 2001; Wu *et al.*, 2008; Hamnett and Butler, 2011; Logan *et al.*, 2012; Williams, 2012; De la Fuente *et al.*, 2013; Rangel, 2013; Williams and Wang, 2014).

On the other hand, some research, like that by Talen (2001), explored the relationship between pupils' travel distance and pupils' socioeconomic traits, but it failed to find a consistent relationship between them across the USA, which is because the context effects of where pupils live are ignored. As pupils are nested within the neighbourhood where they live, using a single level model, which neglects the context effects of neighbourhood and hierarchical structure, will lead to limited analyses. Accordingly, a multilevel structure that creates these inequalities should be adopted. Specifically, how the socio-spatial context of pupils' residential areas (e.g. neighbourhood, sub-district) and their individual level socioeconomic characteristics jointly influence the pupils' access to education resources (e.g. school) and this should be clearly illustrated. However, there is a lack of research on education equality linking context effects with individual level characteristics in China.

2.4.4 The gaps in existing research

Some gaps in current research on education inequalities from different geographical scales in China can now be summarised.

Firstly, under the changing policy regimes, accurately measuring national education inequalities and evaluating the effectiveness of policies targeted at reducing national educational inequalities is central for government to establish national education and financial policies of the next stage. There has been no composite metric used to comprehensively measure regional education inequalities in China, despite the increasingly serious regional education disparities under decentralisation in the field of education. Secondly, due to the absence of data in small areas, there is almost no research exploring the relationship between socio-spatial structure and pupils' access to education resources. Moreover, there is no research using a multilevel approach to reveal how socio-spatial structure and pupils' characteristics jointly influence pupils' access to education resources.

2.4.5 Institutional background of China

In order to develop a specific analysis framework, the institutional background of China will be presented from the three aspects: the geographies and the administrative division system in China; the controversial hukou system and characteristics of education sector in China.

2.4.5.1 The geographies in China

In China, a three-level administrative system divides the country into provincial level units, counties and townships respectively (National People's Congress, 2004). There are 34 provincial level units in China (Figure 2.3), including provinces (23), autonomous regions (5), municipalities directly under the Central Government (4), and special administrative regions (2) (The Central People's Government of PRC, 2005). The analyses in this research only focus on the mainland of China, which consists of 31 provincial level units; Taiwan and two special administrative regions, Hongkong and Maco, are not considered.

Figure 2.3 Provincial level administrative divisions of China

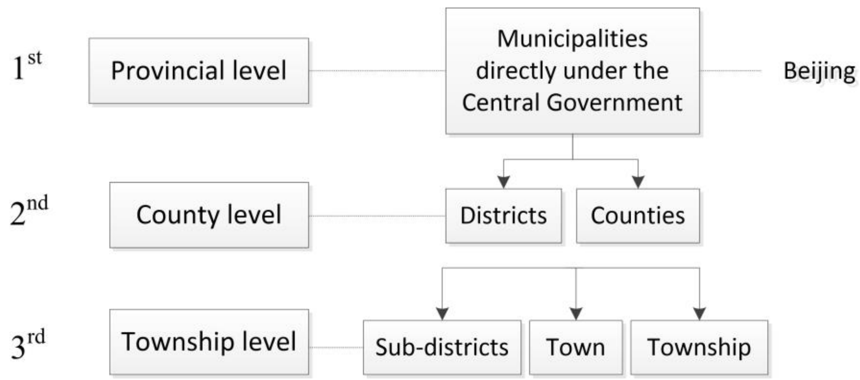


Figure 2.4 Hierarchy of administrative divisions of China

Let us take Beijing as an example to illustrate the three level administrative system (Figure 2.4). Beijing is a provincial level unit, but also a municipality directly under the control of Central Government, which can be subdivided into counties or districts. In China, a municipal district can establish sub-districts (township level) as its agencies. The sub-district government is the basic administrative unit of local government (Zhao *et al.*, 2010). In addition, the data collection for China's decennial census is at the sub-district level, and thus most detailed census data available in China are published at sub-district level (sub-district, township and town). In this research, regional education inequalities will be explored between different provincial level units, while Beijing is selected as the case study area to explore education inequality within urban areas (local scale).

Beijing, as China's political, economic and cultural capital, is a strong magnet for both population and industry. The city has experienced a large influx of in-migrants in the last 20 years (Figure 2.5) and rapid urbanisation has taken place, with 86.4% of its population classified as urban in 2014 (Beijing Statistical Bureau, 2015). In addition, there has been massive urban renewal, shantytown rebuilding and new town development (Yang *et al.*, 2014). Thus, the study of Beijing has value as a reference for research on other cities which are in the process of rapid urbanisation in China. Moreover, it is a transitional city that is governed by a socialist regime but influenced by marketisation and globalisation, and therefore provides a unique and contrasting context *vis á vis* western cities.

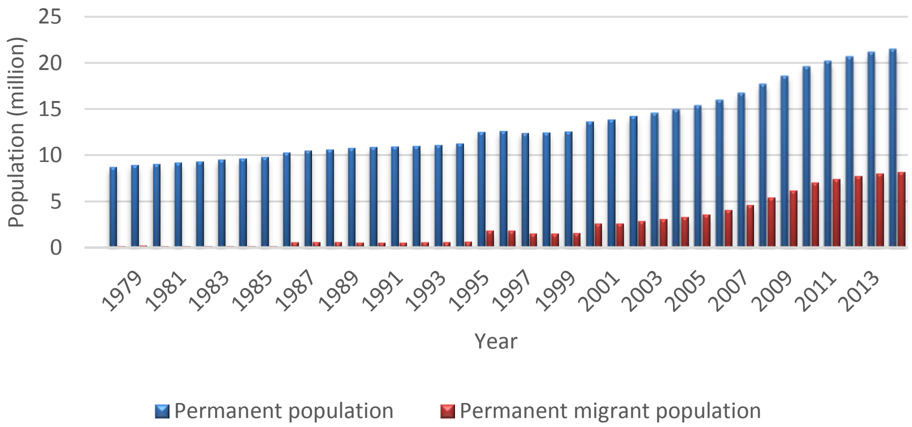


Figure 2.5 Permanent population composition of Beijing, 1978-2014

Source: Beijing Statistical Yearbook, 2015

The city of Beijing is divided into four major zones (Figure 2.6): the Capital Core Functional Area; the Urban Function Extension Area; the Urban Development New Area; and the Ecological Conservation Area. As the boundaries between the final two types are not all consistent with the district boundaries, they are combined in Figure 2.6a. The Capital Core Functional Area and Urban Function Extension Area are usually referred to as 'the city zones of Beijing', which includes the districts of Dongcheng, Xicheng, Chaoyang, Fengtai, Shijingshan and Haidian (Figure 2.6b). The 136 sub-districts of these zones constitute Central Beijing and are selected for this research.

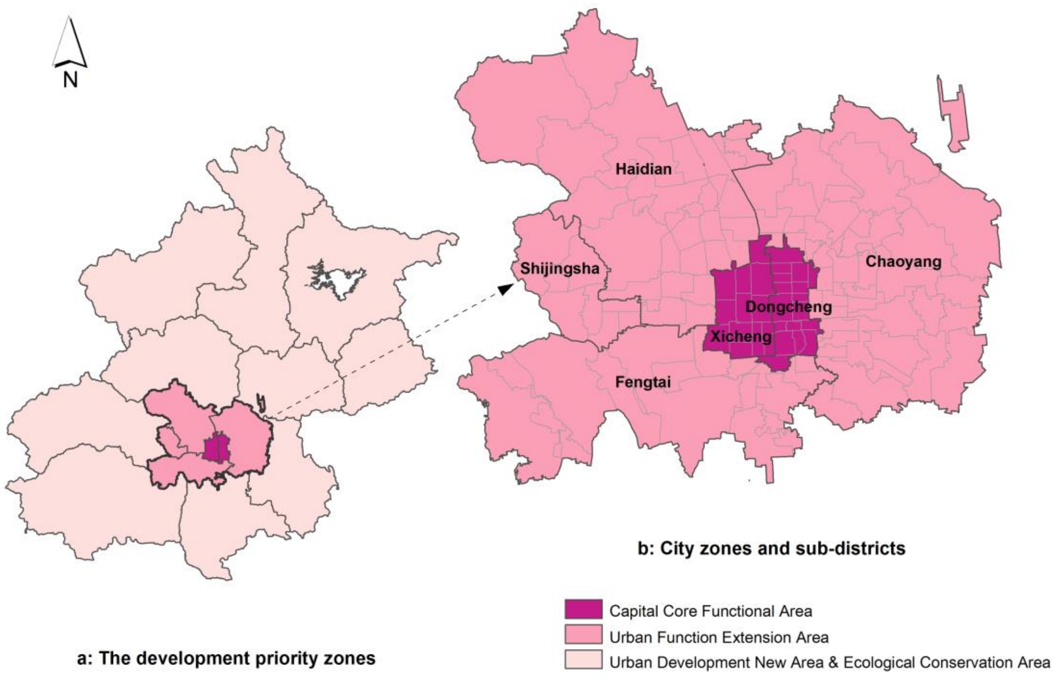


Figure 2.6 The development priority zones, city zones and sub-districts of Central Beijing

Source: National Administration of Surveying, Mapping and Geo-information

2.4.5.2 Education sector in China

The public education system in the mainland China is run by the Ministry of Education. Children in the state education system start at infant school, go on to primary education, and then to a secondary school before entering higher education. The nine years of compulsory education in China normally includes six years of primary education and three years of junior secondary education (middle school), taking pupils from the age of six to the age of 15 (Table 2.1). It is shorter than 13 years compulsory education (from age of 5 to 18) in England.

Table 2.1 Educational system in China

Typical Age (years)	Education		Compulsory
18+	University BA, MA, PhD	Vocational college	No
15 - 18	Senior secondary school education (High school)	Vocational school or Technical school	No
12 - 15	Junior middle school education		Yes
6 - 12	Primary school education		Yes

After nine years of compulsory education, children can either go to senior secondary school (high school) or secondary vocational schools or technical schools, which is the preparation stage before them receiving higher education. Some developed coastal areas (e.g. Shenzhen) in China adopt 12-year compulsory education, which takes three years of high school education as compulsory. However, it is not widely implemented and the compulsory education stipulated by law is still nine years. There is no entrance examination for school-age children to receive compulsory education.

Schools in China are officially divided into two main types because of different funding sources: state schools and 'minban' (literally 'people-run') schools. The state schools are financed by the government and are supposed to be tuition-free (Wu, 2013; 2014), while the 'minban' schools are funded by parents, organisations or other channels (Wu, 2013). The 'minban' schools are subdivided into two types: one type is a non-profit school and the other type expects to have a reasonable return on investment (Wu, 2014). The former is established for poverty-stricken areas or migrants' children in urban areas, while the latter is usually called a private school, and is allowed to charge high tuition fees to make profits (Wu, 2013). Most children in China tend to go to public schools, because there are only a few 'minban' schools available and are considered less attractive compared to good public schools (Zheng *et al.*, 2015); for example, only 65 out of a total 1,040 schools in Beijing were 'minban' schools in 2015 (Beijing Municipal Commission of Education, 2015). Accordingly, this research only considers education provided by state schools.

Although the compulsory education law stipulates that compulsory education schools shall not be ranked (Standing Committee of the National People's Congress, 2006e), state schools are usually classified into two major types: regular schools and key schools (Wu, 2014). Usually, key schools have more qualified teachers, advanced facilities and top-performing pupils compared to regular schools (Wu, 2014). Thus, they are mostly over-subscribed.

2.4.5.3 Household registration system (hukou)

Despite the Compulsory Education Law of PRC stipulating that children should enjoy equal rights to receive compulsory education regardless of sex, ethnicity, economic background and religions, etc. (Standing Committee of the National People's Congress, 2006b), migrant children still cannot have equal educational opportunities within urban areas, which is because of the hukou system.

The hukou system was established in 1958 as a means to control population movement in China. China's population management policy formulation and implementation were based on the hukou system. The basic unit of this system is a household (The National People's Congress of PRC, 1958). Every household in China is required to register with a community which also authorises certain rights to each household.

The hukou system separated the whole population into two types: people with urban hukou and people with rural hukou. In the pre-reform era (before 1978), rural residents with rural hukou were not allowed to freely migrate to the urban areas and thus could not benefit from comparatively higher living standards within urban areas. This system formed 'an invisible wall' among local urban residents and rural migrants (Chan, 1994). This policy played a positive role in the early stages of the 'new' China; for example, it sustained urban social stability during the economic recession and political crisis periods in the 1960s and guaranteed comparatively high employment rates in urban areas. However, it attracted more and more criticism because of its negative social effects, i.e. it led to increasingly more structural social inequalities (Zhao and Howden-Chapman, 2010; Li and Placier, 2015; Wang *et al.*, 2015). The household's hukou has always functioned as the basis for goods and services distribution, with urban hukou holders consuming more goods and services than those with rural hukou (Zhao and Howden-Chapman, 2010).

Although the division of rural and urban hukou was abolished by the State Council (2014), the difference between local hukou and non-local hukou (migrants) status still exists. Residents with local hukou are normally granted priority access to urban services and facilities, while migrants with non-local hukou gain limited access to the same facilities and serveries, such as public housing, health care facilities and schools (Zhao and Howden-Chapman, 2010). Thus, in-migrants are disadvantaged in receiving education in urban areas, an issue which will be explored in detail in later chapters of the thesis.

2.4.6 A three-dimensional framework for this study

Based on the research gaps that have been identified and China's specific institutional background and education system, this research has addressed educational inequalities using the three-dimensional framework (Figure 2.7), which is adapted from the framework introduced in Section 2.2.3 (Figure 2.2).

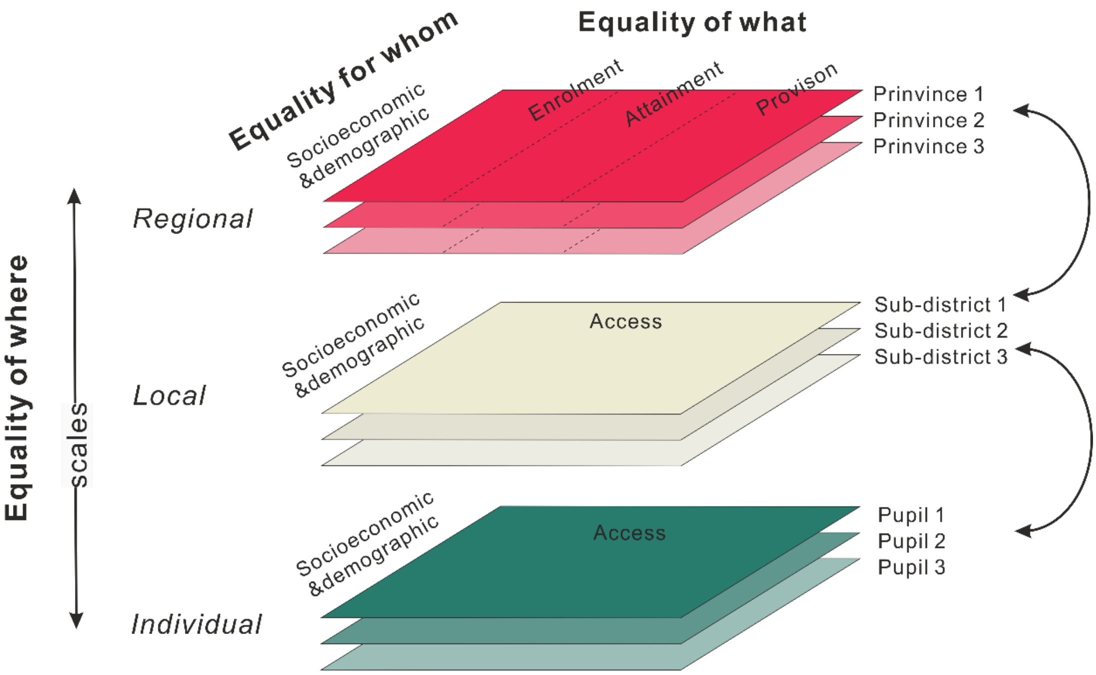


Figure 2.7 A three-dimensional framework for the thesis

Education inequality has been measured at three scales: regional (or provincial), local and individual. In terms of regional education inequalities, a multidimensional index has been introduced to measure education inequalities between different provinces and evaluate the effectiveness of current educational policies. Three facets have been identified to build the above index, that is, education enrolment, attainment and provision.

Within urban areas, access to education resources (distance and quality) is chosen as the focal variable for a comparison of education inequalities at the local and individual level. In addition, multiple socio-economic factors (e.g. income, occupation, qualification and housing type) and demographic factors (e.g. hukou status) have been used to explore the pattern of socio-spatial access to education resources within Beijing. Moreover, the research has linked the influence of sub-district context effects and individual level characteristics on pupils' access to schools (travel distance) by using a multilevel approach.

2.5 Summary

In this chapter, a three dimensional framework, ('equality of where?', 'equality for whom?' and 'equality of what?') of thinking, defining and analysing education equality has been proposed on the basis of reviewing previous concepts and theories of equality and education equality. Although there are numerous concepts or theories relating to education equality, they invariably avoided any consideration of geographical aspects. However, the definition or analysis devoid of geography is inadequate since the spatial aspects of education can greatly influence education inequalities. Accordingly, this research incorporates geography into theories of education equality and offers a framework to define and measure education inequalities.

Based on the framework, research on education inequalities at different geographical scales, (international, national and local scale), has been reviewed. Previous research in China from a geographical perspective is limited and the gaps within this research are summarised. The specific institutional background and education system in China has been introduced in the chapter to provide the necessary background information for analysis in the chapters that follow. Finally, a specific research framework for education equality has been proposed.

The rest of the thesis will explore education inequalities in China from different geographical scales, beginning with the regional (provincial level) and then local scale (within urban areas). It worth noting that the thesis focuses on formal education and basic education (primary and secondary education) in China, while informal education and higher education will not be considered in this study. However, before proceeding any results, it is necessary to explain the methodology and datasets used in subsequent chapters. This is the subject of the next chapter.

Chapter 3
Methodology and Data

3.1 Introduction

Based on the overall research framework presented in the previous chapter, this chapter introduces the research methods and corresponding data sources used in this thesis. As indicated in the research framework, there are three major analytical parts in this research, thus multiple techniques and data sets are used to address different research objectives. Firstly, one objective identified by Chapter 2 is that it is crucial to construct a composite index to measure educational inequalities and to monitor the effects of implemented policies. Therefore, the theoretical framework underpinning the comprehensive composite index is presented in Section 3.2.

Secondly, due to the paucity of direct educational data for small areas in China, there is an absence of research which depicts the patterns of socio-spatial differentiation and educational disparity within urban areas in China. These patterns are determined by and associated with multiple factors in the context of China. Therefore, geodemographic classification has been applied as a multidimensional analysis framework to identify spatial variations in population types and potential education inequalities in access to education resources in Central Beijing. This approach is introduced in Section 3.3.

Individuals are intrinsically nested within geographical areas where they live, thus individuals' access to educational opportunities is influenced by the characteristics of areas, where they live. On the other hand, there is heterogeneity in the characteristics of residents; thus, residents would have different access to educational resources even if they live in the same geographical areas. In this context, multilevel modelling is used to link meso scale (e.g. sub-district) context effects with individual level characteristics to analyse the disparities in pupils' access to schools (e.g. travel distance). This method is elaborated in Section 3.4.

Multiple data sets are used according to the requirements of the research objectives and research methods. Specifically, the Educational

Statistics Yearbook of China and China Educational Finance Statistical Yearbook (Section 3.5.1) are used to build the composite index to measure education inequalities and assess current policies. As there is limited educational data in small areas in China, recent Chinese census data (Section 3.5.2) are used to identify potential inequalities in access to compulsory education. The fifth Comprehensive Household Travel Survey (CHTS) (Section 3.5.3), provides individual level data that are used in the multilevel analysis. In order to avoid biased sampling in CHTS and confirm the characteristics and patterns revealed by CHTS, another larger dataset, the Pupil Travel Survey (PTS) is also used in this research.

3.2 Index of Regional Education Advantage (IREA) based on the capability approach

A composite index, which can compare the performance of different areas, identify trends and highlight particular issues within policy, is used to measure regional education inequalities under changing policy regimes in this study. It is based on a set of indicators where an indicator is a measure (quantitative or qualitative) which is derived from observed facets that reflect a given area's relative position in wide-ranging fields, like education (OECD, 2008). As education is a multidimensional issue, which is unable to be captured by a single indicator, a composite index is more appropriate. Composite indices are increasingly recognised as useful tools and are widely used in policy analysis (Bandura, 2008), which can provide simple comparisons of areas and illustrate complex issues (OECD, 2008). There are several advantages to using a composite index in the context of this research: it can summarise multi-dimensional realities without losing underlying information; it is easier to interpret and use to support policy decisions than separate indicators; and it can identify educational advantage and disadvantages in areas and assess their progress of education development (Saltelli, 2007).

The quality of a composite index not only depends on the methods applied in its construction, but primarily on the quality of the framework and data used. Using a weak theoretical background or data containing measurement errors to construct such an index can lead to misleading messages for policy makers. Accordingly, the starting point when constructing a composite index is to provide a sound theoretical framework. In this section, a new multidimensional measure of nationwide education inequality named Index of Regional Education Advantage (IREA) is introduced. Capability theory, proposed on the basis of criticisms of other human well-being evaluation approaches, was used to develop the analysis framework.

3.2.1 Capability approach and its application in education

Education is a key social factor among the non-economic dimensions that measure the well-being of an area (Jorda and Alonso, 2017). In this research, Sen's capability approach, a normative framework for assessment and evaluation of well-being, social arrangements and policy design (Sen, 1992; Sen, 2001; Robeyns, 2005), has been adopted as a theoretical framework to underpin the index. As Sen only sets out a general framework and deliberately leaves the capability component under-specified (Walker, 2005), this approach requires further adaptation to the specific context.

Capability refers to the ability or level of freedom of an individual to choose or achieve something that he/she has reason to value being or doing (Walker, 2005). In terms of education, children are not yet mature enough to make their own choices, so their freedom is constrained by compulsory education and the freedom considered here is more about the level of freedom they will have in the future (Saito, 2003), given that education is crucial for them to develop other relevant capabilities later in the life course.

'Agency' is also one of the central concepts of the capability approach, referring to responsible individuals or groups who make their own choices and shape their own valued lives (Walker and Unterhalter, 2007). Although capability normally regards each individual as an agent, in this research, the capability approach has been applied in the context of education within a region and thus all the people within that region are taken as one agent; in other words, this research looks at the *average* capabilities within each region (Robeyns, 2005).

'Functioning' is the other core concept in Sen's approach that is relevant here. This refers to the achieved outcomes of education. Previous evaluations of education have usually focused on the achieved outcomes, i.e.

the functioning of systems, measured by educational attainment (Qian and Smyth, 2008; Lopez-Acevedo, 2009; Tomul, 2009). Average years of schooling (AYS) has been frequently used as a proxy for educational attainment by researchers (Herrero et al., 2012) and organisations like the World Bank and (The United Nations Educational, Scientific and Cultural Organisation) UNESCO. However, evaluating only outcomes or functioning provides little information about the process and context. There may be different stories lying behind equal achievement; however, the underpinning differences are germane to the discussion of equality and policy implications (Terzi, 2007). The capability approach emphasises the potential to achieve functioning which requires us to not only simply evaluate the current functioning but also the opportunities and real freedoms available to achieve what people value (Walker and Unterhalter, 2007).

Resource-based approaches are also frequently used in education disparity assessment; these consider the individual or group being equally 'well off' when they have the same amount of resources (Sen, 1992) and are defined without considering the substantial variation in the ability to achieve conversion from capability into functioning across individuals and societies (Koo and Lee, 2015). Individuals or groups may achieve different levels of functioning with the same resources. In contrast, the capability approach looks not only at the resources people have at their disposal but also the freedom to achieve the functioning combinations they value (Sen, 2001).

Due to the deficiencies of existing simplistic measures, this research presents a new analytical framework for education in China based on Sen's capability approach (Figure 3.1). The proposed capability approach-based framework that is used for accessing regional education disparity in China involves three dimensions: enrolment, attainment and provision, each of which is influenced by social context, including education policies, the environment and social norms. The achieved educational attainment and enrolment rates work as the conversion factors. The enrolment rate indicates the available opportunity for children to participate in education, as only enrolled children have the chance to be well educated. Attainment refers to the current outcome (functioning of the past) and the educational foundation which will influence the ability of a region to convert resources into functionings in the future (future education attainment). Education provision indicates the availability of educational resources (e.g. schools, teachers), which are normally related to the regional economic context (i.e. per capita GDP), and their quality and degree of sufficiency will influence the capability to achieve functionings. In addition, the social norms and traditions that form people's preferences within a region will consequently influence their aspirations and effective choices. The achieved functionings, in turn, will influence, through feedback, the region's future resource conversion and capability. The Sen's capability approach, therefore, offers the theoretical justification for a comprehensive and multi-dimensional method to evaluate real regional education advantage or disadvantages. This assessment framework fills the current theoretical void and provides a basis for inequality measurements using the new synthetic index.

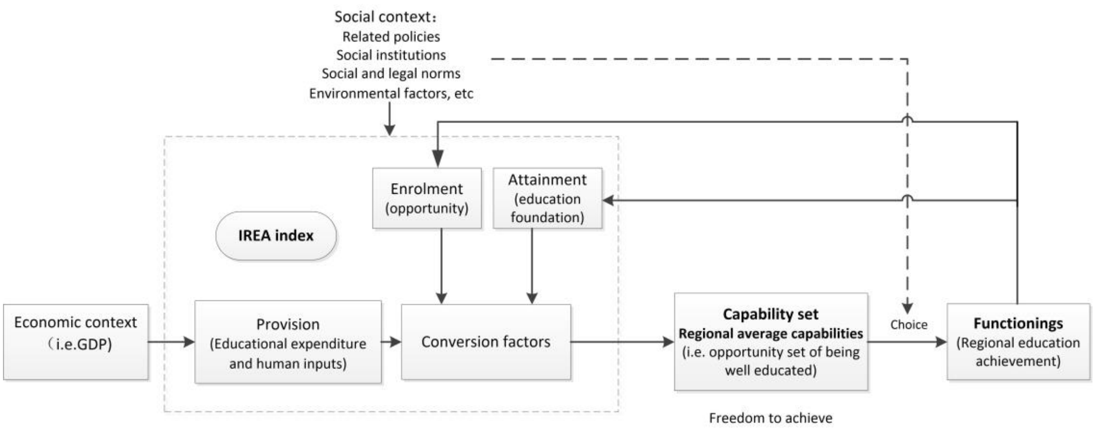


Figure 3.1 The theoretical framework for the IREA index, adapted from Robeyns (2005)

3.2.2 Index of Regional Education Advantage (IREA)

According to the above theoretical framework, an Index of Regional Education Advantage (IREA) has been created specifically to enable the comparisons of the characteristics of education found in different regions for different years. The IREA provides a numeric measure that represents the magnitude of educational advantage or disadvantage and takes account of three facets (influential factors) which follow the conceptualisation of educational inequality measurement in the last section.

The proposed theoretical framework outlined and guided how each variable and facet should be interpreted, prioritised and aggregated in the capability centred evaluation as shown in Figure 3.2. Firstly, the indicators of each facet of the IREA have been processed and combined to produce a score for that facet, a facet index. Then, the scores for enrolment and attainment have been aggregated as the composite score of the conversion factors, which is crucial for converting the resources that are provided into a region's potential capability to realise better education achievement. The importance of conversion factors has been discussed in detail in the above section. After that, the score for the provision and the corresponding conversion factors can then be integrated into an overall measure, which enables the level of educational inequality to be compared between regions (provinces). In addition, the IREA will not only be used to assess regional education performance but also can provide a means to assess education development over time. The process of constructing the IREA comprised five stages (Figure 3.2): (1) data on a set of indicators representing the three different facets were collected from sources available; (2) all the indicators were normalised; (3) the scores of each facet were obtained by geometric mean or arithmetic mean; (4) the enrolment score and attainment score was combined to get a composite score to reflect the condition of the conversion factors; and (5) the IREA score was acquired through calculating the geometric mean of the provision score and the scores of the conversion factors. The detailed input data, results and analyses of the IREA will be presented in Chapter 4.

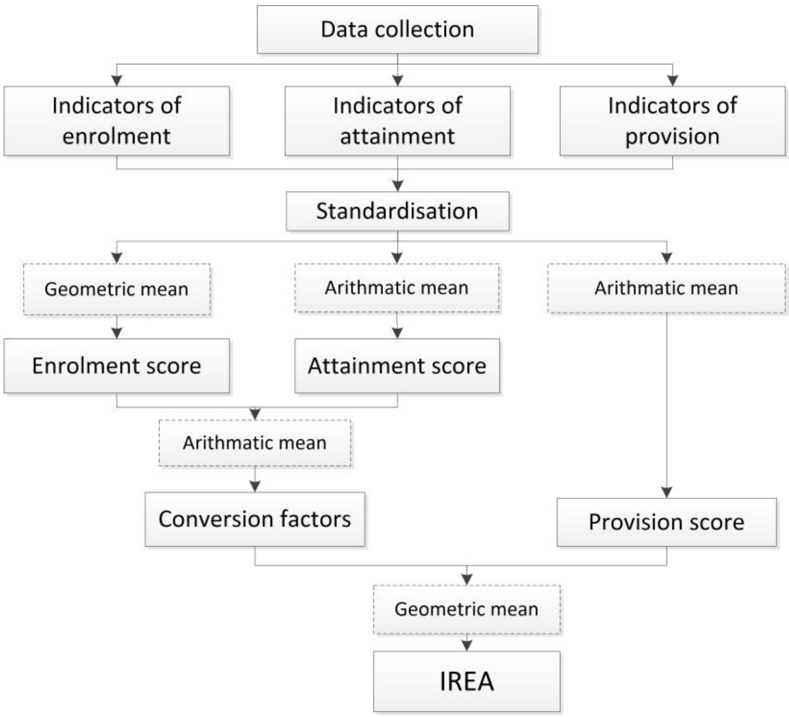


Figure 3.2 The process of constructing the IREA

3.3 Geodemographics

Geodemographics is a form of analysis which combines ‘geo’, implying the places where people live, with ‘demographics’, indicating the demographic characteristics of households or individuals (Leventhal, 2016). Residents and places are dependent on and interact with each other. Places usually pre-date their residents and tend to attract certain population groups because of their characteristics. Geodemographic classification assigns a category to similar neighbourhoods according to the characteristics of their residents. This allows the synthesis of diverse and complex attribute data. The behaviour and characteristics of the residents will further shape the evolution of an area (Harris *et al.*, 2005). This relationship indicates that measurement of physical and socio-economic properties of settlements can generate useful information of residents’ preferences and characteristics within these settlements (Harris *et al.*, 2005). Geodemographics, therefore, provides a framework for understanding the characteristics and behaviour of people who live in these neighbourhoods.

Modern geodemographics in the UK date back to the 1970s and the release of small area national census data in 1971 (Leventhal, 2016). Richard

Webber developed a national classification for research on inner city deprivation and his classification was taken up by a company called CACI to undertake market analysis and renamed ACORN (A Classification of Residential Neighbourhoods). Since then, many area classifications have been developed as market-specific systems, such as CAMEO, Mosaic, Personicx or Sonar, and have been sold to service industries and retail chains (Leventhal, 2016). These classifications are applied to identify focus geodemographic categories to which future advertising, promotional mailings and new store openings, would be most appropriate (Harris *et al*., 2007). Benefiting from increased cooperation between commercial data vendors, governmental organisations and public sector researchers, it is re-entering areas of social research where it originated (Ashby and Longley, 2005; Williamson *et al.*, 2005). For example, in the 2000s, the Office for National Statistics (ONS) in England and Wales produced a general-purpose classification, the 2001 Output Area Classification (OAC), which is based on 2001 census data (Vickers and Rees, 2007) and this has been superseded by the 2011 OAC (Gale *et al.*, 2017).

In social research, geodemographics is usually regarded as a multidimensional technique to measure socio-spatial differentiation (Singleton, 2010). There is an increasing number of applications using geodemographics in sectors such as health (Tickle *et al.,* 2000; Stafford and Marmot, 2003; Abbas *et al.,* 2009), crime (Craglia *et al.,* 2001; Ashby and Longley, 2005) and migration (Dennett and Stillwell, 2011). The methodology has also been applied in the education sector, where inequality is influenced by numerous factors such as demographics (e.g. age, sex, race), socioeconomic status (e.g. poverty level, single parent households, home ownership, income) and housing conditions (e.g. overcrowded households, basic facilities) (Wang, 2012a; Williams and Wang, 2014). Within this research, geodemographic classification is applied as a framework to simplify a range of education-related attributes (variables) of the resident population and households into distinct spatial clusters (Gale *et al.,* 2017). Moreover, the level of access to education services can be strongly influenced by the geographical location and social context of where they live (Hamnett and

Butler, 2011). Thus, geodemographics can be used to investigate potential socio-spatial education inequalities in terms of access to education.

In terms of social and spatial access to education resources, Harris (2008; 2012; 2016; 2017) has paid a considerable amount of attention to social and ethnic segregation within the education system. Segregation within schools or neighbourhoods is mainly measured by the index of dissimilarity, index of isolation and concentration (Harris, 2012; 2016). Free school meal (FSM) eligible pupils are frequently used as the indicator for measuring social segregation in schools (Harris, 2012) and the ethnic composition is used as the indicator to measure ethnic segregation in schools and neighbourhoods (Harris and Johnston, 2008). However, in reality, social and ethnic segregation are easily confounded and overlap to some extent. Using a single indicator based index cannot capture the multi-dimensional socio-spatial differentiation. Moreover, these single-number indices all use a single number to summarise segregation in target area. This will lead to considerable information loss and provide nothing about spatial variation within given regions (Johnston *et al.*, 2006). Compared to the above segregation indices, a classification-based approach, geodemographics, can capture multi-dimensional socio-spatial differentiation and allow the spatial patterns of different types of neighbourhoods to be explored (Harris *et al.*, 2017).

Although scholars have employed geodemographic classifications to assist education-related research (Tonks and Farr, 1995; Tonks, 1999; Williamson *et al.*, 2005; Harris *et al*., 2007; Singleton and Longley, 2009; Singleton, 2010), most of the applications involve access to higher education. Harris (2007) used a general purpose geodemographic system, the 2001 OAC, as a framework to analyse whether pupils from varying types of neighbourhoods in Birmingham (United Kingdom) were likely to attend their nearest state secondary school. In addition, the 2011 OAC has been used to predict the neighbourhoods where have the highest proportions of Asian pupils in London (Harris and Feng, 2016). However, using a general purpose classification may not best reflect education-related problems. Singleton (2010) therefore devised an education-specific OAC in the United Kingdom (UK), but for higher education rather than compulsory school age education. Singleton and Longley (2015) have further argued that the OAC, as a national model, can smooth away important characteristics of the regional socio-spatial structure and thus for applications in urban and regional research, there may be significant advantages in building discrete urban or region-specific geodemographic classifications. Furthermore, no neighbourhood classification has been constructed for use in the public sector in China. This thesis is the first to apply a geodemographic classification methodology in the context of compulsory education in China and to select Central Beijing as a geographical system of interest in order to gain a better understanding of the spatial diversity of potential education inequalities. The methodology is explained in detail in Chapter 6.

3.4 Multilevel modelling

Multilevel modelling is also known as ‘random component modelling’, ‘random effects modelling’, ‘mixed-effects modelling’ and ‘hierarchical linear modelling’ (Hox, 2010), an approach with the capabilities to simultaneously handle different levels of dependency, contextuality and heterogeneity within a set of data. It can be applied to various hierarchical datasets and allows for analysis of complex relationships and patterns which vary both at micro and macro levels of the data structure. As spatial data can also be treated as hierarchical data, in which individuals are inherently nested within a specific level (e.g. a spatial area where they live), multilevel models have been widely applied in geographical research (Jones, 1991).

In terms of geographical or spatial analysis, the differences between the two main components are identified: heterogeneity between individuals and heterogeneity between places (Bullen *et al.*, 1997). As place heterogeneity may vary differently in different contexts, it is unreasonable to reduce the complex variations of ‘geography’ to a single map. In addition, there may not only be variation in higher level contexts, but variation at the individual level may also be too complex to be reduced to a simple ‘average’. The advantage of using multilevel modelling is that this works by specifying models at the macro (e.g. neighbourhood) and micro (e.g. individual) levels which are then combined into an overall hierarchical model (Owen *et al.*, 2016).

Compared to a general regression model, which assumes the relationship between the predictors and response variables holds constant everywhere, the relationship is allowed to vary from place to place in a multilevel model, which potentially shows how ‘geography matters’. As multilevel modelling analyses contextual factors while keeping the micro scale of individuals as objects of study, it avoids both the ecological fallacy and the atomistic (individualistic/exception) fallacy (Subramanian *et al.*, 2009). Multilevel modelling has been widely used to analyse contextual or neighbourhood effects, thereby increasing knowledge about various social research issues including, for example, the influence of schooling on educational attainment (Goldstein, 1987; Courgeau, 2003). It is also used to understand and assess spatial inequalities in health research when introducing interventions and policies to improve social well-being (Owen *et al.*, 2016).

In this research, multilevel modelling will be used to analyse the factors that influence the distance that pupils travel to school in Beijing. Two-levels will be used to estimate the multilevel models. A typical multilevel strategy usually starts with an individual level multiple regression model, which is thereafter developed into a multilevel model by including the complex variations at the second or higher level (between places) and then considering heterogeneity at the first level (among individuals) (Jones, 1991). In order to confirm that there is evidence of variation between places or contextual effects, a null model (without independent variables) will be fitted first to look at the variances at level 2 (between places) relative to the differences between individuals at level 1 (Owen *et al.*, 2016). Explanatory variables in the single-level multiple regression model are then added. Following this, if there is still significant variance at a higher level, then it is regarded as evidence of a contextual effect. The general specification and estimation process of the model applied in this research is illustrated in the following section, where the dependent or response variable is the distance that pupils travel to school.

3.4.1 Single-level model

3.4.1.1 A continuous predictor

All regression models have the same underlying structure and can be expressed by the verbal equation:

$$\text{Response} = \text{Fixed parameters (systematic component)} + \text{Random parameters (fluctuations)} \quad (3.1)$$

In the case of a simple regression model (simplified single-level bivariate model), the general equation can be written as:

$$\text{Response} = \underbrace{\text{Intercept} + (\text{Slope} \times \text{Predictor})}_{\text{fixed}} + \underbrace{\text{Residual}}_{\text{random}} \quad (3.2)$$

or expressed in standard algebraic notation as:

$$y_i = \beta_0 x_{0i} + \beta_1 x_{1i} + (e_i) \quad (3.3)$$

where y_i is the travel distance of pupil i;

x_{1i} is an individual predictor variable for pupil i;

β_0 is an intercept term, which is associated with the constant term, x_{0i};

x_{0i} is a set of ones;

β_1 is a fixed slope term, which represents the relationship between the travel distance and individual predictor variable; and

e_i is a random (residual) term.

The single-level regression model thus consists of two parts: the fixed part and random or residual part (with the random part indicated in parentheses). The fixed part represents the systematic relationship between pupil travel distance and explanatory variables. As its name suggests, this part is fixed and does not change or vary from pupil to pupil or from place to place. The random part is the difference between the individual predictions and the observed values and allows for fluctuations around the fixed part (Bullen *et al.*, 1997). In the single level regression model, the residuals (or the random part) represent all of the idiosyncratic features of the individuals concerned, which have not been included in the systematic part of the model. In this case, the travel distance of pupils is allowed to vary around the intercept, and thus the model can be written more completely as:

$$y_i=\beta_0 x_{0i}+\beta_1 x_{1i}+(e_i x_{0i}) \quad (3.4)$$

In the single level model, the underlying variation for all individuals is assumed to be the same (Jones and Duncan, 1996) with an Independent and Identical Distribution (IID) (Kennedy, 1979) and represented by a single variance term, σ_e^2.

3.4.1.2 Categorical predictors

If one or more predictors are categorical variables, special codes need be assigned to include these in the regression model (Goldstein, 1987). Including a categorical explanatory variable in a model can be treated as fitting a separate model for each category. Thus various coding methods are used to combine the categories into one model giving equivalent estimations in terms of fixed parameters (Jones and Duncan, 1996). Two of the most commonly used recoding methods are called ‘contrast coding’ and ‘separate coding’, and both are used with the simplest dummy coding (0 and 1). Hukou registration status, which includes three categories, is taken as an example of a categorical variable and can be included in the single level model. The three types of hukou status are: hukou within a district (WD); hukou outside a district but within a municipality (WM); and hukou outside a municipality (OM).

Contrast coding

In contrast coding, there is a constant for all categories and the most common or frequent category will be set as the base category. As the hukou status of WD is the most frequent category for this dataset, WD is treated as a base category in the following equation:

$$y_i=\beta_0 x_{0i}+\beta_1 x_{1i}+\beta_2 x_{2i}+(e_i) \quad (3.5)$$

where x_{0i} is also defined as a set of ones;

x_{1i} is set to 1 when a pupil’s hukou status is WM, and 0 otherwise;

x_{2i} is set to 1 when a pupil’s hukou status is OM, and 0 otherwise;

β_0 is the average travel distance for the base category, WD;

β_1 is the differential premium for WM; the average travel distance for WM is given by β_0+β_1; and

β_2 is the differential premium for OM; the average travel distance for OM is given by β_0+β_2.

The strategy of contrast coding is shown in Table 3.1. This replaces the categorical variable (n categories) with n-1 new dummy variables. For example, when dummy variables x_1 and x_2 are both 0, it indicates the pupil's hukou status is WD.

Table 3.1 Contrast coding for hukou status

Type	Dummy 1 (x_1)	Dummy 2 (x_2)
WD (base category,0)	0	0
WM (1)	1	0
OM (2)	0	1

Separate coding

With separate coding, all of the categories of a categorical variable are replaced and re-labelled by indicators.

$$y_i=\beta_0 x_{0i}+\beta_1 x_{1i}+\beta_2 x_{2i} + (e_i x_{0i}) \tag{3.6}$$

where x_0 is set to 1 when pupil i's hukou status is WD, and 0 otherwise;

x_1 is set to 1 when a pupil i's hukou status is WM, and 0 otherwise;

x_2 is set to 1 when a pupil i's hukou status is OM, and 0 otherwise;

β_0 is the average travel distance for WD;

β_1 is the average travel distance for WM; and

β_2 is the average travel distance for OM.

The strategy of separate coding is shown in Table 3.2. This replaces the categorical variable (n categories) with n new dummy variables.

Table 3.2 Separate coding for hukou status

Type	Dummy 1 (x_0)	Dummy 2 (x_1)	Dummy 3 (x_2)
WD (1)	1	0	0
WM (2)	0	1	0
OM (3)	0	0	1

Contrast coding is usually used in the fixed part and level-2 random part of models, while separate coding is generally used in the level-1 random part of models (Bullen *et al.*, 1997).

3.4.2 Multilevel model

A single level model is constructed based on an assumption that all of the observations are independently and identically distributed (IID), which indicates individuals are expected to be completely independent of one another and come from an unstructured random sample (Thomas *et al.*, 2014).

However, many social systems, education, in particular, have a clustered or hierarchical structure in which 'units' of one level are nested within 'units' at a higher level. For example, individuals are nested within the neighbourhood where they live (Goldstein, 1987). The individuals from the same neighbourhood tend to be more alike than individuals chosen randomly from the whole population (Wodtke *et al.*, 2011) (similar to the principles that underpin geodemographics). The individuals are supposed to interact with the neighbourhood or social contexts to which they belong; that is, the neighbourhoods or groups are influenced by the component individuals, and the individuals are in turn influenced by the contexts or neighbourhoods which they inhabit. Accordingly, the individuals and neighbourhoods are conceptualised as a hierarchical system with individuals nested within neighbourhoods, with individuals (e.g. pupils) and neighbourhoods (e.g. sub-districts) defined as separate levels in a hierarchical system (Hox, 2010). The dependence or clustering within data, therefore, can be accounted for in multilevel models. Using the single level model, which ignores this dependency and assumes a hierarchical structure does not exist, will lead to inferential error and impoverished analyses and increase the risk of finding relationships or differences where none exist.

This can be remedied by applying a multilevel model which considers the dependency in the hierarchical structure within a dataset and operates at more than one level. In addition, the effects of cluster-level variables, on individual travel distance are allowed to be analysed in multilevel models. Thus, simultaneously, the fitted systematic relationship is allowed to vary between individuals but also vary from place to place in the multilevel model (Bullen *et al.*, 1997). It allows us to explore how individual travel distance varies across neighbourhoods and how the effects of covariates vary across neighbourhoods.

Although the contextual differences can be considered through a fixed effects approach by including dummy variables for neighbourhoods, this limits the further inference of contextual characteristics (level-2 or neighbourhood level variables), which may be of substantial analytical relevance and importance for explaining variations between neighbourhoods in the response. All the neighbourhood level characteristics will be confounded with the fixed effects of each neighbourhood, which makes the identification of the influence of each neighbourhood level variable mathematically impossible (Fielding, 2004). In addition, the model including the neighbourhoods in the fixed effects can be redundant when the number of neighbourhoods is large. In contrast, multilevel modelling is a random part expansion method, which assumes the neighbourhoods in the samples are randomly selected from a larger normally distributed neighbourhoods. The stochastic expansion enables the neighbourhoods to be treated as a separate level with normally distributed between neighbourhood residuals (Jones and Bullen, 1994; Thomas *et al.*, 2014). In order to specify the multilevel models used in this research, a two-level multilevel model is specified in the following section. It handles the macroscale of sub-districts (neighbourhoods) and microscale of pupils (individuals). The models are now introduced, beginning with the simple random intercept models and then adding more complexities to extend these models into more complex random coefficient models.

3.4.2.1 Random intercept models

In the random intercept multilevel models, the sub-districts are treated as a separate level. The residuals between sub-districts are assumed to be defined by their intercepts, vary randomly around the overall grand mean (β_0) (grand mean for continuous predictors and the mean of base category for categorical predictors) and can be summarised by a single variance term (σ_{u0}^2). In this research, pupils or individuals (level 1) are nested within sub-districts of residence (or schools), which is a hierarchical structure. The multilevel models simultaneously specify regression equations at the micro (individual) and the macro (sub-districts) levels. The micro part of the model can be defined as:

$$y_{ij}=\beta_{0j}x_{0ij}+\beta_1 x_{1ij}+(e_{ij}x_{0ij}) \quad (3.7)$$

where y_{ij} is the logarithm of the travel-to-school distance of pupil i in sub-district j;

x_{0ij} is a constant, a set of ones;

x_{1ij} is an individual predictor variable for pupil i in sub-district j;

β_{0j} is the estimated intercept for sub-district j;

β_1 is the estimated slope;

e_{ij} is the level-1 residual; and

The estimated intercept for each sub-district is calculated as:

$$\beta_{0j} = \beta_0 + u_{0j} \quad (3.8)$$

where u_{0j} is known as the random or differential intercept, estimating the negative or positive differences that sub-district j has compared to the grand mean. For example, in this research, as most of the predictors will be categorical variables, u_{0j} is the difference in travel distance for pupils in sub-district j from the mean travel distance for all categories.

The macro model (equation 3.8) can be substituted into the micro model (equation 3.7), creating a fully explicit multilevel random intercepts model as follows:

$$y_{ij}=\beta_0 x_{0ij}+\beta_1 x_{1ij}+(u_{0j}x_{0ij} + e_{0ij}x_{0ij}) \quad (3.9)$$

$$[u_{0j}] \sim (0, \sigma^2_{u0}) \tag{3.10}$$

$$[e_{0ij}] \sim (0, \sigma^2_{e0}) \tag{3.11}$$

where $(\beta_0 x_{0ij} + \beta_1 x_{1ij})$ is the fixed part and $(u_{0j} x_{0ij} + e_{ij} x_{0ij})$ is the random part. The random terms in the model are assumed to be independent of each other and follow a normal distribution. Thus, the random part of the model involving between sub-district (level-2) variance is represented as σ^2_{u0}, while the random part within a sub-district involving variations between individuals is represented as σ^2_{e0}. It is worth noting that the subscript '0' has been added to the group effect, u_{0j}, and individual residual, e_{0ij}, which is beneficial to add further complex random effects (Steele, 2008b). There is a special case of the random-intercept model, variance component model, which includes no covariate.

Estimating group effects

In order to make comparisons among level-2 units, the residuals, which are the 'neighbourhood effect', the latent variable at level-2, need to be estimated. In single level models, the residuals can be estimated simply by subtracting the individual predictions from the observed values. However, it is more complex in a multilevel model, which partitions residual variance into level-2 and level-1 components. The raw residual for level-1 units is computed as:

$$r_{ij} = y_{ij} - \hat{y}_{ij} \tag{3.12}$$

where, y_{ij} is the observed travel distance, while $\hat{y}_{ij}$ is the predicted travel distance; and the estimated, precision weighted residual for level-2 units (the random effects) is:

$$\hat{u}_{0j} = \frac{\sigma^2_{u0}}{\sigma^2_{u0} + {\sigma^2_{e0}}/{n_j}} . r_j \tag{3.13}$$

where r_j is the mean of r_{ij} for sub-district j, which is also called the raw residual;

σ^2_{u0} is the between sub-district variation;

σ^2_{e0} is the within-sub-district, between pupil variation; and

n_j is the number of observations in sub-district j.

The multiplier in the above equation is always less than or equal to 1, so that the estimated residual $\hat{u}_{0j}$ is usually less in magnitude than the raw residual, r_j. Thus, the raw residual is multiplied by a shrinkage factor and the estimated residual is also sometimes called a shrunken residual. The sub-district level residuals actually have been shrunken towards the overall average.

The idea of shrinkage in multilevel models is used to minimise the unreliability of the model estimation when the sample size is imbalanced. For example, when there are only a few samples in some groups or units, using the information from other groups can increase the reliability of estimation. In this research, as the samples in each sub-district are limited and there is no complete information about each sub-district, it is better to use information from the other groups to improve the guesstimate. The sub-district level residual therefore will be less sensitive to outlying elements of the group. There will be more shrinkage if there are not many pupil samples in this sub-district, if the sub-district level variance is small or if the individual level variance is large. Therefore, the calculation of shrunken residuals is preferred to directly using the raw residuals.

The estimation of the group effects, $\hat{u}_{0j}$, is achieved after fitting the model and based on the estimates of parameters σ_{u0}^2 and σ_{e0}^2 in the model (Steele, 2008a). The variance, σ_{u0}^2, which reflects the between-group (e.g. sub-district) variance, is estimated by the mean raw residuals, r_j. After obtaining the level-2 shrunken residuals, the level-1 residuals can be acquired by subtraction:

$$r_{ij} = e_{0ij} = y_{ij} - (\beta_0 + \beta_1 x_{1ij}) - \hat{u}_{0j} \qquad (3.14)$$

where the individual level residual is the pupil's travel distance minus the modelled pupil's travel distance $(\beta_0 + \beta_1 x_{1ij})$ and then minus the estimated level-2 (sub-district) shrunken residual $\hat{u}_{0j}$.

Interpreting variance components within multilevel models

There are several ways to interpret the variance components in multilevel models, such as coverage intervals, Variance Partition Coefficients (VPC) and Intraclass Correlation Coefficients (ICC). The coverage interval is used to interpret the absolute magnitude of the variance component while VPC and

ICC allow us to interpret their relative magnitudes in the matrix of the response variable(Leckie and Bell, 2013).

The coverage interval is an approach to interpret and evaluate the effects and importance of variance attributed at a higher level by computing their 95% coverage intervals. Given the assumption that the random effects are normally distributed, the 95% coverage interval for a variance component lies in the range -1.96σ_u to 1.96σ_u and allows the researcher to get a handle on the additional influence of context (Thomas *et al.*, 2014).

According to Goldstein *et al.* (2002), the VPC is a measure of the contribution of each level to the total variation. For the two level random intercept model, the VPC (ρ) reports the proportion of variation at sub-district level (level-2) out of the total variation and is expressed by the following equation:

$$\rho = \frac{\sigma_{u0}^2}{\sigma_{u0}^2 + \sigma_{e0}^2} \tag{3.15}$$

This statistic can also be interpreted as the degree of similarity in the travel distance between two randomly selected pupils within a sub-district. If $\rho = 0$, this indicates that there is no difference between sub-districts (level-2) or no contextual dependency and all variations are located at the micro level. If $\rho = 1$, this indicates the maximum difference between sub-districts and complete similarity within sub-districts.

Calculating ICC is the other way of interpreting the relative magnitude of the variance components. ICC statistics can be used to measure the expected degree of homogeneity (similarity) between individuals living within a given sub-district. The correlation between two pupils i and i' is presented in the following equation,

$$\text{corr}(y_i, y_{i'}) = \frac{\text{cov}(y_i, y_{i'})}{\sqrt{var(y_i)}\sqrt{var(y_{i'})}} \tag{3.16}$$

where

$$\begin{aligned}\text{cov}(y_i, y_{i'}) &= \text{cov}(\beta_0 + u_{\text{subdistirct}\,(i)} + e_i, \beta_0 + u_{\text{subdistirct}\,(i')} + e_{i'}) \\ &= \text{cov}(u_{\text{subdistirct}\,(i)}, u_{\text{subdistirct}\,(i')}) + \text{cov}(e_i + e_{i'})\end{aligned} \tag{3.17}$$

=cov($u_{\text{subdistirct }(i)}$, $u_{\text{subdistirct }(i')}$)+0

=cov($u_{\text{subdistirct }(i)}$, $u_{\text{subdistirct }(i')}$)

and

$$var(y_i) = var(y_{i'}) = \sigma_u^2 + \sigma_e^2 \quad (3.18)$$

If the two pupils belong to the same higher level units (e.g. sub-district), their correlation will be $\rho = \frac{\sigma_{u0}^2}{\sigma_{u0}^2+\sigma_{e0}^2}$, whilst individuals from the different sub-districts are assumed to have zero correlation or to be uncorrelated, having controlled for the predictors (Thomas *et al.*, 2014).

The correlation structures can be used as a way to present the dependency and clustering assumptions and hierarchical structure of a multilevel model in summary form. In order to make a comparison, an example of the correlation structure associated with the single-level models (equation 3.4) is displayed in Table 3.3. There are ten individuals within three sub-districts. As was mentioned in Section 3.4.1.1, having controlled for the predictors, the model assumes that all observations are IID, namely, all the individuals are expected to be completely independent of one another. Thus, the leading diagonal of the correlation structure represents the correlation of individuals with themselves and are equal to one, while the correlations between any pair of different individuals are assumed to be zero.

Table 3.3 Full correlation structure of a single level model

Sub-district		1	1	1	2	2	2	2	3	3	3
	Pupil	1	2	3	1	2	3	4	1	2	3
1	1	1	0	0	0	0	0	0	0	0	0
1	2	0	1	0	0	0	0	0	0	0	0
1	3	0	0	1	0	0	0	0	0	0	0
2	1	0	0	0	1	0	0	0	0	0	0
2	2	0	0	0	0	1	0	0	0	0	0
2	3	0	0	0	0	0	1	0	0	0	0
2	4	0	0	0	0	0	0	1	0	0	0
3	1	0	0	0	0	0	0	0	1	0	0
3	2	0	0	0	0	0	0	0	0	1	0
3	3	0	0	0	0	0	0	0	0	0	1

Adapted from Jones and Gould (2014)

In contrast, as with the equation discussed previously, in a two-level random intercept model, there are correlations, ρ, between individuals within a higher-level unit (i.e. sub-district), while individuals from different sub-districts are still assumed to be uncorrelated (Table 3.4). In other words, the strength of the correlation between two individuals depends on whether they belong to a higher unit and the relative importance of higher level variations.

Table 3.4 Full correlation structure of a two-level random intercept model

Sub-district		1	1	1	2	2	2	2	3	3	3
	Pupil	1	2	3	1	2	3	4	1	2	3
1	1	1	ρ	ρ	0	0	0	0	0	0	0
1	2	ρ	1	ρ	0	0	0	0	0	0	0
1	3	ρ	ρ	1	0	0	0	0	0	0	0
2	1	0	0	0	1	ρ	ρ	ρ	0	0	0
2	2	0	0	0	ρ	1	ρ	ρ	0	0	0
2	3	0	0	0	ρ	ρ	1	ρ	0	0	0
2	4	0	0	0	ρ	ρ	ρ	1	0	0	0
3	1	0	0	0	0	0	0	0	1	ρ	ρ
3	2	0	0	0	0	0	0	0	ρ	1	ρ
3	3	0	0	0	0	0	0	0	ρ	ρ	1

Adapted from Jones and Gould (2014)

3.4.2.2 Random coefficient models

In the above random intercept models, the effects of each covariate are constrained to be the same in all level-2 units, while this assumption is relaxed in the following random coefficient models. The random coefficient model, which also called a random intercept and slope model, is a model in which the slopes and intercepts are both allowed to vary and are extended from the random intercepts model by incorporating random slopes. The micro component of the model is defined as:

$$y_{ij}=\beta_{0j}x_{0ij}+\beta_{1j}x_{1ij}+(e_{ij}x_{0ij}) \quad (3.19)$$

where the subscript j of β_{1j} indicates that the slope term is allowed to vary at level-2 and the macro part can be specified as:

$$\beta_{1j} = \beta_1 + u_{1j} \quad (3.20)$$

where β_{1j} is the estimated sub-district specific slope term related to the level-1 predictor, x_{1ij}; u_{1j} is called the random slope or the differential slope, representing the negative or positive differential contribution that sub-district j has compared to the average slope term, β_1 .

Similar to the random intercepts model, the random intercept and random slope model can be defined by combining equations 3.19 and 3.20 and equation 3.8, such that:

$$y_{ij}=\beta_0 x_{0ij}+\beta_1 x_{1ij}+(u_{0j}x_{0ij}+u_{1j}x_{1ij}+\ e_{ij}x_{0ij}) \quad (3.21)$$

$$\left(u_{0j,}\, u_{1j}\right)\sim(0,\Omega_u),\ \ \Omega_u=\begin{bmatrix}\sigma_{u0}^2 & \\ \sigma_{u01} & \sigma_{u1}^2\end{bmatrix} \quad (3.22)$$

$$\left[e_{0ij}\right]\sim(0,\sigma_{e0}^2) \quad (3.23)$$

The term, $u_{1j}x_{1ij}$, is another sub-district level random term and can be summarised by σ_{u1}^2. For categorical predictors, the σ_{u0}^2 term represents the differences between sub-districts for the pupils in the base category. As it is unnecessary to assume that the two level-2 random terms are independent of each other, there is a covariance term included, σ_{u01}, to allow the random slopes and random intercepts to covary according to a sub-district level joint distribution. Thus, the combined variability of two random variables, the between sub-district variance for the contrast category, is a more complex quadratic function and is given by following equation:

$$var(u_{0j}x_{0ij}+u_{1j}x_{1ij})=\sigma_{u0}^2 x_{0ij}+2\sigma_{u01}x_{1ij}+\sigma_{u1}^2 x_{1ij}^2 \quad (3.24)$$

3.4.2.3 Random coefficient models with complex differences between individuals

For the models explained so far, only the complex heterogeneity between places are considered with a constant variance function, $\sigma_{e0,}^2$ at the individual level. However, there are also other possibilities that exist, e.g. the travel distance of pupils with OM may be longer and also may show more variability compared to pupils in other groups. Therefore, level-1 heterogeneity needs also to be accepted and explicitly modelled. The model which considers two levels of heterogeneity simultaneously is defined as:

$$y_{ij}=\beta_0 x_{0ij}+\beta_1 x_{1ij}+(u_{0j}x_{0ij}+u_{1j}x_{1ij}+e_{0ij}x_{0ij}+e_{1ij}x_{1ij}) \quad (3.25)$$

where the model contains one more term, $e_{1ij}x_{1ij}$, in comparison with equation 3.21. As a result, there are two random terms for level 1 and their variance can be specified by a quadratic equation:

$$var(e_{0ij}x_{0ij}+e_{1ij}x_{1ij})=e_{e0}^2 x_{0ij}+2\sigma_{e01}x_{1ij}+\sigma_{e1}^2 x_{1ij}^2 \quad (3.26)$$

The sign (-/+) and magnitude of variance-covariance allow a wide variety of complex level-1 heterogeneity. However, there are more additional complications for the model with categorical predictors (Bullen *et al.*, 1997). Take the categorical variable, hukou status, as an example. In order to make the specification simple, it is assumed that there are only two categories of hukou status, hukou within a municipality (WM) and hukou outside a municipality (OM), and WM is taken as the base category.

When hukou status, a level-1 categorical covariate, is included in the level-1 random part, there is no covariance term as no pupil simultaneously has hukou within a municipality and outside this municipality. Thus, the variability of OM is being constrained to be larger than or equal to the base category, WM. This is because there is no covariance term and the differentials can only be positive or zero if contrast coding is used for the level-1 random part.

One of the most commonly used methods of solving this problem is to apply separate coding for the level-1 random part, when contrast coding is still used for the fixed part. The previous multilevel model can be therefore adapted as follows:

$$y_{ij}=\beta_0 x_{0ij}+\beta_1 x_{1ij}+(u_{0j}x_{0ij}+u_{1j}x_{1ij}+e_{1ij}x_{1ij}+e_{0ij}x_{2ij}) \quad (3.27)$$

where x_{1ij} is set to 1 when pupil's hukou is OM, 0 otherwise; and

x_{2ij} is set to 1 for WM, 0 otherwise;

The separate coding is used in the level-1 random part. The variance for the level-1 residual of each type of hukou status is directly given by the following equation:

$$\text{Variance between pupil's hukou is WM} = \sigma_{e0}^2 x_{2ij}^2 \quad (3.28)$$

$$\text{Variance between pupil's hukou is OM} = \sigma_{e1}^2 x_{1ij}^2 \quad (3.29)$$

Thus, the variances of the two hukou statuses are defined independently from each other.

3.4.4 Model estimation

The most commonly used method to estimate the regression coefficients and variances of the random effects (intercept variance and slope variances) is the maximum likelihood (ML) method. The ML method is a general estimation procedure, which estimates the values of parameters that maximise the probability of observing the data which are actually observed, given the model (Eliason, 1993; Hox, 2010). As the most commonly applied method in multilevel modelling, its apparent advantage is its overall robustness and production of consistent estimates. The ML estimation proceeds by maximising a likelihood algorithm. There are two main likelihood algorithms that are applied in multilevel modelling, which are Iterative Generalised Least Squares (IGLS) and Restricted or Residual IGLS (RIGLS). The former method (IGLS) includes both the regression coefficients and variance components in the function, while the latter (RIGLS) only includes the variance components in the likelihood function and estimates the regression coefficients in the second estimation step. RIGLS is preferred when there is only a limited number of higher level units, as using the IGLS tends to create downward biased higher level variance with few clusters. However, the limitation of the RIGLS method is that the deviances cannot be compared for models under different fixed effects specifications.

The maximum likelihood estimation involves an iterative procedure and the first iteration of the IGLS can be summarised as the following steps:

Step 1: The starting values for the various parameters are generated by the computer program based on the single-level regression estimation.

Step 2: The random part structure will be defined on a set of indicators and initial parameters of the random parts will be produced.

Step 3: The starting values of the fixed part parameters will be improved to obtain revised estimations using the above random part estimates.

After the initial iteration, the computation procedure of steps 2 and 3 will be repeated in further iterations to revise the fixed and random part parameters until the consecutive estimates are adequately close, that is, achieving convergence; then the estimates of the multilevel model are produced if the random part parameters are normally distributed (Goldstein, 1986; Thomas *et al.*, 2014). As the IGLS algorithm is proven to provide fast and reliable estimation, and it is more suitable for the continuous response with relatively simple random effects (Goldstein, 1986; Goldstein, 1989), it has been used in this research.

3.4.5 Diagnostics and significance testing

There are a number of ways to assess the significance of the fixed and random part parameters of a model and model fit. For fixed part estimates of a model, z-ratio (or z test), p values and Wald tests can be used to assess their significance. A z-ratio is calculated by dividing the estimates of parameters by their standard error (if the z-ratio of one parameter estimate is greater than critical value of 1.96, the parameter is tested to be significantly different from zero (Leckie and Bell, 2013)), while p-values of the fixed part parameter estimates are also used to test whether the parameters are significantly different from zero (a null hypothesis of zero). The Wald test is another way to test the significance of individual parameters. The individual Wald test, which is also called a univariate Wald test, is employed to test whether the estimates associated with specific parameters are significantly different from zero (Jones and Subramanian, 2011). In addition, the Wald test can also be used to test sets of parameters in the fixed part of the model, which is called a grouped parameter (or multivariate) Wald test and can deal with the contribution of sets of parameters, such as to test a significant difference between the differentials for the non-base (non-reference) categories of a set of categorical predictors. As the differences between parameters are involved and the variance-covariance of their estimates needs to be considered, the univariate Wald test or other tests are inappropriate, a multivariate Wald test thus is in favour in this context.

For the estimation of the random part parameters (the variance component), z-ratios and p-values, which are usually directly reported by the software package, should not be employed to test the significance of these parameters. This is due to these tests all being based on asymptotic normality assumptions which may not be fulfilled. It is possible to use the Wald test to test the specific parameters of the random part of the model if the number of higher level units are large. However, the Wald test is also markedly influenced by the normality requirement of the sampling distribution for the point estimates (Jones and Subramanian, 2011). Thus, the other statistical test, the Likelihood Ratio Test (LRT), which is not based on the asymptotic normal sampling distribution assumption, is preferred to the Wald test when testing the hypothesis on variance components (Leckie and Bell, 2013).

The LRT is also a procedure to test the overall model fit and check the improvement in the model fit, which is based on deviance change from one model to another. The deviance, which is defined as minus twice the natural logarithm of the likelihood (-2*log-likelihood), is a function of the number of observations and estimates to reflect the badness of fit of the model and indicates the amount of remaining unexplained information. Smaller deviance represents less unexplained observations, thus LRT can be used to compare models by calculating the reduction in deviance to assess whether one model is significantly a better fit than another model for the data. The LRT statistic follows a Chi-square distribution and thus it is possible to calculate the p-value to assess the significance of the model improvement (the reduced degree of freedom is equal to the number of extra parameters in the more complex model).

In conclusion, this section has provided a brief overview of multilevel modelling. Hierarchical, nested or clustered data structures are very common in social and behavioural science (e.g. education research) systems. As the IID assumption of traditional single level models cannot be fulfilled, multilevel models are developed for analysing the hierarchical data accounting for the dependency and clustering found in the data. In addition, multilevel models also offer a flexible framework for analysing geo-referenced data and allow us to simultaneously analyse the effects of individual characteristics (e.g. hukou, income) and neighbourhood characteristics (e.g. school density) on the response variables. Based on the above discussions in this chapter, multilevel modelling is chosen as the most appropriate approach for analysing pupils' travel distance in Chapter 7.

3.5 Data sources

The datasets used to conduct the empirical analyses are introduced in following sub-sections.

3.5.1 Educational Statistics Yearbook of China and China Educational Finance Statistical Yearbook

The Educational Statistics Yearbook of China has been published by the Development Department of the Ministry of Education of China annually since 1987. This yearbook includes provincial level education data, such as the net enrolment ratio of primary school, the illiterate population aged 15 and over, education attainment, number of schools, teachers and students for each education stage (e.g. primary school education, middle school education and high school education), teachers' academic qualification and professional rank. The China Educational Finance Statistical Yearbook is also published annually and specifically provides the information about educational funding investment and educational expenditure (e.g. per-pupil educational expenditure of primary school education) for various education stages and types of education in each provincial level unit. In addition, some educational data are also divided into rural areas and urban areas, like the number of pupils, teacher and teachers' academic qualification(s) and professional rank, however some data, such as the enrolment rate for primary schools and educational attainment do not have separate data for rural and urban areas. As the information provided by the two datasets are most relevant and reliable, they are used to construct IREA in Chapter 4.

3.5.2 Census data

The neighbourhood classification has been designed solely using data at the sub-district level from China's latest decennial census in 2010. It collected data using a short census form and long census form. The short census form is legally required to be answered by all people and includes basic information such as gender, age, hukou, education, and housing type. The long census form includes more detailed information about socio-demographic attributes, but only 10% of the population in each sub-district were selected to complete it (Ma *et al.*, 2015). Although it has some problems in continuity, the census is, by far, the most important data source, providing accurate, complete and consistent data on people's characteristics. In addition, there is no other dataset containing such comprehensive geographic courage.

There are 136 sub-district units in Central Beijing. Two sub-districts have been excluded (one because of incomplete information; the other due to it being a mining area and actually outside Beijing) and data for four other sub-districts have been merged into the sub-districts which contained them. Thus, the census data of 130 sub-districts in Central Beijing have been used in Chapter 6 to construct the geodemographic classifications.

3.5.3 Two pupil travel related survey datasets

The two travel surveys in Beijing, the fifth Comprehensive Household Travel Survey (CHTS) and the Pupil Travel Survey (PTS) datasets have been used in Chapter 6 and Chapter 7 of this research. Both datasets were launched by the Beijing Municipal Commission of Transport in 2014. The CHTS is a household survey which has been carried out regularly since 1986 and conducted every four to five years since the second Travel Survey in 2000. With the purpose of determining the situation with regards to urban traffic, it is the basis for proposing new transport policies. The fifth CHTS was conducted from September to December of 2014 and includes travel information about 40,003 households and 101,815 individuals (Table 3.5). The sample covers 0.52% of the total population of Beijing across all districts and counties. The areas of focus for this survey are contained within Central Beijing; however, samples from the main new towns and urban fringe areas were also collected.

Table 3.5 A comparison of the two travel survey datasets

	CHTS	PTS
Time	September to December, 2014	November and December, 2014
Form	Household survey	School-based online survey
Direct executor	Beijing Municipal Commission for Transport	Beijing Municipal Education Commission
Aim and objectives	Current situation of urban traffic, the basis for proposing new transport policies	Ease traffic congestion around schools and provide safe traffic environment for pupils
Sample size	40,003 households and 101,815 individuals (all age groups)	460,306 pupils (6-18 years old)
Coverage	Beijing (households)	Central Beijing (schools)

In contrast, the PTS was carried out for the first time in 2014 and is one of 16 specific surveys that supplemented the CHTS. This survey was administered by the Beijing Municipal Commission of Transport but directly conducted by the Beijing Municipal Education Commission, most recently in November and December of 2014. It is an online survey undertaken by the schools where pupils study and aims to obtain the trip patterns of school pupils so as to find out the influence of school enrolment policy on local traffic environments. It is also used to provide data to support government initiatives to ease traffic congestion around schools and provide a safer environment for children moving between residences and schools every day. Unlike the CHTS, which covers the whole of Beijing, the PTS is undertaken in schools within six districts in Central Beijing, including Dongcheng, Xicheng, Chaoyang, Fengtai, Shijingshan and Haidian. Although this is a school-based and pupil oriented survey within Central Beijing, the origins (residences) of pupils are spread over the whole of Beijing. There were 482,581 pupil responses to the online PTS in 2014. After the data had been cleaned by the Beijing Municipal Commission of Transport, the number of valid records were reduced to 460,306. A more detailed comparison and analysis have been discussed in the following sections.

3.5.3.1 The fifth Comprehensive Household Travel Survey (CHTS)

Given that the focus of this research is basic education, only the commute trips of 6-18 years old respondents are selected. Survey respondents' activities over a 24 hour (3 am to 3 am) were recorded. These form the sub-dataset for the analysis which follows, representing school pupils' daily commute journeys between their homes and their schools.

There are 18 original variables in this dataset (Table 3.6), including travel origin and destination, travel purpose, departure time, arrival time and travel distance, the respondents' relationship to the head of the household; the socio-economic traits of the household, such as hukou status, qualification and occupation; the last year income of household; housing type; car ownership; and floor area of residence. Most of these variables are categorical except for floor area of residence, travel distance, departure and arrival time. Distance is measured from the centroid of the community where a pupil lives to the school that he/she attends. To preserve the confidentiality of each pupils' information, accurate home addresses and school locations are not released; the home address and school location information are aggregated into the sub-districts where they are located.

Table 3.6 Variables in the CHTS

Variables	Type
1. Relationship with the head of household	Categorical variable
2. Hukou status of the head of household	Categorical variable
3. Highest qualification of the head of household	Ordinal categorical variable
4. Occupation of the head of household	Ordinal categorical variable
5. Industry of the head of household	Categorical variable
6. Household income of last year	Ordinal categorical variable
7. Year of birth	Continuous variable
8. Car ownership	Categorical variable
9. Housing type	Categorical variable
10. Floor area of residence	Continuous variable
11. Travel number	Categorical variable
12. Travel purpose	Categorical variable
13. Sub-district of origin	Categorical variable
14. Sub-district of destination	Categorical variable
15. Departure time	Date and time
16. Arrival time	Date and time
17. Travel mode	Categorical variable
18. Travel distance	Continuous variable

The total sample contains 13,537 observations (6-18 years old) made by 5,655 individuals, including 5,377 individuals whose travel purpose is the school commute involving 12,341 school commute trips. The detailed summary is shown in Table 3.7 and only the school commute trips will be discussed in the following section. There are 3,265 respondents (accounting for 61% of total interviewees) living in Central Beijing with 6,601 trips (accounting for 54% of total trips) made including trips from home to schools and from schools to home. Given that some pupils may go home to have lunch they have to commute between home and schools four times, however, most pupils only make two trips.

Table 3.7 Summary of the sample size of the CHTS

	No. of trips	**No. of pupils**
All school commutes in sample	12,341	5,377
Go to school	6,305	5,372
Go home	6,036	5,270
Home within Central Beijing	6,601	3,265
The proportion of the sample in Central Beijing	54%	61%

3.5.3.2 Basic information about the Pupil Travel Survey (PTS)

Out of the 459 primary schools in Central Beijing, 445, or 96.9% of them have taken part in this survey (Table 3.8). The number of valid observations of primary school pupils is 303,289, accounting for 62.5% of all primary pupils (484,912). Meanwhile, 319 secondary schools participated in the survey, accounting for 96.1% of all secondary schools (332). The number of valid questionnaires for secondary school pupils is 157,017, representing 53.5% of all secondary (middle school and high school) pupils.

Table 3.8 PTS Survey sample size details

	Number of primary schools	Share (%)	Number of primary school pupils	Share (%)
All	459	100	484,912	100
Sample	445	96.9	303,289	62.5
	Number of secondary schools	Share (%)	Number of secondary school pupils	Share (%)
All	332	100	293,671	100
Sample	319	96.1	157,017	53.5

Source: Beijing Municipal Commission of Transport

All the variables included in this survey dataset are shown in Table 3.9. This dataset has been supplied directly from the Beijing Municipal Commission of Transport. The full school name information is not released; instead, every school name is coded into an identifiable ID as shown in row 2. Pupils are asked to provide the time when they leave home and arrive at school, so the travel time in this dataset is calculated by subtracting the latter from the former. The distance measurement is same as that of CHTS, and also the sub-districts where home and schools are located are released.

The sub-district codes beginning with 1 and 4 refer to Xicheng district; codes 2 and 3 indicate Dongcheng district and codes 5,6,7,8 represent Haidian, Chaoyang, Fengtai and Shijingshan district respectively. The tenure includes ‘lodging’, owned and rented accommodation. Lodging indicates students who do not live in their own home but in accommodation which is entirely free or has very low fees, which is different from rented accommodation.

Table 3.9 Variables in the survey dataset with example records

Variables	Example records			
School ID	1013001	…..	…..	…..
Education stage	Primary school	Middle school	High school	
Grade (1-12)	3	…	…	…
Travel time	00:05:00	…	…	…
Commute mode	Car	Motorcycle	Metro	…..
Tenure	Owned	Rented	Lodging	
Sub-district code of school	207	207	…	…
Sub-district code of home	613	207	…	…
Distance	11.959	1.177	…	…

Source: Beijing Municipal Commission of Transport

The dataset supplied (460,306 cases) has some deficiencies, so further data cleaning was needed to remove false or unwanted values, change dubious entries or estimate missing values. For this research, the first task for data cleaning was to recode and estimate the various commute modes which will influence further analysis. After that, school and home location data cleaning was conducted.

Commute mode recoding

In order to meet the requirements of the research, data cleaning was conducted to remove the redundant data on the mode of travel or to estimate missing data. Twelve records were deleted because they lacked the essential information for analysis (e.g. commute mode and tenure). The commute mode values of 1,395 records have been recoded as they contain repeated information, impossible values or trivial answers that would affect the subsequent analysis.

Some respondents did not read the options for the commute mode very carefully and have overlooked the given options, so they selected "Other" and filled in the blank space on the survey form with their modes. In this case, their commute mode is generalised and merged into the existing answers given by respondents. For example, if the pupils' commute mode was community shuttle bus, military region shuttle bus, Peking University shuttle bus, rent shuttle bus, vehicle of embassy, etc., they are all recoded as "shuttle bus of work unit". The answers such as Jeep, Lamborghini, etc., are all recoded as "car" (Table 3.10).

Table 3.10 Numbers of recoded commute mode records

Commute mode	Number of records recoded into each commute mode	Commute mode	Number of records recoded into each commute mode
Shuttle bus of work unit	102	School bus	9
On foot	1	Bus	2
Car	146	Three-wheel electric bicycle	29
Bicycle	12	Motorcycle	2
Two-wheel electric bicycle	17	Rental Bicycle	0
Taxi	0	Other	1,075
Metro	0	Total	1,395

The "Other" commute mode contains very rich and diverse information, so the observations containing useless, incorrect or nonstandard information are all recoded. For example, "Other: indeterminate", "Other: unknown" are all recoded into "Other"; "Other: elephant", "Other: rocket", which are considered as impossible answers, are all recoded as “Other: incorrect answer”; “Train: Other” is reordered as “Other: train”.

School and home location data cleaning

After matching the sub-district code with the sub-district name, it is found that there are two schools located outside the districts of Central Beijing. One school with 91 records is in the Yizhuang sub-district within the Daxing district; the other with 1,502 records is in the Beiqijia Town within the Changping district. The 1,593 observations of the two schools located outside Central Beijing are saved in a separate file, which probably will be helpful in revealing the pupil flow commuting out of Central Beijing but not used in this research (Table 3.11), whereas the 458,498 records in the main dataset will be used to investigate the population flow commuting within Central Beijing. There is also a school for which there is no matched district name. This may be because of a mistake in coding the school location, so the 203 cases in this school have been deleted.

Table 3.11 Summary of PTS data after data cleaning

	No. of pupil records	No. of schools
School located within Central Beijing	458,498	761
School located out of Central Beijing	1,593	2
Deleted information	215	1

In conclusion, after data cleaning, the sample contains 458,498 pupil records relating to 761 schools within Central Beijing, 1,593 records for two schools located out of Central Beijing, and 215 cases have been deleted, which include 203 cases without a corresponding location code and 12 observations without necessary commute information.

3.5.3.3 Further comparisons of the two datasets

Further comparisons between the two pupil commute datasets have been conducted (Table 3.12). Both of the two datasets provide travel information for primary, middle and high school pupils, like the sub-district of origins and destinations, travel distance, travel time and travel mode. Some information contained in the two datasets is similar but appears in different formats; for example, the pupils' grade and education stage are given in the PTS, whereas there is information of pupils' year of birth in CHTS; three types of tenure information of pupils is provided in the PTS: owned, rented and lodging, while more detailed housing information is provided in the CHTS, for example, owned housing is divided into commercial housing, purchased public housing, economically affordable housing/price-capped housing and the countryside homestead and rented housing is divided into private rental housing and cheap rental housing/ public rental housing. Only travel time is provided in PTS. However, CHTS also contains the departure and arrival time.

Compared to the PTS, the CHTS contains more information about pupils' socio-economic background, like hukou status, parents' qualification and occupation; the last year income of household; housing type; car ownership; and floor area of residence. The PTS only contains one journey for each pupil, but there are several trips for each pupil in CHTS.

Table 3.12 Comparison between the two pupil commute related survey datasets.

	CHTS	PTS
Same	Sub-district of home	
	Sub-district of school	
	Travel distance	
	Travel mode	
Similar	Year of birth	Grade and education stage
	Housing type (Purchased public housing, cheap rental housing, etc.)	Tenure (Owned, rented, lodging)
	Departure time and arrival time	Travel time
Differences	12,341 trips for 5,377 pupils	458,498 trips for 458,498 pupils
	Beijing	Central Beijing
	Relationship with the head of household	School type
	Hukou status of the head of household	
	Highest qualification of the head of household	
	Occupation of the head of household	
	Industry of the head of household	
	Household income of last year	
	Car ownership	
	Floor area of residence	
	Travel number	
	Travel purpose	

3.6 Summary and conclusions

This chapter has presented the applied research methods and corresponding data sources used in this thesis, which serve as a basis for the subsequent analysis in the following chapters. Firstly, an evaluation framework based on Sen’s capability approach has been proposed to offer theoretical justification for building a multi-dimensional index to comprehensively measure regional education advantage or disadvantages and assess educational policies. All the data used to construct the index are from two specific statistic yearbook, Educational Statistics Yearbook of China and China Educational Finance Statistical Yearbook. The detailed construction process of the index and its implications on assessing regional policies are discussed in detail in Chapter 4.

Secondly, geodemographics has provided a multi-dimensional framework to assess potential inequalities at a local scale, within urban areas.

It can reveal potential advantages and disadvantages in assessing education resources through depicting the education related socio-spatial structure within urban areas. The area classification of Central Beijing's sub-districts has been designed using data from the 2010 census, and its application and analysis results are given in Chapter 6.

Finally, the multilevel analysis offers a flexible statistical method, allowing us to link the contextual (place) differences with individual level determinants. The individual level data sets, CHTS and PTS, fulfil the requirement of multilevel modelling and the CHTS, which have more detailed attributes, will be employed in Chapter 7 to analyse the individuals' economic traits and residential origin's influence on pupils' travel distance. The dataset with larger sample size, PTS, is used to validate the results generated from the CHTS.

Chapter 4
Measuring and Assessing Regional Education Inequalities in China under Changing Policy Regimes

4.1 Introduction

State policies relating to the Chinese education system have undergone considerable changes in the last thirty years based on the different development stages that the country has experienced and the various objectives associated with each stage set by Government. The education system also has experienced significant decentralisation and (re)-centralisation trends. The education related policies can be classified into two broad periods according to their main purpose. The policies of the first stage, from 1985 to 2005, were focused on mobilising all the resources to enhance the efficiency of education funds during a period of financial and administrative decentralisation. However, they were gradually tempered due to serious regional education disparities and; after 2005, attention has shifted from efficiency to equality, with Central Government playing an increasingly more important role in eliminating inequalities. As indicated in Chapter 2, no proper measure has been used to evaluate the effects of these policies on reducing regional education inequalities. Thus, a composite index of education inequality underpinned by Sen's capability approach is proposed in this chapter as a more robust measure and used to investigate temporal changes as well as geographical disparities at the regional (or provincial) scale across the whole of mainland China.

This chapter is structured into five sections. The next section sets the scene by reviewing previous and current education related policies with a focus on their influence on education equality; the following section explains the construction process of the new Index of Regional Education Advantage (IREA); the results of the analysis are reported in the fourth section whilst the fifth section discusses the effects of related policies and makes a series of policy suggestions. Conclusions are then drawn in the final section.

4.2 Policy background

4.2.1 Financial and administrative decentralisation, 1985-2005

Education in China experienced a particularly dramatic disruption in the chaotic social movement referred to as the Great Proletarian Cultural Revolution (1966-76) (Unger, 1984). Political loyalty and labour were valued above academic attainment and the overarching goal was to eradicate social differences through a radical socialist agenda. The focus for education was not on education equality but on class struggle. The national examination system for higher education was abolished in this period. Thus, the connection between occupational achievement and education was removed (Unger, 1984) and the disparities in educational indicators experienced a short-term reduction during that period (Hannum *et al.*, 2007).

With the shift from a planned to a market oriented economy in the late 1970s and early 1980s, the importance of science and technology for economic transition and development was reiterated by the Chinese Government (Hannum *et al.*, 2007; Huang *et al.*, 2015). The Decision on the Reform of the Education System (DRES) made by the Central Committee of the Communist Party of China (Zhonggong zhongyang guanyu jiaoyu tizhi gaige de jueding) was issued in May, 1985. Its main aim was to produce a qualified labour force for promoting market reform and economic modernisation. The nine-year compulsory education framework was confirmed in the DRES and, in addition, the financial management and administration functions of the education system were to be decentralised so as to increase efficiency (Hannum *et al.,* 2007). These measures were implemented through the Compulsory Education Law of the People’s Republic of China in 1986 (Sun, 2010).

After these reforms were introduced, management and financial responsibilities for education provision were transferred to local government. This meant that local governments were given the primary responsibility for providing most of the funding for schools, including investment in the construction or reconstruction of school buildings, education facilities, teachers’ salaries, and all recurrent expenditures (Sun, 2010). In rural areas, primary, middle and high schools were sponsored by local authorities in villages, townships and counties respectively, while primary and secondary education in urban areas was sponsored by the district and municipal governments respectively (Tsang, 1996; Sun, 2010). For these above local governments, there were two main sources of public funds for education: budgetary allocation from local government and very limited categorical grants from higher levels of government (Tsang, 1996). Central Government thus had almost no role in the financing of basic education under the new system (Hu, 2012). In order to complement insufficient budgets, local authorities were allowed to collect education levies and surcharges as extra-budgetary funds to support education within the same locality. However, this tax income proved insufficient to cover the related education expenditure and was an unstable source of income. Therefore, schools needed to raise funds through different methods to meet expenditure. Schools raised extra money (known as 'non-budgeted item' or yusuanwai) both to pay for non-recurrent expenditure and to raise teachers' income through charging miscellaneous fees, running enterprises and receiving individual donations and generating income from school-run industries (Sun, 2010; Huang *et al*., 2015).

The diversification of sources of funding for education is a distinctive characteristic of this period. The mobilisation of non-government resources was broadened and intensified at the school level. The extra-budgetary resources grew sharply and became increasingly important sources of funding for basic education (Tsang, 1996). In 1997, the State Council encouraged enterprises, social organisations or individuals to manage or establish non-state schools by introducing the 'Regulations for Running a School by Social Force', which allows for the exploration of multiple funding sources for education and also transfers responsibilities to non-state stakeholders (Huang *et al.*, 2015). Gradually, education service became a valued commodity and access to education became increasingly linked to the consumers' ability to pay (Whitty, 1997).

With the rapid growth of the economy, huge progress in education was also made after 1985 with 98.5% of the counties in China introducing the nine-year compulsory education system. Moreover, the conditions of school buildings, education equipment and teachers' qualifications also improved dramatically (Tsang, 2000). The policies in this period were successful in mobilising additional government and non-government resources but they also exposed significant inefficiencies and glaring inequalities. It became apparent that the decentralised administration and financial system was limiting the Central Government‘s ability to reduce regional disparities (Tsang, 1996), and the allocation of resources for regional education services had been directly linked to their economic development (Zhu and Peyrache, 2017).

Therefore, due to the uneven regional economic development occurring in China in these years, different areas had varying abilities to invest in education (Zhang *et al.*, 2012). Given the continuing inadequacy of national investment in education, this situation led to certain areas becoming seriously disadvantaged. In poor rural areas, the weak tax base of local governments, meagre household incomes and impotence to mobilise non-governmental resources imposed strong limits on the amount of budgetary and extra-budgetary funds that could be collected for basic education. Furthermore, as a result of the worsening financial circumstances in some areas, teachers' payments were delayed or stopped. As a consequence, poor and remote areas had very low enrolment and completion rates for basic education as well as higher proportions of dilapidated school buildings (Tsang, 1996). In contrast, wealthier areas became capable of mobilising their affluent non-public resources to improve their education services. This situation increased regional education disparities and family educational expenditures (Hannum *et al.*, 2007).

In the 1990s, China's Central Government gradually realised the limitations of decentralisation and responded with a series of policies to promote education equity and expand access to education in disadvantaged areas. For example, the West Development Campaign was designed to increase public investment in the undeveloped western areas of China (Hannum *et al.*, 2007). Education law in 1995 stipulated that county level governments (and above) should set up an educational special fund to support basic education in disadvantaged areas (Tsang, 1996). A transfer scheme was also introduced to transfer resources from affluent governments to authorities in relatively poor areas. In addition, some policies even exempted disadvantaged areas from education levies or levied taxes at a lower rate

(Tsang, 1996). However, the effectiveness of these policies has been very limited according to Hu (2012).

Since late 2003, the Ministry of Education of China began to criticise the tendency of educational commercialisation and shift the focus of their attention from economic efficiency to solving compulsory education problems for disadvantaged areas and groups. Since 2001, the “Two Waivers and One Subsidy” policy, involving exemption from textbook charges and miscellaneous fees and subsidising the living costs for boarding pupils, has been implemented to help children from poor families in rural areas receive compulsory education. From 2001 to 2005, the basic sources of funding for compulsory education had moved from the township government to county government, which better assured educational expenditure, especially on teachers’ pay in poor villages. However, there were still substantial impoverished counties with insufficient finance to supply adequate funds for education. Thus, education in undeveloped areas remained in the predicament of having a funding shortage and more policies were required to solve this problem after 2005.

4.2.2 Education equality and unified planning at the provincial level, after 2005

The policies in the last ten years have attempted to reduce education inequalities among different groups and different regions by implementing some substantive measures. The five most important policies discussed in this research, in chronological order, are: (1) notifications of the reform of the funding guarantee system for rural compulsory education (NR) (December 24, 2005); (2) the Compulsory Education Law of the People’s Republic of China (June 29, 2006); (3) notifications of the State Council on exemption of tuition and miscellaneous fees for compulsory education in urban areas (NU) (August 12, 2008); (4) National Medium and Long-Term Educational Reform and Development Plan Outline (2010-20) (NMLERD) (July 29, 2010); and (5) notifications of further improvement of the funding guarantee system for urban and rural compulsory education (NUR) (November 25, 2015).

It is important to recognise that a macro-region division has frequently been used by the Chinese Government when implementing policies. Mainland

China is divided into three economic zones: eastern (eleven provincial level units); central (eight provincial level units); and western (twelve provincial level units) (Figure 4.1), based on their economic development level and geographic location (Li and Wei, 2010). The eastern areas include 11 provincial level units; that is, Beijing, Tianjin, Hebei, Liaoning, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong and Hainan. These areas are all coastal areas with flat terrain and each plays a leading role in China's economic development. The central areas contain Shanxi, Jilin, Heilongjiang, Anhui, Jiangxi, Henan, Hubei and Hunan, eight inland provincial level units. The main grain production bases and 80% of China's coal resources are located in these areas. The western areas refer to Inner Mongolia, Guangxi, Chongqing, Sichuan, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Ningxia, Qinghai and Xinjiang, 12 provincial level units with comparatively high terrain. The topography and environment of these areas are very diverse, with plateaus, basins, deserts and prairies. Their economic development has come much later than the eastern areas. By adjusting the size of each province according to their per capita GDP (PCGDP), the inset cartogram (Figure 4.1) displays the extent of economic inequality. The macro region with highest PCGDP is the eastern region, while the western region has the lowest PCGDP. However, within each region, the variations of PCGDP are very significant such that the areas of Beijing and Shanghai extend far beyond their actual size, but that of Shandong and Hebei are much smaller. This regional division is frequently used by the Chinese Government when implementing policies. The financial policies for education which will now be described are based on this spatial partitioning.

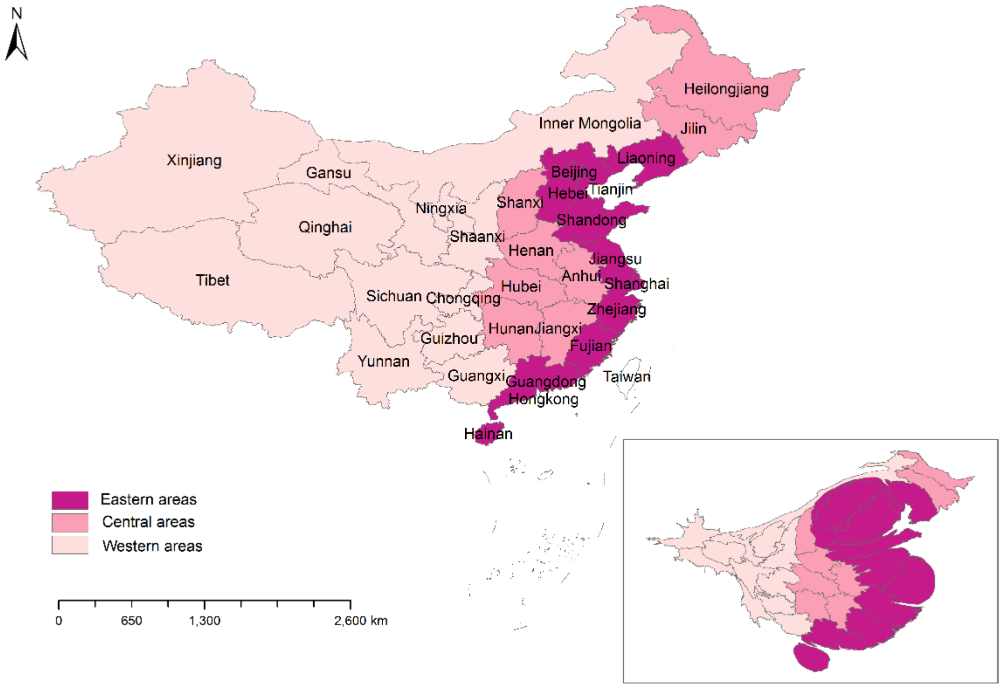


Figure 4.1 Three region division of China and per capita GDP with inset cartogram, 2014

Since 2005, a new funding guarantee system of education has been gradually introduced that provincial level governments are required to make overall plans for the provision of education, and the role of county level governments has changed from providing funding to administering funding for education. In December 2005, NR was issued by the State Council (Table 4.1), indicating that all the rural areas in western China were exempt from tuition and miscellaneous fees from 2006, and all rural parts of central and eastern areas were exempt from these fees from 2007. In addition, the new policy stipulated that the basic standard for per pupil public funding in rural areas in each provincial level unit should be formulated by provincial level governments. Accordingly, the public funds for each pupil should not be lower than the amount of this basic standard. The funds for waiving tuition and miscellaneous fees and basic public funds for rural areas are shared by Central Government and local governments on the basis of the items and proportions as prescribed by the State Council: 80:20 for the west; 60:40 for central; the proportions for eastern provinces, except for municipalities directly under the control of Central Government, were determined by their financial position respectively. Central Government provides all the funds for free textbooks in western and central areas, whereas these fees are guaranteed by local governments in the eastern region. Building renovation expenses for all primary and middle schools in rural areas are jointly sponsored by Central Government and local government (50:50) for provincial level units in western and central areas, while these funds are provided entirely by local government in the eastern region. The provincial level governments should enhance the amount of transfer payments to ensure the salaries of teaching staff in rural areas.

Table 4.1 The summary of latest policies for compulsory education.

Policy		NR (2005)			NU (2008)	NUR (2015)		
Eligible areas		Rural			Urban	Urban and Rural		
		Western	Central	Eastern		Western	Central	Eastern
Free tuition and miscellaneous fees	Eligible groups	All rural			All urban	All		
	Funding sources	80:20*	60:40	D**	L	80:20	60:40	50:50
Free textbooks	Eligible groups	Needy families			Minimal assurance families	All		
	Funding sources	C***	C	L	L	C	C	C
Cost of living allowance	Eligible groups	Boarding pupils from needy families			Boarding pupils from minimal assurance families	Boarding pupils from needy families		
	Funding sources	L			L	50:50	50:50	50:50
Public funds	Funding sources	80:20	60:40	D***	L	80:20	60:40	50:50
Construction expenditure	Funding sources	50:50	50:50	L	L	50:50	50:50	L
Teachers' salaries	Funding sources	L			L	L		

* 80:20 refers to the funds shared by Central Government and local government, that is, the number on the left indicates the proportion shared by Central Government, while the number on the right indicates that of local government

** D refers to where the proportions depend on the financial position of each provincial level unit respectively, except for municipalities directly under the control of Central Government.

*** C refers to the Central Government and L indicates the local government.

Exemption from tuition and miscellaneous fees for compulsory education was confirmed by the Compulsory Education Law in July 2006 and therefore extended across the country as a whole (Standing Committee of the National People's Congress, 2006a). It also stipulated that the funding for compulsory education should be fully guaranteed by public finances from national and local government to fundamentally solve the problem of insufficient educational funds. It also proclaimed that compulsory education should be administered by the county or higher-level authorities. Each level of government should establish separate funding for compulsory education, and these funds should be equally distributed, except for the extra funds provided to rural areas and low-performance schools (Standing Committee of the National People's Congress, 2006d). It is clear that the principles emphasised by Central Government and imposed by law were to allocate education resources rationally, improve the education condition of disadvantaged areas and promote the balanced development of education. At the same time, the local authorities in urban areas were still responsible for providing funds for their own compulsory education and were only partly supplemented by Central Government through limited grants and transfer payments. Although the Compulsory Education Law had already claimed that no tuition or miscellaneous fees should be charged for the provision of compulsory education, it had been applied to urban areas only after the release of NU in 2008 (Table 4.1) and the related funds were still solely provided by the local (provincial level) government.

From 2004 to 2010, the Government’s educational budget (zhengfu yusuannei jiaoyu bokuan) increased from £40.3 billion[3] to £134.9 billion, an increase of 235% (Ministry of Education of PRC, 2005; 2011). Moreover, the funds from the Central Government increased from £2.99 billion to £25.47 billion, or 7.5 times (Ministry of Education of PRC, 2011). Central Government improved its ability to reduce education inequalities and to support education in rural and western areas. Furthermore, the NMLERD was published in 2010 (The State Council of China, 2010; Sun, 2012). The word “equity” appears in the document 17 times (Hu, 2012) and the text indicates that fairly large regional and rural-urban inequalities continue to exist (Sun, 2012). More

---

3 A conversion rate of 10 Chinese *yuan* to £1 is used in this research.

accessible and equitable education which benefits everyone was posited as the most important objective; the plan aims to achieve equal basic public education for everyone and to narrow disparities (Thomas and Peng, 2010; Sun, 2012).

In 2015, the NUR was issued to unify the funding guarantee system for compulsory education in urban and rural areas. The education expenditure proportions shared by Central Government and local governments are unified in this new policy, while in previous policies the Central Government mainly supported education in rural areas and local governments in urban areas had to take responsibility for the funds for their own compulsory education services. In addition, the basic standard for per pupil public funds per year were also unified: £60 for primary school pupils and £80 for middle school pupils from western and central areas; £65 for primary school pupils and £85 for middle school pupils in eastern areas. Specifically, the public funds are guaranteed by Central Government and local governments in the proportion 80:20 for western areas, 60:40 for central areas and 50:50 for eastern areas (Table 4.1).

All in all, it is clear from this synopsis of policy that the Central Government has put more emphasis on promoting education equality by supporting disadvantaged groups and regions. However, before addressing the question of how policies have affected inequalities within the education system, it is necessary to accurately monitor the impact of policies that have been implemented already. In order to achieve these objectives, the evolution of regional education inequalities is evaluated in the following section using a new index.

4.3 Constructing Index of Regional Education Advantage (IREA)

In this section, a new multidimensional index of provincial level education inequality, the IREA, is constructed according to the framework proposed in Section 3.2. The IREA has been created specifically to enable comparison of the characteristics of education found in different regions for different years. This is to reflect the development of education across China and the way in which educational inequalities have evolved before and after the introduction of new policies since 2005. IREA scores for 2004, 2009 and 2014 were calculated in this research. The year 2004 was selected as a reference year and the analysis of data for 2009 and 2014 enable an evaluation of the effectiveness of the related policies and provide a background for the implementation of NMLERD (2010) and NUR (2015).

4.3.1 Input data

According to the conceptualisation discussion in Chapter 3, IREA combines 17 education-related variables (Table 4.2), each of which is related to three facets of education – enrolment, attainment and provision – in a single score that quantifies the extent of education inequality between different provinces in China.

The availability, compatibility and applicability of data have been considered and all the input data are data from Educational Statistics Yearbook of China and China Educational Finance Statistical Yearbook that relate to geographical provinces. In this case, data from Educational Statistics Yearbook of China and China Educational Finance Statistical Yearbook are better than data from the decennial census, since only 2000 and 2010 Census data are available. In addition, most of the variables follow the official definitions (Ministry of Education of the People's Republic of China, 2015a) and measure education in the same direction such that a higher score for each indicator represents a preferable situation.

Table 4.2 Indicators for the three facets of education inequality at the provincial scale.

Facets	No.	Data / Indicators	Influence on education
Enrolment	1	Net enrolment ratio for primary school	Education opportunity
	2	Primary to middle school pupil transfer rate	
	3	Middle to high school pupil transfer rate	
Attainment	4	Literacy rate	Current foundation for education development
	5	Average years of schooling (AYS)	
Provision	6	Teacher-pupil ratio for primary school	Quantity and quality of education provision
	7	Teacher-pupil ratio for middle school	
	8	Teacher-pupil ratio for high school	
	9	Proportion of teachers’ attainment is equal and above associate bachelor for primary school	
	10	Proportion of teachers with higher professional title for primary school	
	11	Proportion of teachers’ attainment is equal and above undergraduate for middle school	
	12	Proportion of teachers with higher professional title for middle school	
	13	Proportion of teachers’ attainment is equal and above undergraduate for high school	
	14	Proportion of teachers with higher professional title for high school	
	15	Per-pupil educational expenditure of primary school	
	16	Per-pupil educational expenditure of middle school	
	17	Per-pupil educational expenditure of high school	

4.3.1.1 Enrolment

Enrolment rates in education in different regions demonstrate the variation in educational opportunity, which will provide children with the chance to become qualified to a specific level of education. Based on the previous literature on access to schools (Connelly and Zheng, 2007a; 2007b; Yanqing, 2012), the net enrolment ratio for primary school, the primary to middle school pupil transfer rate and the middle to high school pupil transfer rate were chosen as indicators for enrolment. The net enrolment ratio of primary school in an area (province) refers to the percentage of the number of enrolled primary school age children account for the total number of school age children, and its equation is as follows:

$$V_{1i} = \frac{Number\ of\ school\ age\ children\ enroled\ in\ area\ i}{Total\ number\ of\ school\ age\ children\ in\ area\ i} \times 100\% \qquad (4.1)$$

where the subscript i indicates province i. The children of primary school age before 1991 refer to 7 to 11-year-old children, however, the definition of school age children varies for different areas in the statistics after 1991 given the different primary school enrolment age and the length of schooling in each area (Ministry of Education of the People's Republic of China, 2015b). The higher the net enrolment rate indicates school children have more chance of receiving education, and its maximum value, in theory, is 100%. However, due to the early or late admission, not reaching 100% does not indicate these pupils are not enrolled.

As there is no net enrolment rate for middle and high schools, these two indicators are replaced by two proxies, the rate of pupil transfer from primary school to middle school: V_{2i} (equation 4.2) and the rate of pupil transfer from middle school to high school: V_{3i} (equation 4.3) where the subscript i indicates province i:

$$V_{2i}(\%) = \frac{New\ student\ enrolment\ in\ middle\ schools\ in\ area\ i}{Graduates\ of\ primary\ schools} \times 100\% \qquad (4.2)$$

$$V_{3i}(\%) = \frac{New\ pupil\ enrolment\ in\ high\ schools\ in\ area\ i}{Graduates\ of\ middle\ schools} \times 100\% \qquad (4.3)$$

It is noteworthy that the focus of this study is on basic academic education, and thus the enrolment for vocational education, vocational junior secondary schools, vocational-technical training schools, vocational high schools and secondary vocational schools are not considered. As there is no separate data on middle school enrolment in 2014, the data on junior secondary schools (including information of middle schools and vocational middle schools) were used in the calculations for 2014. However, only 0.0182% of junior secondary stage pupils were enrolled in the vocational junior secondary school in 2014, so it will not significantly influence the consistency of the data or the results.

4.3.1.2 Attainment

The evaluation and estimation of educational attainment can reflect current education conditions and human capital stock in an area. In this study, educational attainment will be measured by the literacy rate and average years of schooling (AYS) in different areas (provinces). Literacy is always viewed as the process of individuals obtaining cognitive skills, which will contribute to socio-economic development and better social awareness (UNESCO, 2006). The literacy rate used in this study is described in equation 4.4:

$$V_{4i} = 100 - \frac{Illiterate\ and\ semi - illiterate\ population\ in\ area\ i}{Pop_i} \times 100\ \% \qquad (4.4)$$

where Pop_i means the population aged 15 and over in area i.

AYS is also considered to be a suitable proxy for education attainment and human capital and is widely used in analysing the evolution of education inequalities by scholars and organisations such as the World Bank and UNESCO (Barro and Lee, 1993; Tomul, 2009; Herrero *et al.*, 2012; van Leeuwen and Foldvari, 2014; Jorda and Alonso, 2017). Two types of information are needed to estimate the AYS: the number of people in each level of education and the duration of each education stage. Data on the first type were collected from a 1% sample survey of population change and the level of education of the population aged six and over was classified into five types: illiterate; primary school; middle school (junior secondary school); high school (senior secondary school); and college and higher level, which depends on the last year of education a pupil had attained. There is, however, no unified length for each education stage; for example, although more than 90% of primary schools are six-year primary schools, there are still a lot of five-year schools, especially in Shanghai, where 98.25% of newly enrolled primary pupils were enrolled in five-year primary schools in 2014 (counted from Educational Statistics Yearbook, 2014).

Thus, the duration of different levels of schooling was modified from previous research by Qian and Smyth (2008) and Wang and Yao (2003). Since the educational attainment of a pupil was defined as the highest level of education they had attended but not necessarily completed, a person who has not finished at primary level or who has dropped out of primary school will still be deemed to have reached the primary school level of attainment. Therefore, the education duration for each level of education used is slightly shorter than the full duration. Average education duration was assumed as five years for primary school level, eight years for junior secondary school level (includes middle school and vocational junior secondary school), 10.5 years for senior secondary school level (includes high school and vocational high school) and 14.5 years for college and above. AYS,V_{5i}, is calculated by the formula proposed by Qian and Smyth (2008) as follows :

$$V_{5i} = \frac{5p_i + 8m_i + 10.5h_i + 14.5u_i}{Pop_i} \times 100\ \% \qquad (4.5)$$

where p is the number of people whose highest level of attainment is primary school education, and m , h and u represent middle school, high school and college education and above respectively and Pop_i is the population aged six and over in area i in that year.

4.3.1.3 Provision

In order to measure the quantity and quality of education provision, the teacher-pupil ratio, the teachers' professional and academic qualifications, and per pupil education expenditure were considered.

Teacher-pupil ratio: Previous studies suggest that a smaller size class, that is, a higher teacher-pupil ratio, has advantages in education, as pupils can get more attention from teachers and are more likely to interact with classmates, which is beneficial for positive development (Becker and Powers, 2001; Cheung and Chan, 2008). Thus, the teacher-pupil ratio in this research is used to detect the regional disparities in human input for education. Although there are also some substitute teachers, part-time teachers and administrative personnel in each region, only full-time teachers are counted. The teacher-pupil ratio at primary school level is shown in equation 4.6, and those of middle school and high school are similar to this equation:

$$V_{6i} = \frac{Total\ primary\ school\ full - time\ teachers\ in\ area\ i}{Total\ enrolments\ in\ primay\ schools\ in\ area\ i} \times 100\ \% \qquad (4.6)$$

This indicator mainly reflects the adequacy of teachers, whereas other factors such as teachers' quality, teaching experience, professional background and teaching methods will all have impacts on the teaching quality.

Teachers' education quality: Education quality has been made a top-level priority for China's educational policymakers. The quality of education that pupils receive is largely decided by the quality of the teachers, who are the direct knowledge providers. Teacher quality is measured by each teacher's attainment and professional title. In this research, the percentage of teachers with undergraduate and above qualifications is used to measure qualified teacher availability except for primary schools in which comparatively lower teacher qualifications are required. The proportion of qualified primary school teachers is shown in equation 4.7:

$$V_{7i} = \frac{Graduate_i + Undergraduate_i + Associate\ bachelor_i}{Total\ primary\ school\ full - time\ teachers\ in\ area\ i} \times 100\% \quad (4.7)$$

where $\mathrm{Graduate}_i$, $\mathrm{Undergraduate}_i$ and $\mathrm{Associate\ bachelor}_i$ represent the number of teachers with graduate, undergraduate or associate bachelor qualifications in area i respectively.

Moreover, each school teacher needs to have a quality ranking which is based on an assessment conducted by a county level education bureau, including the assessment of educational background, teaching skills, work ethic and professional accomplishments (Ding and Lehret, 2007). The ranking titles, in increasing order, are: undecided; third grade; second grade; first grade; and senior. These titles partly decide how much salary a teacher earns and also reflects the teacher's performance (Ding and Lehret, 2007). 'Senior' is the best in this ranking system, so the percentage of teachers with senior and first grade professional titles is treated as one indicator for assessing the qualified teacher proportion in each province.

Teacher rank is highly related with school quality, and key schools normally have a higher percentage of teachers with the senior grade; thus the percentage of teachers with senior and first grade ranking is applied in this research to reflect the educational quality of each area (province). According to a notice issued by the Ministry of Human Resources and Social Security of the People's Republic of China and Ministry of Education (2015), the ranking system of primary and secondary schools was unified in that the senior teacher of a primary school is equal to a first grade teacher in a middle school. The proportion of qualified teachers is based on the teachers' ranking system for primary and middle schools and shown in equations 4.8 and 4.9

respectively. The equation for high school is the same as that for middle schools.

$$V_{8i} = \frac{Secondary\ school\ senior\ teacher_i\ +\ Primary\ school\ senior\ teacher_i}{Total\ primary\ school\ full-time\ teachers\ in\ area\ i} \times 100\% \qquad (4.8)$$

$$V_{9i} = \frac{Secondary\ school\ senior\ teacher_i\ +\ Secondary\ school\ first\ grade\ teacher_i}{Total\ secondary\ school\ full-time\ teachers\ in\ area\ i} \times 100\% \qquad (4.9)$$

Teachers' professional and technical positions are important manifestations of teachers' teaching ability, but they are not the only aspects. Moreover, there are differences in the criteria for obtaining certain professional and technical posts for teachers in different places.

Per-pupil educational expenditure (PPEE): Expenditure on education directly determines the condition of physical assets and human capital for education, and thus influences pupils' academic achievements. For example, inadequate education investment will lead schools to lack the latest technology infrastructure to assist pupil study and teachers' ability to teach (Cheung and Chan, 2008). PPEE is a comprehensive indicator, which reflects the overall amount of education funding sponsored by local and Central Government, such as public funds, construction funds, teachers' salaries and students' subsidies. The data for PPEE are available in the China Educational Finance Statistical Yearbook.

In summary, 17 indicators are selected to create the IREA, which can summarise information from three facets of education to reflect education inequality. These indicators need to be standardised before constructing the IREA.

4.3.2 Standardisation

As each variable has a different measurement unit, standardisation is required to convert the indicators into a common metric to allow aggregation (OECD, 2008). The most commonly used methods, z-score and max-min standardisation, are not feasible for the requirements of further analysis and data characteristics of this research. The z-score is not suitable because all the values needed for the later analysis (e.g. cartogram) should be positive. Drawbacks in the max-min method have been pointed out by Herrero *et al.*

(2012); this approach increases the variations for some variables, which will arbitrarily increase the variance of these indicators. For example, the average net enrolment rate for primary schools in 2009 was 99.51%, with a minimum value of 96.59% and maximum value of 100%, which means that the enrolment rate for primary schools is very high in every province. It indicates that this indicator is very weak in reflecting education differences. However, if the max-min standardisation method is used, the differences will be improperly enlarged. An alternative method, the distance to a reference for area i, I_i, is proposed for use in this study, and defined as:

$$I_i = \frac{x_i}{x^*} \tag{4.10}$$

where x_i is the value of a variable x for area i and x^* is a reference value of variable x for all areas. In order to enable the index to be comparable across years, the mean of each variable in 2009 was used as the common reference values for all three years. Using the mean of 2009 is helpful since this enables the index to be comparable across years. Once standardisation has been undertaken, the data are ready for aggregation in the next step.

4.3.3 Weighting and aggregation

As the IREA is a composite and single summary measure, it is often desirable to assign weights to indicators based on their perceived importance (Jiang and Shen, 2013). In addition, the IREA adopts a hierarchical indicator system so the score of each facet should be obtained before calculating the final composite score. There are several approaches that can be used to determine the weights of indicators, such as equal or arbitrary weights, expert opinion weights, Principal Component Analysis (PCA) and factor analysis (Decancq and Lugo, 2013). Whilst all methods have their advantages as well as disadvantages, the decision about which method to adopt depends on the research purpose, the data type and the data characteristics (Deutsch and Silber, 2005). We have used different weighting and aggregation methods for assigning weights to facets and indicators.

For the first facet, enrolment, there are three indicators: the net enrolment ratio for primary schools, the primary to middle school pupil transfer rate and the middle to high school pupil transfer rate. The three variables for measuring enrolment (Table 4.2) are not independent and not perfectly substitutable; for example, the enrolment in primary schools will influence the volume of students transferring from primary school to middle school. The geometric mean is therefore used to obtain the enrolment score.

There are two indicators for the second facet, attainment: literacy rate and AYS. The weights for the two indicators of attainment are determined according to the principle of frequency-based weights (Deutsch and Silber, 2005; Decancq and Lugo, 2013). The indicators which are weak in reflecting education differences should be given relatively lower weights. As the literacy rate is already very high (Table 4.3) and AYS has been emphasised in previous research (Qian and Smyth, 2008), a weighting of 1 was given to the literacy rate, while a weight of 2 was assigned to AYS.

Table 4.3 The descriptive statistics of indicators for attainment, 2009

	Literacy rate	Average years of schooling
Mean	91.07	7.23
Std. Deviation	6.67	1.02
Minimum	62.23	4.03
Maximum	96.89	9.75

Weightings, which represent relatively subjective approaches, proved difficult to establish for the final facet, provision of education, where the assessment involves a large number of indicators (twelve variables); therefore a more objective, mathematical approach (PCA) was considered since it is a useful statistical technique to simplify a large set of multidimensional variables and was originally designed as a dimension-reducing technique (Jiang and Shen, 2013; Pan *et al.*, 2017). It avoids arbitrariness (Pan *et al.*, 2017) and can also take into account the (multi)collinearity between variables, the so-called double counting problem (Decancq and Lugo, 2013). PCA is a useful statistical technique to simplify a large set of multidimensional variables and was originally designed as a dimension reducing technique (Jiang and Shen, 2013; Pan *et al.*, 2017). It transforms a set of initial input indicators into an equal number of uncorrelated linear combinations. The weights can be acquired either from the first main component that explains the largest proportion of the total variance or by using the weighted combination of several main components (Decancq and Lugo, 2013).

In order to make the results comparable over time, the data used in PCA were standardised by the above normalisation method. PCA works best when a high correlation is observed among most variables and it will be a waste of time using PCA if there are only a few variable correlations above 0.3 (Desai and Shah, 1988; OECD, 2008); therefore correlation analysis is conducted before applying PCA. Table 4.4 shows the results of correlating each of the 12 variables (6-17) with one another and high correlations (>0.7 marked in bold) among certain variables are evident. In addition, other statistical tests, such as the Kaiser-Meyer-Olkin Measure of Sampling Adequacy, also demonstrate the applicability of using PCA.

Table 4.4 The correlation matrix using normalised data in 2009

Correlation Matrix

Variables		6	7	8	9	10	11	12	13	14	15	16	17
Correlation	6	1.00	.71	.40	.48	.36	.46	.42	.43	.30	.41	.35	.29
	7		1.00	**.75**	**.76**	.54	**.76**	.65	.66	.53	**.75**	**.74**	.63
	8			1.00	.64	.36	**.72**	.48	.53	.49	**.87**	**.88**	**.88**
	9				1.00	.54	**.93**	.66	**.87**	.50	**.74**	.69	.62
	10					1.00	.51	**.94**	.57	**.74**	.37	.37	.31
	11						1.00	.65	**.90**	.51	**.79**	**.76**	**.70**
	12							1.00	.66	**.84**	.51	.51	.45
	13								1.00	.44	.62	.59	.54
	14									1.00	.50	.52	.48
	15										1.00	**.98**	**.96**
	16											1.00	**.97**
	17												1.00

Table 4.5 gives the variance explained by each component after applying PCA. The first principal component explains 65.5% of the variance in the dataset, which accounts for the largest variability and provides the most information about the dataset.

Table 4.5 The total variance explained by components

	Total Variance Explained					
	Initial Eigenvalues			Extraction Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	7.9	65.5	65.5	7.9	65.5	65.5
2	1.6	13.7	79.2	1.6	13.7	79.2
3	1.0	8.1	87.4	1.0	8.1	87.4
4	0.8	6.3	93.7			
5	0.2	1.9	95.6			
6	0.2	1.6	97.2			
7	0.1	1.1	98.3			
8	0.1	0.8	99.1			
9	0.0	0.4	99.5			
10	0.0	0.2	99.7			
11	0.0	0.2	99.9			
12	0.0	0.1	100.0			

Extraction Method: Principal Component Analysis.

For the purpose of weighting, a commonly used approach is factor loading of the first component to weight the indicators related to education provision (Jiang and Shen, 2013). The indicators with more unequal distributions will be assigned higher weights in PCA, while the ones with low standard deviations would be given lower weights (Vyas and Kumaranayake, 2006). The factor loading of the first component is shown in Table 4.6. The teacher-pupil ratio of primary school and proportion of teachers with a higher professional title for the primary school were assigned lower weights because their variations are not as significant as with other indicators. Although PCA is a very objective approach and has a lot of advantages, it is based on the statistical nature of the data, not the real-world structure and theoretically guided assumptions (Jiang and Shen, 2013); thus it is not applied to all variables.

Table 4.6 The factor loading for the first component

Variables	Component 1
Teacher-pupil ratio for primary school	0.56
Teacher-pupil ratio for middle school	0.88
Teacher-pupil ratio for high school	0.84
Proportion of teachers' attainment is equal and above associate bachelor for primary school	0.88
Proportion of teachers with higher professional title for primary school	0.66
Proportion of teachers' attainment is equal and above undergraduate for middle school	0.91
Proportion of teachers with higher professional title for middle school	0.79
Proportion of teachers' attainment is equal and above undergraduate for high school	0.81
Proportion of teachers with higher professional title for high school	0.69
Per-pupil educational expenditure of primary school	0.89
Per-pupil educational expenditure of middle school	0.88
Per-pupil educational expenditure of high school	0.83

After acquiring the scores for three facets, the enrolment and attainment scores were combined to produce a score for conversion factors with a weight of 1:2, given that attainment represents educational grounding for a region and is a more important conversion factor. According to the theoretical discussion, the conversion factors will influence the agents' capability to achieve valued functioning with given education provision, and will jointly affect the results. Thus, the IREA score was acquired by calculating the geometric mean of the scores of the provision and conversion factors. This fits with the theoretical framework and, in addition, it avoids the perfect substitutability feature of the arithmetic mean and penalises the dispersion of the variables that are aggregated. The marginal utility of an increase would be much higher for the variable with a lower score; thus, this geometric aggregation method provides a greater incentive for policymakers to address the problems within facets with low scores (OECD, 2008; Herrero *et al.*, 2012).

4.4 Analysis of results

Overall IREA scores have been calculated by ranking all areas by the value of their IREA across all the three years and then dividing the rank order into categories of equal class interval so that quintiles are produced allowing comparison over time (Figure 4.2a, b and c). Thus, from 2004 to 2014, if a province's IREA changed quintile, this can be interpreted as the conditions of education having worsened or improved (Norman and Darlington-Pollock, 2018). Quintiles 1-5 cover the range from lowest to highest scores. The spatial patterns of IREA scores across the three years display the overall regional educational inequalities and their evolution. The IREA was also decomposed into the scores for different facets to reveal the detailed variations and rationale behind the regional education disparities.

4.4.1 Spatial patterns of the IREA

Unlike previous regional educational studies which all focus on the coast-inland education dichotomy in China (Qian and Smyth, 2008) or inequalities among the regional divisions, a different spatial pattern is apparent. Using the macro division between the north-east and south-west provinces shown by the dotted line, the IREA scores of the northeast provinces are significantly higher than those of the south-west provinces (inset maps of Figure 4.2), although the specific distributions have changed over time. The inequalities are more distinctly exhibited in the cartograms, in which the size of each region is adjusted according to its IREA value. In 2004 (Figure 4.2a), the IREA scores for Shanghai and Beijing were significantly higher than all other areas and most of the south-west provinces were in the worst quintile. In 2009 (Figure 4.2b), the distribution appears different, but in reality, the advantaged areas were still the north-east provinces, while the comparatively disadvantaged areas were located in the south-west. The distribution in 2014 (Figure 4.2c) is also consistent with this pattern. Thus, most provinces to the north-east of the dotted line have education advantages; the education conditions in provinces on this line are mixed and intermediate; the educationally disadvantaged provinces are concentrated to the south-west of this line. Figure 4.2d shows the improvement of IREA for each province during the period from 2004 to 2014. We can observe that the improvement of western areas, especially the northern part, has been dramatic, like Inner Mongolia and Xinjiang. Education in the eastern coastal areas also has developed considerably, like Jiangsu and Fujian, while the development of south and middle areas has been comparatively slow, like Guangdong and Hunan.

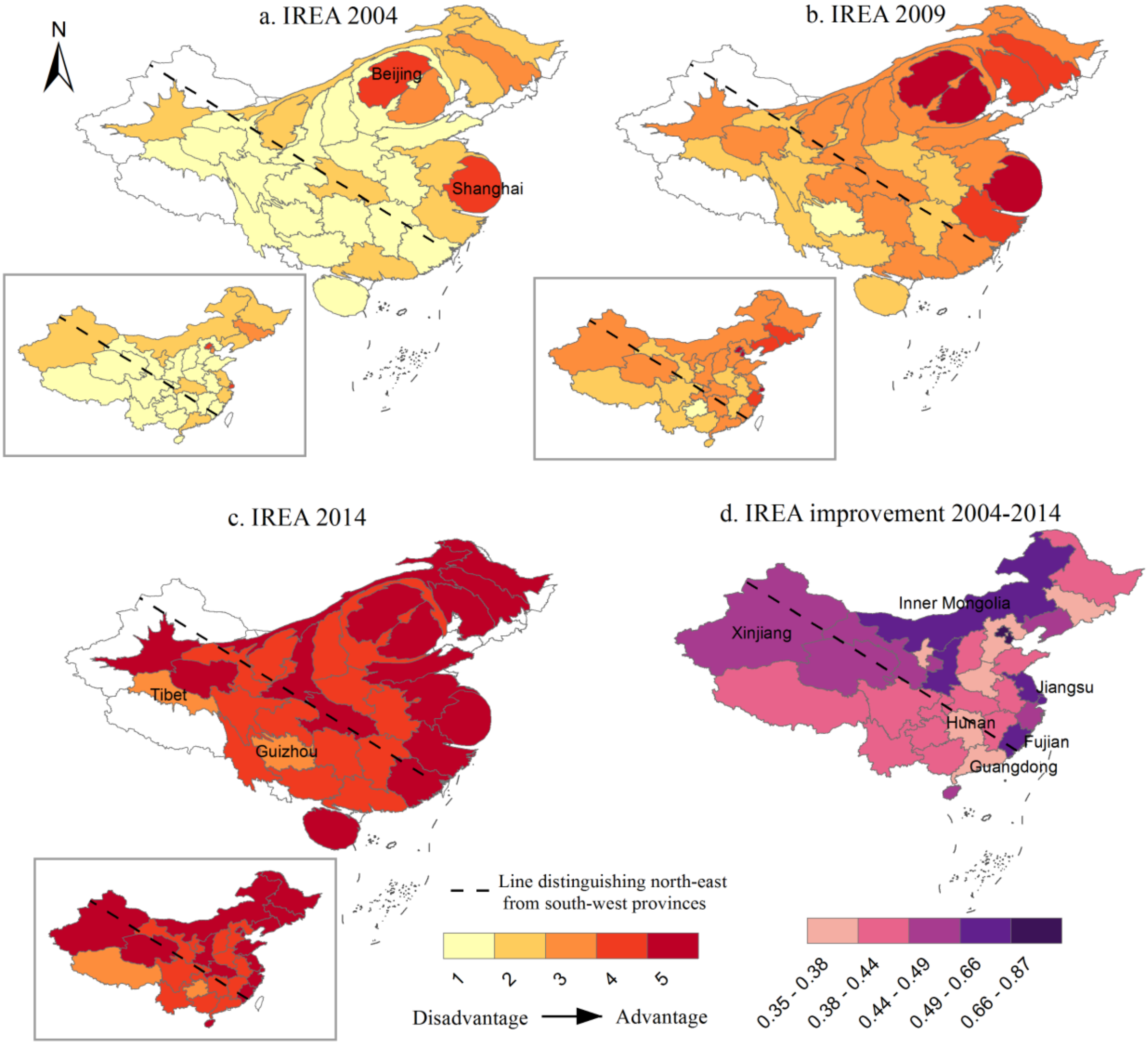


Figure 4.2 IREA quintile distribution by province, 2004, 2009 and 2014

4.4.2 Decomposition of the IREA

The IREA is a composite score which comprehensively reflects the condition of regional education inequalities. However, a single score conceals information and detailed variations between different component facets. For example, the two areas with the same IREA score may contain different development status of education attainment and provision. In order to reveal the potential processes and detailed variations behind this index, the IREA was decomposed into its three component parts: index of enrolment, index of attainment and index of provision. These indices were calculated by aggregating indicators within each facet (Figure 4.3).

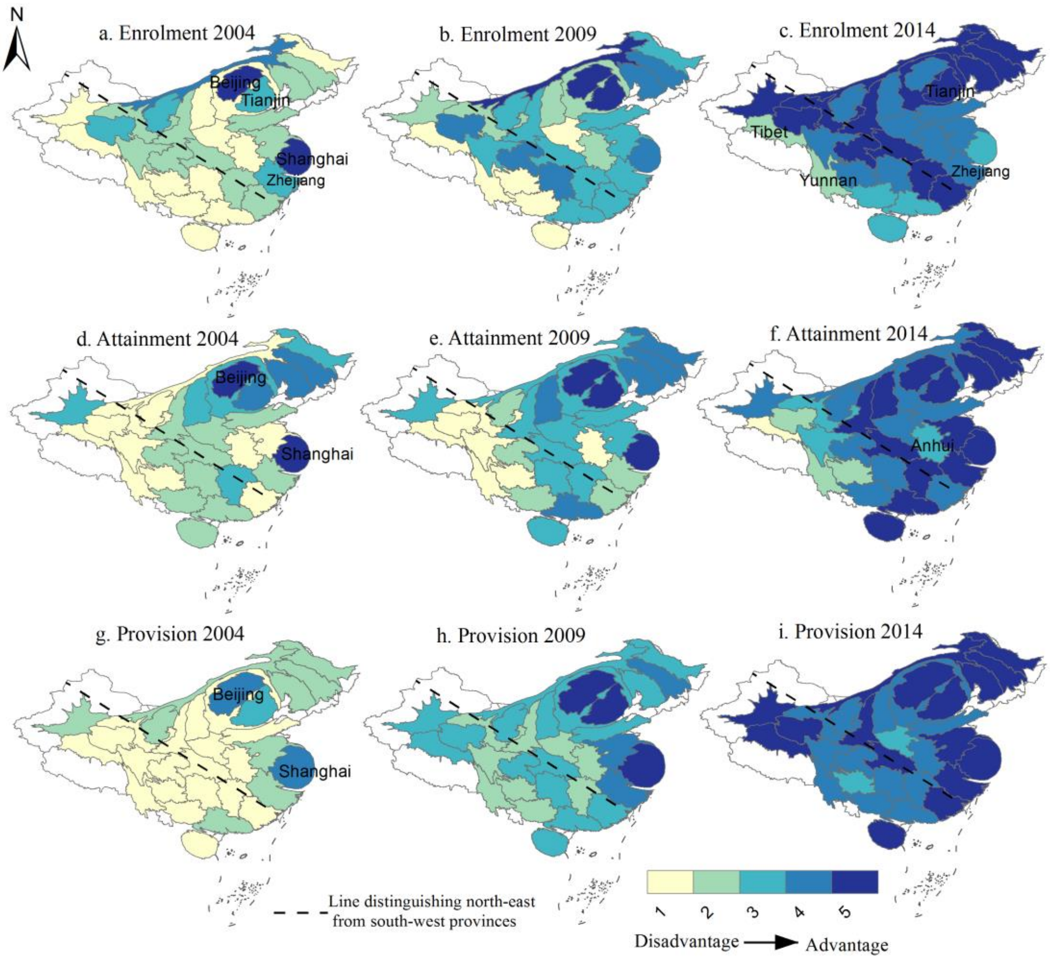


Figure 4.3 Cartograms for enrolment, attainment and provision by province, 2004, 2009 and 2014

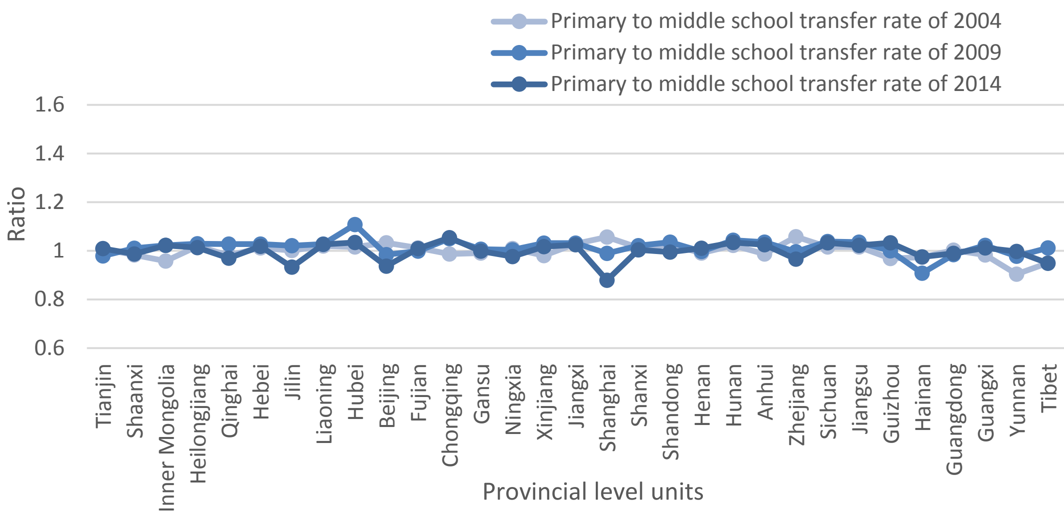
4.4.2.1 The spatial pattern of enrolment

In 2004 (Figure 4.3a), the enrolment scores for Beijing and Shanghai were significantly higher than other provinces and located in the first quintile, followed by Tianjin and Zhejiang which are provinces next to Beijing and Shanghai. The enrolment scores of the remaining provinces were very low in 2004. After the implementation of the Compulsory Education Law in 2006 and Central Government's support to the western regions, the situation changed dramatically. In the map of 2009 (Figure 4.3b), it can be seen that the enrolment rates of provinces on the dotted line and to the north-east were evidently higher than that in 2004, although with different rates of improvement. The enrolment rate for Inner Mongolia, which is an autonomous region, appears in the most advantaged quintile in 2009, while the enrolment score of Shanghai and Zhejiang had decreased. In 2014, the enrolment rates in most areas on the dotted line and to the north-east improved to the fifth quintile. The enrolment scores of areas located in the south-west of China were still very low compared to other areas, especially Tibet and Yunnan. The scores of Beijing and Shanghai had declined further by 2014 (Figure 4.3c), probably due to the enrolment restrictions of hukou, which limits migrant children transferring to middle school and high schools in these megacities.

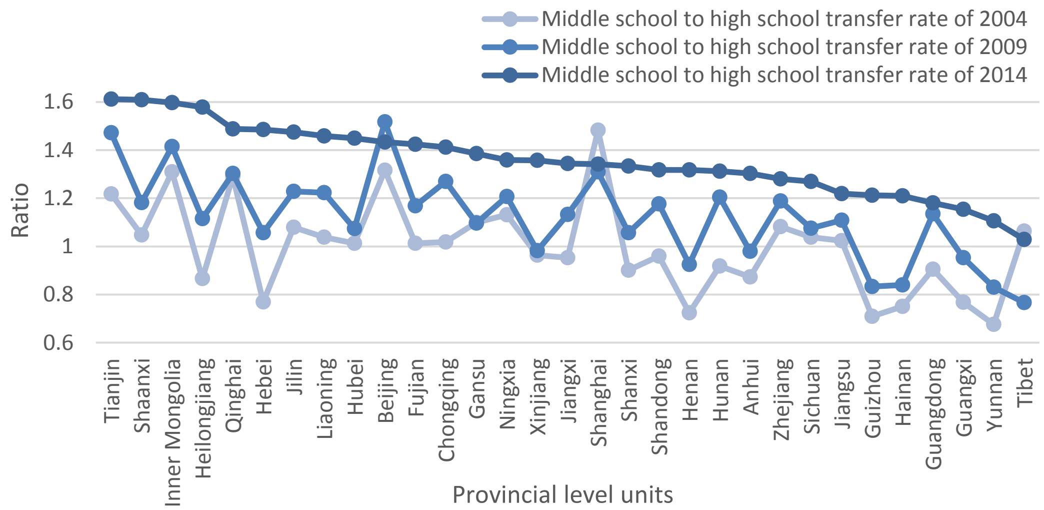
This can be further explained by the line charts in Figure 4.4. Education enrolment and transfer rates at each education stage for every province in 2004, 2009 and 2014 are normalised by dividing by their average score in 2004 respectively, so the average value of each variable in 2004 is 1. All the provinces are ranked by their transfer rates from middle school to high school in 2014. The net enrolment rates of primary schools in each province were all close to the average, 1, thus they are not shown in Figure 4.4. Only those rates for Qinghai, Sichuan and Tibet were slightly lower in 2004 but had already improved to the average level by 2014.

The primary to middle school transfer rates for provincial units (Figure 4.4a) also fluctuated above and below the average, 1. For the less developed areas such as Yunnan and Guizhou, the primary to middle school transfer rate experienced a continuous marginal increase. In contrast, this transfer rate for the most developed areas such as Beijing, Shanghai and Zhejiang decreased significantly over the ten years, partly due to migrant children returning to their registered homes to attend middle school education after graduating from the primary school in these areas. This occurs because they do not have the chance to enrol in a local high school, so it will be helpful for them to study in their hometown and prepare for their high school enrolment examination. Even recently, they still do not have the opportunity to enrol in local high schools and are only eligible to enrol in a local vocational school after finishing their compulsory education.

Compared to the variation of enrolment rates at the primary and middle school level, the provincial differentiations in the middle to high school transfer rates (Figure 4.4b) are very substantial in 2014. The high school transfer rate for Tibet was only around 0.76 of the average in 2014. From 2004 to 2014, the middle to high school transfer rate increased dramatically, especially from 2009 to 2014. However, the rate in Beijing increased from 2004 to 2009 but decreased from 2009 to 2014. The rate in Shanghai decreased significantly from 2004 to 2009 but slightly increased from 2009 to 2014. The rates in Qinghai, Hubei, Gansu, Xinjiang and Shanxi show almost no improvement from 2004 to 2009, but a dramatic improvement from 2009 to 2014. The primary to middle school transfer rate in Tibet decreased to the value of 2004 in 2014 after a slight increase from 2004 to 2009 (Figure 4.4a). Tibet's middle to high school transfer rate increased to the value of 2004 in 2014 after a substantial decline from 2004 to 2009 (Figure 4.4b). The exceptional decrease in developed areas, like Beijing, and in less developed areas, like Tibet, will be discussed in detailed in the next section.



a. Primary to middle school transfer rate



b. Middle to high school transfer rate

Figure 4.4 Normalised primary to middle school and middle to high school transfer rates, 2004, 2009, 2014

Source: Computed from China Educational Statistical Yearbook, 2004, 2009 and 2014

4.4.2.2 The spatial pattern of attainment

For education attainment, in 2004, the most eastern and central areas are educationally advantaged areas, especially Beijing and Shanghai, while western areas are educationally disadvantaged areas, except for Xinjiang (Figure 4.3d). From 2004 to 2014 (Figure 4.3d, e and f), the attainment level improved countrywide and the spatial pattern gradually changed from variations between eastern and western regions to differentiation between north-east and south-west areas. The south-west provinces remained disadvantaged in terms of educational attainment. Among the eastern provinces, attainment in Anhui was always lower than its surrounding areas.

Figure 4.5 displays the average years of schooling (AYS) in each province in the three years ranked by their value in 2004. Beijing and Shanghai have the highest AYS, not only because of the high concentration of top universities and research institutions in these areas but also because of their economic development and their ability to attract excellent well-educated labour from elsewhere. The other provinces' AYS focused on the range from six to eight years. However, the AYS of Tibet was less than four years, which was defined as "education poverty" in the Education for All, Global Monitoring Report 2010 and regarded as the minimum education period to acquire basic literacy skills (UNESCO, 2010). Moreover, Tibet is the only province to experience a slight decrease from 2009 to 2014, which is related to the changes in its primary to middle and middle to high school transfer rates in this period. Anhui and Jiangxi experienced a slight decrease from 2004 to 2009, which is probably due to insufficient support from the government or the substantial outflow of talent.

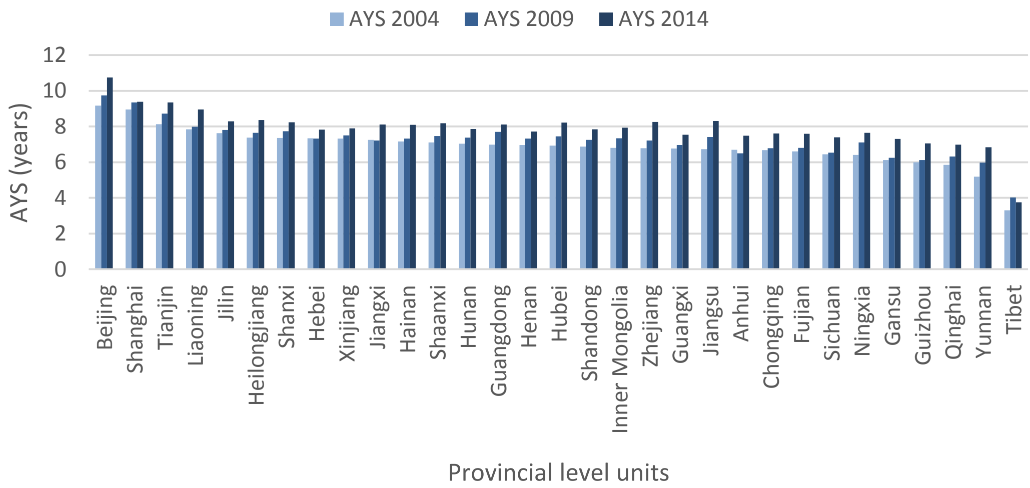


Figure 4.5 Average years of schooling by province, 2004, 2009 and 2014

Source: Computed from China Educational Statistical Yearbook, 2004, 2009 and 2014

4.4.2.3 The spatial pattern of provision

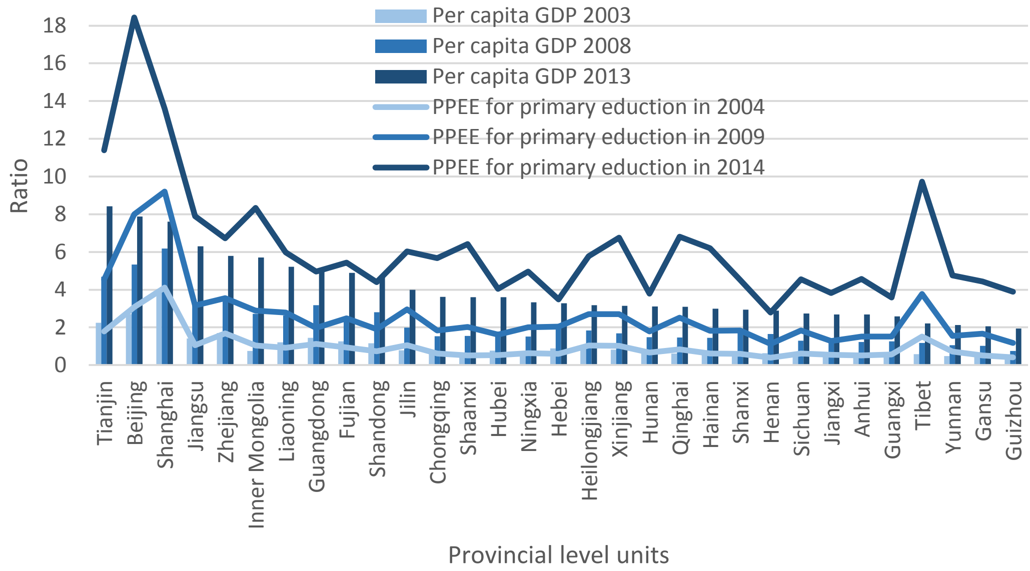
For education provision, only Beijing and Shanghai had comparatively high education provision scores in 2004 (Figure 4.3g). In addition, the scores of southern provinces were lower than those of the northern and coastal ones. From 2004 to 2014, as the western areas acquired more support from the Central Government, the pattern changed such that middle and southern areas became neglected with lower provision scores.

Figure 4.6 reveals the relationships between per-pupil educational expenditure (PPEE) in 2004, 2009 and 2014 and per capita GDP (PCGDP) in 2003, 2008 and 2013 by province. All the PPEE values were normalised by dividing the average PPEE of each education stage in 2004 and are represented by the continuous lines. In order to be consistent with PPEE, the PCGDP was divided by the average PCGDP in 2003 and represented by the clustered bar chart (Figure 4.6). Both the average PPEE in 2004 and PCGDP in 2003 were one and the two graphs are all ranked by PCGDP in 2013.

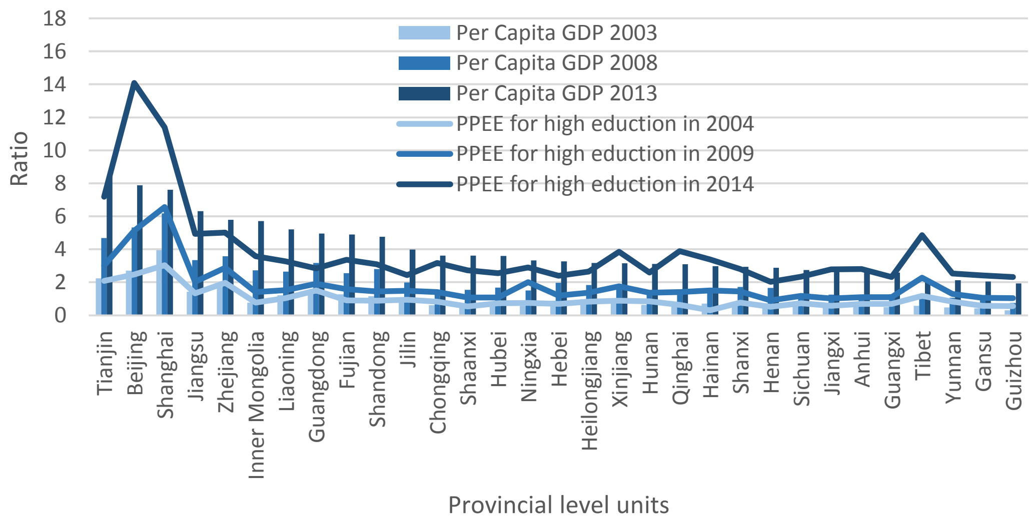
Previous empirical research has shown that PCGDP has a dominant influence that drives up PPEE inequalities (Wang, 2014). Similarly, in Figure 4.6a, Tianjin, Beijing and Shanghai, with high PCGDP in 2003, also have relatively high PPEE in 2004. These local governments all had greater ability to invest more in their educational development because of their better economic development and adequate tax income.

Owing to more support from Central Government, the Pearson correlation coefficient for PPEE for primary schools and the PCGDP have decreased from 0.94 to 0.75 from 2004 to 2014. The PPEE for Tibet, Qinghai and Xinjiang increased dramatically over the ten year period and were also relatively high despite their PCGDP being comparatively low. The high PPEE in these areas in 2014 is probably because of funding support from the Central Government. Except for the above special areas, which are supported by the Central Government, there is a high correlation between PCGDP and PPEE. In addition, the PPEE for primary schools increased much faster than PCGDP, due to the guidance of policy.

The PPEE for high schools (Figure 4.6 b) also shows the same trend, but we can find that the growth of PPEE for high schools is much lower than the growth of PCGDP in most regions, except for Beijing, Shanghai, Qinghai, Xinjiang and Tibet.



a. PPEE for primary school and per capita GDP



b. PPEE for high school and per capita GDP

Figure 4.6 The relationship among per-pupil educational expenditure of primary school and high school and per capita GDP by province, 2004, 2009 and 2014

Source: Computed from China Educational Finance Statistical Yearbook, 2004, 2009 and 2014 and China Statistical Yearbook of 2004, 2009 and 2014.

4.4.2.4 The explanations of spatial inequality in China under the capability approach

In this research, the regional education inequalities were explained against a background framework of Sen’s capability approach. This approach alerts us not to narrowly define education inequalities in terms of access, attainment or the distribution of resources but shifts the axis of analysis to evaluate the regions’ conditions or abilities to achieve better educational results, which they have reason to value. The framework based on the capability approach also outlined the way in which each facet might be prioritised and interpreted. The IREA provides us with a way to track progress and inequalities in education for China, whereas the decomposition of this index revealed the stories behind the index.

In 2004 (before the relevant policies had been implemented), attainment, which is the achieved function of a previous stage, showed variations between eastern and western areas. For enrolment, the scores of the provinces to the north-east of the IREA line are relatively better. In addition, the provision scores of southern provinces were lower than those of the northern and coastal ones. Beijing and Shanghai had comparatively high scores for the three facets in 2004. The educational resources provided will be converted into the regions’ capability under the influence of the conversion factors, enrolment and attainment. Accordingly, in 2004, some of the northern and coastal provincial level units have a higher capability (IREA score) to achieve better education in the future.

Consequently, education attainment in 2009 displays a pattern of differentiation between north-east and south-west China. In addition, the enrolment score shows a clear north-east versus south-west pattern. At this stage, there is more financial investment from Central Government devoted to support education in the western areas. Education provision in the northern, eastern and part of the western areas is higher than the remaining areas. Accordingly, the IREA score, which measures the capability, also shows a north-east and south-west gradient.

From 2004 to 2014, the attainment level improved countrywide and the spatial pattern gradually changed from variations between eastern and western regions to differentiations between north-east and south-west areas. In 2014, the enrolment rates in most areas along the dotted line and to the north-east improved and moved into the fourth or fifth quintiles. However, it worth mentioning that the enrolment scores of Beijing and Shanghai experienced a decline from 2004 to 2014. As the western areas acquired more support from the Central Government, the pattern of education provision changed such that middle and southern areas became neglected and showed lower provision scores.

In addition, in 2014, in the northern areas where there was more support from Central Government and a solid education foundation, and eastern coastal areas with their higher economic development levels, the capability of education (IREA) was higher, indicating that these areas will have more chance to be advantaged, while the south-west provinces, especially Tibet and Guizhou, will remain disadvantaged in education in future. Therefore, it is likely that education inequalities between the south-western and north-eastern areas will continue in the near future if the relevant policies remain unchanged.

4.5 Discussions

From the above analysis, we can conclude that basic education in China had experienced considerable development in the ten year period during which investment in the public budget for compulsory education increased from £24.3 billion to £119.9 billion with an average annual growth rate of 19.4% (Ministry of Education of PRC, 2015a). In addition, China’s education expenditure achieved its target of accounting for 4% of the Gross National Product (GNP) in 2012 (Figure 4.7).

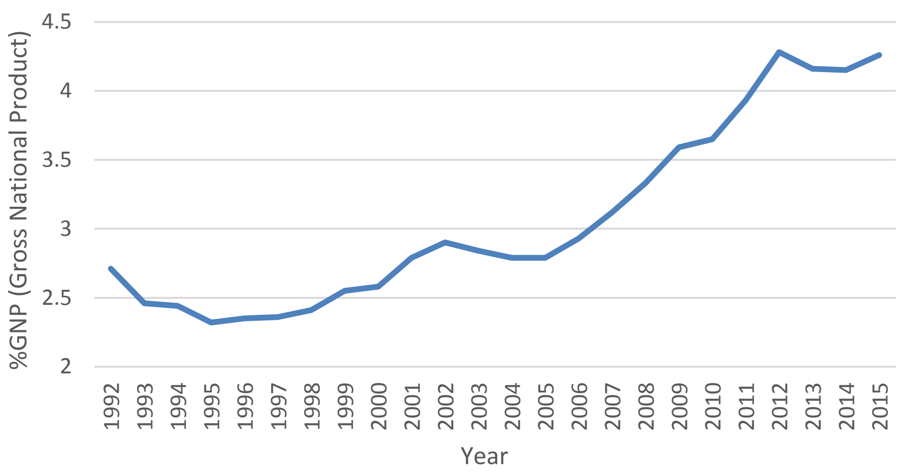


Figure 4.7 The government appropriation for education as a percentage of GNP, 1992-2014

Source: China Statistical Yearbook

Central Government has played an increasingly important role in promoting the development of education and in reducing education inequalities in China. The differences between provincial level units with highest and lowest PPEE had decreased from 9.38 to 5.60 times for primary education and from 9.31 to 5.96 times for high school education, so the gap of education expenditure among these areas is narrowing. However, the IREA index and its decomposition also indicate certain problems with the current policies, which will be discussed in the following four sub-sections: (1) the unreasonable regional division; (2) unequal education opportunities for rural child migrants; (3) lagging high school education; and (4) the poor education environment in the Tibet autonomous region.

4.5.1 Spatial pattern of IREA and area partitioning of policy implementation

It is not reasonable to implement education policies, especially the fiscal policies like NR, based on dividing China into eastern, central and western regions (Figure 4.1) to promote education equality. From current analysis, no matter what the patterns of IREA or the patterns of decomposed facets are, they do not match with the regional division adopted by Central Government. Education related policies should be implemented based on the evaluation of education-related indicators, whereas this division is based on economic development levels and natural environmental conditions. Thus, the new IREA which was developed based on direct indicators of education can better capture the differences in education development for each region across a number of facets.

Furthermore, even the economic development measured by PCGDP, cannot be generalised and fully revealed by this rough partition. Undeniably, this regional division is helpful to clearly define the fiscal responsibilities of Central Government and local governments; however, education inequalities vary spatially and temporally, and it is not reasonable to set policies without considering the variations within regions and their different development trajectories. This will decrease the effectiveness of these policies in reducing the provincial level and rural-urban educational disparities and cause a waste of education funds to some extent. Unfortunately, the newly issued policies of NUR in 2015 are still based on these problematic regional divisions.

4.5.2 Education of child migrants

Figure 4.3a and c illustrate that the enrolment scores of Beijing and Shanghai had decreased over the ten year period, this was due to their decreased rates of transfer from primary to middle school and from middle to high school. This is most likely caused by the strict limitations on the influx of migrants receiving education in these areas. Figure 4.8 displays the proportion of in-migrants in each provincial level unit, which is calculated by using the population whose registered places are in other provincial level units to divide by the total population and multiply by 100. The proportion of in-migrants in Shanghai and Beijing were the highest among all areas in 2010. Thus, there are extremely strict limitations on migrants in terms of receiving education in these areas. For example, migrants cannot enter local high schools and this has forced migrants' children to go back to their hometowns to receive middle and high school education, which has hindered population mobility and lead to education inequalities for child migrants. In other areas, even if the pupils are allowed to enrol in local schools, they need to provide a lot of documents, such as parents' certificate of employment and income, and the schools with good quality are normally not available to them.

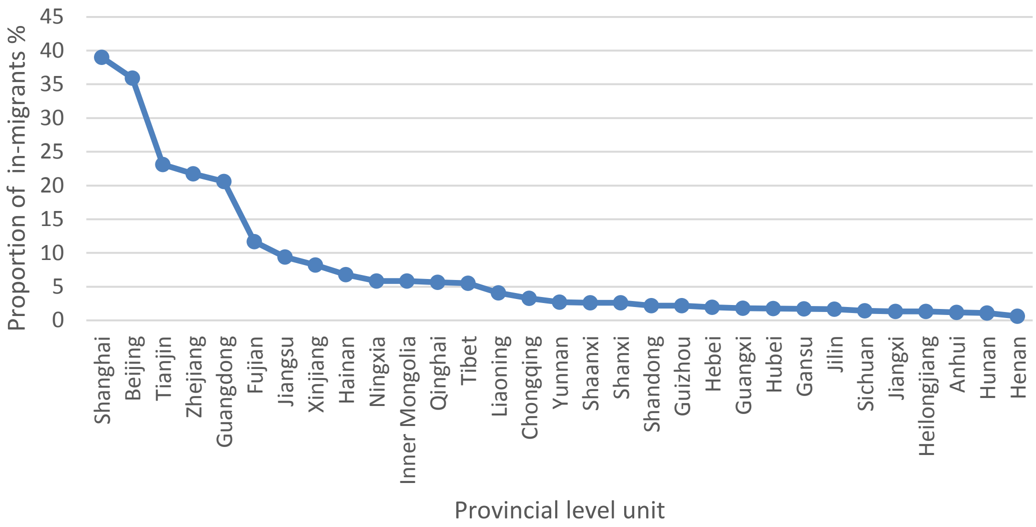


Figure 4.8 The proportion of in-migrants by provincial level unit, 2010

Data source: The sixth census data (2010)

Before 2015, only education in rural areas was supported by Central Government, whereas local authorities in urban areas had to cover almost all their education expenditure themselves, including free tuition and miscellaneous fees. These policies are very helpful to balance and improve education conditions in undeveloped rural areas but they have not been adapted to the situation of current rapid urbanisation and large-scale population movement. In 2014, migrant workers' children accounted for more than 25% of all primary school pupils and nearly 23% of all middle school pupils (calculated from data of China Statistical Yearbook in 2015). Providing education to substantial numbers of migrants from rural areas will largely increase the financial burden on urban governments, therefore urban authorities usually display a negative attitude to offering equal education to migrant children. Quality education is not really available for migrant children (Li and Placier, 2015; Xiang *et al.*, 2018). Thus, most migrants have to leave their children with their grandparents in rural areas. The so-called 'left-behind' children, accounting for 58.9% of primary school age migrants' children and 68.7% of middle school age migrants' children in 2014 (calculated from data of China Statistical Yearbook in 2015), suffer from developmental issues (e.g. juvenile delinquency and suicide caused by mental problems) and this is harmful for the education of the next generation of migrants. With continuing urbanisation, increasing numbers of rural residents will move to urban areas.

How to ensure that urban authorities provide equal education for migrants is a major problem confronting the Chinese Government.

The latest policy in 2015, NUR, is aimed at solving the above problems through unifying the share of Central Government's support via urban and rural education expenditure. Central Government will cover all the funds for free textbooks. Other expenditures will also be shared by local and Central Government based on certain proportions, regardless of whether they refer to urban or rural areas, except for the teachers' salaries. This policy will reduce urban governments' fiscal pressure on offering education to children of migrants from rural areas, as these funds will be largely sponsored by the Central Government. However, it worthy of note that the proportion shared by local governments and Central Government still vary by the regional divisions. A higher proportion of education expenditure is shared by local governments in eastern areas, which are the main target areas for migrants. In addition, teachers' salaries are still totally covered by local governments. Accordingly, this new policy will be helpful in solving the population movement between rural and urban areas, but its effects will vary between regions.

4.5.3 Lagging high school education

During the calculation process, we noticed that the enrolment of compulsory education is already very high (with the minimum 98.5% in 2014) and the variation is small among different areas at present. However, the provincial differentiation in middle to high school transfer rates is very substantial in 2014, with the lowest value of 43.9% (Tibet) and the highest value of 68.8% (Tianjin), the latter value being 1.5 times the former. In addition, high school education is the bridge between compulsory education and higher education. The comparatively lower transfer rate for high school indicates there may be a relatively lower percentage of high quality individuals entering the labour force in the future. The percentage of persons (aged from 20-65) with high school and above qualifications only accounts for 27.5% of the population, as measured using data from the sixth census in 2010. This suggests that the amount of high quality labour in China is still inadequate.

The apparent spatial variations in the middle to high school transfer rate occur arguably because of the lower percentage of education investment.

The growth rate of PPEE for high school education (3.2) is much lower than that of primary (6.5) or middle school education (6.9) from 2004 to 2014 (calculated from data in China Educational Finance Statistical Yearbook of 2004 and 2014). The PPEE of each education stage increases from primary to high schools, as higher-level education needs more funding for better facilities and teachers. However, the public financial expenditure on education per high school pupil (£872) was lower than that per middle school pupil (£954) in 2014 (Figure 4.9). In addition, the proportion of the public budget for high schools was only 68%, while that of primary and middle schools accounted for more than 80% in 2014. Thus, the education expenditure on high school education left more of a financial burden on families. More importantly, the expenditure on construction per pupil for high schools in 2014 only accounted for 0.64 of that in 2004. This means that less funding was used to construct infrastructure for high schools in 2014 compared to 2004, whereas construction funding for primary and middle schools more than doubled. The Pearson correlation coefficient of PPEE for primary schools and the PCGDP decreased from 0.94 to 0.75 from 2004 to 2014 owing to the Central Government financial support for less developed areas. However, that of high schools increased from 0.94 to 0.95, which indicates the education investment in high school education is still highly related to economic development.

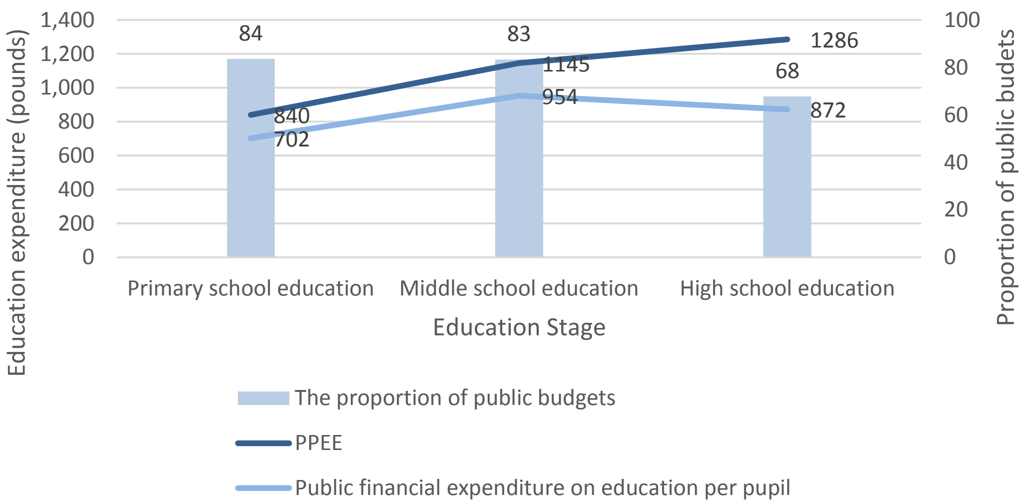


Figure 4.9 The comparison of education expenditure by each education stage, 2014

Source: Author’s calculations using data from China Educational Finance Statistical Yearbook of 2014

At present, compulsory education has acquired more policy attention, such as NR and NUR. However, after the popularisation of compulsory education, the more crucial task for China’s education system is to provide high quality labour to help China complete its industrial structural transformation. Thus, more emphasis should be put on high school education and policies are needed to specify a standard share of government expenditure for high school education. The financial burden on families and local government should be reduced and more efforts should be used to support the senior secondary education development in less developed areas. Furthermore, more policies should be implemented to increase the availability of high school education for migrant children, while most emphasis by the government at present has been on helping the children of migrants receive compulsory education locally.

4.5.4 Education in the Tibet Autonomous Region

The framework of the capability approach also helps us to explain why attainment falls short in some regions, despite the increased resources that have been allocated. Tibet is a case in point. Although an improvement of education in Tibet has taken place and its education provision has increased very rapidly in the last ten years (the PPEE for primary schools in Tibet is £1,661, twice as much as the national average of £840 in 2014), education in Tibet is still worse than other provincial units, especially in terms of the enrolment rate and attainment (for example, the middle to high school transfer rate was only 43.9% compared to the national average 56.4% in 2014; the AYS was only 3.8, while the national average was 7.9). Except for its historically low enrolment and attainment, its ability to convey educational resources into capability has been influenced by its different context, including its natural physical and social environments.

In 2004, the average population in Tibet was 2.6 persons per square kilometre and about 80% of residents lived in rural and nomadic livestock breeding areas (Postiglione *et al.*, 2011). The service radius of primary schools in rural areas is around 15 to 20 kilometres, and the situation is even worse for nomadic areas, reaching up to 100 to 150 kilometres (Postiglione, 2009). Moreover, it is also hard to attract qualified teachers to these remote nomadic areas. In addition, educational progress is hampered by parents' cultural perspective on education (Postiglione *et al.*, 2011). In addition, Tibet possesses a complex religious tradition and distinctive culture. The low enrolment rate in Tibet is not only due to its rugged geography, harsh climate, scarcity of resources and dispersed population, but also because of parents' cultural perspective on education (Postiglione *et al*., 2011). Some parents, especially those living in nomadic livestock breeding areas, are not willing to provide financial support for their children's schooling, because the curriculum that children are taught in school can be vastly different from their experiences in everyday life. In addition, they have not experienced many examples of economic improvement led by education and have not recognised its long-term value. Moreover, as a result of the influence of religion, some of the children have unfavourable attitudes to education since education content is incompatible with their own culture. (Gyatso *et al.*, 2005; Postiglione *et al.*, 2011).

4.6 Conclusions

A multidimensional index, the IREA, has been proposed in this chapter to evaluate the effectiveness of policies targeted at reducing regional educational inequalities in China since 2005. Education equality has been conceptualised by adopting Sen's capability approach to provide theoretical justification for the measurement of educational disparities and fill the current theoretical void for education inequality measurement. Education in the north-east areas appears better than in the south-west part of China, which is different from the area division adopted by the Central Government as the base of policy implementation. In addition, the temporal comparison (2004, 2009 and 2014) of the IREA helps us to explain how the pattern of education inequalities has evolved over time. Furthermore, key issues such as the equal education of child migrants and the obvious regional variation within high school education were all highlighted and some suggestions for improvement were proposed.

As all of the above policies are not only aimed at reducing regional education variations but also narrowing the education gap between rural and urban areas, rural and urban education inequalities in China is the focus of the next chapter in the thesis.

Chapter 5
The Rural-urban Education Gap in China

5.1 Introduction

China's population management policy formulation and implementation was based on the hukou system for a long time. As introduced in Chapter 2, the whole nation was divided into two groups; that is, people with urban hukou and with rural hukou. Explanations for the introduction of the hukou system can be traced back to the period of the establishment of the People's Republic of China (PRC) when China followed the model of Soviet Russia with accelerated industrialisation in urban areas at the cost of the development of rural areas (Fu and Ren, 2010). This policy is aimed at producing agricultural goods in rural areas and concentrating limited economic resources in urban areas to foster rapid industrialisation. It is also served as a constraint on rural-urban migration to ensure there were adequate rural labourers to produce agricultural goods and to limit the amount of urban residents who enjoyed low-price food (Cai, 2007). Therefore, the rural population suffered from these urban development centred policies(Fu and Ren, 2010).

During the 1980s and 1990s, the educational decentralisation tightened the link between local economic circumstances and school resources (Hannum, 2003). As discussed in Chapter 4, the lower level local governments struggled to collect enough funding for compulsory education, especially in the impoverished rural areas. Local governments in rural areas tend to transfer the responsibility for compulsory education to rural peasants and have relied heavily on informal levies (extra-budgetary sources) to compensate for insufficient state funding (Wang, 2014). Tuition fees, miscellaneous fees, and rural educational surtax paid by rural parents became the major sources of funding for compulsory education in rural areas between 1990 and 2000 (Fu and Ren, 2010). Consequently, decentralisation led to an excessive financial burden for rural residents who were already suffering from urban-centric development (Fu and Ren, 2010; Wang, 2014).

Important adjustments by the Central Government had been implemented to improve the unequal relationship between urban and rural areas since 2000. Over the past 15 years, the Chinese Government has been constantly increasing the input for rural education, including transfer payments and investments in special projects (e.g. Project Hope (Li and Liu, 2014)) and reducing the financial burden of rural residents. Education equality between rural-urban areas is an important part of education equality, which has been given great importance in the National Medium and Long-term Educational Reform and Development Plan Outline (2010-20) (NMLERD) in 2010 (The State Council of China, 2010; Sun, 2012). However, what has been the effects of these policies? Has the rural-urban education gap been reduced? Which areas have experienced a widening of the rural-urban educational gap? Thus, the purpose of this chapter is to quantify the education gap between rural and urban areas across all the provincial level units and evaluate the effects and influences of related policies.

This chapter is structured as follows. The following section gives a brief background to policies related to rural areas. The third section describes the education attainment gap between rural and urban areas using the 2000 and 2010 census data and discusses inequalities in educational provision between rural and urban areas. Some further discussions on rural-urban educational gap are presented in the fourth section. Finally, the fifth section summarises and concludes the chapter.

5.2 Policy background

Three aspects of the related policy background will be introduced: taxation reform and supportive policies for rural areas; the implementation of 'School Consolidation' in rural areas; and the varied rural-urban definitions.

5.2.1 The taxation reforms in rural areas

Since 2001, the Central Government of China has initiated a series of rural taxation reforms to relieve the excessive fiscal burden on rural residents. "Tax-For-Fee" was implemented as the first stage with all fees collected previously in villages and townships, including education surcharges, being abolished and replaced by agricultural taxes and related surcharges (Wang, 2014; Wang and Zhao, 2014). In 2004, the rural taxation reforms moved to the second stage as the Central Government decided to phase out the agricultural tax and abolish it entirely in 2006 (Wang, 2014; Wang and Zhao, 2014). The implementation of rural taxation reforms not only reduced previous financial burdens on rural families but also led to the establishment of a new financing mechanism for rural education with the provincial governments and Central Government playing an increasingly important role (Wang, 2014).

In addition, the “Two Waivers and One Subsidy” policy was introduced to further reduce the educational fiscal burdens on rural residents with financial difficulties in 2001. Under this policy, the Government provided poor pupils in rural areas with free textbooks, exemption from miscellaneous fees and a living cost subsidy for boarding pupils with financial difficulties (Wang, 2014; Rao and Ye, 2016). The rural public schools had shifted from mainly being funded mainly by local rural inhabitants to being a public-run education system, which was beneficial to the long-term development of rural education (Rao and Ye, 2016). In 2004, a “one-charge system” was initiated nationwide, allowing schools to collect a standard charge from pupils which was set by provincial governments. However, this policy has had limited influence because of the enactment of the NR policy in 2005, which extended the targets of “Two Waivers and One Subsidy” from rural families with financial difficulties to all rural residents.

In the meantime, Central Government shifted its main funding sources from township to county in 2001 and then to provincial government in 2006, ensuring adequate education spending in rural areas, while the administrative responsibility was shifted to county level government in 2001 (Wang and Zhao, 2012).

5.2.2 School Consolidation in rural areas

In the initial period of establishment of the People's Republic of China (PRC), rural residents felt pride in being the masters of the nation and they committed themselves to develop rural education, motivated by the goal of realising a socialist country. The primary and middle schools had witnessed a dramatic increase in rural communities from 1949, when the PRC was established, to 1978 (Rao and Ye, 2016). A principle of ‘at least one primary school in each village, and one middle school in each town or township’ was established in rural areas (Mei *et al.,* 2015). In addition, plenty of education departments at county level stipulated that schools should be built within 2.5 kilometres of a village to accommodate rural pupils, thus allowing the rural pupils to attend schools located near to their homes (Rao and Ye, 2016).

Since the 1980s, the number of rural school-age children had decreased sharply due to the strict one child policy and the rapid urbanisation process (Rao and Ye, 2016). The birth rate and school age population have continued to decline because of the strictly implemented 'one child policy'. Due to loose migration controls, a large number of surplus rural labour flooded into urban areas searching for work and their children accompanied them seeking to study opportunities. Moreover, affluent rural families tended to send their children to urban schools to receive better education (Zhao and Parolin, 2014). Correspondingly, The Decision on the Reform and Development of Compulsory Education was issued by the State Council (2001), explicitly requiring that the school plans and educational resources should be adjusted and integrated to improve the education quality of rural schools (Mei *et al.*, 2015; Rao and Ye, 2016). After that, 'School Consolidation', also called Rural School Mapping Adjustment (Rao and Ye, 2016), was implemented by the local government. The village-level schools which were considered to be costly and inefficient were withdrawn or merged and boarding schools were promoted in rural areas; accordingly, the location of primary schools shifted from within villages to townships and the location for the secondary schools changed from townships to counties (Rao and Ye, 2016). Simultaneously, there were more investments in merged schools which were normally located above the township level to improve the teaching and physical conditions of these schools (Wang and Zhao, 2012).

The main purpose of School Consolidation was to save costs, enhance administrative efficiencies and improve education quality in sparsely populated rural areas (Zhao and Parolin, 2014). However, a lot of problems transpired due to unreasonable school mergers and unreasonable educational planning, such as wasted rural educational resources, oversized schools, long commuting distances and increased education-related financial burdens on rural families. In 2012, the School Consolidation policy was suspended by the Ministry of Education (Mei et al., 2015).

5.2.3 The discrimination of rural-urban division and rural-urban education

Before undertaking a rural-urban education comparison, the definition of rural and urban education needs clarification. According to regulations about rural and urban divisions stipulated by the National Bureau of Statistics of the PRC, there are three categories: urban; town (or county); and rural areas. The statistical definition adopted by the China Statistical Bureau (1999) indicates that the urban and town both belong to the 'urban' category (Fu and Ren, 2010; Zhao and Parolin, 2014). Figure 5.1 displays the spatial distribution of the proportion of the urban population in each provincial level unit in 2005 and 2014. As the data for 2004 are not available in the China Statistical Yearbook, the data for 2005 are used for comparison with those for 2014. The average proportion of the urban population by the end of 2005 was 43% indicating that the urban population was still less than that of the rural population (Figure 5.1a). However, this proportion varied significantly among different regions and decreased from coastal to inland areas. The urban population of Shanghai and Beijing accounted for as much as 89.1% and 83.6% of the total population respectively, while that of Tibet was only 20.9% in 2005. In 2014, the national percentage of the urban population had increased to 54.8% (Figure 5.1b). From 2005 to 2014, the proportion of the urban population increased countrywide; however, these proportions varied spatially. The urban proportion in Shaanxi grew by 15.3% from 2004 to 2014, while that of Shanghai and Beijing only increased by 0.5% and 2.7% respectively. By 2014, the proportion of the urban population in most regions was more than 40%, except for Tibet, where it remained below 30% at 25.7%.

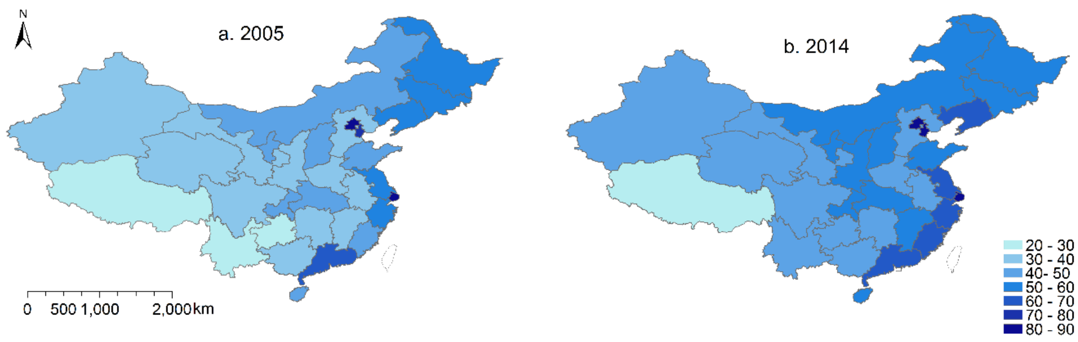


Figure 5.1 The proportion of population that is urban by provincial level units, 2005 and 2014

Source: China Statistical Yearbook, 2013 and 2016

Although 'town' and 'urban' are both regarded as urban areas in the rural-urban division mentioned above, the education in 'towns' will be treated as rural education in this research. Normally only the primary schools are located in rural areas, but the secondary schools tend to be located in townships or counties even before School Consolidation (Mei *et al.*, 2015). Secondly, due to a large number of village-level primary schools that were withdrawn or merged under the School Consolidation policy, the location of primary schools moved from within villages to townships and secondary schools moved from townships to counties (Rao and Ye, 2016). Table 5.1 illustrates that the number of primary and middle schools had decreased by 61.8% and 53.5% respectively in rural areas between 2004 and 2014 (actually, the number had experienced an increase after the suspension of the School Consolidation in 2012), while the number of primary and middle schools in towns increased by 38.9% and 44.5% from 2004 to 2014. The number of high schools decreased both in rural areas and towns.

Table 5.1 School decrease in rural areas and school increase in towns

a. School decreases in rural areas

Rural areas	Primary schools	Middle schools	High schools
2004	337,318	38,095	2,454
2014	128,703	17,707	667
Decrease rate	61.8%	53.5%	72.8%

b. School increases in towns

County(town)	Primary schools	Middle schools	High schools
2004	33,420	16,218	7,169
2014	46,414	23,429	6,164
Increase rate	38.9%	44.5%	-14.0%

Furthermore, as shown in Figure 5.2, the total rural primary school, middle school and high school pupil ratios were 32.3%, 17.1% and 3.3% respectively in 2014. These proportion of rural pupils are significantly lower than the percentage of the rural population (45.2%), which means a large number of rural pupils were not studying in rural areas but in towns or urban areas. Thus, it is not appropriate to employ the above rural-urban division to conduct a rural-urban education comparison, especially for middle and high school education.

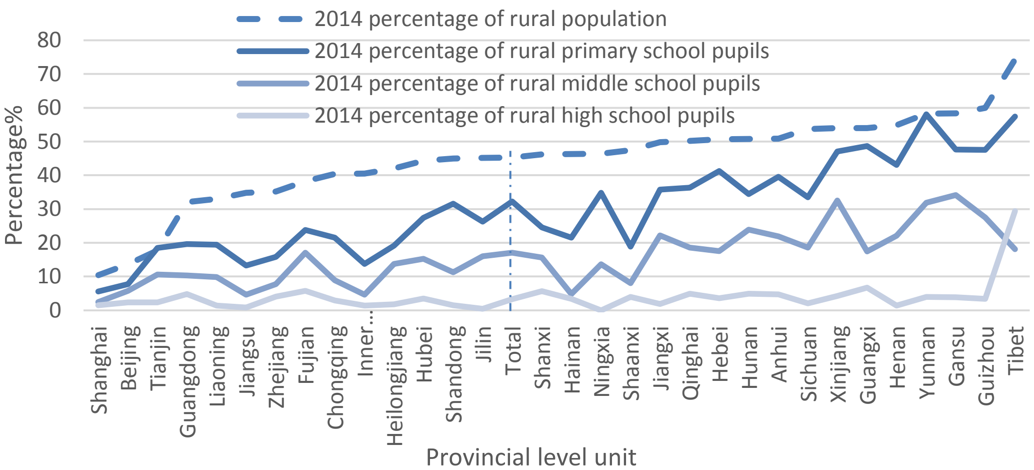


Figure 5.2 The percentage of rural population and of pupils studying in rural areas, 2014

Source: Computed from China Statistics Yearbook and China Educational Statistics Yearbook, 2014

The scope of rural education used in most of the official educational development reports does not exactly coincide with the definition of rural areas based on the rural and urban division, such as in the definition used in educational statistics yearbook, the report on national education development (Ministry of Education of PRC, 2016) and the Statistical Indicator System of China's Educational Monitoring and Evaluation (Ministry of Education of PRC, 2015b). In these official education statistics, rural education refers to the education of 'big rural' areas, which includes education in rural areas and towns. Rural education is defined from the perspective of the service object; as only some of the primary schools are located in rural areas, the middle schools and high schools which also provide education to rural population are mainly located in towns or counties. Consequently, rural education is regarded as education in rural areas and towns (counties) in this research, which is consistent with the rural-urban education division used by most educational statistics.

However, education in county seats is excluded from the concept of rural education in the official educational statistics, like the Educational Financial Statistical Yearbook, and the transfer payments of Central Government for rural education will not be allocated to these areas. Thus, for education policies and education fiscal policies, rural education refers to education in 'towns' and 'rural areas', and excludes the county seats. Most of the data used in this research are from the China Educational Statistical Yearbook, where the data are classified into the three categories: urban, town and rural areas. Thus, the data on rural education are obtained by combining data on 'rural areas' and 'towns'; however, the data of county seats cannot be excluded. Thus, the rural pupil and school proportions will be slightly larger than their actual values in the rural-urban comparison which now follows.

5.3 The rural-urban education comparison

After introducing the related policies and confirming the rural-urban education definition, rural-urban education has been compared from both the demand (attainment) and supply (provision) sides. As there is no separate enrolment data for rural and urban areas, only education attainment and education provision will be compared in this section.

5.3.1 Education attainment variation between rural and urban areas

As no time-series data are available on education attainment for rural and urban areas in each province in the Educational Statistics Yearbooks, the data used for assessing rural and urban education attainment disparities are derived from the 2000 (the fifth) and 2010 (the sixth) population censuses. Given that only national level data are available for use in comparing rural and urban education attainment in 2000, the temporal variation can only be compared at the national scale. As mentioned in the last section, the data for ‘rural’ and ‘town’ are also combined to give data for rural areas.

Figure 5.3 displays the qualification composition of the population living in rural and urban areas in 2000 and 2010. It is clear that there exists a gap in educational attainment levels between rural and urban areas in both years. For example, in 2000, those who received education at senior secondary school (high school and vocational secondary school) and above (college, undergraduate and postgraduate) accounted for 36.1% in urban areas and only 9.3% in rural areas. Owing to the implementation of the Compulsory Law in 2006 and the “Two Waivers and One Subsidy” policy of NR (2005), the proportion of people with middle school and above qualifications had increased significantly in rural areas by 2010. However, the educational attainment in rural areas in 2010 was still significantly lower than that of urban areas and even lower than that of urban areas in 2000. The supply of high-quality labour in rural areas remains insufficient; for example, only 4.1% of rural residents had achieved a college and above education in 2010, while the proportion was 21.5% for urban areas. In order to compare the educational attainment gap between rural and urban areas, the AYS was calculated in the following section.

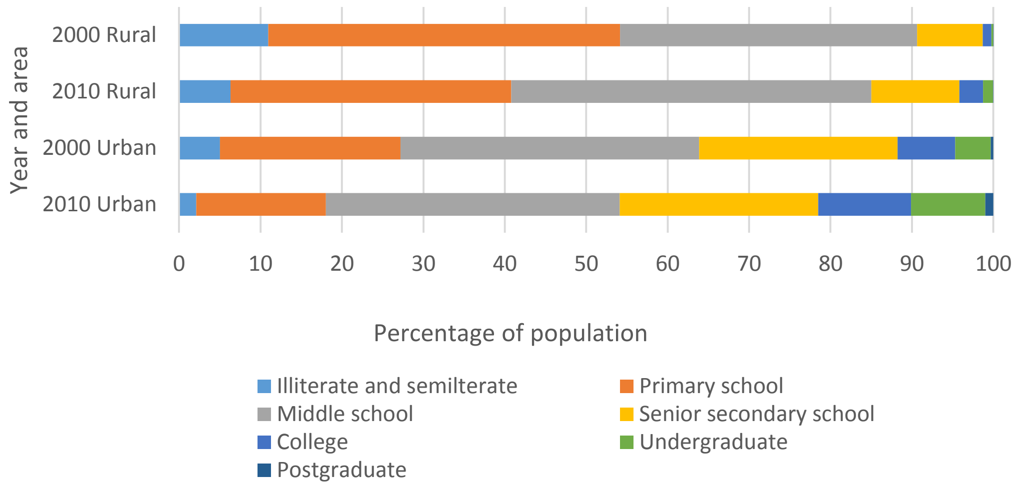


Figure 5.3 The composition of residents' qualifications in rural and urban areas, 2000 and 2010

Source: Computed from 2000 and 2010 census data

Average years of schooling (AYS) is used as a proxy for comparing rural and urban education attainment in 2000 and 2010. The population aged six and over was classified into eight categories on the basis of their qualifications in the 2000 census data, which also depends on the last year of education a pupil had attained: illiterate; primary school; junior secondary school (includes middle school and vocational junior secondary school); high school (senior secondary school); vocational secondary school; college; undergraduate; and postgraduate. The education duration for each level of education used is slightly shorter than its full duration. Average education duration was assumed as five years for those completing primary school (p), eight years for middle school (m), 11 years for high school (h), 10 years for vocational secondary school (v), 14 years for college (c), 15 years for undergraduate (u) and 19 years for those completing postgraduate study(po). Pop_i refers to population aged six and over in area i.The AYS in 2000 in area i,V_{2000i}, is calculated by the formula modified from Qian and Smyth (2008) as follows:

$$V_{2000i} = \frac{5p_i + 8m_i + 11h_i + 10\mathrm{v}_i + 14\mathrm{c}_i + 15u_i + 19po_i}{Pop_i} \qquad (5.4)$$

In the 2010 census, the above qualifications were classified into seven types. High school and vocational secondary were combined into high school, thus the average education duration for high school in 2010 is 10.5. The formula of AYS in 2010 therefore slightly different to equation 5.4:

$$V_{2010i} = \frac{5p_i + 8m_i + 10.5\mathrm{h}_i + 14\mathrm{c}_i + 15u_i + 19po_i}{Pop_i} \quad (5.5)$$

Table 5.2 reveals the overall attainment gap between rural and urban areas across the 10 years. The data of both 2000 and 2010 revealed a very substantial gap between rural and urban areas, although many policies were introduced to try to narrow this gap. The conclusion from the comparison of AYS in Table 5.2 is that these policies failed to narrow the gap entirely between rural and urban areas in terms of attainment from 2000 to 2010.

Table 5.2 The national level AYS in rural and urban areas, 2000 and 2010.

	2000 AYS	2010AYS	2000 Urban rural gap	2010 Urban rural gap
Urban	8.35	9.39	2.22	2.4
Rural	6.13	7		

Source: Computed from 2000 and 2010 census data

The provincial level distributions of rural and urban education qualification levels are shown in Figure 5.4. The distribution of AYS in rural (Figure 5.4a), urban areas (Figure 5.4b) and the rural-urban AYS gap (Figure 5.4c) all follow the pattern mentioned in Chapter 4. The educational attainment in provinces which are to the north-east of the dotted line is generally better than those to the south-west of the line.

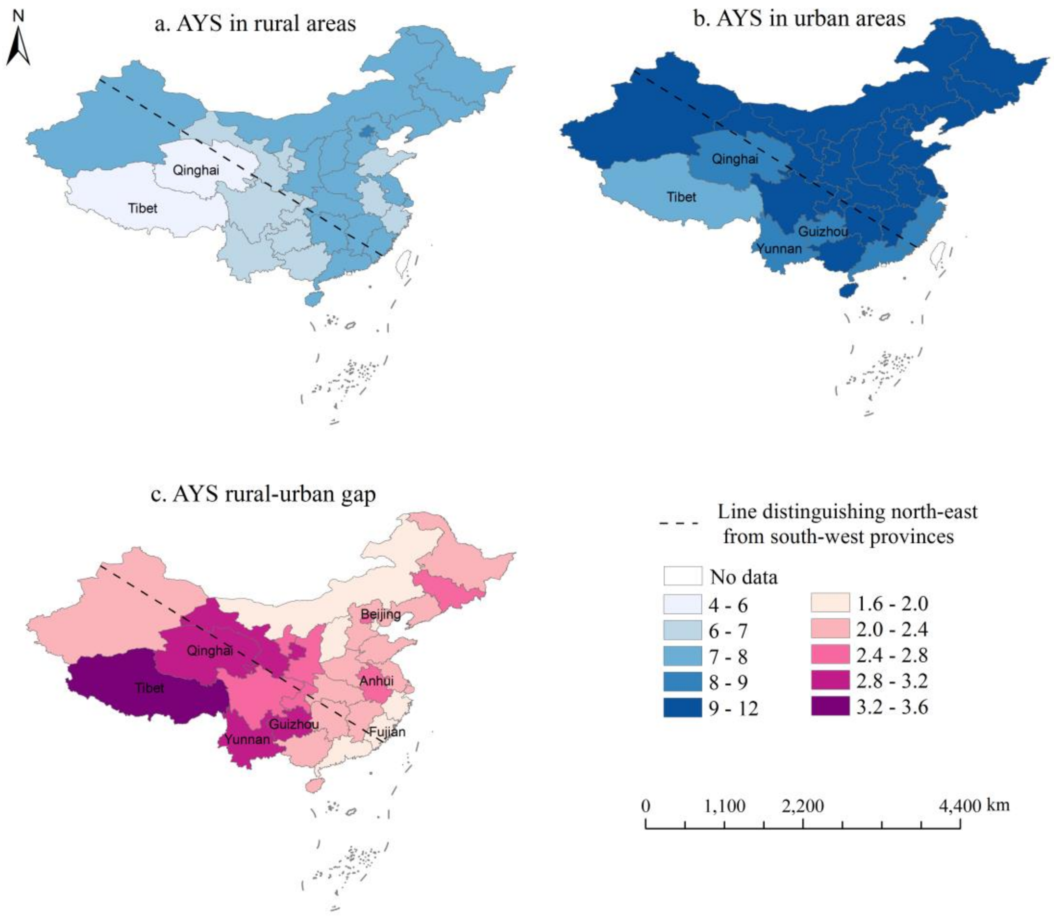


Figure 5.4 AYS of rural-urban areas in provinces in 2010

Source: 2010 census data

Figure 5.4c reveals the variations in the rural and urban gap across the provinces. The south-east coastal areas have the smallest gap (e.g. 1.69 years in Fujian), while the less developed western inland areas, and have a comparatively larger gap (e.g. 3.56 years in Tibet). When comparing Figure 5.4c with Figure 5.1, the provincial level units with higher proportions of the urban population normally have a lower rural-urban education attainment gap and vice versa, except for Beijing and Shanghai. For example, although the AYS in rural Beijing in 2010 is the highest among all rural areas, the rural-urban gap is still very high at 2.51 years.

5.3.2 Inequalities in education provision between rural and urban areas

An examination of the rural-urban education provision has been conducted from four perspectives: per-pupil educational expenditure (PPEE); the teacher-pupil ratio; teachers' qualifications; and teachers' professional rank.

5.3.2.1 PPEE variation between rural and urban areas

Only data on average PPEE for rural-urban and rural areas are available from the China Educational Finance Statistical Yearbooks; there is no separate PPEE data for urban areas, so the rural and urban PPEE disparities will be revealed by comparing average PPEE for rural areas with that of the rural-urban average. Moreover, these data are only accessible for primary and middle school education; there is no rural PPEE data available for high school education. In 2004, the PPEE disparity between rural and urban areas was very significant, especially for middle school education (Figure 5.5 and 5.6). From 2004 to 2014, the rural and rural-urban average PPEE ratio (RA) for primary and middle schools had increased dramatically (Figure 5.5). The PPEE of rural areas was only marginally lower than that of the national average in 2014 (3% lower for primary schools and 4% for middle schools). Thus, the gap in education investment between rural and urban areas was narrowing.

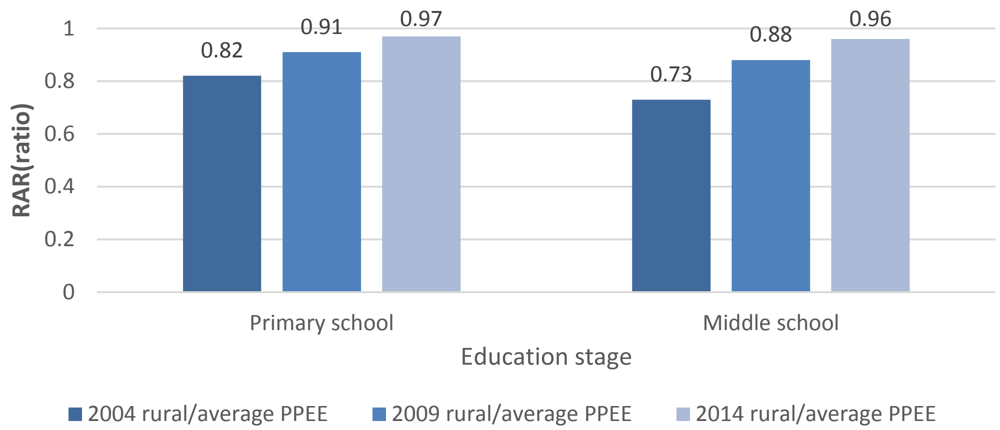


Figure 5.5 Rural and rural-urban average PPEE ratios (RA), 2004, 2009 and 2014

Source: Author's calculations using data from the China Educational Finance Statistical Yearbook of 2004, 2009 and 2014

The RA distribution patterns of primary schools and middle schools are different (Figure 5.6). The RA for primary schools were lower in the south-east areas in both 2004 and 2014, which means the rural-urban disparities of PPEE were more significant in these areas. In contrast, owing to more investment from Central Government since 2001 and NR policy in 2005, the most western areas had increasingly smaller rural-urban PPEE differentials and some western rural areas even had slightly higher PPEE in 2014, for example, Inner Mongolia (1.50), Shanxi (1.23) and Tibet (1.06). This is very beneficial for reducing the current rural-urban education gap.

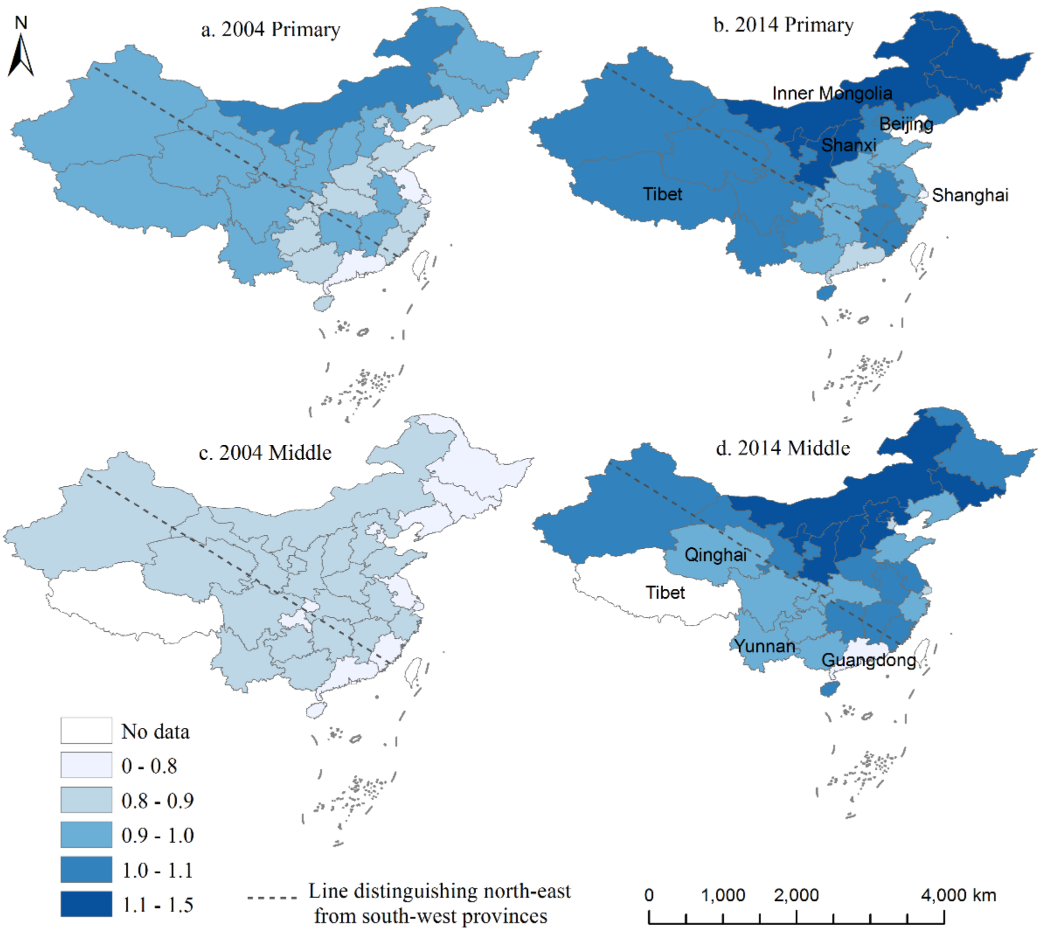


Figure 5.6 Rural and rural-urban average PPEE ratios (RA) for primary and middle schools by province, 2004 and 2014

Source: Author's calculations using data from China Educational Finance Statistical Yearbook of 2004 and 2014

The spatial variations in the rural-urban PPEE gap for middle schools are different from that of primary schools in both 2004 and 2014 and the gap was more substantial. The areas to the south-west of the IREA line still have lower RA (Figure 5.6d) in 2014, such as Yunnan and Qinghai. This will most likely lead to the education quality of middle schools further lagging behind, which means more financial support for middle schools is required in these rural areas.

Although RA were very low in some developed areas, like Beijing and Shanghai, the rural PPEE in these areas was already greatly higher than the national PPEE in rural areas. For example, the RA for primary schools is only 0.78 (ratio) in Shanghai in 2014, however, the PPEE in rural areas of Shanghai is already 2.23 times that of the national average PPEE in rural areas. This is because the local authorities in these developed areas normally have more funds for investment in urban education. Even though there is already quite a lot of educational investment in these rural areas, the PPEE is still comparatively low compared to the extremely high urban PPEE. Guangdong is an exceptional in having a lower RA and the PPEE in rural areas is also lower than the national average. There is no PPEE data for middle schools in Tibet.

5.3.2.2 Comparing teacher-pupil ratios between rural and urban areas

Teacher-pupil ratio, which has been mentioned in Section 4.3.1.3, is an indicator of education provision. For example, the average teacher-pupil ratio of middle schools in rural areas was 0.052, while that for urban areas was 0.062 in 2004. As a higher teacher-pupil ratio indicates that each pupil can expect more attention from teachers, thus urban middle school education had clear advantages when compared to rural middle school education in 2004. Rural-urban teacher-pupil ratios (RUTP) are obtained by dividing the teacher-pupil ratio in rural areas by that of urban areas. Figure 5.7 displays the changes in the national level RUTP from 2004 to 2014. In 2004, the RUTP for all education stages was lower than one, which means that the teacher-pupil ratio was higher in urban areas in 2004. However, this situation then transformed dramatically. In 2014, the RUTP in primary and middle schools were higher than one, which indicates a higher teacher-pupil ratio in rural areas. However, the RUTP for high school was still below one.

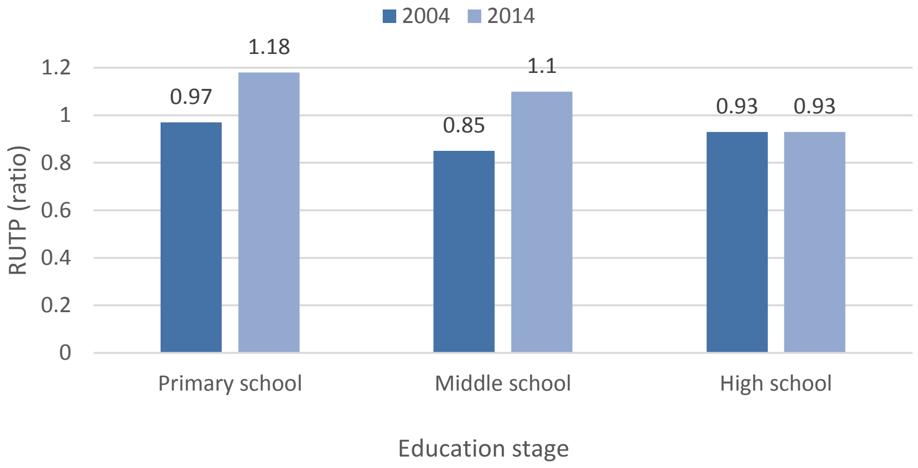


Figure 5.7 Rural-urban teacher-pupil ratios (RUTP), 2004 and 2014

Source: Computed from China Educational Statistical Yearbook of 2004 and 2014

The RUTP in Figure 5.8 has increased significantly in almost all provincial level units from 2004 to 2014 for primary and middle schools. Most provincial level units had an RUTP of more than one in 2014 for primary and middle schools; that is, rural areas' teacher-pupil ratios were higher in these areas. The dramatic increase of teacher-pupil ratios in rural areas for primary and middle school education was likely to have occurred because of increased payments for rural teachers, the urbanisation process with increasingly more pupils moving away from rural and into urban areas, and the large rural and urban gap in education quality which has led more pupils to go to urban areas to receive education. However, the ratios in certain provincial level units, such as Tibet (0.88), for middle schools were lower than one, which means significantly lower teacher-pupil ratios in rural areas in these areas.

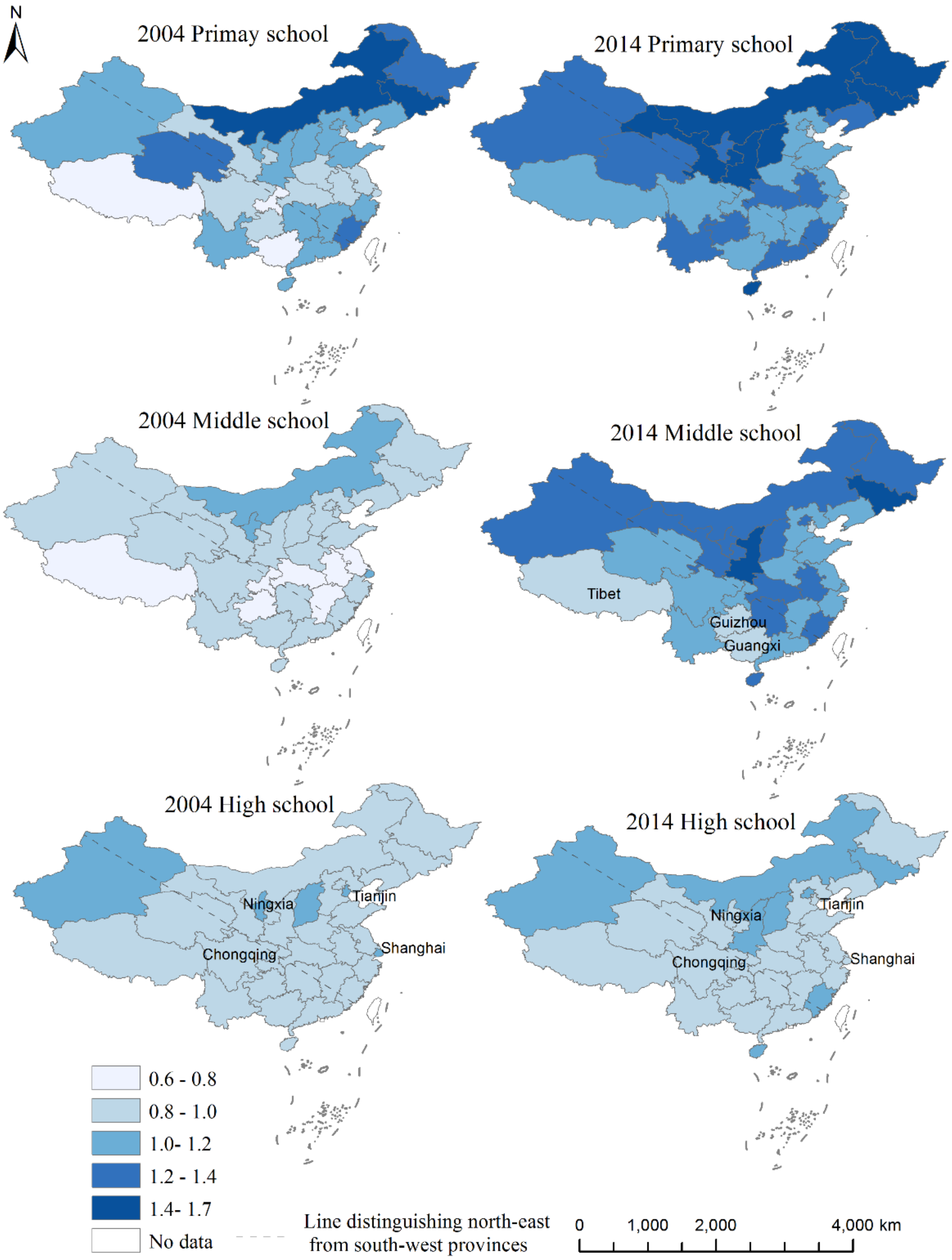


Figure 5.8 Rural-urban teacher-pupil ratios (RUTP) by province, 2004 and 2014

Source: Computed from China Educational Statistical Yearbook of 2004 and 2014

The change of RUTP for high schools is minimal and the values in most areas are still lower than one. The RUTP in some provincial level units even decreased over the ten year period, like Tianjin, Shanghai and Ningxia. The

RUTP in Chongqing were all around 0.87 in both years. The lower teacher-pupil ratio for high schools indicates that there is still a lack of sufficient input in terms of teachers for rural high school education. Since the migrants' children are not allowed to attend the university entrance examination locally, they are forced to go back to their hometown to receive high school education. This is most likely the reason why the RUTP for high school education is different from that of compulsory education (primary and middle school education).

5.3.2.3 Teachers' qualification comparison between rural and urban areas

In order to compare teachers' qualifications between rural and urban areas, the rural-urban qualified teacher ratio (RUQT) was calculated by using the qualified teacher proportion in rural areas divided by that of urban areas. Figure 5.9 reveals that there were very dramatic rural-urban disparities for teachers' qualifications in 2004, especially for middle schools. However, the national average RUQT had increased dramatically by 2014, which indicates that the teacher quality gap between rural and urban areas was shrinking. The ratio of qualified teachers in middle schools had improved the most, by around 0.43, but was still lower than one, which means there was still a teacher qualification gap between rural and urban areas for middle schools in 2014. The qualified teachers in rural primary and middle schools were 0.90 and 0.84 that of schools in urban areas, while the teacher qualifications for high schools in rural and urban areas were very similar in 2014.

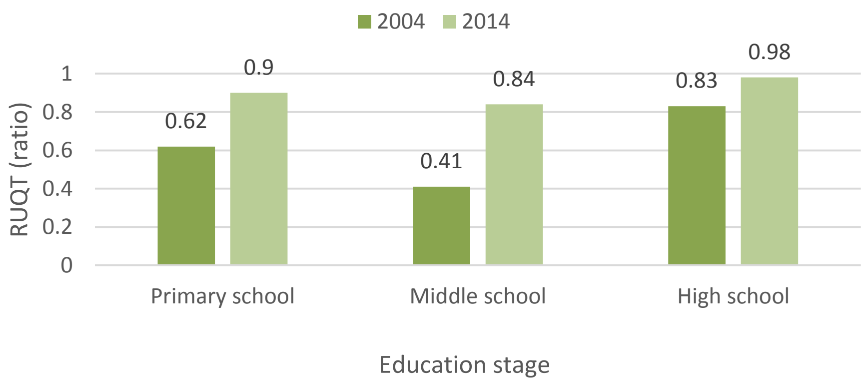


Figure 5.9 Rural-urban qualified teacher ratios (RUQT), 2004 and 2014.

Source: Compute from China Educational Statistical Yearbook of 2004 and 2014.

The spatial distribution of RUQT by the provincial unit is shown in Figure 5.10. The rural-urban teachers' qualification gap in 2004 was largest for middle schools and followed by primary schools, but these gaps at each education stage had decreased very considerably between 2004 and 2014, especially for middle schools. In 2004, the RUQT for middle school in a lot of central and southern provinces were less than 0.4; however, the ratios for most of the provincial level units improved to between 0.8 and 1.0 in 2014. The spatial patterns of RUQT in 2004 for each education stage were very diverse but the pattern was more unified in 2014. The urban-rural gaps were mostly between 0.8 and 1.0 and the RUQT in a lot of provinces were very similar. The primary school teachers' qualifications in Tibet were even higher than those in urban areas.

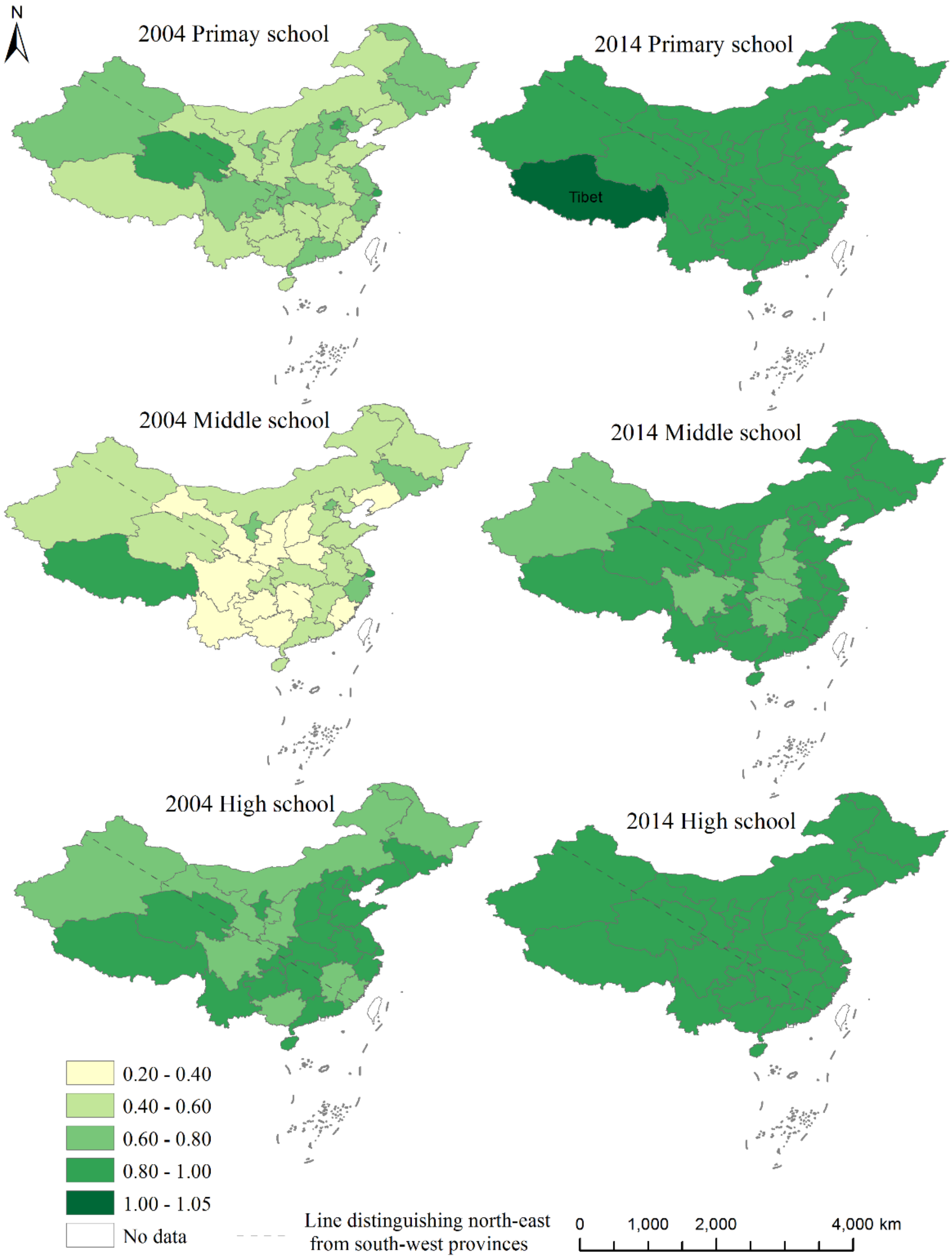


Figure 5.10 Rural-urban qualified teacher ratios (RUQT) by province, 2004 and 2014

Source: Computed from China Educational Statistical Yearbook of 2004 and 2014

5.3.2.4 Teachers' professional rank comparison between rural and urban areas

Apart from teacher qualifications, the other indicator reflecting teaching quality is teachers' professional rank. A higher percentage of teachers with higher ranked professional titles indicates better educational quality. For example, rural areas were disadvantaged in terms of education quality for middle school education in China in 2004, as the percentage of teachers with senior grade and first grade ranking in rural areas (36.0%) was lower than that of urban areas (55.9%). The rural-urban higher professional rank ratio (RUPR) was obtained by comparing the proportion of teachers with higher professional rank in rural areas with that of urban areas (e.g. the RUPR of middle school education was 0.64 in 2004). Figure 5.11 illustrates that the gap between rural and urban areas is shrinking, but the gaps are still considerable including high school education.

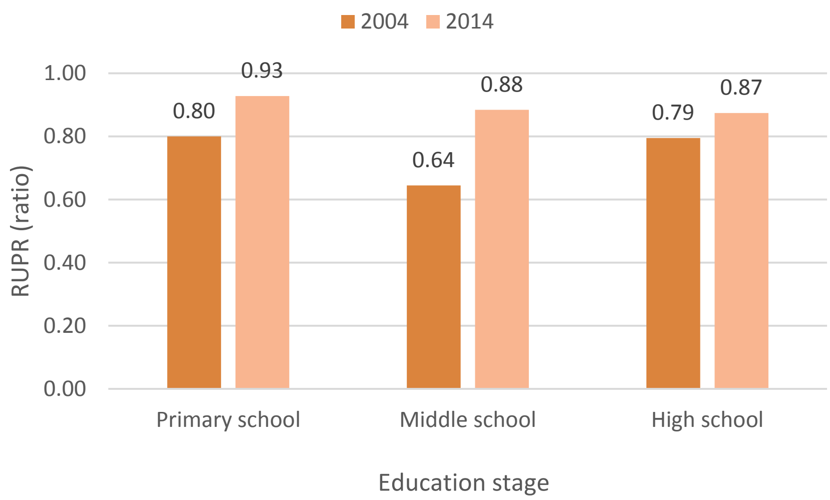


Figure 5.11 Rural-urban teachers with higher professional rank ratios (RUPR), 2004 and 2014

Source: Compute from China Educational Statistical Yearbook of 2004 and 2014

Although the RUQT and RUPR are both used to mirror the education quality or teaching quality gap between rural and urban areas, they display slightly different patterns. The RUPR displays significant rural-urban gaps that vary spatially (Figure 5.12). Compared to the relatively unified rural-urban gap distribution reflected by RUQT, the RUPR indicates there were still dramatic gaps in rural areas in terms of teachers with higher professional rank in 2014. Teachers' professional rank is based on the assessment of their educational background, teaching skills and professional accomplishments. Compared to teacher qualifications, teacher rank is a more comprehensive indicator and is also influenced by experience and years of working. If the policies are implemented to recruit teachers with higher qualifications in rural areas, the RUQT can be improved significantly in a short time. However, the improvement of the RUPR will be slower than RUQT. In addition, the RUPR reflects the rural-urban gap in terms of high quality teaching resources. It is apparent that the high quality education resources gap between rural-urban areas was still significant in 2014, especially in western regions.

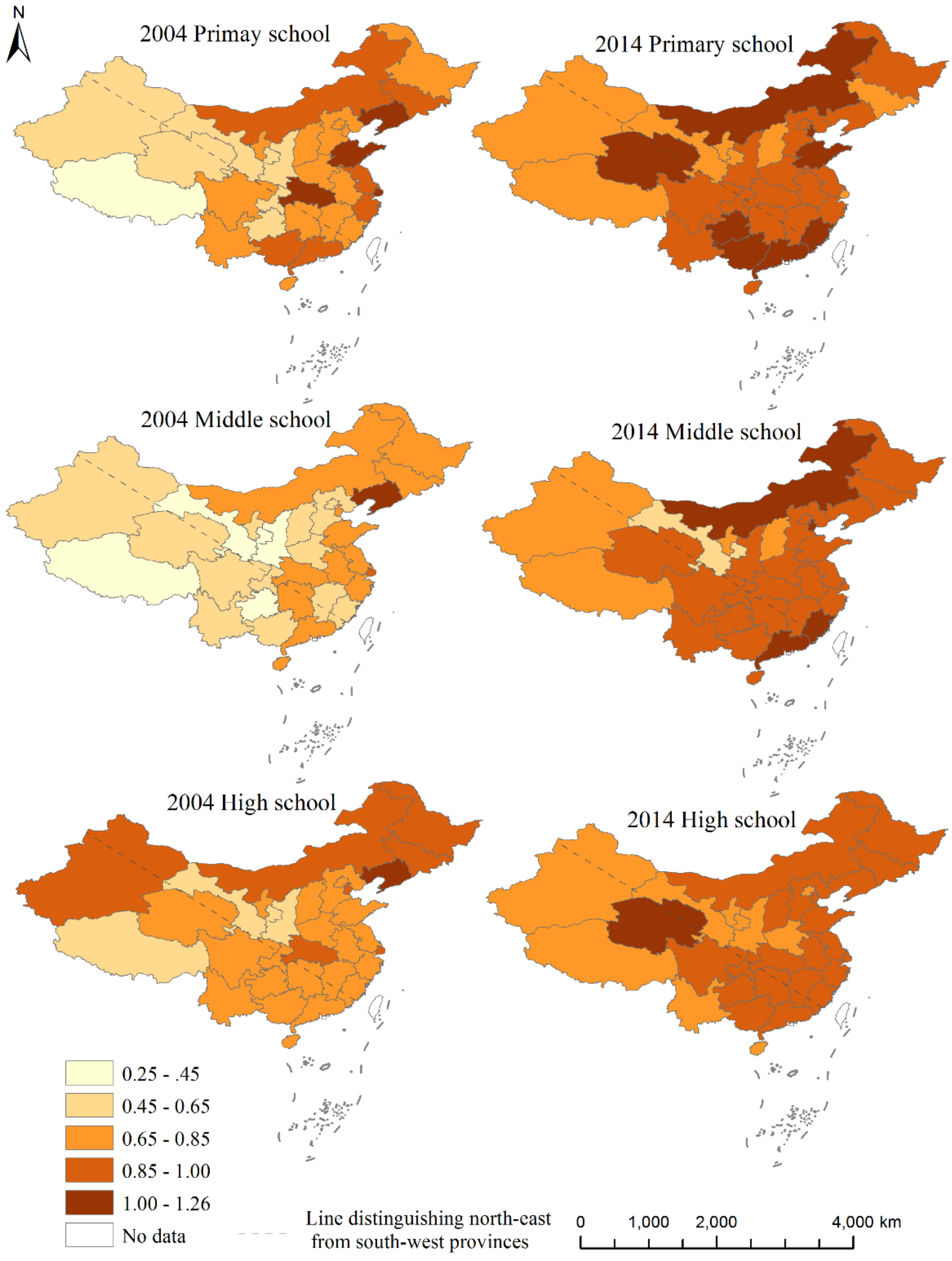


Figure 5.12 Rural-urban teachers with higher professional rank ratios (RUPR) by province, 2004 and 2014

Source: Computed from China Educational Statistical Yearbook of 2004 and 2014

5.4 Discussions

Owing to the policies for promoting education development in rural areas, the educational attainment in rural China overall has improved substantially; however, the rural-urban gap of educational attainment has not been reduced between 2000 and 2010. It is necessary to identify and rethink the failures within the current policies to realise the targets of NMLERD. Related discussions are presented in the following sections.

5.4.1 More resources should be allocated to rural areas

It is apparent that the overall PPEE and teachers' quality in rural areas have increased significantly and the rural-urban gaps of PPEE and teachers' quality have narrowed dramatically between 2004 and 2014. However, the overall educational provision in terms of PPEE, teachers' qualification and teachers' rank in rural areas is still slighter lower than that of urban areas. According to the analysis framework proposed in Chapter 3 (Section 3.2.1), education in each area is not only influenced by the amount of education resources (education provision) at their disposal but also will depend on the conversion factors, that is, their ability to convert educational resources into education attainment. As the educational foundation (attainment) in rural areas was lower in 2004, so the conversion ability of rural areas will be lower than urban areas even if the same amount of educational resources is provided. Accordingly, so as to reduce the rural-urban educational gap and allow people in rural and urban areas to enjoy the same chance of success in education, more resources should be allocated to rural areas compared to urban areas. If the Government continues to offer the same or a similar amount of resources to rural areas, the existing rural-urban gap cannot be narrowed; it can only prevent the gap between rural-urban areas becoming larger. But the provision (existing teachers and facilities) in rural areas was still lower than that of urban areas in 2014. This is why the overall rural-urban (attainment) educational gap has not been reduced in 2010. Moreover, given the RUQT (Figure 5.10) and RUPR (Figure 5.12), we find that there is still a long way to go to improve teaching quality in less developed rural areas. More measures should be conducted to improve the teachers' teaching skills and the teaching quality of existing teachers.

5.4.2 The spatial variations of rural-urban educational gap

However, the above analysis only focuses on the overall or national average rural-urban gap. Actually, the rural-urban educational gap within each provincial level unit varies spatially. Although the overall PPEE for rural areas is still slightly lower than that for urban areas, the PPEE in rural areas of many provincial level units is already higher than that of urban areas in 2014 (Figure 5.6). Most of the west or north-west provincial units have higher PPEE in rural areas for primary school education, while most of the areas to the north-east of the dotted line (educational advantaged areas) have higher PPEE in rural areas for middle school education. The higher educational investment in these areas is beneficial for them to narrow the rural-urban gap within these areas. In general, for the north part of China, most provincial level units have higher PPEE for primary and middle school education, however, the south-west undeveloped areas have high PPEE for primary school education but lower PPEE for middle school education in 2014.

In terms of educational attainment, the spatial patterns of rural education, urban education and the rural-urban education gap are all consistent with the pattern identified by the dotted line in Figure 4.2. North-east areas are advantaged areas for urban education and rural education, and have a smaller rural-urban educational gap (Figure 5.4) and a higher proportion of the urban population (Figure 5.1). This indicates that the education development of provincial level units seems still closely related to their own economic development, in spite of increasing efforts of Central Government.

In addition, the unreasonable area partitioning for policy implementation has decreased policy effectiveness. For example, Figure 5.4 indicates that the rural-urban gaps for the south-west areas which are below the dotted line are very considerable. These patterns are different from the Government's area partitioning. As discussed in Section 4.5.1, it unreasonable to implement policies and allocate resources simply based on dividing China into three regions; the specific development level of each area should be evaluated and the different needs of each area should also be identified. In addition, although the rural-urban PPEE gap in western regions for primary school education was already small and most western regions already have higher PPEE in rural areas in 2014 (Figure 5.6b), this gap was still considerable for middle school education in the south-west regions (Figure 5.6d). It is apparent that the development of different education stages also varies and therefore funding support should also take the specific development level of each education stage into consideration. More efforts by Central Government should be taken to promote educational development in the south-west less developed areas, especially for middle school as well as high school education.

The results of the comparison of PPEEs, teacher-pupil ratios and teacher qualifications also reveal that the spatial variation in the rural-urban gaps for middle and high school education are much more significant than for primary school education. In addition, the lower teacher-pupil ratio in rural areas for high schools indicates that there is still a lack of enough financial input for rural high school education. Most of related policies only concern the development of compulsory education in rural areas; however, in order to improve the overall education in rural areas, especially that for acquiring high quality labour, Central Government should put more effort into improving high school education in rural areas as well. It is apparent that high quality labour is still rare in rural areas (Figure 5.4).

5.4.3 The negative effects of School Consolidation

Despite the School Consolidation policy playing a positive role in enhancing administrative efficiency, saving costs and improving education quality, it has led to very substantial negative effects for rural education. Because of the adoption of the policy, the previous policy of ‘At least one primary school in each village, and one middle schools in each town or township’ was changed into one primary boarding school in each township, and secondary schools in each county (Rao and Ye, 2016). Some pupils were compelled to enrol in schools located in counties or cities, leading to an enlarged gap between rich and poor (Rao and Ye, 2016). It is undeniable that the School Consolidation policy has further encouraged the centralisation of basic education in urban areas and put the impoverished villages at a greater disadvantage. It is noteworthy that AYS in Tibet decreased from 4.03 to 3.75 from 2009 to 2014

(Chapter 4, Figure 4.5) and Tibet was the only provincial level unit that had experienced a decline. The AYS for Tibet had increased from 3.30 to 4.03 from 2004 to 2009 owing to more investment in education by Central Government which was specified by the NR. However, the rural primary schools in Tibet underwent problematic ‘School Consolidation’ from 2009 to 2014 and the number of primary schools in rural areas declined from 754 to 657 while the number of primary schools in urban areas increased (Figure 5.13). This consolidation did not consider the already long travel distance for pupils in rural areas and led to higher dropout rates in rural Tibet, thus counteracted the positive effects of more financial investments from Central Government.

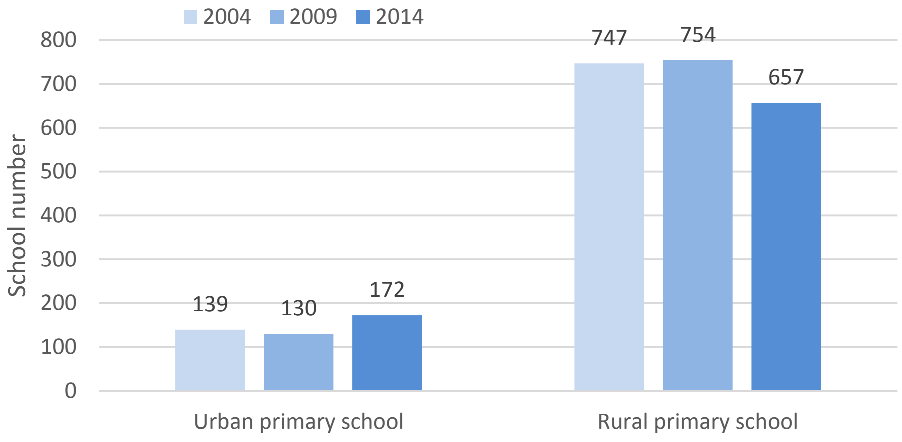


Figure 5.13 The comparison of school number of urban and rural primary schools in Tibet in 2004, 2009 and 2014

Source: China Educational Statistical Yearbook, 2004, 2009 and 2014

School Consolidation is probably one of the reasons why the rural-urban attainment gap had not narrowed from 2000 to 2010. Moreover, the return of rural youths to rural areas has been prevented, as the current education system is focused on cultivating talent for industrialisation and urbanisation in urban areas. The focus on urban-oriented education has channelled rural pupils away from rural communities and their main motivation to receive education has been to leave rural areas and agriculture practices to become urban residents. Rural education should be given high priority as a means of cultivating talent for the development of rural areas and the revitalisation of rural culture (Rao and Ye, 2016).

School Consolidation has also caused a good deal of wasted capital. Although Central Government increased educational public financial investment to reduce the disparity between the cities and rural countryside through sharing transfer payments and projects, a large part of these funds were used to adjust school layouts and merge schools (Rao and Ye, 2016). Furthermore, due to School Consolidation, the accessibility of education for rural pupils was considerably reduced. Additional costs of accommodation and transportation, caregiving, and mental burdens (children's mental health problems) were imposed on rural residents, and thus the economically disadvantaged groups in rural areas were even more vulnerable to these changes (Rao and Ye, 2016). From 2004, the Central Government had invested £1 billion to construct and extend boarding middle schools for rural pupils. However, the boarding schools have separated pupils from their parents and the local community, which has exacerbated disadvantage in rural areas and more pupils from poverty-stricken families may have dropped out of education due to the growing inconvenience (Wang and Zhao, 2012; Mei *et al.*, 2015). Simultaneously, unreasonable consolidation led to a huge waste of existing education resources in some villages whereas certain merged schools were overburdened. Some schools became increasingly large and super-sized classes (>100 pupils) appeared so that teaching quality was extremely hard to guarantee (Mei *et al.*, 2015).

The value of the 'urban oriented and urban priority' policy is the fundamental reason for the failure of the School Consolidation approach and has taken its toll on the rural areas. Although the State Council suspended the School Consolidation policy in 2012, it is extremely hard to mitigate the negative impact of current school layouts on rural areas. Small rural schools play a very important role in delivering education in rural areas in China; therefore they should be retained and equipped with appropriate facilities to fulfil their roles (Zhao and Parolin, 2014). In addition, local rural residents should have more chance to participate in the decision process involving School Consolidation.

5.4.4 The discrimination for the rural migrants' children

In 2014, the overall teacher-pupil ratios in rural areas for primary school education and middle school education are already higher than those of urban areas. This is probably because of more education investment and a higher payment level in rural areas. But, it is more likely to be because more children have moved to urban areas to receive education under the process of urbanisation.

As discussed in the early section, it is not always easy for migrants' children in urban areas to acquire local hukou as migrants' children. Therefore, they have to go back to their homes to attend entrance examinations for university (and in some megacities, such as Beijing, they are even not allowed to attend the entrance examination for normal high school). These migrant children with non-local hukou tend to go back to their hometown to receive a high school education. It is perhaps the reason why the teacher-pupil ratios for high schools in rural areas in most provincial level units are different from those of primary and middle schools, which are slightly lower than those of urban areas, except for some northern provinces.

The discussions in the previous chapter show that the migrants' children are suffering from the strict limitation of receiving education locally and more than half of them become left behind in rural areas. Even though they can receive compulsory education in urban areas, they still have to go back to hometown to receive high school education. Due to different areas having different curriculums and different requirements, it is apparent that the rural migrants' children are restricted by the education policy based on hukou and cannot enjoy complete education.

5.5 Conclusions

This chapter has examined the educational inequalities between rural and urban areas and evaluated the policies aimed at reducing rural-urban disparities. The rural-urban educational gap reflects the historical rural-urban socio-economic development differences, partly due to urban and efficiency priority strategies (Mei *et al.*, 2015). Education in rural areas has experienced tremendous development due to the policies aimed at reducing inequalities between rural and urban areas. The PPEE, teacher-pupil ratio and teacher qualification and rank gaps between rural and urban areas have been reduced between 2004 and 2014. However, the overall rural-urban attainment gap has not been narrowed from 2000 to 2010. Given the existing gap between rural-urban areas in terms of education environment, facilities and teaching quality, even if the education investment gaps between rural and urban areas are successfully narrowed, there will still be a rural-urban education gap. In a sense, only more investment in the education of rural areas can de facto eliminate the rural-urban gap.

Although the overall rural-urban educational gap has not been reduced, there are significant spatial variations in the rural-urban educational gap in terms of PPEE. The higher PPEE in rural areas in some provincial level units has probably narrowed the rural-urban educational gap in these areas and vice versa. Therefore, the allocation of financial resources should be based on the specific development level of each area and the stage of education that each area has reached. Moreover, some of the closed rural schools or teaching posts should be returned to service to reduce the negative effects caused by the School Consolidation policy. The policymakers should begin to pay more attention to high school education in rural areas. Furthermore, rural education should cultivate talent for the development of rural areas and the concept of urban priory should be transformed in education. Thus, although some success has been achieved by the previous policies that have been implemented, more adjustments are needed to realise education equity between rural and urban areas in China.

Due to the rapid urbanisation process and the large amount of in-migrants in China, the education of migrants' children is one of the main concerns of both regional and rural-urban education. The next chapter will discuss the socio-spatial education inequalities within urban areas, part of which are the inequalities of migrants' children.

Chapter 6
An Education-related Geodemographic Classification of Sub-districts in Central Beijing and Its Implications for Access to Compulsory Education

6.1 Introduction

Given the dramatic regional inequalities and rural-urban education disparities revealed in the last two chapters, this chapter investigates education inequalities at the local scale. As discussed in the introduction (Chapter 1) and the literature review (Chapter 2), educational research at small area scale is hindered by limited available data. There is a lack of research on education inequality within urban areas and its spatial diversity in China. Thus, this chapter will make full use of the available data in China to explore the potential influence of 'where they live' (socio-spatial differentiation) on pupils' access to education resources under the current criteria associated with the nearby enrolment policy. 'Where they live' has two meanings; that is, the geographical location and the socio-economic and demographic composition of pupils' residential areas. In addition, the socio-spatial differentiation in Chinese cities is likely to be determined by multiple factors rather than a single factor, including hukou status, income, occupation, etc. Accordingly, a geodemographic classification using the data available from the most recent Chinese census provides a multi-dimensional analysis framework to identify spatial variations in population types in Beijing and, in combination with survey data, allows us to better understand the inequality in access to compulsory education that exists in the city.

This chapter is organised as follows: education inequality within urban areas and the Chinese 'nearby enrolment' policy are discussed in Section 6.2; the process of constructing the area classification for sub-districts of Central Beijing is elaborated on in Section 6.3; the results of inequalities in the demand for and access to education, partly due to the failure of the Governments' nearby enrolment policy, are discussed in Section 6.4, together with the corresponding 'action' measures. This chapter has been published as a paper (Xiang *et al.*, 2018).

6.2 Hukou system and 'nearby enrolment' policy

As mentioned in Chapter 1, the urban population of China exceeded its rural population in 2011 and the rural population is set to decline by half by 2050. This rapid urbanisation means there will be increasingly more migrants to the big cities from rural areas whose children require compulsory education (Li and Placier, 2015). Within urban areas in China, there are substantial variations in school enrolment, access to educational facilities and academic or vocational achievement predominantly between the children of local residents and in-migrants. Studies of these disparities have become prominent in the academic literature (Fu and Ren, 2010; Zhang *et al.*, 2015). These in-migrants are often disadvantaged in terms of education because of their non-local hukou registration status (Liu and Jacob, 2013).

As discussed in Chapter 2, the hukou system was established in China as a means to control population movement; everyone is registered with a community which authorises certain rights. The local urban citizens are granted priority access to urban services and facilities, while migrants with a non-local hukou normally have limited access to the same services and facilities, including schools (Zhao and Howden-Chapman, 2010). Consequently, they are discriminated against by policies based on the hukou system, such as the 'nearby enrolment' policy for compulsory education[4]. The Compulsory Education Law of the People's Republic of China (PRC) stipulates that local governments should ensure that school-age children enrol in schools near the places where their residence is registered (Standing Committee of the National People's Congress, 2006c) except for boarding schools and private schools (Ministry of Education of PRC, 1992).

Local governments have put forward different measures for implementing this so-called 'nearby enrolment' policy. The school district is one of the most important units associated with this policy and is based on the household registration system. Its boundaries are usually consistent with sub-

[4] As indicated in Chapter 2, compulsory education in China normally includes six years of primary education and three years of junior secondary education (middle school), taking pupils from the age of six to the age of 15.

district administrative boundaries and it is specified as a legal spatial range for enrolment containing several primary schools and middle schools (Bi and Zhang, 2016). In Beijing, primary schools are supposed to enrol children who live within or whose hukou is within their ‘attendance zones’ (Zheng *et al.*, 2015). Thus, in China, there is a two-tier system for primary education: school attendance zones are located within each school district (Bi and Zhang, 2016) (Figure 6.1). The Local Education Authority (LEA) will specify every primary school attendance zone[5] by naming the exact residential buildings or building complexes. The middle school enrolment range is normally within the school district.

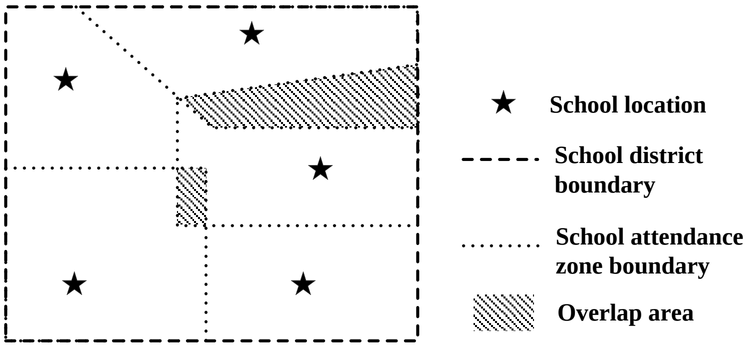


Figure 6.1 The relationship between attendance zones and a school district for primary education

It is recognised that enrolment in the UK and the Netherlands, where there is greater freedom of parental choice, tends to lead to longer commuting distances (Hamnett and Butler, 2013). However, the school enrolment system in China, which is strictly constrained by the location of hukou or property, is struggling to implement ‘nearby enrolment’. For example, the public funding for education in China is from the provincial or municipal general fund rather than from local property or council taxes as in the USA (Zheng *et al.*, 2015). Thus, in some highly populated Chinese cities like Beijing, only property owners or people with local hukou are entitled to enrol in the nearby public school (Zheng *et al.*, 2015). The children of families who rent their homes are not guaranteed enrolment under the nearby enrolment policy. Most in-

[5] It may be the case that there is overlap between attendance zones of neighbouring schools. For example, pupils living in the same building but on different floors may be assigned by the LEA to different primary schools.

migrants who do not have local hukou or do not own property are disadvantaged by this ‘nearby enrolment’ policy and often their children have to commute long distances to school.

However, families with sufficient economic capital are not always restricted by nearby enrolment policy. Although school places should be allocated following the principle of nearby enrolment and school choice is officially prohibited by governments at all levels, in reality, there are numerous ways for high income groups to realise school choice if they are not satisfied with their nearby schools. Firstly, as property ownership also can be used as the proof for nearby enrolment, buying a house which is within the catchment area of a high performance school is most frequently used means of acquiring a school place in the preferred school. This phenomenon leads to sky-high housing prices in the catchment areas of oversubscribed schools (Wu *et al.*, 2016). Secondly, some pupils with a high income background can secure school places in a certain high performance school through paying for an extra very high ‘school choice fee’. This is due to insufficient investment from government, so these schools have to use their educational resources to collect additional income to make up for the shortage of funds (Wu, 2013). Moreover, the students having special talents (e.g. with high level prizes or certificates) are also able to enrol in high performance schools. However, the high cost of hiring a tutor and taking extracurricular classes can only be afforded by high income families. As a consequence, these school choices have exacerbated the educational disparities that already exist in China.

Combined with the recent termination of the ‘one child’ policy (Bi and Zhang, 2016), many cities, especially megacities such as Beijing, are facing massive challenges in the provision of education services for the children of in-migrants (Zhang, 2011) as well as those of deprived native residents (Bi and Zhang, 2016). Since education resources are inadequate, local governments need to ensure that additional resources are directed to those in greatest need. A geodemographic classification will be constructed in this research to identify the disadvantaged groups and their distributions and to detect potential education inequalities in access to compulsory education and its spatial pattern.

6.3 A geodemographic classification of Central Beijing's sub-districts

6.3.1 Input data

Variables from the 2010 Census have been used to create a sub-district classification for Central Beijing. Based on related literature (Talen, 2001; Harris *et al.*, 2007; Williams and Wang, 2014) and after initial correlation analysis, 33 census variables have been selected to construct the geodemographic classification (Table 6.1). Although some of the chosen variables are highly correlated, they were retained as their potential correlations can re-enforce important dimensions in the classification (Singleton and Longley, 2009; Singleton, 2010) and add descriptive and predictive power (Gale *et al.*, 2017). All variables were measured in percentages to normalise the data across sub-districts and avoid the influence of unequal populations. Each variable has a direct or indirect influence on education (as indicated in Table 6.1) and is associated with one of four domains: demographic structure, household composition, housing characteristics and socio-economic traits. As the domains of household composition and socio-economic traits are likely to be familiar to readers, only the demographic structure and housing characteristics, which contain unique variables in China, are explained in subsequent sections.

Table 6.1 Selected census variables and their influence on education

Variable number	Domain	Variable description	Influence on education
1	Demographic structure	Persons with hukou within a sub-district (%)	Enrolment
2		Persons with hukou outside a sub-district but within a district (%)	
3		Persons with hukou outside a district but within a municipality (%)	
4		Persons with hukou outside a municipality (%)	
5		Non-han nationality*	Culture difference
6	Household composition	Child dependency ratio[6]	Family burden and education requirement
7		Divorce rate	Psychological and economic
8	Housing characteristics	Households in cheap rent housing (%)	Dwelling enrolment and family socio-economic status
9		Households in other rented housing (%)	
10		Households who own self-built housing (%)	
11		Households who own commercial housing (%)	
12		Households who own second hand housing (%)	
13		Households who own economically affordable housing (%)	
14		Households who own purchased public housing (%)	
15		Households whose residence is without tap water (%)	Residence condition of children
16		Per capita housing area is equal or less than 16 square metres (%)	
17		Per capita housing area is from 17 to 39 square metres (%)	
18		Per capita housing area is equal or more than 40 square metres (%)	
19		Households whose rent is less than 200 yuan (%)	Household income
20		Households whose rent is 200 to 500 yuan (%)	
21		Households whose rent is 500 to 1000 yuan (%)	
22		Households whose rent is 1000 to 2000 yuan (%)	
23		Households whose rent is more than 2000 yuan (%)	
24	Socio-economic traits	Persons aged over 6 with lower qualification (%) **	Educational level of parents
25		Persons aged over 6 with middle qualification (%) ***	
26		Persons aged over 6 with higher qualification (%) ****	
27		Employed persons who are managers or directors of government offices, party organisations, enterprises or institutional organisations (%)	Parents' occupation and occupation based social class grading
28		Employed persons with professional occupations (%)	
29		Employed persons with administrative or secretarial occupations (%)	
30		Employed persons with commercial or service occupations (%)	
31		Employed persons with agricultural occupation (%)	
32		Employed persons in manufacturing or transportation (%)	
33		Unemployment rate (%)	

* Han ethnic group is the dominant group in China, which comprises 91.6% of the country's population

in 2010 census, correspondingly, non-han nationality refers to minority groups.

** No academic or professional qualifications, primary school qualification, and middle school qualification.

*** High school certificate, which also includes vocational high school, secondary specialised school, technician training school; and junior college certificate.

**** Bachelor degree and postgraduate degree certificates.

6.3.1.1 Demographic structure

The first set of variables measuring demographic structure includes proportions with ethnic and hukou registration status. The census data are collected at an individual's current usual residence. A different hukou status occurs when the current usual residence of an inhabitant is not consistent with his or her registration location. There are four categories of residential status: hukou within a sub-district; hukou outside a sub-district but within a district; hukou outside a district but within a municipality; and hukou outside a municipality. Figure 6.2 illustrates the influence of different types of hukou on a child's school enrolment based on the current commonly applied enrolment policy in Beijing. As mentioned previously, it is normal that the school district boundaries coincide with sub-district boundaries and each school district consists of one or more sub-districts. In the following illustration, it has been assumed that each sub-district is equal to a school district, so the large round circles in each component of Figure 6.2 not only represent a sub-district boundary but also indicate the range of each school district. The rounded rectangles and the rectangles outside the circles represent district and municipality boundaries, respectively. The small black dot shows a resident's current place of residence. The unshaded hexagons refer to schools performing badly, whereas the shaded hexagons represent schools performing well. The lines connecting the black dots and hexagons represent the preferred school choice of a pupil under each status of hukou, as explained below.

[6] The child dependency ratio is calculated by using the population of children aged 0-14 divided by the population aged 15-65.

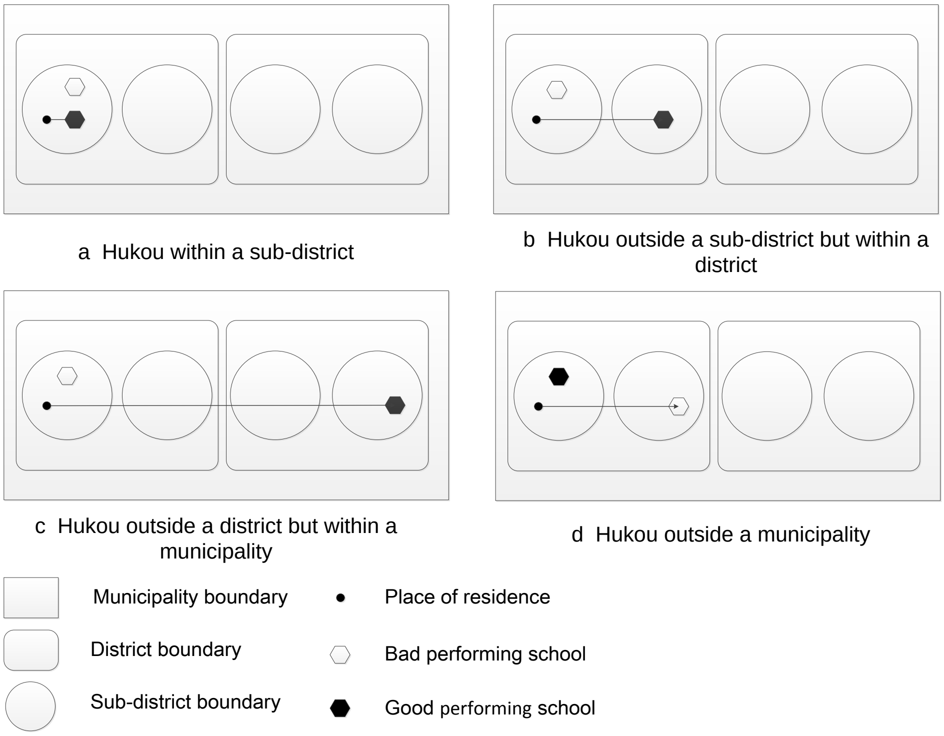


Figure 6.2 Schematic diagram of school enrolment with different hukou status

Hukou within a sub-district (Figure 6.2a): In this instance, the pupil's current residence is consistent with the registration sub-district. According to the general admission policy within Central Beijing, children in this category are given priority to enrol in a nearby school and thus the commuting distance for this group may not be very far if they choose to enrol in nearby schools. However, whether there is an advantage in terms of education quality for them depends on the quality of the nearby school.

Hukou outside a sub-district but within a district (Figure 6.2b): In this case, a pupil lives in one sub-district but has registration in another sub-district but within the same district of residence. Usually, the housing ownership certificate and the household registration record are both used as proof for nearby enrolment (Wu, 2013). Thus, if the parents are homeowners of the current residence, they will have the right to enrol in a school either near their current residence or their place of registration. However, if there is a better school near their place of registration (Figure 6.2b), the child is likely to choose this school and commute a longer distance. This is one of the important reasons that explain the long commuting distances made by children with a particular status.

Hukou outside a district but within a municipality (Figure 6.2c): In this situation, a pupil's place of registration is in another district but within Beijing. Similar to the above description, if there is a school near the pupil's place of registration which is better (Figure 6.2c), the result will probably be an even longer commute to school compared to the pupil shown in Figure 6.2b, as in this case, the commute to school will cross a district boundary.

Hukou outside a municipality (Figure 6.2d): If a pupil who does not have hukou within Beijing wants to enrol in a public school, the family needs to provide many documents, such as parents' job contracts and temporary residence permits to meet the requirements of each district. Many migrant children have been deprived of formal education because of these rigorous requirements. Even if they can provide the correct documents, schools will meet the demand from local people first, so children with the non-local hukou status are not guaranteed a school place through the 'nearby enrolment' policy; they can only enrol in schools where there are places left. They do not have the chance to compete with local residents for places in oversubscribed 'key' schools; the schools where they will be able to enrol are not nearby and are most likely to be of lower education quality.

Thus, a pupil's hukou status is a crucial influence on his or her accessibility to schools and, consequently, on the education quality he or she receives. The influence of the other domains will be briefly introduced.

6.3.1.2 Household composition

The child dependency ratio is defined as the ratio of the number of juveniles to the working age population. It measures the working-age people's average burden for supporting juveniles. In addition, it can partly reflect the family burden and education requirement of certain areas.

6.3.1.3 Housing characteristics

There are four aspects used in this research to measure housing characteristics: housing type, per capita living space, facilities and rent. The living conditions and family socio-economic status of children will be reflected by these variables; they will significantly affect a child's achievements as shown by previous research (Logan *et al.,* 2012; Williams and Wang, 2014).

Seven housing types were included in the set of selected variables (Table 6.1). In addition to common residence categories, such as commercial housing, there are also some unique housing categories in China, such as purchased public housing and economically affordable housing (EAH) (jingjishiyongfang in Chinese). Purchased public housing is also called privatised public housing (Yang *et al.*, 2014) and includes houses previously owned by state-owned enterprises (SoEs) that have been sold to their employees at considerably reduced prices. Economically affordable housing (EAH) is aimed at low- and medium-income households, while cheap rent housing (CRH) targets the bottom low-income households with a local hukou. The price of EAH is lower than that of similar quality commercial housing, which only covers construction costs and with limited profits and makes the housing affordable for low- and middle-wage households. This is subsidised by local and Central Government. Eligible candidates for EAH in Beijing require local resident registration and must meet the demarcation line standards for total income and total wealth, e.g. their current living floor space (including self-owned and rental housing) per head must be below 10 m^2.

Since CRH targets the bottom low-income households with local registration in Beijing, the number of eligible households is rather limited. In 2009 and 2010, only 4,000 CRH units were built in Beijing. In addition, there is a small amount of 'self-built' housing which refers to housing units financed by households and usually located in the collectively owned land in rural areas on the peripheries of cities (Yang *et al.*, 2014; Fang *et al.*, 2015; Fu *et al.*, 2015). Moreover, the Chinese census requires residents to specify the size of their accommodation (square metres) (Harris *et al.*, 2005). The per capita housing area reflects the degree of residential overcrowding.

6.3.1.4 Socio-economic traits

The highest level of qualifications of residents (six and over) is divided into seven categories by the State Statistics Bureau of China. These seven grades are aggregated into three levels:

Lower qualifications: No academic or professional qualifications, primary school qualification, and middle school qualification.

Middle qualifications: High school certificate, which also includes vocational high school, secondary specialised school, technician training school; and junior college certificate

Higher qualifications: Bachelor degree and postgraduate degree certificates

It is worth noting that the qualifications of residents aged six and over will be influenced by the high proportion of school age (primary school or middle school) children in an area. The qualification of residents aged 16 and over would be more credible, however, there are no such data available in census.

6.3.2 Pre-processing the data

6.3.2.1 Transformation

Before clustering can be undertaken, the data have to be pre-processed. An analysis of skewness indicated that some variables, such as ‘Non-han nationality (%)’ and ‘Employed persons with agricultural occupation (%)’, were significantly positively skewed. However, clustering procedures work better on normally distributed data, while highly skewed variables would create uneven cluster sizes (Vickers and Rees, 2007). Therefore, the skewed variables were transformed to a log scale to reduce the influence of outliers at the high end of the value scale (Vickers and Rees, 2007). Since some variables in the dataset had the value 0, which will cause an error when doing log transformation, 1 was added to each value which needs to be transformed to solve this problem. After transformation, the ends of the data series are squashed and the middle was expanded. There is no very highly skewed variable and all the skewness is at an acceptable scope.

6.3.2.2 Standardisation

Before clustering, variables need to be standardised to ensure all have the same equal importance and their variations can be captured equally. As all clustering methods work on the basis of variables’ resemblance or dissimilarity, it will be problematic if variables have different scales or magnitudes, because variables with bigger values or larger variance will impact more on the clustering process. Thus, standardisation should be used to ensure every variable is equally represented.

Range standardisation has been conducted for all the variables to ensure each has the same importance and their variation can be captured equally (Vickers and Rees, 2007). Compared to other standardisation methods, such as using z-scores, range standardisation is helpful for coping with extreme outliers (Vickers and Rees, 2007). The formula for range standardisation (0-1 normalisation) is defined as follows:

$$R_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \qquad (6.1)$$

where R_i is the standardised value of variable x for zone i, x_i is the initial value of variable x in zone i, x_{min} is the minimum value of variable x, while x_{max} is the maximum value. The value of each variable will range from a minimum value 0 to a maximum value of 1 after this standardisation, ensuring that the similarity or dissimilarity is measured on the same numeric scale.

6.3.3 Clustering the data

After being pre-processed, the data were clustered by applying the k-means clustering algorithm.

6.3.3.1 Clustering method

There are many techniques that can be used to construct a neighbourhood classification, from some relatively traditional clustering algorithms to more sophisticated ones such as the self-organising map (Openshaw, 1994), fuzzy classification (Voas and Williamson, 2001) and decision trees (Quinlan, 1986). The k-means clustering method was selected for this research. It is a commonly used approach in geodemographics (Harris *et al*., 2005). It is a non-hierarchical and highly efficient way of clustering a large volume of data. The iterative allocation-reallocation method (k-means where k is the number of clusters) first creates clusters by imperfectly arranging zones into the final number of predefined groups and reallocating the assignments iteratively to obtain a better fit. This approach tries to minimise variability within clusters and maximise variability between clusters (Vickers and Rees, 2007). The clustering criterion, the total Within-cluster Sum of Squares (WSS), is minimising the objects' Euclidean sum of squared deviations to the cluster mean. This clustering criterion can be defined as:

$$WSS = \sum_{k=1}^{K} \sum_{i \in C_k} \sum_{j=1}^{P} (x_{ij} - \overline{x}_{kj})^2 \qquad (6.2)$$

where x_{ij} is the value of observation i of variable j;

$\overline{x}_{kj}$ is the average of values of variables j in cluster k;

C_k is the collection of members within cluster k;

P is the number of variables for each observation;

and K is the number of clusters.

It can be applied for continuous data variables, so long as an appropriate distance measurement can be defined (Leventhal, 2016). Thus, it is an ideal approach to be used in this research.

In terms of software implementation, when using SPSS, the final cluster solution may be highly influenced by the starting values (the original cluster centre or initial seeds). The clustering method adopted in the 2001 OAC in the UK (Vickers and Rees, 2007) was challenged by Dennett and Stillwell (2011) since this algorithm generates different solutions depending on the order in which the data are sorted. The rigid initial seeds assignment is not helpful to get an optimal global solution. The parameter, 'nstart', in R was used as a solution to this problem running the clustering algorithm multiple times (50 in this instance) with a different set of initial cluster centres each time. Although the optimum global solution is impossible to determine, the cluster scheme created by comparing the results using different initial seeds can get very close to an optimum global solution compared with the use of rigid initial cluster centres.

6.3.3.2 Number of clusters

One feature of this method is that the selected number of clusters must be specified before running a k-means algorithm rather than being derived automatically. As the number of clusters needs to be suitable for the research purpose and useful for further analysis, it should be as close to the perceived ideal as possible (Vickers and Rees, 2007). As there are 130 sub-districts within central Beijing, six may be a target number of clusters for this research to enable ease of visualisation and labelling. Thus, solutions in the range of four to eight clusters were investigated. The R package ‘NbClust’ from Charrad *et al.* (2014) was used to determine the optimum cluster scheme. This package provides an exhaustive list of 30 indices to determine the number of clusters for a given dataset. For each index, NbClust proposes the best cluster scheme, so that the results of all indices can be compared to get the best solution. Using NbClust, 14 out of 30 indices suggested five as the best number of clusters. In addition, two of the top performance indices suggested by Milligan and Cooper (1985), the Duda index and the Beale index, also proposed five clusters as the best solution and the one that was subsequently adopted.

6.3.4 Naming and describing the clusters

After running the k-means clustering and obtaining the results, the clusters were labelled and given descriptions, as shown in Figure 6.3, based on the principal and distinctive variables. The locations of primary and secondary schools are also shown on the map in Figure 6.3 to indicate the school distribution in each cluster. The geodemographic clusters display a pattern of concentric rings, although some parts of these rings are discrete. The cluster labelled ‘Long settled Beijing society’ (cluster 3) is concentrated in the central areas of Beijing, including the Dongcheng and Xicheng districts and areas within the second ring road. Areas in cluster 2, labelled ‘Employees of state-owned enterprises and white collar workers’, are located in the middle of inner Beijing, mainly located between the second and fifth ring roads, although some are located further out. The distribution of sub-districts classified and labelled as ‘High income internal migrants’ (cluster 1) is more widely spread, but these areas are mainly located in the outer area of inner Beijing. Areas in the cluster labelled ‘In-migrants living in deprived suburbs’ (cluster 4) are found outside and around the above three clusters. Cluster 5, whose areas are farthest from the city centre, is called the ‘Urban-rural transition belt’. In addition, the typical neighbourhood of each cluster has been presented in Figure 6.4.

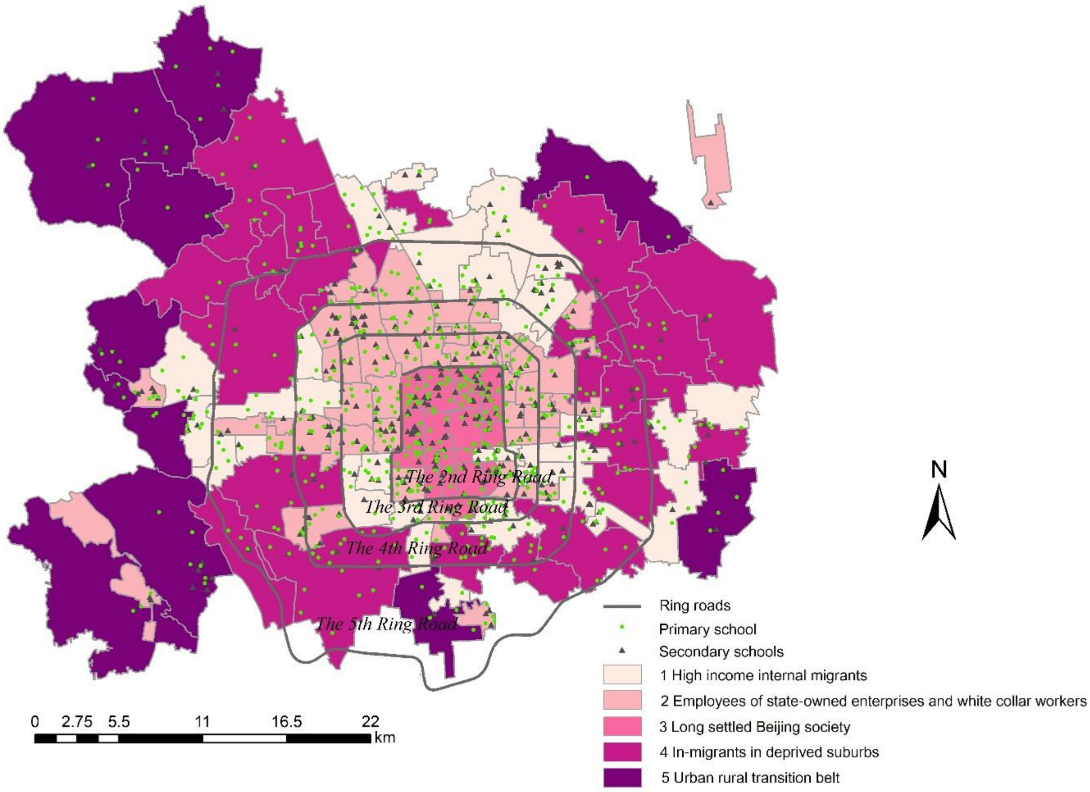


Figure 6.3 Education related geodemographic classification of Central Beijing by sub-district and the distribution of primary schools and secondary schools

Source: Geographical Information Monitoring Cloud Platform

a

b

c

d

e

a Cluster 1 – Luxury villas outside fifth ring road and next to Olympic Sports Center	b Cluster 2 – Work-unit compound in petroleum sector within the third-ring road
c Cluster 3 – Hutong-courtyard neighbourhood in the inner city	d Cluster 4 – Migrants' neighbour in Xibeiwang and beyond the third ring road
e Cluster 5 – Rural residential area in northwest of Haidian district	

Figure 6.4 Typical neighbourhood of each cluster.

Source: Baidu streetscape.

The principal features of each cluster can be summarised and displayed using the radial plots below (Figure 6.5-6.9). Z-scores were used in this research to assess the distinctive variables, because identification of variable effects is facilitated if the value sits above or below the study area mean, 0. The variables in each cluster were sorted by z-score and displayed anticlockwise (highest to lowest). The difference in the value of each variable from the mean (0) is represented by the length, direction and colour of each bar and the range of the y-axis is from -1.5 (the centre of the circle) to 3. The highest and lowest values of variables are immediately distinguishable and reflect the principal features of the cluster in terms of its demographic structure, household composition, housing and socio-economic features.

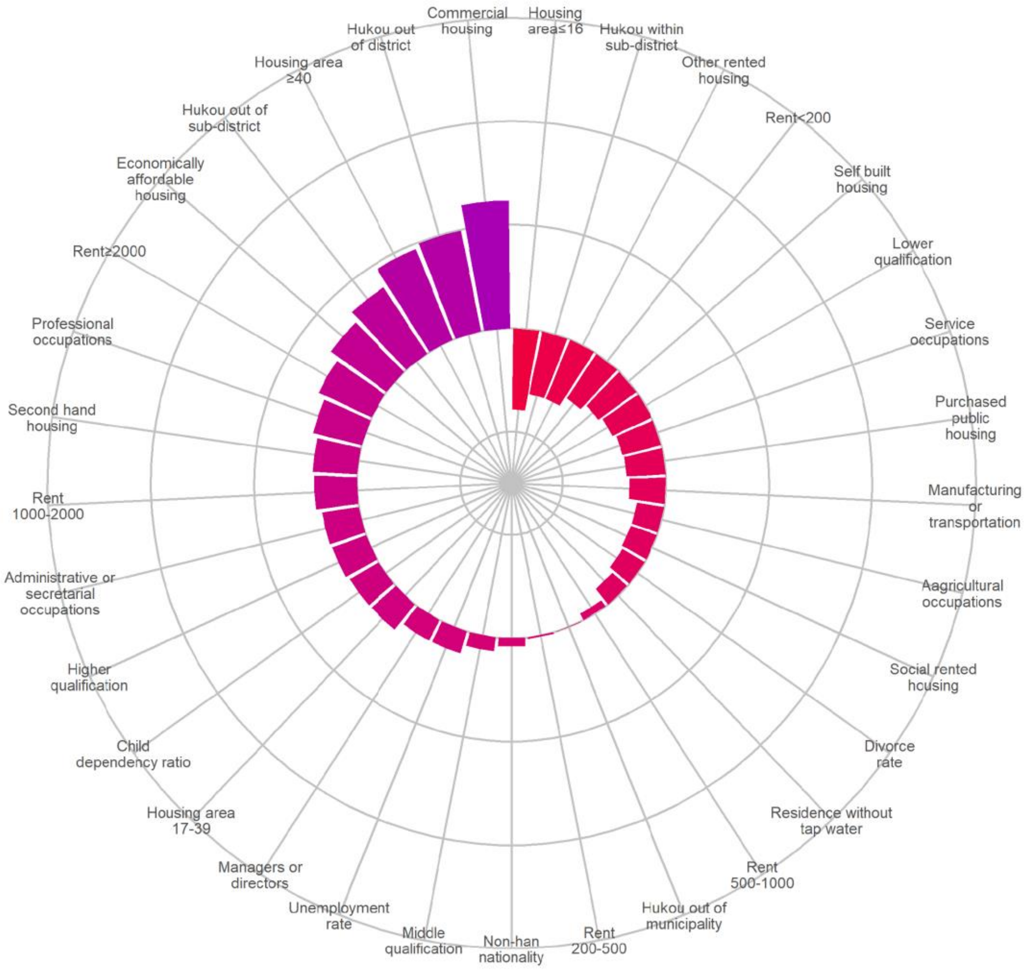


Figure 6.5 Radial plot for cluster 1: High income internal migrants

The distinct characteristics and education related descriptions of each cluster have been described in detail by the following pen portraits. In reality, due to the drawbacks of geodemographics, such as the ecological fallacy (Burrows *et al.*, 2005; Graham, 2005; Leventhal, 2016) and the Modifiable Areal Unit Problem (MAUP) (Openshaw, 1994), within each sub-district, these characteristics will show a degree of variability and applicability.

6.3.4.1 Cluster 1: High income internal migrants

The areas in this cluster have a high percentage of spacious and owner occupied commercial housing (Figure 6.5), which include the most luxury villas (Figure 6.4a). This means their economic status is comparatively high. There is a high proportion of residents with their hukou outside a district but within the municipality, or with their hukou outside the sub-district but within the district. The population in these areas is likely to contain many residents who have migrated within Beijing. In view of Beijing's development process and housing reform policy, it is likely that the hukou of these residents is within inner Beijing, where they previously lived. To avoid the crowded living environment in central areas, and with the encouragement of the commercial housing development in the outer zone of Central Beijing, this sub-group contains people who changed their residence and moved to an outer zone with spacious housing and a more pleasant living environment. However, the schools with a good reputation historically are mainly located in the inner areas of Central Beijing; there are only a few high quality schools located in the newly developed outer areas. As the place of hukou registration of these inhabitants tends to be in inner Beijing, their children still have the opportunity to enrol in a better school in an inner area as discussed in Section 6.3.1.1 (Figure 6.2b and Figure 6.2c). Chinese parents prefer to choose a school with higher teaching quality and commuting distance seems to be less important if their economic status permits (Bi and Zhang, 2016). As their economic status is high, it can be speculated that some of the children living in these areas would commute quite a long distance to enrol in a better school rather than enrol in a nearby school, resulting in an increase in the traffic burden and more congestion.

6.3.4.2 Cluster 2: Employees of state-owned enterprises and white collar workers

More people in this group own purchased public housing and have their hukou in the same sub-district as where they live (Figure 6.6). Their living conditions are moderate (neither crowded nor very spacious) (Figure 6.4b). The high proportion who own purchased public housing indicates that quite a lot of people in these areas are working in state-owned enterprises that are the legacy of the danwei (work unit) system which is now defunct. Prior to the housing reform, employees of state owned companies were provided with welfare housing and comprehensive services including medical supplies and education facilities by the danwei unit, which is a unique organisational form of the socialist system. Since the 1990s, a market based housing reform was implemented to transform housing from a welfare item of the socialist system to a commodity in the post-socialist market (Yang *et al.*, 2014). In 1998, residents were encouraged by the Government to buy commodity housing from the market or buy public housing which was owned by the state, the so-called state-owned enterprises (SoEs), at considerably reduced prices (Yang *et al.*, 2014).

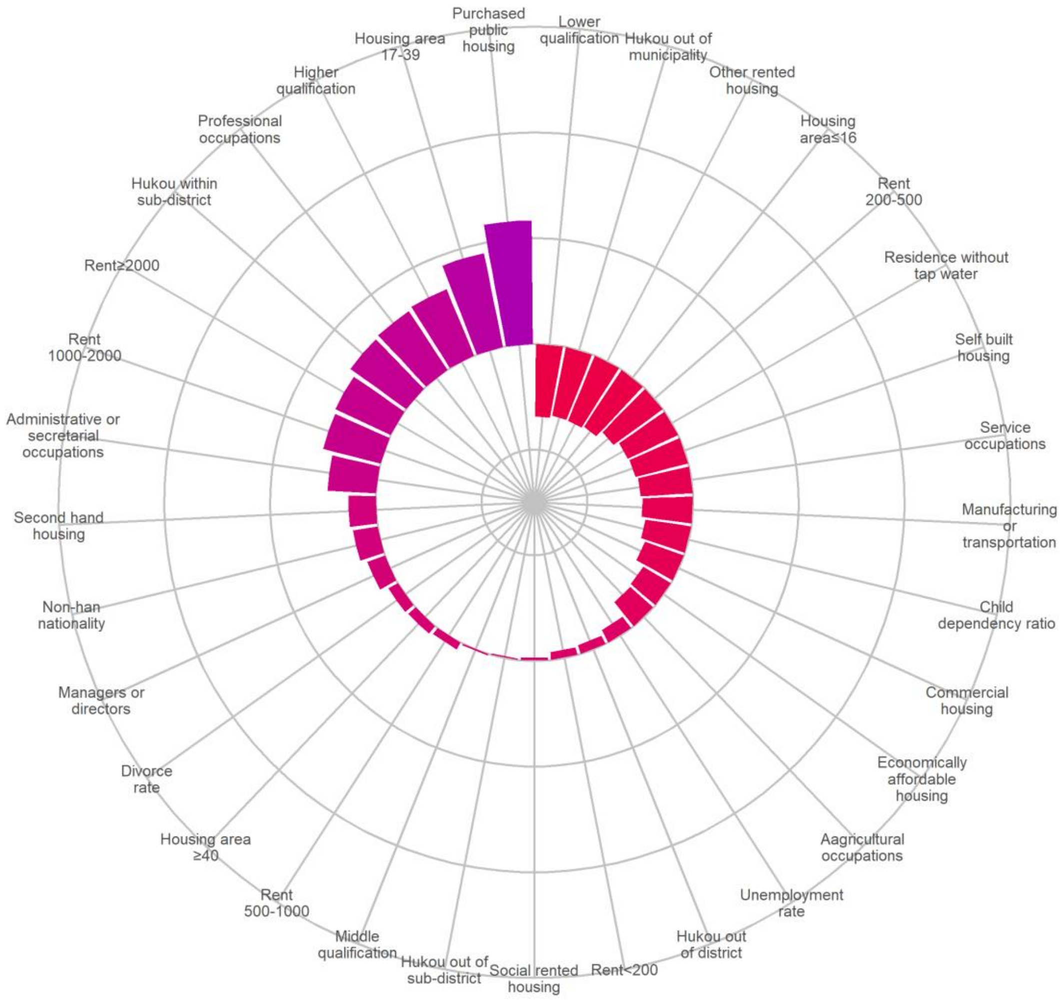


Figure 6.6 Radial plot for cluster 2: Employees of state-owned enterprises and white collar workers

As each danwei usually had its own schools for employees' children, the education provision in these residential areas was normally sufficient (Chen *et al.*, 2015). In addition, in the danwei system, the matched schools were normally located close to or even within employees' residential areas, which is also indicated by the school densities shown in Figure 6.3. Although the danwei system has disappeared, the schools have remained, so education resources for employees of SoEs are relatively adequate, there are several schools with high quality (key schools), and children's commuting distances are not very long.

Other distinctive characteristics of this cluster are its high percentage of the population with higher qualifications and with professional white collar occupations. The best universities, scientific research institutions and high-tech firms in China are concentrated in these areas, including Tsinghua University, Peking University and the technology hub, Zhongguancun (known as 'China's Silicon Valley'). Existing research has demonstrated that parents' education is important for predicting children's achievement (Davis-Kean, 2005). Parents with high qualifications will not only have an influence on what their children achieve but find it easier to get the limited local hukou quota, so they have more chance to enrol their children in a nearby quality school. Thus, children in this cluster are likely to have a significant advantage in receiving education not only because of the shorter commute distance and better education quality but also because their parents' qualifications and occupations are superior.

6.3.4.3 Cluster 3: Long settled Beijing society

Most people living in cluster 3 have their hukou in the same sub-district (Figure 6.7), although more people are living in crowded and less desirable environments. Few residents in these areas are living in spacious accommodation. Areas in this cluster are found inside the second ring road (Figure 6.3), where there is a concentration of Beijing's hutong and quadrangle courtyard areas that have existed for hundreds of years (Figure 6.4c). Beijing's quadrangle dwellings were designed for large single families, but they have been subdivided into smaller quarters and shared by a lot of different families since 1950; thus the living environment of residents in this cluster is densely populated. Most buildings are comparatively low-rise constructions (no more than six storeys) and some of them are falling into disrepair. However, these inner areas are not only occupied by the poor; there are also residents of the affluent class still residing in these areas because of redevelopment that has occurred (Fang *et al.*, 2015). High quality education resources are historically located in these areas (Bi and Zhang, 2016), so some rich people stay there to give their children a better education.

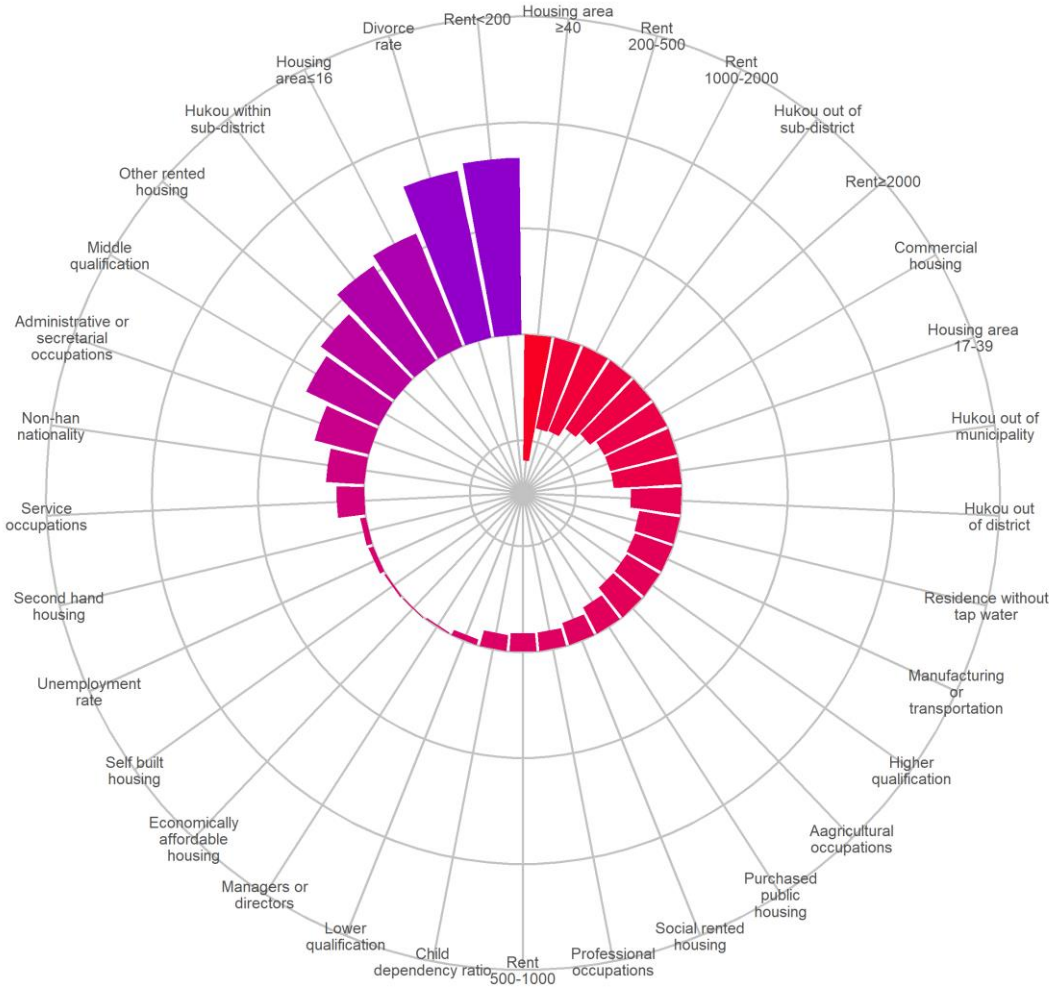


Figure 6.7 Radial plot for cluster 3: Long settled Beijing society

As the proportion of pupils with their hukou within the sub-district in this cluster is high, the commute distance of children is likely to be relatively short. In terms of education quality and distance to school (school density is high), children in this cluster should not be considered as being very deprived; however, the unpleasant living environment and high divorce rate may affect pupils' success.

6.3.4.4 Cluster 4: In-migrants in deprived suburbs

Compared to other clusters, there is a large group of in-migrants who do not have a Beijing hukou, living in low rent accommodation, without the necessary facilities and with low qualifications (Figure 6.8). As only people with a Beijing hukou are eligible to apply for social rented housing, there is a high percentage of residents living in other rented housing paying 200 to 500 yuan (around £20-50) per month, which is a low price for private rental accommodation (Figure 6.4d). Their residential environment is unpleasant, crowded and lacking basic facilities. Due to the strict household registration system, migrants only have access to commodity housing which is expensive. Moreover, they are disadvantaged in the urban housing market by requiring various types of certificates for purchasing housing. According to a survey by Logan (2009), most migrants are living in market rented rather than owner occupied housing, as shown by the census data.

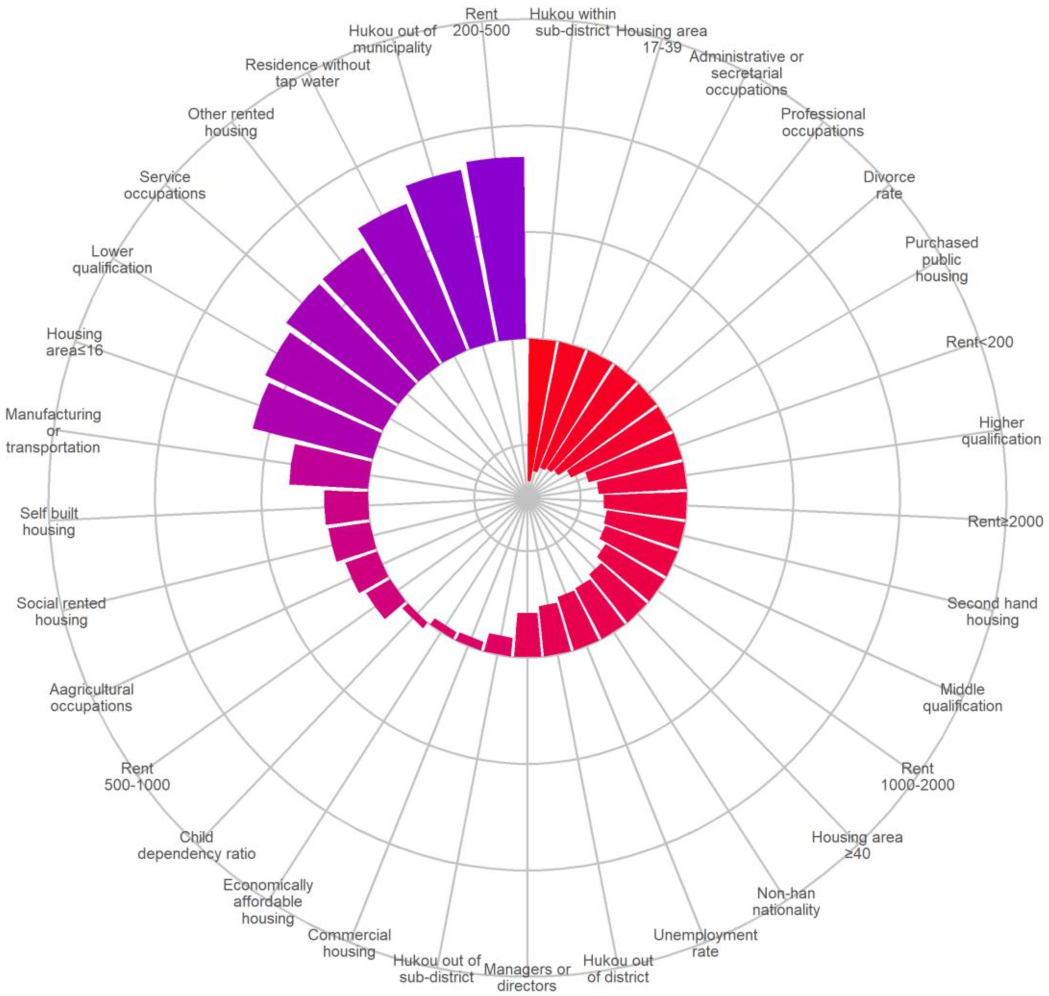


Figure 6.8 Radial plot for cluster 4: In-migrants in deprived suburbs

A high proportion of people in this cluster type are employed in commercial and service occupations and have primary or middle school education. As shown in Figure 6.2d, for those who can provide the required documents, the children of migrant families usually cannot enrol in nearby schools because they lack a Beijing hukou and do not own property in Beijing.

They will be allocated to less popular schools within the range of the district of residence. As the areas of these outer districts are very large, pupils are likely to commute a longer distance to low-performance schools. Therefore, these areas will potentially be the most educationally disadvantaged areas in the city.

6.3.4.5 Cluster 5: Urban-rural transition belt

Areas in this cluster tend to have a large amount of spacious and self-built housing (Figure 6.9 and Figure 6.4e), which is usually located in rural and suburban conjoint areas (Figure 6.3). There are high proportions of residents with agricultural occupations and very low qualifications. Scholars have demonstrated a significant gap in the average years of schooling, enrolment and graduation rates between the urban and rural population in China (Zhang *et al.*, 2015). The unbalanced education provision between rural and urban areas potentially leads to educational disadvantage for residents in these areas. If they have local hukou, they can enrol in a nearby school but this normally has lower education quality compared to inner areas of Beijing. The education support from parents is comparatively weak because of their poor qualifications. In addition, the children dependency ratio is also very high in this cluster, which is likely to be due to a large number of 'left behind' children, whose parents went to urban areas to find jobs.

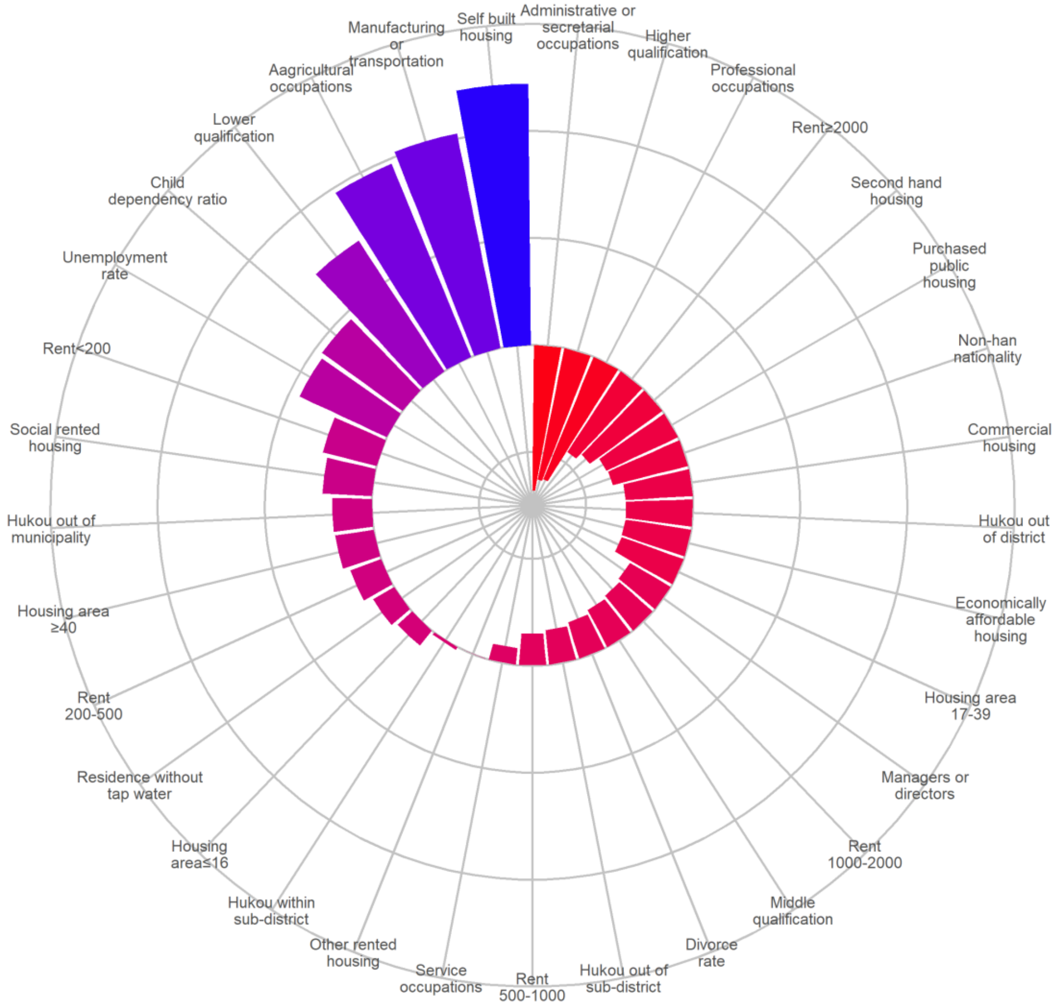


Figure 6.9 Radial plot for cluster 5: Urban-rural transition belt

6.3.5 Application and validation

To evaluate how well this classification performs and validate the above descriptions of potential education inequality, data from 2014 Pupil Travel Survey (PTS) have been used. Most primary schools (96.9%) and secondary schools (96.1%) in Central Beijing have taken part in this survey, and the valid observations account for 62.5% of all primary pupils and 53.5% of all secondary pupils. Information such as the sub-district name of pupils' home and school, and travel time and distance from home to school are included in this survey and can thereby be linked to the clusters described above. The average travel distances of pupils living in each cluster are shown in Figure 6.10. Although differentiation by the educational quality of schools is not possible, the results of the 2014 PTS validate the assumptions about pupils' commuting distance in Section 6.3.4.

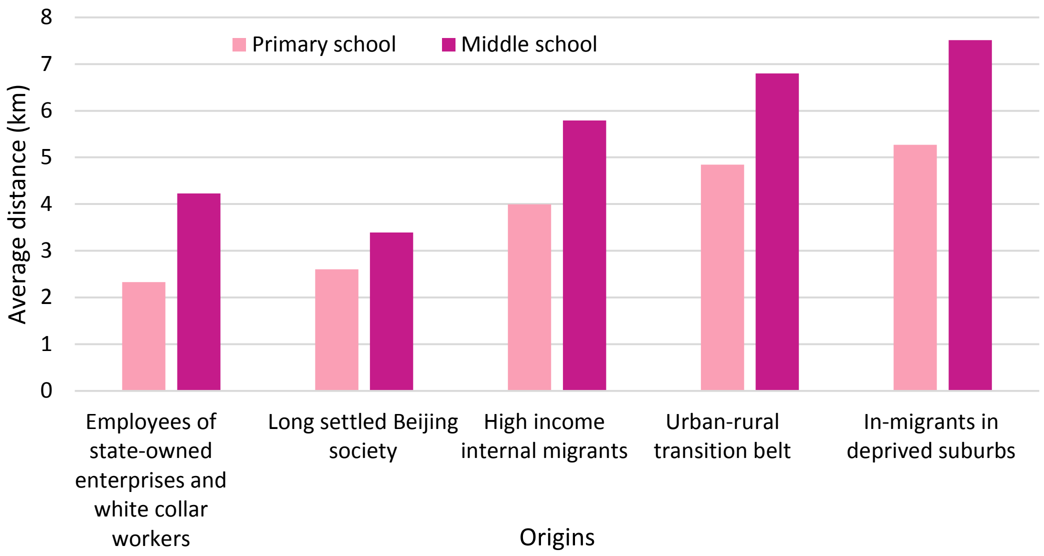


Figure 6.10 Average travel distance to school within each cluster by education stage

The cluster with the shortest average commute distance to primary school is 'Employees of state-owned enterprises and white collar workers', especially at the primary school level. The mean commute distances to primary and middle schools in 'Long settled Beijing society' are very short compared to other clusters, especially middle schools. In addition, 'Urban-rural transition belt' and 'In-migrants in deprived suburbs' are shown to be disadvantaged groups from the perspective of accessibility to education services. According to a specification issued by the former state education commission, the distance between a pupil's residence and school should be within 3km (Ministry of Education of PRC, 2006). However, apart from the first two clusters, the average commute distances for primary schools of the other clusters are all over 3km, indicating that nearby enrolment policy is not working in some places. In addition, the primary school commute distance of 'In-migrants in deprived suburbs' exceeds 5km, which is regarded as too far to commute for primary school pupils (McDonald, 2010).

6.4 Conclusions

This research provides the first geodemographic classification relating to a public sector service in China. It is also the first time that a compulsory education-specific area classification has been used to assess potential inequalities in access to schools. This study offers a valuable data exploration approach that yields new ideas and insights about potential inequalities in education, providing a useful point of reference for areas where educational research is limited by the availability of data for small areas. In terms of substantive findings, potential inequalities in access to schools due to the nearby enrolment policy for different types of area are revealed in Central Beijing. This is associated with housing type, parental occupation, qualification and, in particular, registration status which has mediated the influence of market sorting and maintained the advantage of local households over migrants with similar or even higher incomes (Fang *et al.*, 2015). This area classification provides some crucial education-related implications and is constructive in clarifying the main issues of education development within each type of area revealed by the cluster analysis and in formulating the corresponding 'action' measures proposed below.

Areas with 'In-migrants living in deprived suburbs' are likely to be the most educationally disadvantaged places in the city. Compared to local low-income households in the clusters labelled 'Long settled Beijing society' and 'Urban rural transition belt', those migrants living in areas labelled 'In-migrants living in deprived suburbs' may be far more disadvantaged in terms of education because of their non-local hukou status. The clustering and decentralisation of migrants occur because they do not have access to social rented housing in inner areas and the only cheaper rented accommodations available for them are in the outer areas, some of which are provided by local rural villagers (Wu and Treiman, 2004; Song *et al.*, 2008). Their children normally have to commute long distances to enrol in relatively unpopular schools, and the long-distance commutes have been validated by the analysis of data from the 2014 PTS. Moreover, it is the case that for most migrants, due to not having the required documents such as parents' job contracts, temporary residence permits and social insurance, their children do not have the right to enrol in schools in Beijing at all. They have to leave their children with their parents in rural areas as the so-called 'left-behind' children. Education inequality for these migrant children will not only have an influence on their social mobility, but will also be harmful to the city's human capital accumulation and hinder its long-term economic development (Zhang *et al.*, 2015).

According to the above disadvantages identified by the area classification, local authorities in the problematic sub-districts should undertake specific measures to address these apparent educational disparities for in-migrant children. First, the education of migrant children should be properly integrated into the system of planning for education needs and the education provision for migrant children should be included in the performance evaluation of local government in these areas. Second, the residents with rented properties in these areas should enjoy equal nearby enrolment rights as those with local hukou or property ownership. Moreover, better quality schools, which are normally over-subscribed and therefore unavailable for migrant children, should provide some places for these pupils. Third, the government should develop schools in those areas labelled as 'In-migrants living in deprived suburbs' for these migrant children to meet the demand of the increasingly concentrated migrant population in these areas. As these areas are mostly in outer parts of the city, these measures can solve not only the problems of difficulties in access to schools for in-migrant children and reduce travel distances, but can also facilitate the population dispersal from congested inner Beijing and the formation of in-migrant concentration areas where the environmental conditions are more healthy.

People living in areas labelled 'Urban rural transition belt' are also likely to be deprived in terms of education due to few high performance schools being located in rural areas. Moreover, the lower density development of rural areas means that the commute distance in rural areas tends to be longer than in urban areas. Increasing financial and human capital resources should be allocated to these areas to promote the education quality of schools and increase accessibility by providing more school buses or adding more teaching points. With the population dispersal from inner Beijing, more in-migrants will tend to move further away from the inner city and reside in the rural-urban transition areas. The long-term education provision in these areas should be proactively prepared in the light of this trend. Furthermore, the education and mental health of the 'left behind' children indicated by the high child dependency ratio in these areas should also be given attention.

This research demonstrates that the 'nearby enrolment' policy can be inefficient. This is due to the strict hukou restrictions and the requirement of property ownership during the enrolment process. From this analysis we recognise that migrant children are excluded from consideration by the 'nearby enrolment' policy because of their non-local hukou, but even for local residents, the implementation of the 'nearby enrolment' policy is still distorted because of their hukou status. Children in 'High income internal migrants' areas also tend to commute long distances to schools of high quality because their current places of residence and those of their hukou registration (where high quality schools are located) are different, except for some areas where the policy is even more rigorously applied. In the Dongcheng district in Beijing, only children who can provide both property ownership certificates and have local hukou are eligible to enrol in nearby schools (Xinhuanet, 2015).

Moreover, some of the high income groups are capable of 'buying their way' into prestigious school catchment areas in the inner city, while living in 'High income internal migrants' areas where there are more spacious properties (Wu *et al.*, 2016). This not only exacerbates the traffic congestion but also has negative implications for the nearby enrolment rights of the more disadvantaged residents. This phenomenon is not unique to China; in the western world, research has shown that wealthy middle-class parents can 'work' the system effectively no matter whether there is total freedom of parental choice, limited parental choice or strict nearby enrolment policy (Hamnett and Butler, 2013). However, unlike the actual residence-based nearby enrolment in western countries, such as London (Hamnett and Butler, 2011), the hukou and the property-based enrolment policy strengthens the advantages of high income groups further in receiving high quality education under China's specific institutional background. Thus, decreasing the privilege of Beijing hukou and providing similar quality education services citywide are key of reducing the long commute distances in 'High income internal migrants' and easing traffic pressure in Beijing.

All in all, this research identifies for the first time the potential socio-spatial education inequality in Beijing using geodemographic classification. Within a wider education inequality context, this work also provides new theoretical perspectives for research on school enrolment policy and will be a reference for the development of education equality in other countries. It is acknowledged there are some limitations to this research, such as the ecological fallacy issue of using geodemographic classification. Thus, more in-depth and comprehensive qualitative research should be conducted in some sub-districts to investigate the detailed variations and disparities within each cluster and inform on the influences of areas on a disadvantaged group's education. When new census data are available, an education related classification can be reconstructed to thereby monitor changes in the effects of education related policy on education inequalities.

Chapter 7

Educational Equality and Socio-spatial Access to Education Facilities in Beijing: A Multilevel Analysis of Travel Distance

7.1 Introduction

The spatial access to education facilities (e.g. travel distance) is an important aspect of education equality. A long commute to school will lessen the time that pupils have to undertake private study on a day-to-day basis. Another limitation may include internal factors such as physical and mental weariness caused by the long travel distance (Talen, 2001; Hamnett and Butler, 2011). So, what will influence pupils' spatial access to educational resources?

Pupils are nested within the geographical areas (e.g. sub-districts) in which they live. 'Where they live', the pupils' residential location or spatial context, has a crucial impact on the structure of available education opportunities for pupils (e.g. school places, school quality and the enrolment policy) and their access to education resources (Gordon and Monastiriotis, 2007; Hamnett and Butler, 2011). Moreover, 'Who they are', pupils' demographic and socioeconomic background (e.g. hukou, housing type and income), also has an influence on pupils' access to education facilities (Talen, 2001; Wu *et al.*, 2008; Logan *et al.*, 2012; Williams, 2012; De la Fuente *et al.*, 2013; Rangel, 2013; Williams and Wang, 2014). Thus, 'Where they live' influences the spatial separation between schools and children, while 'Who they are' influences the mobility of children.

From a social justice and policy-making perspective, it is important to explore how social and spatial factors influence pupils' travel distance, whether long commuting distances are a particular problem for specific individuals or groups with specific characteristics, and whether long commutes are more pronounced in particular areas. However, Talen (2001) suggested the relationship between pupils' socioeconomic traits and travel distance was inconclusive and not consistent across countries, which was described as "unpatterned inequality". The failure of Talen's study to discover the relationship between pupils' socio-economic characteristics and travel distance is due to the neglected influence of spatial context on pupils' travel distance and the hierarchical nature of the system. Thus, a more appropriate method, multilevel modelling, is used in this research to make up for the deficiency.

Multilevel modelling is a flexible statistical methodology which acknowledges the structure of hierarchical data (Goldstein and Silver, 1989), allowing the introduction of contextual (place) differences without ignoring the compositional variations (individual determinants) (Gould and Fieldhouse, 1997; Owen *et al.*, 2016). As the education system has a hierarchical organisation, multilevel modelling has been frequently applied in educational research to investigate the effects of school resources (Steele *et al.*, 2007), ethnic composition of schools (Fekjær and Birkelund, 2007) and neighbourhood effects (e.g. deprivation in the home neighbourhood) (Garner and Raudenbush, 1991) on pupils' educational attainment. In terms of school commuting, multilevel modelling has been used to examine the influences of predictors at different levels on active school commuting (Easton and Ferrari, 2015) and school choice (Harris *et al.*, 2007). Although the commute distance between home and work (Shuttleworth and Gould, 2010) and distance moved by migrants (Thomas *et al.*, 2015) has been explored by multilevel modelling, there is a lack of research relating to the travel distance of school commuting.

In the last chapter, an education related geodemographic classification has been constructed to analyse pupils' access to schools by where they live. However, the geodemographic clusters have been criticised because of the potential internal heterogeneity (Harris *et al.*, 2007). Accordingly, geodemographic classification has been used as a framework rather than a tool to make sense of the socio-spatial patterns encoded within complex census data. Also, as there is a hierarchical structure in geodemographic classifications indicating that people reside in neighbourhoods which are classified into neighbourhood clusters, multilevel modelling can be employed as a natural complement to geodemographic classification (Harris and Feng, 2016). Adopting multilevel modelling allows us to assess whether pupils from different types of neighbourhood (i.e. geodemographic clusters) in Beijing will travel longer or shorter distances and to simultaneously consider the influence of individual level characteristics, the internal heterogeneity, on pupils' travel distance.

Using individual observations from a comprehensive and geographically identifiable pupil travel survey, the latter parts of this chapter focus on the adoption of multilevel modelling to assess the influence of context factors (spatial factors) in determining pupils' travel distance in Beijing, simultaneously taking account of possible confounding predictors (social factors) related to the pupils at individual level (e.g. hukou, household income, and parents' qualifications).

This chapter is organised as follows: Section 7.2 introduces the distribution of pupils and education resources in Beijing; the preliminary descriptive analysis of the Comprehensive Household Travel Survey (CHTS) and Pupil Travel Survey (PTS) datasets are explained in Section 7.3; the modelling strategy and specifications of model estimation are described in Section 7.4; the discussion and policy implications are outlined in Section 7.5; and the concluding remarks are contained in the final section.

7.2 Distribution of educational resources in Beijing

Education provision in Beijing is briefly introduced from both demand and supply perspectives in the following sections.

7.2.1 Imbalance in the distribution of education resources in Beijing

The distribution of education resources in Beijing is very unequal. As shown in Figure 7.1, there is a general pattern whereby school density decreases significantly from inner areas to outer areas in 2014. The density distribution of primary and secondary schools is quite consistent and there are only slight differences between them.

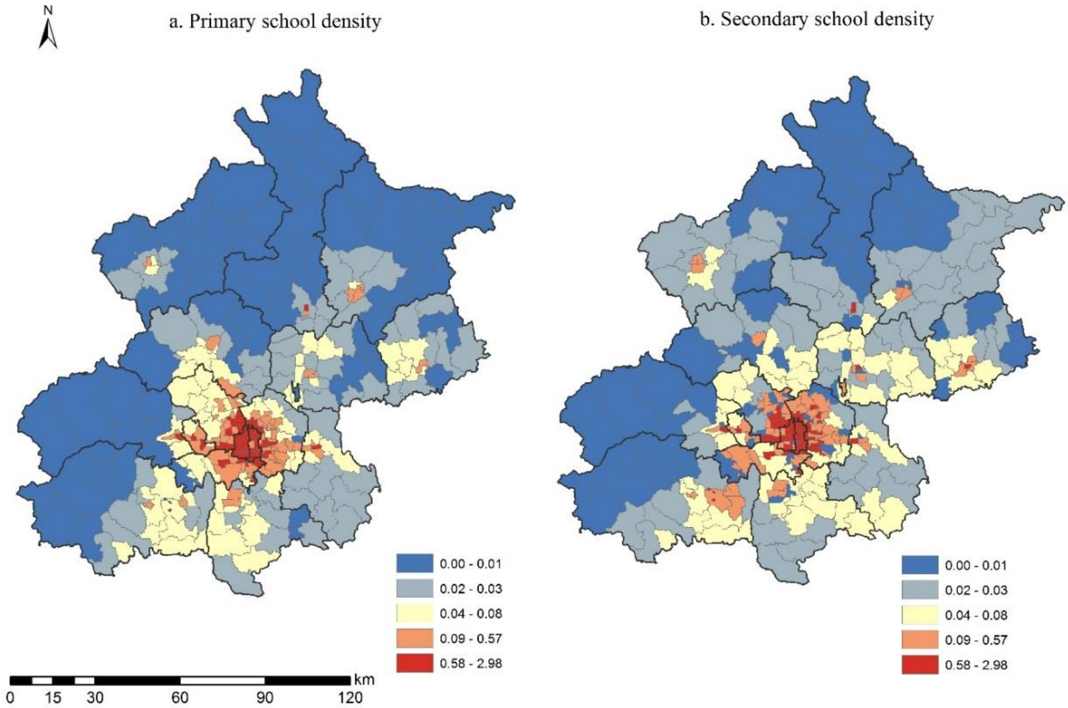


Figure 7.1 The school density quintile distribution of primary and secondary schools by sub-district in Beijing, 2014[7] (per km^2)

7.2.2 The change of population and pupils (school places) in Beijing

Table 7.1 displays the change of population and pupil number in Beijing from 2005 to 2014. The number in the table is calculated by using the value in 2014 to divide that in 2005 and then multiply by 100%. For example, the number in the first cell, 140%, is calculated by using the permanent population in Beijing in 2014, 21,516,000 to divide by the population of 15,380,000 in 2005 and then multiply by 100%. The bold numbers are values higher than the global average for Beijing.

It can be seen that the increase in the registered population, the people with local hukou, has been marginal, but the in-migrant population has risen to 230% of its 2005 total. From the supply side, the number of school pupils can be used to indicate the number of available school places. The number of

[7] The school density is calculated by dividing the number of schools located in each sub-district by the area of each sub-district.

primary school pupils has increased the most (170%), while the number of pupils attending middle school stayed almost the same and high school pupils have decreased significantly (60%). The migrant children are not allowed to receive high school education (they can only receive vocational education) in Beijing, and thus most of the high school pupils are local residents, except for those in private high schools. Due to the influence of the ‘one child policy’, the number of high school pupils has decreased dramatically. The major increase in primary school pupils is because most of the migrant children have been allowed to receive primary education in Beijing since 2007. However, in order to receive consistent education and attend entrance examinations for high school in their hometown, some pupils choose to go back to their hometown to receive middle school education, although they are permitted to attend middle school in Beijing. This explains why the number of middle school pupils has not increased to the same level as that of primary school pupils or decreased similar to high school pupils.

Table 7.1 The change of population and pupils (school places), 2005-2014

		Demand (%)			Supply (%)		
		Permanent population	Registered population	In-migrant population	Primary school pupils	Middle school pupils	High school pupils
Beijing		140	113	229	170	100	60
Central Beijing	Dongcheng	106	101	139	130	80	60
	Xicheng	109	111	155	160	90	60
	Chaoyang	**140**	**119**	214	**270**	**150**	60
	Fengtai	**147**	**116**	**233**	**210**	**140**	60
	Shijingshan	124	108	142	**180**	**130**	**70**
	Haidian	**142**	**124**	204	**180**	**130**	**90**
Outer Beijing	Fangshan	119	105	**224**	120	70	50
	Tongzhou	**156**	112	**282**	**210**	80	50
	Shunyi	**141**	109	**249**	160	80	60
	Changping	**244**	**121**	**458**	**240**	**150**	**90**
	Daxing	**174**	**115**	**299**	**180**	80	**70**
	Mengtougou	110	105	120	110	80	60
	Huairou	118	103	196	100	70	60
	Pinggu	102	101	221	70	40	40
	Miyun	109	102	206	90	60	**70**
	Yanqing	113	102	180	70	70	50

Source: Statistical Yearbook of Beijing, 2005 and 2014

Detailed changes in every district are also shown in Table 7.1. The change in the permanent population is not even across the whole of Beijing. The population of the inner areas of Central Beijing, where there is limited space for in-migrants (e.g. Dongcheng and Xicheng) and outer suburbs of Beijing (e.g. Pinggu and Miyun) have changed slightly. There are more migrants moving to the outer districts of Central Beijing and the districts which are adjacent to these outer districts.

7.2.3 School places condition in each district

Whether there are sufficient education resources is decided by a comparison of the available school places (supply) and the number of school age children (demand) and displayed by the bar charts in Figures 7.2 and 7.3. As there are no data available for primary school or middle school age children in each district, children aged 0-14 in each district are used as a proxy for school age children. The equation for reflecting sufficient or insufficient school places for primary schools in each district is as follows:

$$S_i = P_i - P * \frac{C_i}{C} \quad (7.1)$$

where S_i is the sufficiency index for district i;

P_i is the number of primary school pupils in district i;

P is the total primary school pupils in Beijing;

C_i is children aged 0-15 in district i; and

C is the total of children aged 0-15 in Beijing.

If the value of S_i is positive, it means there are some extra school places for those from elsewhere, which is represented by upward bar (Figure 7.2 and 7.3); if the value is negative, it means there is not enough school places, which is represented by downward bar.

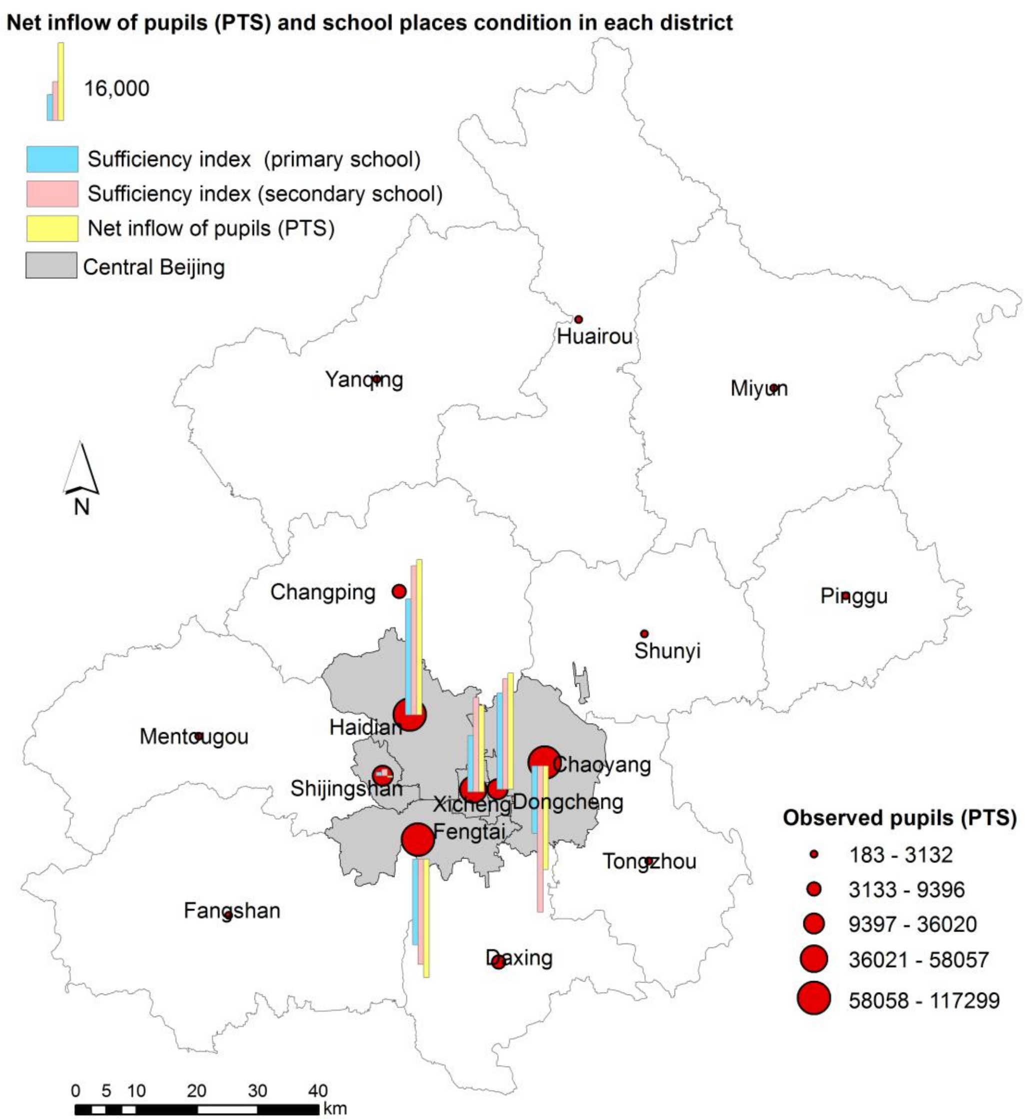


Figure 7.2 Distribution of observed pupils in the PTS by district and the condition of school places in each district

Source: Beijing Municipal Commission of Transportation

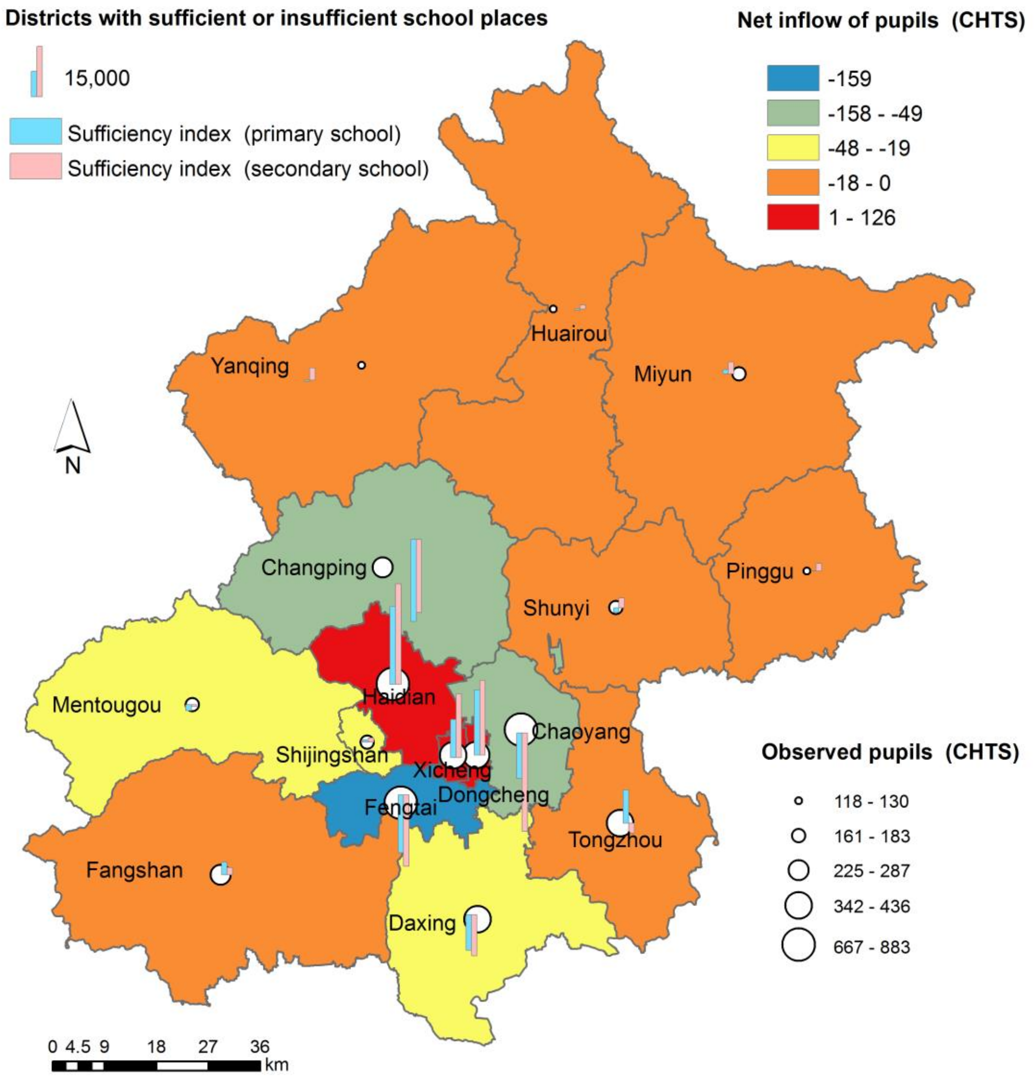


Figure 7.3 Distribution of observed pupils in the CHTS by district and the condition of school places in each district

Source: Beijing Municipal Commission of Transportation

The equation for secondary school pupils is similar. Although the number of primary school pupils per 1,000 permanent residents in Dongcheng and Xicheng district is the highest, the spare school places are highest in Haidian district. Chaoyang is the district most lacking school places for secondary schools, while Fengtai district is lacking the most in primary school places. The spatial variation in the sufficiency index is more uneven for secondary schools than for primary schools.

7.3 Descriptive analysis of datasets

Two datasets, the fifth Comprehensive Household Travel Survey (CHTS) and the Pupil Travel Survey datasets (PTS) (Section 3.5.3) have been used in analyses in the following section. A comprehensive data description and validation has been elaborated in Chapter 3. Given that the purpose of the research is to consider access to education services (schools), only school commute trips of 6-18 year-old respondents in CHTS have been selected, including 5,377 observations involving 12,341 school trips (i.e. trips to and from school). As the response variable in the models is travel distance, only a one-way trip is selected. Thus, 5,373 pupils making 5,373 trips (four pupils were excluded from the sample because their travel distance is zero, which cannot be processed in the later log transformation) from home to school were selected. In order to avoid biased sampling in the CHTS and confirm the characteristics and patterns revealed by CHTS, the other larger dataset, the PTS, with 458,498 respondents (more than half of the pupils within Central Beijing) and a few variables in common with the CHTS, has been used in this research.

Descriptive analysis is a helpful precursor to exploring the dataset before the more sophisticated multilevel modelling (Shuttleworth and Gould, 2010). Thus, some basic descriptive analyses have been conducted and are reported in this section to give an overview of the datasets and general patterns within them. The common variables in the two datasets will be introduced in the first sub-section (7.3.1) and the similar variables will be discussed in the second sub-section (7.3.2), while the unique variables in the CHTS will be analysed in the final sub-section (7.3.3).

7.3.1 Analysis of common variables in the CHTS and PTS

7.3.1.1 Origins and destinations

The distributions of pupils who are in the PTS and CHTS within each district are shown in Figures 7.2 and 7.3 respectively. The PTS is a school-based travel survey in Central Beijing, an area which is shown in grey in Figure 7.2. The red graduated symbols in Figure 7.2 display the distribution of respondents by their home locations in each district in Beijing. Although most respondent pupils live in Central Beijing, there is still a large number of pupils not living in Central Beijing. As there are no data for pupils studying in the outer districts of Beijing, Figure 7.2 only displays the net inflows of pupils in each district of Central Beijing, which are shown by the yellow bars. The upward yellow bars represent more pupils coming to study in these areas rather than leaving these areas to study elsewhere (net inflow), while the downward yellow bars indicate there is a net outflow of pupils to areas outside the district.

Similarly, the white graduated symbols in Figure 7.3 indicate the distribution of respondents in the CHTS by their home locations. Most of the samples are in Central Beijing except for the Shijingshan district (Figure 7.3); the remaining districts have relatively smaller samples. The choropleth map (Figure 7.3), shows that there are more pupils studying in Haidian, Dongcheng and Xicheng districts than living there, whereas the net inflow is negative in other districts. There are significant net pupil inflows into these three districts, which contain the best education resources in Beijing, while the other districts are places with net outflows.

Both the PTS and CHTS show that Haidian, Xicheng and Dongcheng are districts with large numbers pupil inflows, which is probably because of their large provision of school places (a high proportion of schools in Beijing are located in these districts).

Figures 7.2 and 7.3 display a consistent relationship between the sufficiency index and the inflow of pupils in the two datasets. This relationship also has been validated using a Pearson correlation coefficient (r). The net inflow of pupils is highly correlated with the sufficiency index in each district in both datasets. In the large dataset, the PTS, the r values are as high as 0.997 and 0.987 for primary and secondary schools respectively, which indicates that the school commutes across district boundaries are due to the imbalance in education resources from district to district. In CHTS, the Pearson correlation coefficients for primary and secondary school pupils are 0.856 and 0.888, which also suggests a high correlation.

7.3.1.2 Nearby enrolment (enrol in the same district)

In this section, the nearby enrolment rate (district), which is defined as the rate of pupils' residence and school being in the same district, in Central Beijing will be examined using both datasets, the PTS (431,669 pupils) and the CHTS (3,262 pupils). There are 80.2% of pupils in the PTS and 82.9% of pupils in the CHTS enrolled in the same district in 2014 in Central Beijing. The enrolment rates revealed by the PTS and CHTS show nearly the same trend except for Shijingshan and Chaoyang district (Figure 7.4). The nearby enrolment rates in Haidian, Dongcheng and Xicheng are the highest, which is consistent with their relative wealth of educational resources. Fengtai and Chaoyang have the lowest nearby enrolment rates probably due to their insufficient education resources.

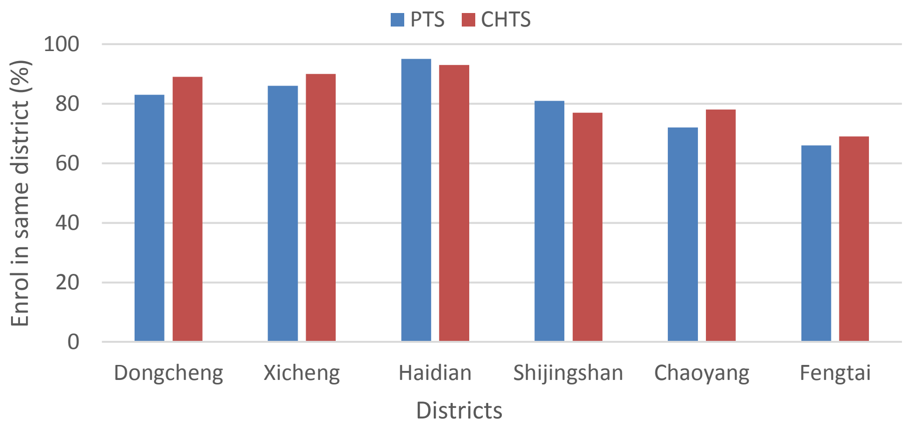


Figure 7.4 Nearby enrolment rates (enrol in same district) by district, PTS and CHTS, 2014

Figure 7.5 displays the nearby enrolment rate by geodemographic cluster (see Chapter 6) for the PTS and CHTS and their distributions show a very similar pattern. 'High income internal migrants' is the cluster with the lowest nearby enrolment rate, which suggests a high proportion of cross-district boundary commuting to school is taking place.

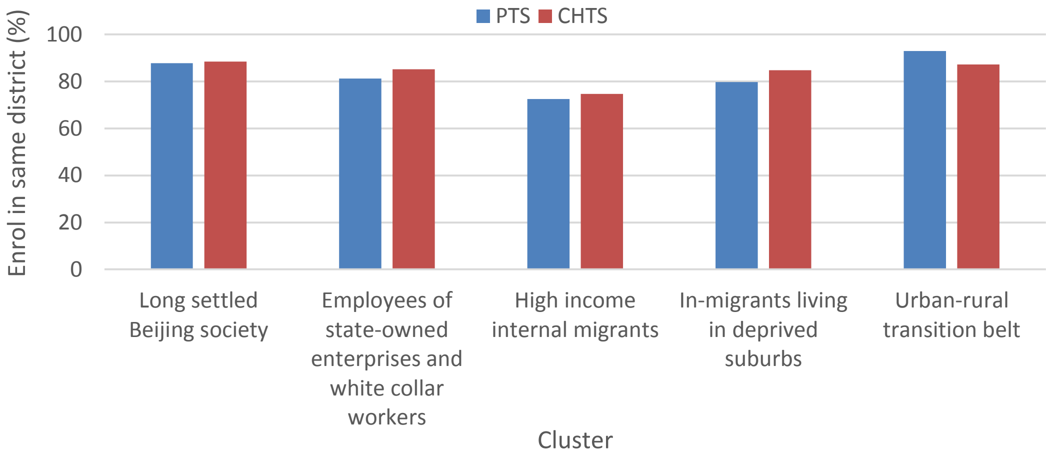


Figure 7.5 Nearby enrolment rates by geodemographic cluster, PTS and CHTS, 2014

There is no direct education stage information available in the CHTS, so only the PTS is used to reveal the nearby enrolment rate for each education stage in each district (Figure 7.6). The nearby enrolment rates are lower for middle and high school (secondary school) pupils. For the three districts which have sufficient pupil places, the enrolment rates only experience slight decreases. However, this rate declined dramatically for the three districts where education resources are insufficient. Figure 7.6 suggests that the education resources distribution for secondary education is more uneven than that for primary school education. The significant shortage of education resources for high school education in Chaoyang and Fengtai districts leads to very low nearby enrolment rates of 29% in the former and 40% in the latter.

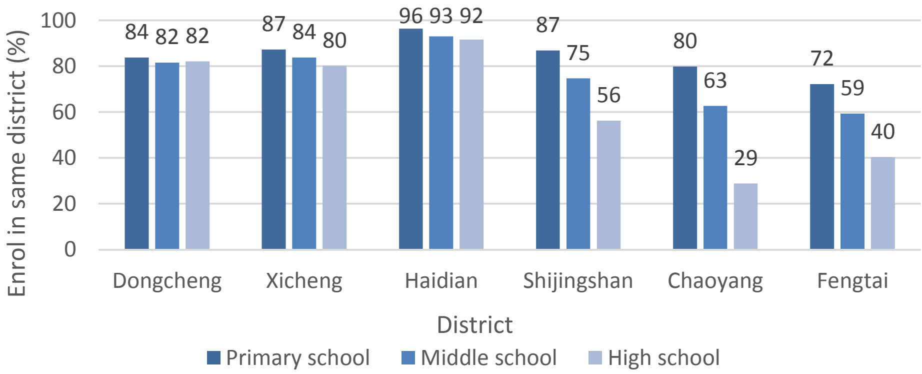
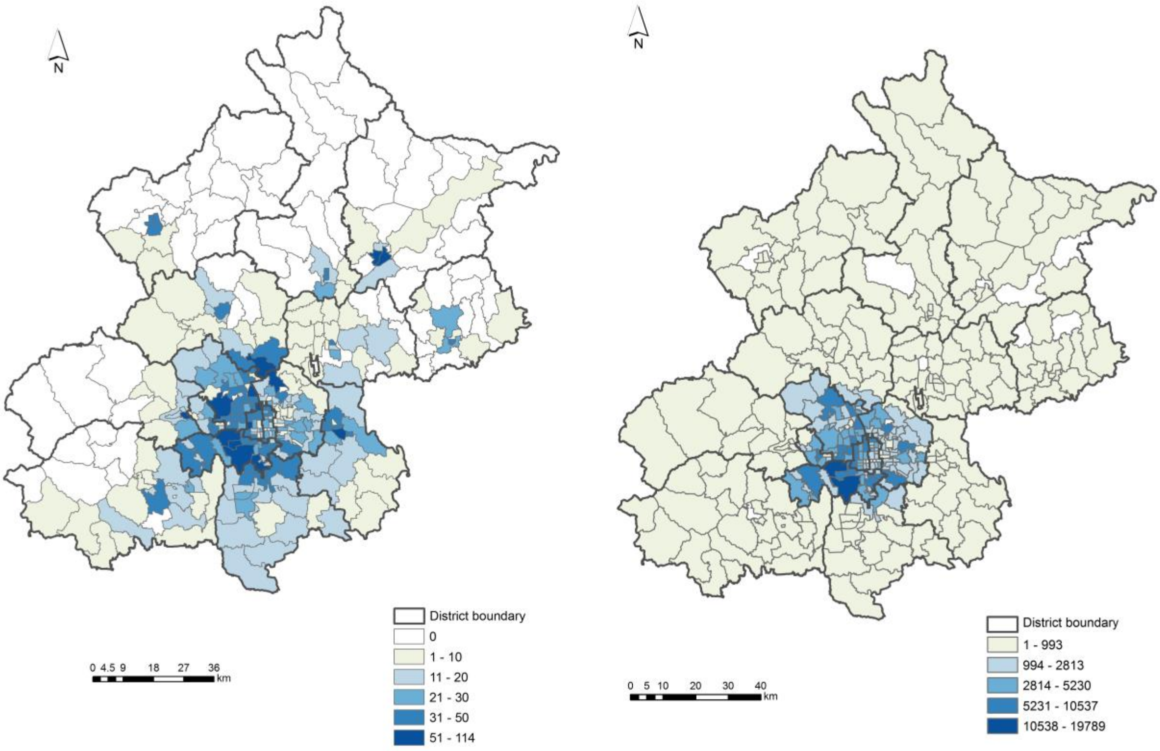


Figure 7.6 Nearby enrolment rate (enrol in the same district) for each education stage in PTS, 2014

The nearby enrolment rate for the pupils with hukou outside a municipality is higher (89%) in CHTS than other categories of hukou status (within a district 83%, outside a district but within a municipality 61%) , which is likely to occur because pupils in this category rarely have the chance to enrol in a school in another district.

7.3.1.3 Data distribution by sub-district

Maps a and b in Figure 7.7 display the detailed distributions of pupils from each survey living in each sub-district, which are very uneven. The sample distribution of the CHTS is largely consistent with the distribution of communities. The places with no pupils in the sample (CHTS) are normally in the mountainous areas without large communities.



a. Pupils resident by sub-district (CHTS) b. Pupils resident by sub-district (PTS)

Figure 7.7 Distribution of pupils in Beijing by sub-district from the PTS and the CHTS

7.3.1.4 Travel mode

In the CHTS, the 17 ‘commute to school’ travel modes in the dataset were combined into six main types of travel mode to facilitate analysis: on foot; bike; bus; car; metro; and other. Walking and cycling are the two leading travel modes for the whole of Beijing, followed by bus and car. The dataset can be divided into four categories based on the pupils' home and school locations: home and school both in Central Beijing (3,233), home and school both outside Central Beijing (1,975), home in Central Beijing but school outside Central Beijing (29) and home not in Central Beijing but school in Central Beijing (136).

Compared to pupils whose home and school are both outside Central Beijing (Table 7.2), more pupils living and studying in Central Beijing have to travel to school by bus and metro and less travel by bike. This is most likely to be due to the more complete public transportation network in Central Beijing. For the pupils who both live and study outside Central Beijing, a higher proportion choose to go to school by bike and lower proportion use the bus and the metro, which is likely due to the lower public transportation density or shorter commute distance. Table 7.2 also shows the commute mode chosen by pupils who commute between Central Beijing and outer areas of Beijing. Due to the long travel distance, most of them choose to commute to school by car, bus or metro.

Table 7.2 Travel mode comparison by location of home and school (CHTS and PTS)

	CHTS n=5,373				**PTS** n=458,498	
n=	3,233	1,975	29	136	431,669	26,829
	Home and school both within Central Beijing (%)	Home and school both outside Central Beijing (%)	Home in Central Beijing but school outside (%)	Home outside Central Beijing but school within (%)	Home and school both within Central Beijing (%)	Home outside Central Beijing but school within (%)
On foot	34.2	35.7	17.2	0.7	32.2	13.3
Bike	23.9	33.0		6.6	21.9	9.5
Bus	21.8	15.1	41.4	25.0	16.3	24.0
Car	15.5	15.3	13.8	47.8	25.7	42.0
Metro	4.0	0.3	27.6	19.9	3.7	10.1
Other	0.6	0.7			0.3	1.1

As the PTS is a school-based survey, the dataset is divided into two subsets: pupils with home and school both in Central Beijing (431,669) and pupils with home outside Central Beijing but school inside Central Beijing (26,829). Compared to the group with both home and school located in Central

Beijing, the pupils living outside Central Beijing tend to commute by car, bus and metro and are less likely commute by bike or on foot.

When comparing those commuting within Central Beijing in the CHTS and the PTS, there appears to a higher percentage commuting by car in PTS (PTS 25.7%, CHTS 15.5%), while fewer pupils are commuting by bus (PTS 16.3%, CHTS 21.8%) (Table 7.2). For pupils whose home is outside Central Beijing but attend school within Central Beijing, there are fewer pupils choosing the metro in the PTS (PTS 10.1%, CHTS 19.9%), but more pupils walk to school (PTS 13.3%, CHTS 0.7%). However, there are still some common characteristics in the two datasets: most pupils commuting within Central Beijing go to school on foot, while most pupils living outside Beijing but studying in Central Beijing will choose to go to school by car or bus. It indicates longer commute distances for pupils commuting across the boundaries into Central Beijing districts, as expected.

In the CHTS, the in-migrant pupils have a significantly higher percentage going to school on foot (46%). In addition, there is a higher percentage of pupils travelling to school by car for those parents with higher qualifications and higher incomes. However, there is also a higher proportion of pupils choosing to go to school on foot for those higher income groups but they are less likely to go to school by bus. This may indicate that higher income groups would prefer to send children to school by car or live close to a school which is convenient for walking, or they are more aware of the benefits of active travel.

The commute modes of pupils at different education stages in the PTS indicates that the percentage of pupils who go to school on foot decreases from primary school to high school (38%, 20% and 13%), while there are more higher grade pupils travelling by bus (primary 11%, middle 29% and high 28%) or metro (primary 1.6%, middle 6.1% and high 14.2%), which is because the commute distance for middle and high schools is normally longer and not always walkable.

7.3.1.5 Travel distance

Table 7.3 shows a summary of the distribution of travel distances of respondents to the samples living in Beijing and Central Beijing. The median travel distances are all around 2 kilometres. However, the distances captured by the CHTS are all smaller than those identified from the PTS.

Table 7.3 Summary of travel distance (km)

	CHTS		PTS	
n=	5,373	3,262	458,498	431,669
	Beijing	Central Beijing	Beijing (school in Central Beijing)	Central Beijing
1st Qu.	1.1	0.97	1.3	1.3
Median	2.2	2.2	2.8	2.6
3rd Qu.	4.9	4.7	5.6	5.9

It can also be concluded from Figure 7.8 that the median value for travel distance and travel time both increase from primary schools to high schools in PTS. The median value is the line in the middle of each box in each education stage; the left edge of each box represents the lower quartile ($Q1$), this means 25% of data fall below this edge, while the right edge represents the upper quartile ($Q3$), indicates that 75% of data fall below this edge; so 50% of data values are located in each box.

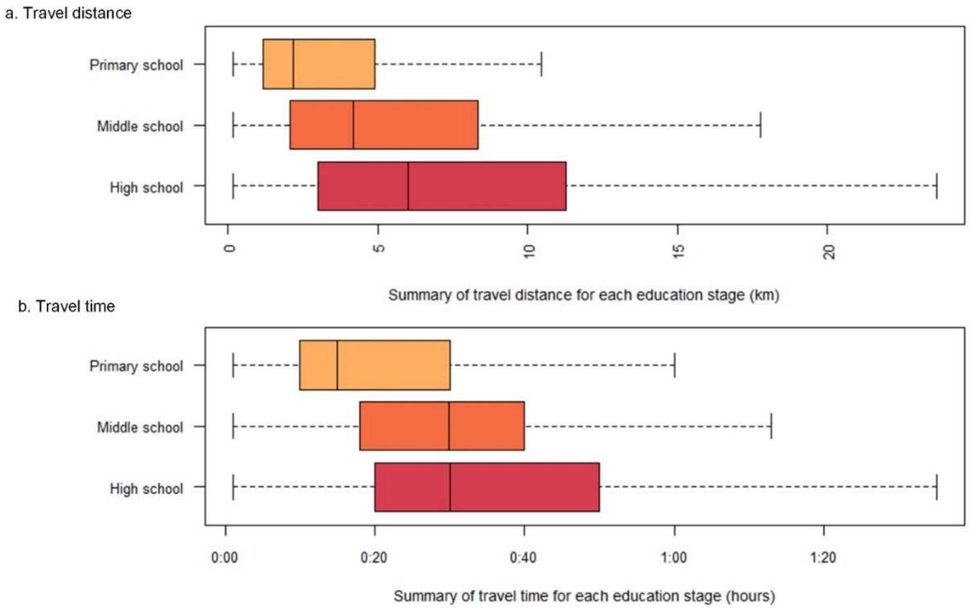


Figure 7.8 Summary of travel distance and travel time by education stage (PTS)

Two horizontal lines, called whiskers, extend from the left and right of the box. These left and right vertical lines represent the lower whisker and upper whisker. The range of the lower and upper whiskers in this research is based on using the default range setting in R. The equations of how to get the upper whisker and lower whisker are as follows:

$$\text{Upper whisker} = \text{Min}\,[\max(x), Q3 + 1.5 * \text{IQR}] \quad (7.2)$$

$$\text{Lower whisker} = \text{Max}\,[\min(x), Q1 - 1.5 * \text{IQR}] \quad (7.3)$$

$$\text{IQR} = Q3 - Q1 \quad (7.4)$$

where IQR is the interquartile range.

In this research, the minimum values are all higher than the value of $Q1 - 1.5 * \text{IQR}$, so the lower whisker is the minimum value in the data, whereas the upper whisker is $Q3 + 1.5 * \text{IQR}$. All the data located out of the range of the lower and upper whiskers are not displayed in Figure 7.8 and regarded as outliers. The travel distance of primary school pupils is mostly within 5km with a median distance around 2km (Figure 7.8a). Half of the middle school pupils commute in the range from 2km to 8km, with a median around 4km. The median value of high school pupils' travel distance is 6km and 50% need to commute from about 3 to 11.5 km. The extreme value for high school pupils is reached at around 24km. The variation in the length of the boxplot indicates that the travel distance distribution gets more varied from primary to high school.

The summary of pupils' commute time (Figure 7.8b) tells us a similar story, but with slight differences. For example, the distribution of primary school pupils' commute time is as dispersed as that for middle school pupils, which is different from that of commute distance. The median travel time of primary pupils is around 15 minutes, while that of middle school and high school pupils is both about 30 minutes. However, the commute time distribution for high school pupils is wider than for middle school pupils.

The travel distance of 'In-migrant children living in deprived suburbs' is the longest (Figure 7.9) in CHTS, followed by those living in the 'Urban rural transition belt' and is consistent with our assumptions (Xiang *et al.*, 2018). In addition, the travel distance for these in the 'Long settled Beijing society', and 'Employees of the state-owned enterprises and white collar workers' clusters are significantly shorter than the median distances for the other three clusters.

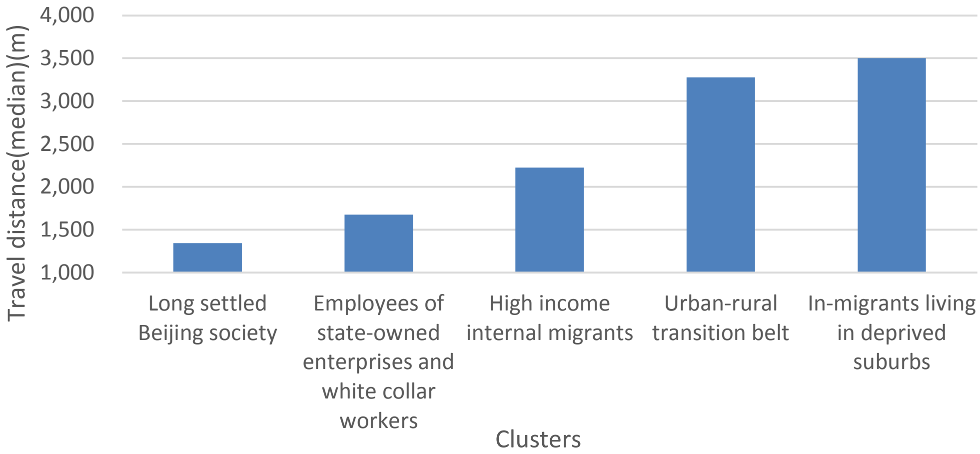


Figure 7.9 The median travel distance of pupils from different geodemographic clusters in Central Beijing (CHTS)

The average travel distances from the PTS in Figure 7.10 show the same pattern as the median travel distance of CHTS, while Figure 7.10 breaks down the data into primary, middle and high school. The average travel distances of high school pupils in the 'urban-rural transition belt' are the longest.

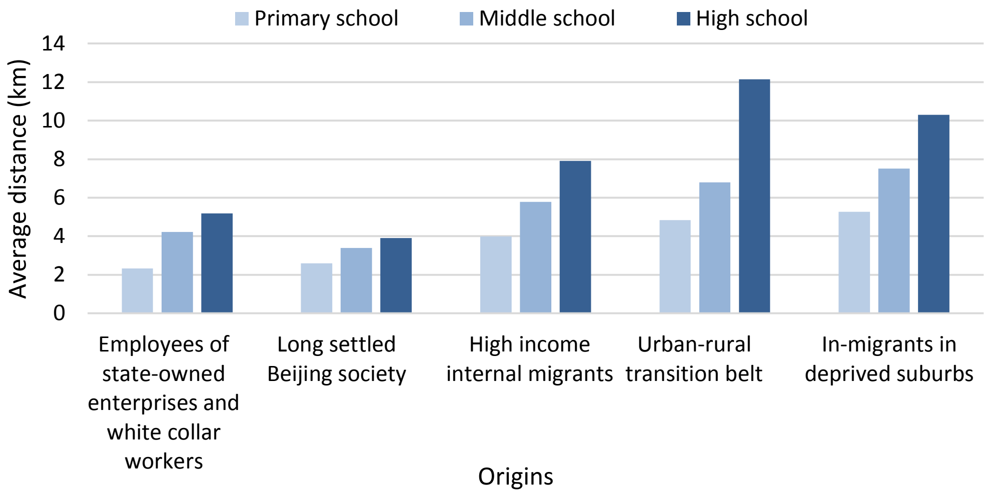


Figure 7.10 Average travel distance to school within each geodemographic cluster by education stage (PTS).

7.3.2 The analysis of similar variables in the CHTS and the PTS

7.3.2.1 Year of birth in the CHTS and grade and education stage in the PTS

In the CHTS in 2014, as the typical age range for primary schools in China is 6-11 (grade 1-6), the pupils who were born in between 2003-2008 are regarded as primary school pupils. Similarly, the pupils who were born in 2000-2002 (grade 7-9) are considered as middle school pupils. Given that the sample of respondents recorded as being born in 1996 is relatively small (n=171) and some high school pupils are 18 years old, the pupils born from 1996 to 1999 (grade 10-13) are all deemed to be high school pupils (Figure 7.11).

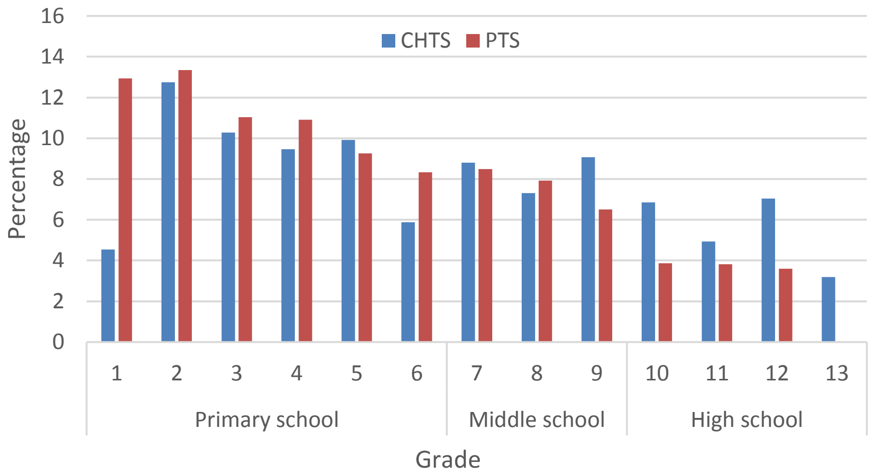


Figure 7.11 Distribution of the observed pupils in Beijing by grade, CHTS and PTS

The cross-tabulation of hukou status and education stage from the CHTS shows that the proportion of school age pupils whose hukou is outside Beijing decreases from primary to high school (23%, 15% and 6% respectively), which is due to the strict school enrolment limitation for in-migrants who are not allowed to attend public high school in Beijing. Thus, most of the high school pupils have Beijing hukou except for those who enrol in private high schools.

In the PTS (Figure 7.11), nearly two thirds (65%) of the sample collected involve primary school pupils (normally grade 1-6), while 22.9% and 11.3% are from middle school (normally grade 7-9) and high school (grade 10-12) respectively. Primary school education in China usually lasts for six years while middle school education is for three years, but there are also a few schools in Beijing that have a different length of schooling, i.e. five years primary education (grade 1-5) and four years middle school education (grade 6-9). There are 200 pupils in this survey who had attended four-year education middle schools where grade 6 is regarded as middle school education. In addition, there are 70, 51 and 42 pupils who are in grade 7, 8 and 9 respectively but they are recorded in the primary school education stage in this survey. This may occur because pupils gave incorrect information for education stage or grade in the survey. The numbers are relatively small, so these cases will be retained for certain analyses.

The population age structure of children in Central Beijing, as represented by the survey, reveals the unusually higher number of children in primary rather than secondary (middle and high) schools due to the continuous influx of migrants of reproductive age and the baby boom in recent years (Bi and Zhang, 2016). In addition, the higher number of pupils in middle school compared to high school observed in Figure 7.11 is because high school education is not part of compulsory education and there is a considerable dropout rate after middle school.

Figure 7.12 shows the changing numbers of pupils enrolling in primary schools (grade 1) in Beijing from 1978 to 2014. Since the implementation of the 'one child' policy in the 1970s, the enrolment of new pupils decreased very dramatically from the late 1970s to the early 1980s. However, the number of new pupils began to increase quickly after 1982 as those born in the 'baby boom' (1950-1957) reached reproductive age. From 1997, because the people who were the first generation of the 'one child' policy had reached reproductive age, the population of new pupils declined quickly again. However, the pupil population began to increase rapidly again after 2006, because the pupils who did not have Beijing hukou were for the first time allowed to attend compulsory education in Beijing in 2007. By 2014, the migrants in Beijing had increased to around 37.6 times the number in 1987. Thus, the influence of migrants on the new enrolment statistics became increasingly significant.

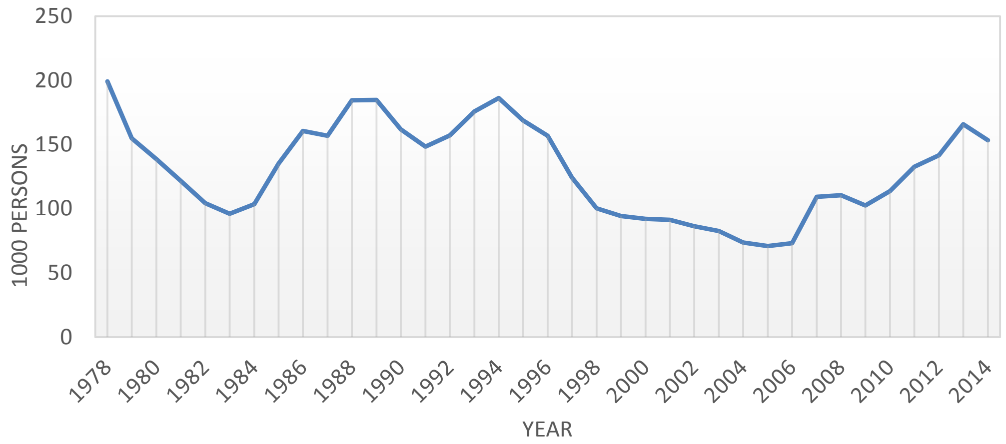


Figure 7.12 The enrolment of pupils in primary schools (grade 1) in Beijing (1,000 persons), from 1978 to 2014.

The grade composition of the sample of pupils in each district in Central Beijing is shown in Figure 7.13. All variables are measured in percentages to normalise the data across districts to avoid the influence of varying populations. Although the overall patterns are similar from district to district, involving a progressive decrease from lower grades to higher grades, there are still very significant differences between them. The proportion of high school students is very low in Chaoyang and Fengtai, while the proportions in Dongcheng and Xicheng, both of which are old city areas, are considerably higher.

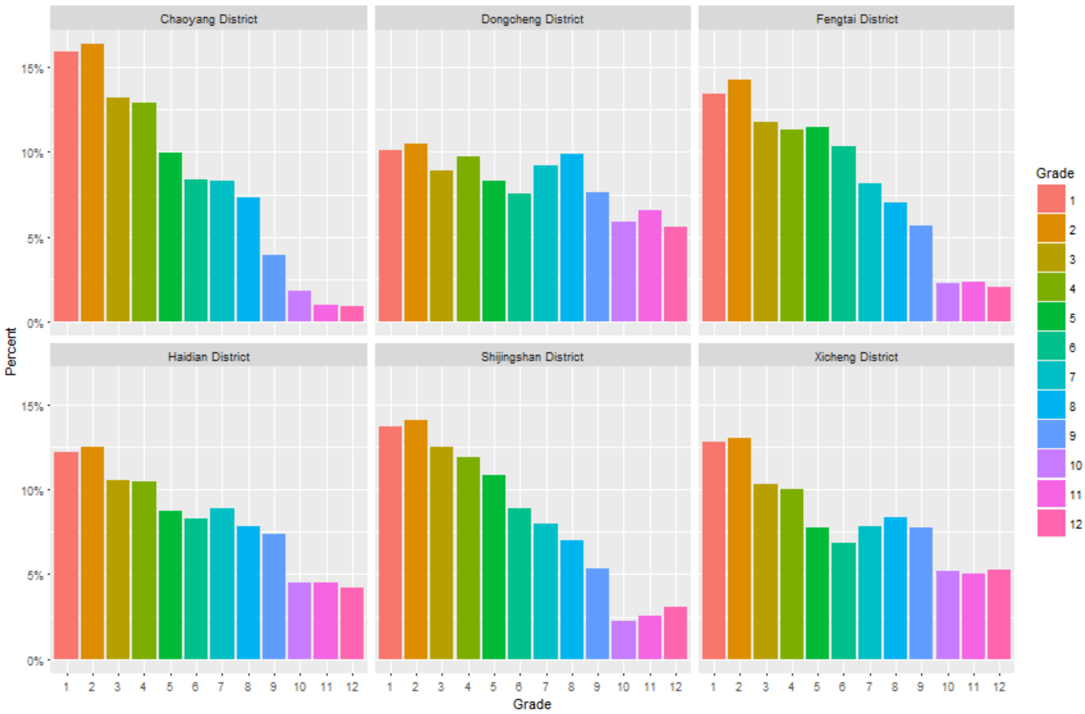


Figure 7.13 The percentage of pupils in each grade within each district, PTS

7.3.2.2 Housing type from the CHTS and tenure from the PTS

Nearly two thirds (64.5%) of pupils in the PTS reside in property owned by their family, and 33% live in rented housing. Due to the incredibly high prices of property in Beijing, most in-migrants in Beijing tend to rent rather than purchase residence (Fang *et al.*, 2015). Thus, some of the pupils in rented properties are very likely to be migrant children, while others may be local residents who have not attended a nearby school. The share of those who are lodging is relatively small (2.6%). Lodging usually means paying no rent and normally occurs with pupils living in houses with grandparents or parents' friends. However, some of the boarders also had selected lodging in this dataset, which had been confused with rented by the survey respondents. In China, not only do boarding schools have boarding pupils but so do some ordinary schools. The pupils in these schools are categorised into 'day pupils' and 'boarding pupils'. Thus, boarders in this survey do not only refer to pupils studying in boarding schools, but also some pupils studying in ordinary schools as boarders. According to the definition of lodging in this survey, boarders should have selected rented as their tenure. However, quite a lot of boarders selected lodging.

In addition to the overall distribution, the tenure composition between different education stages is shown in Figure 7.14. It reveals that the proportion of pupils living in rented housing increases from primary to high school, which is likely to be due to the increased commute distance from primary to high school given that more high school pupils choose to live in nearby rental housings.

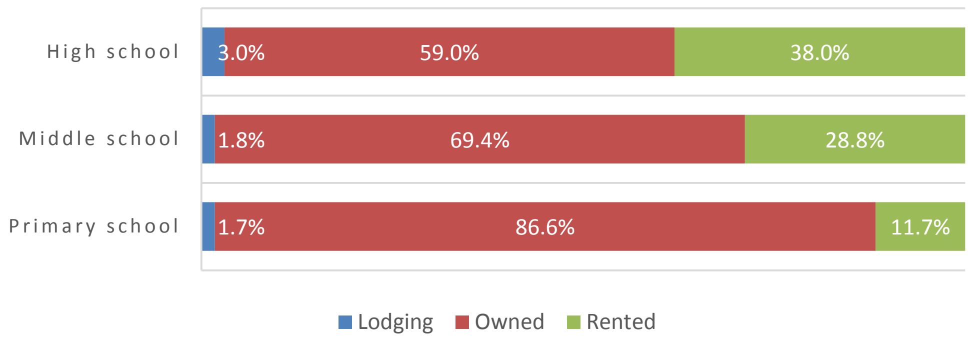


Figure 7.14 Education stage and tenure of pupils, PTS

From the perspective of commuting, the average travel distance and travel time of pupils with rented housing is significantly lower than for those living in properties owned by their parents in PTS, because the aim of renting accommodation is to reduce the commute distance. The choice of rented residence is more flexible, so families tend to rent accommodation close to a school.

The data from the CHTS was disaggregated into two subsets: pupils living outside Central Beijing and within Central Beijing (Figure 7.15). Commercial housing (private housing) is the largest component for both within or outside Central Beijing. However, commercial housing accounts for more than 50% of the sample outside Central Beijing but is only 35.8% within Central Beijing. Purchased public housing (20.2%) and rented housing (16.3%) are the other two most important housing types in Central Beijing, whereas self-built housing is the other most important housing type in areas outside Central Beijing.

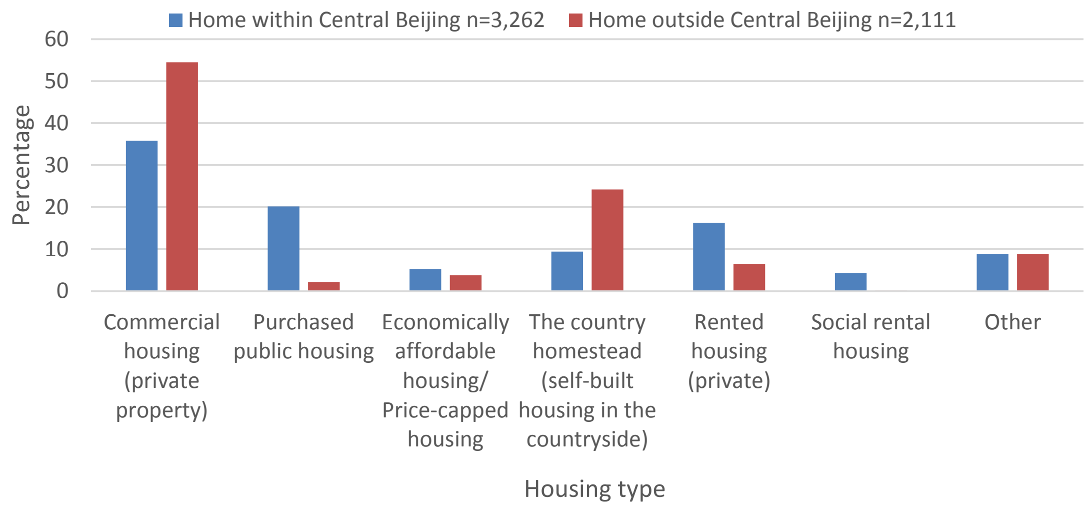


Figure 7.15 Distribution of the observed pupils of Beijing by housing type, CHTS.

In Central Beijing, the people whose hukou is outside a district but within a municipality (49%) have the highest percentage of commercial housing; people whose hukou is outside Beijing have a higher proportion of rental housing (55%); households whose hukou is within the same district have a higher percentage of purchased public housing (24%) and countryside homesteads (12%).

The data also show that a higher proportion of households with higher qualifications and higher income are living in commercial housing, while the lower qualifications and lower income groups are more likely to be living in social rental housing and in countryside homesteads. The difference for rented housing (private) is not very significant among different groups.

7.3.3 Analysis of unique variables in the CHTS

In the following section, the unique variables in CHTS will be introduced.

7.3.3.1 Relationship with the head of household

Most of the survey respondents are children of the head of household, accounting for around 78.8% of all the samples, while 19.7% of pupils are grandchildren or other relatives of the head of household. Moreover, further analyses indicate that more pupils with local hukou who attend primary school live with relatives, while less in-migrant children and high school pupils live with their grandparents. The proportion of pupils commuting to school on foot is higher for those living with grandparents. The qualifications of pupils' grandparents are comparatively lower, but their grandparents' qualifications and occupations cannot comprehensively reflect the socio-economic backgrounds of pupils' parents, which may influence the analysis of pupils' socio-economic background.

7.3.3.2 Hukou status of the head of household

The head of household with hukou within a district, outside a district but within a municipality, and outside a municipality, account for 77.6%, 5.1% and 17.3% in the sample respondents respectively. The latter component in the sample is much lower than that of Beijing in 2014 in Beijing Statistical Yearbook, which is 38.1%. For the pupils who live and study outside Central Beijing, there is a higher percentage with within district hukou. The proportions of those with hukou outside a district but within a municipality are higher for people commuting from outside Central Beijing to Central Beijing to receive education.

7.3.3.3 Highest qualification of the head of household

Some differences exist between the composition of sample according to the highest qualification of the head of household in Central Beijing and outside Central Beijing. The percentage of pupils whose head of household has higher level qualifications (i.e. Bachelor degree and above) in Central Beijing (Figure 7.16) is higher, while the proportion of pupils whose head of household has a lower qualification (i.e. illiterate, primary school and middle school) for those living outside Beijing is higher.

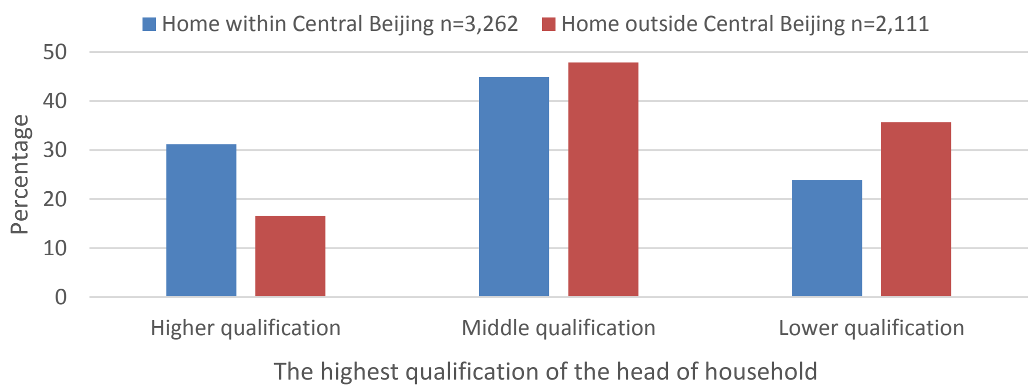


Figure 7.16 Distribution of the observed pupils by the highest qualification of the head of household, CHTS.

The residents who have hukou outside a district but within a municipality have significantly higher proportion of head of household with higher qualifications (38.5%) compared to other types (within a district 26%, outside a municipality 18%).

7.3.3.4 Household income of last year

The households with low annual income (below £5,000), middle income (£5,000-20,000) and high income (above £20,000) account for 25%, 71% and 4% respectively in Beijing. Compared to the whole sample, there are more households with middle income and high income in Central Beijing (Figure 7.17), while there are more households with low income living outside Central Beijing.

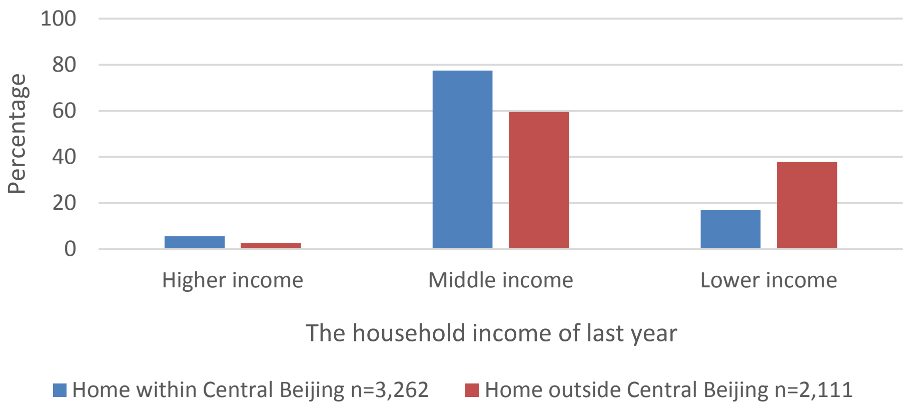


Figure 7.17 Distribution of the pupil sample of Beijing by household income of last year, CHTS

In Central Beijing, the households with hukou within a district include a higher percentage with low incomes (27%), while those with hukou outside a district but within Beijing have a higher proportion of higher income households (11%). Moreover, households with heads of household with high qualifications are likely to have higher incomes.

7.3.3.5 Car ownership

The CHTS survey suggests that more than 60% of pupil households own a car in both the sample for Beijing and that for Central Beijing. The proportion of respondents with different hukou status who own cars is quite similar, while the percentage with cars whose hukou is outside a district but within Beijing is slightly higher, at around 68.3% in Central Beijing. In addition, the data also show that those households with higher income and higher qualifications are more likely to own cars.

7.3.4 Summary

The above descriptive analyses introduce the spatial distribution of data and the statistical summaries of variables in the two datasets, which are essential preparation for the more sophisticated multilevel modelling approach used in the next section. In addition, the larger PTS dataset validates the reliability of the smaller CHTS dataset, while the CHTS comprises more detailed individual information than the PTS. The comparisons between the two datasets make the results acquired from later multilevel models more robust.

7.4 Modelling multilevel variations in pupils' travel distance

7.4.1 Variable selection and treatment

Previous empirical and theoretical work suggests that various factors operating from the microscale through to the macroscale produce multilevel variations in the distance travelled by pupils (Talen, 2001; Williams and Wang, 2014). Thus, a suitable modelling framework needs to deal with such complexities appropriately. In the first instance, it is necessary to apply multilevel modelling, which is an efficient method of dealing with clustering of individuals within higher level units and of estimating and partitioning the variabilities across different levels, groups and classifications (Paterson and Goldstein, 1991; Shuttleworth and Gould, 2010; Thomas *et al.*, 2015).

On the basis of preceding research, the explanatory variables included in multilevel modelling have been selected and are listed in Table 7.4. All the data for micro-level characteristics are selected from the fifth Comprehensive Household Travel Survey (CHTS), including hukou status, income, housing type and education stage. The individual-level data can be supplemented by indicators representing the ecological information of sub-districts from other datasets, which are called contextual variables. The data in the CHTS are hierarchically structured with pupils (level-1 units) (5,373 pupils) nested within the 251 sub-districts in Beijing (level-2 units) where they live. The effects of contextual variables can be referred to as contextual effects and are derived by summarising the characteristics of individuals or schools in specific cluster or areas. In terms of sub-district level contextual predictors, the education related geodemographic classification in Central Beijing (Chapter 6), provides a variable option for operationalising origin (residence) context. The sub-districts are classified into clusters (groups) according to the similarity of their residents' demographic structure, household composition, housing characteristics and socio-economic traits, all of which have a potential influence on pupils' access to education resources. As defined in Chapter 6, there are only five clusters in the bespoke geodemographic classifications in Central Beijing. In this research, sub-districts outside Central Beijing will be treated as the sixth classification. School density is the other context variable that has been considered.

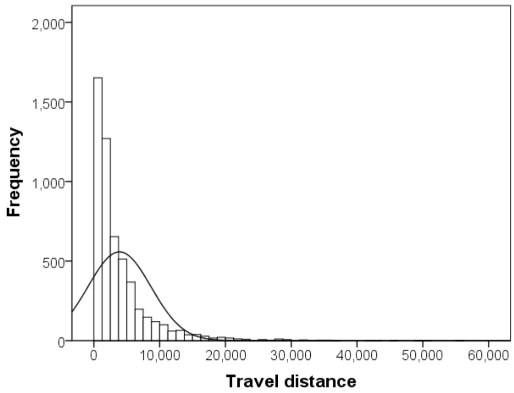
Most of the predictors are categorical variables and recoded using contrast coding (refer to Section 3.4.1.2) and portrayed by a series of dummy variables; that is, k categories of one variable were replaced by k-1 dummies. The characteristics of the base category (against which other categories can be compared) are denoted by italic and bold format in Table 7.4 and are indicated by the 'constant' (a vector of 1s) in the models. The base category in this application represents the pupils whose hukou is within a district, whose parents' are middle income, reside in commercial housing, whose family own cars, who are in primary school, live in cluster 'Employees of state-owned enterprises and white collar workers'. The other characteristics of pupils in the dataset are specified as a set of contrasts, with a 1 given if a pupil has a particular feature, and 0 otherwise.

Table 7.4 Predictors from the CHTS used in multilevel modelling

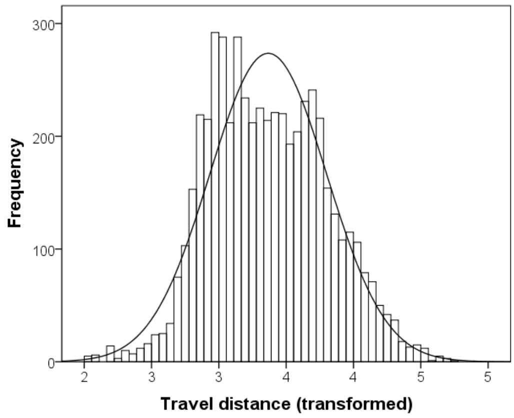
No.	Variables	No.	Categories/ Descriptions	n	%
Individual level variables					
1	The hukou status of the head of household	***1***	***Hukou within a district***	4,170	*77.6*
		2	Hukou outside a district but within a municipality	275	5.1
		3	Hukou outside a municipality	928	17.3
2	Income of household	4	Low income (Below £5,000)	1,350	25.1
		5	***Middle income (£5,000-20,000)***	*3,787*	*70.5*
		6	High income (Above £20,000)	236	4.4
3	Housing type	***7***	***Commercial housing (private property)***	*2,318*	*43.1*
		8	Purchased public housing	704	13.1
		9	Rented housing (private)	670	12.5
		10	Social rental housing	143	2.7
		11	Economically affordable housing/ Price-capped housing	250	4.7
		12	Self-built housing in the countryside	815	15.2
		13	Other	473	8.8
4	Education stage	***14***	***Primary school***	*2,838*	*52.8*
		15	Middle school	1,353	25.2
		16	High school	1,182	22.0
5	Car ownership	17	***Own car***	3371	62.7
		18	No car	2,002	37.3
Sub-district level variables					
6	Geodemographic classification	19	Long settled Beijing Society	*234*	*4.4*
		20	Urban-rural transition belt	629	11.7
		21	In-migrants in deprived suburbs	492	9.2
		22	High income internal migrants	863	16.1
		23	***Employees of state-owned enterprises and white collar workers***	1,044	19.4
		24	Outside Central Beijing	2,111	39.3
7	School density	25	School density of primary and secondary schools in each areas* (continuous variable)		

*The school density is calculated by using the number of primary schools and secondary schools within each sub-district to divide the area of each sub-district.

The pupils’ travel distance from home to school is the response variable, however, it is significantly positively skewed (Figure 7.18a). Methodologically, the response needs to be transformed to the logarithm of the pupil’s travel distance before commencing modelling as including a skewed response in the models is more likely to generate skewed residuals (Shuttleworth and Gould, 2010). Figure 7.18b, which is the distribution of transformed response, shows a nearly normal distribution.



a. Raw data



b. Log-transformed data

Figure 7.18 Histograms of travel distance in CHTS

7.4.2 Multilevel framework for modelling and model strategy

In terms of the modelling strategy, one two-level 'null' model (no predictors) with random intercepts (Model 1) has been specified. It is a constant only model in which pupils (level-1) are nested in sub-districts (level-2) of home. Since the total variability in the logged distance is allowed to be partitioned across the different levels, the null models are used to inspect whether there is strong evidence of between sub-district differences in logged travel distance. Thereafter, compositional differences and context effects are considered by introducing the individual level (level-1) and sub-district level (level-2) covariates into the fixed part of the model respectively. Including the micro-level covariates in the models not only can reveal their influence on variations in pupils' travel distance but also can help identify the context effects by controlling for the compositional differences. Two contextual variables, geodemographic classification and school density within each sub-district have been included in Models 2 and 3 respectively. The sub-district indicators are used to explore part of the influence of the context effects of home sub-district. With the aim of uncovering the varied variability of travel distance for different groups, the level-1 heterogeneity for the individual level predictors, hukou, education stage, income and housing, are modelled in Model 4-7 respectively.

The IGLS algorithm (see Section 3.4.4), based on the maximum likelihood (ML) estimation, is used in this research to provide a fast and reliable model estimation. All the above models are estimated in MLwiN. The Wald test has been used to assess the significance of fixed part estimates, while the likelihood ratio test (LRT) has been employed to test the hypothesis on variance components, test the overall model fit and the improvement in model fit. The detailed principles of IGLS and LRT have been given in Section 3.4.5.

7.4.3 Model results

7.4.3.1 Variance component model

The first model, Model 1, is initially specified as a two-level variance component model for pupils nested within their sub-district of residence (Table 7.5). Although the majority of variation is still found between individuals, there is also strong evidence of contextual variations. In Model 1, the sub-district of home variation is estimated to account for 26.8%[8] total variance in pupils' travel distance. As the micro-level predictors have not been included in the above models, it is hard to conclude that the substantial variations at the higher level revealed by the above model are due to the place based differences (Thomas *et al.*, 2015). Thus, micro-level compositional predictors should be taken into account before any exploration of potential context effects or patterns.

[8] The variance partition coefficients (VPC) (Section 3.4.2.1) are calculated as :

$$\rho = \frac{\sigma_{u0}^2}{\sigma_{u0}^2 + \sigma_{e0}^2}$$

Thus the proportion of variation at sub-district level in Model 1 is:
0.053/(0.053+0.145).

Table 7.5 Results from fitted multilevel models (model 1-3)

Parameter	**Model 1**	**S.E.**	**Model 2**	**S.E.**	**Model 3**	**S.E.**
Fixed part						
Constant	3.426	0.016	3.156	0.034	3.293	0.017
The base category for Model 2 and Model 3 : Hukou within a district; middle income (£5,000-20,000); commercial housing (private property); primary school; own car; 'Employees of state-owned enterprises and white collar workers' (Model 2); mean school density (Model 3)						
The hukou status of the head of household						
Hukou outside a district but within a municipality			0.086	0.023	0.086	0.023
Hukou outside a municipality			-0.027	0.016	-0.028	0.015
Income of household						
Low income (Below £5,000)			-0.012	0.013	-0.012	0.013
High income (Above £20,000)			0.072	0.025	0.071	0.025
Housing type						
Purchased public housing			0.026	0.018	0.023	0.018
Rented housing (private)			-0.025	0.018	-0.025	0.018
Social rental housing			0.032	0.036	0.041	0.036
Economically affordable housing			0.066	0.026	0.065	0.026
Self-built housing in the countryside			0.117	0.022	0.116	0.021
Other housing types			0.021	0.021	0.021	0.021
Car ownership						
No car			-0.063	0.011	-0.063	0.011
Education stage						
Middle school			0.145	0.012	0.143	0.012
High school			0.348	0.013	0.347	0.013
Sub-district variable						
School density					-0.102	0.011
Urban-rural transition belt			0.337	0.074		
In-migrants in deprived suburbs			0.264	0.056		
Long settled Beijing society			-0.058	0.056		
High income internal migrants			0.117	0.05		
Outside Central Beijing			0.223	0.04		
Random Part						
Level-2: Between sub-district of home variance	0.053	0.006	0.037	0.004	0.033	0.004
Level-1: Between-individual variance	0.145	0.003	0.124	0.002	0.124	0.002
Units: Sub-district of home	251		251		251	
Units: Individual (pupils)	5373		5373		5373	
Estimation:	IGLS		IGLS		IGLS	
Deviance (-2*loglikelihood)	5357.709		4458.464		4438.925	

7.4.3.2 Random intercept models with covariates

The individual level covariates (hukou, income, housing type, car ownership and education stage) and sub-district level context covariates (school density and geodemographic cluster) thus have been included in Model 2 and Model 3, leading to substantial model improvement. Models 2 and 3 share the exactly same level-1 predictors but contain different sub-district level covariates. The level-2 covariate employed in Model 2 is the bespoke geodemographic classification, whereas this is replaced by school density in Model 3.

The results for the fixed part of the models have validated most of the expectations in previous descriptive analysis. All the elements are incorporated in the fixed part as additive effects. Most of the fixed parameters are more than twice their standard errors (S.E.), which indicates most of the coefficients of parameters are statistically significant and have been validated by using the Wald test. In addition, the random terms are also significant according to the likelihood ratio test (LRT). In Model 2, the transformed estimate for the constant shows that the base category pupils are expected to travel 1.43km to school. For the individual level predictors, the pupils' education stage provides the largest effects with high school pupils travelling 1.76km and middle school pupils travelling 0.57km further compared to primary school pupils, holding the other covariates constant. Two types of housing also influence pupils' travel distance significantly: self-built housing in the countryside and economically affordable housing. The pupils living in self-built housing travel an extra 0.44km and pupils living in economically affordable housing travel an extra 0.24km. Hukou status also affects pupils' commute distance. The pupils with hukou outside their district of residence but within Beijing travel an extra 0.31km when other effects are controlled for, which is consistent with our assumptions in Chapter 6. However, the effects of having hukou outside Beijing is not very significant (the magnitude of the coefficient (-0.027) is smaller than twice its standard error (0.016) and it is also not significant according to Wald test). This is inconsistent with our expectation and indicates that the distance travelled by pupils who do not have Beijing hukou is not significantly different from the pupils who have hukou within the district.

In regard to the income of the head of household, the pupils in high income families also tend to travel 0.26km further when controlling for other factors. The insignificant effect of low income is due to the inclusion of a highly correlated variable, car ownership. Before including the car ownership variable, the effect of low income is significant and negative (in a model which has not been shown in Table 7.5). However, it becomes insignificant when car ownership is included, which is due to the highly correlated relationship between car ownership and income. As the joint Wald test of categories of income is significant and the coefficient of car ownership (no car) is considerably higher than that of low income (in the model which has not been shown), both income and car ownership have been included in the models.

Except for the pupils' education stage, the largest differential effects on pupils' travel distance are found in the contextual characteristics. For the geodemographic clusters, the pupils living in the cluster distinguished as 'Employees of state-owned enterprises and white collar workers' are in the base category, so the estimates of other clusters are interpreted by comparing them with the distance travelled by pupils living in this cluster. Pupils who live in clusters labelled 'Urban-rural transition belt', 'In-migrants in deprived suburbs', 'Outside Central Beijing' and 'High income internal migrants' travel extra 1.68km, 1.20km, 0.96km and 0.44km compared to pupils living in the base category. The mean distance travelled by pupils in the 'Long settled Beijing society' cluster is not statistically different from pupils in the base category. Overall, these results validated our findings in Chapter 6.

In Model 3, school density replaces the geodemographic clusters as the contextual variable and displays very similar estimation results. As the constant of the model refers to the logged travel distance for the pupils in the base category and the base categories of Model 2 and Model 3 are different ('Employees of state-owned enterprises and white collar worker' in Model 2; mean school density in Model 3), their constants are different. However, the estimates of the fixed part and random parts for the two models are very close. The pupils' travel distance decreases by 0.30 km when the school density in the sub-district of home increases by one unit, holding all the other variables constant. Figure 7.19 shows the boxplot of school density for the geodemographic classifications. The clusters labelled 'Long settled Beijing society' and 'Employees of state-owned enterprises and white collar workers', with the lowest travel distance (Table 7.5) have the highest school density (Figure 7.19). In contrast to the areas labelled 'Urban-rural transition belt', 'In-migrants in deprived suburbs' and 'Outside Central Beijing', with the highest travel distances, have the lowest school densities.

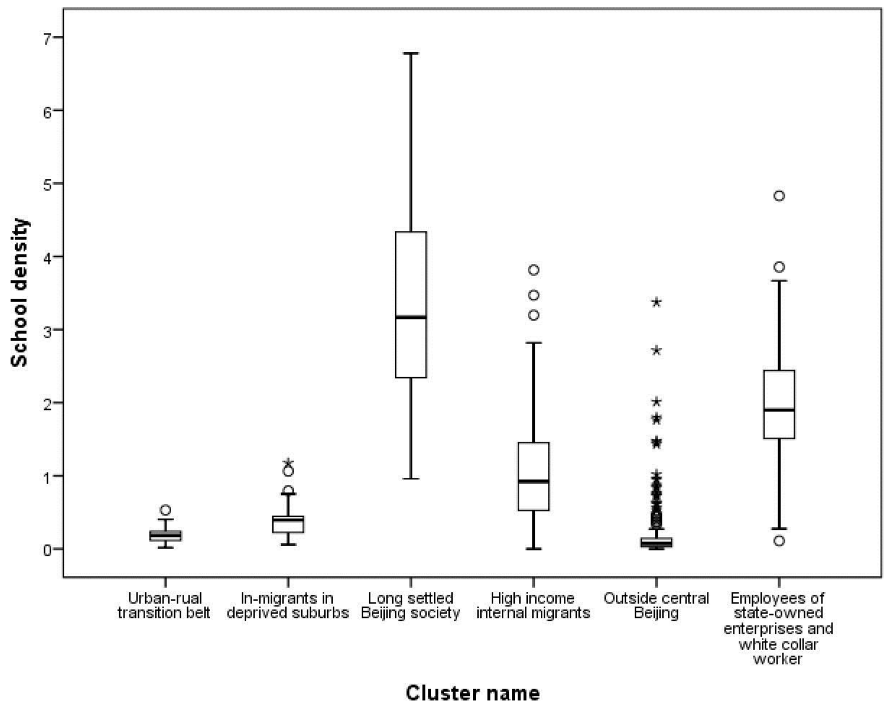


Figure 7.19 Boxplot of school density and geodemographic classification for sub-districts of home

As the fixed parts of the two models are very similar, it is apparent that geodemographic cluster and school density have parallel effects in these models. In other words, the travel distance for each cluster is probably influenced by the school density within each cluster. It is worth noting that the school density distribution for sub-districts outside Central Beijing is comparatively dispersed (Figure 7.19), which is due to this cluster not being defined based on the socioeconomic characteristics of residents within these areas and only refers to the places outside Central Beijing. The coarse definition of this cluster leads to a large number of outliers and the school density of some sub-districts, like new towns, in 'Outside Central Beijing' is already as high as that of the clusters 'Long settled Beijing society' and 'Employees of state-owned enterprises and white collar workers'. Moreover, the deviance of Model 3 (4438.925) is smaller than that of Model 2 (4458.464), which represents fewer unexplained observations in Model 3. It is probably also due to the context effects of the sub-districts within ‘Outside Central Beijing’ which can be better explained by school density than by the generalised cluster. However, the inclusion of geodemographic groups in the model has the advantage of revealing the relationship between the pupils’ travel distance and the socio-spatial characteristics of their residential areas directly.

7.4.3.3 Modelling level-1 heterogeneity between pupils

The differential variability between pupils with different characteristics is explored by fitting the models and relaxing the assumptions that there is equal variance for the level-1 error term (Table 7.6). Models 4-7 permit hukou status, education stage, income and housing type to vary at level-1. As the four models are only different in their level-1 random part and they contain the same fixed part variables, the estimates of the fixed part of the four models are quite close to each other and to those of Model 2. The level-1 heterogeneity between individuals with different hukou types (Model 4) reveals that pupils with hukou outside a district but within a municipality have more variability compared to the other two hukou types. This is likely because the pupils, whose current residence is inconsistent with their places of registration, not only have the chance to enrol in nearby schools where they reside but can also enrol in faraway schools in their places of registration. Thus, their individual level variation tends to be larger than other groups.

Model 5 shows that high school pupils have greater variation in pupils’ travel distance. It is probably not only because the high school pupils are old enough to travel further but also because high school pupils are enrolled across the districts within Beijing and thus have more flexibility in their school choice. Model 6 shows that the pupils with higher income tend to have higher variability at level-1. The pupils from high income households have the highest variability, while the pupils in low income households have the lowest level-1 variation. This suggests that high income groups have more mobility in their pupils’ daily commute which is probably because they are more likely to own cars and have more ‘social capital’ to make school choices for their children. The final model, Model 7, shows that the pupils living in economically affordable housing have the largest variation in their travel distance, whereas the pupils living in self-built housing in the countryside have the lowest variation. The positive fixed part coefficients for the two housing types indicate that the pupils living in these types of housing tend to travel further. However, the individual level variation of self-built housing is small, but that of economically affordably housing is large.

Table 7.6 Results from fitted multilevel models (model 4-7)

Parameter	**Model 4**	**S.E.**	**Model 5**	**S.E.**	**Model 6**	**S.E.**	**Model 7**	**S.E.**
Fixed part								
Constant	3.156	0.034	3.159	0.035	3.16	0.035	3.157	0.036
The base category: Hukou within a district; middle income (5,000-20,000); commercial housing (private property); primary school; own car; ʻEmployees of state-owned enterprises and white collar workers’								
The hukou status of the head of household								
Hukou outside a district but within a municipality	0.086	0.025	0.089	0.023	0.085	0.023	0.083	0.024
Hukou outside a municipality	-0.028	0.016	-0.029	0.015	-0.027	0.016	-0.024	0.016
Income of household								
Low income (Below 5,000)	-0.012	0.013	-0.014	0.013	-0.013	0.012	-0.011	0.012
High income (Above 20,000)	0.071	0.025	0.075	0.024	0.071	0.027	0.07	0.025
Housing type								
Purchased public housing	0.026	0.018	0.016	0.018	0.024	0.018	0.026	0.019
Rented housing (private)	-0.024	0.018	-0.026	0.018	-0.021	0.018	-0.028	0.018
Social rental housing	0.031	0.036	0.029	0.036	0.031	0.036	0.033	0.035
Economically affordable housing	0.066	0.026	0.064	0.026	0.069	0.026	0.07	0.029
Self-built housing in the countryside	0.117	0.021	0.109	0.021	0.118	0.021	0.102	0.02
Other housing types	0.022	0.021	0.018	0.021	0.023	0.021	0.022	0.02
Car ownership								
No car	-0.063	0.011	-0.062	0.011	-0.064	0.011	-0.06	0.01
Education stage								
Middle school	0.144	0.012	0.145	0.012	0.143	0.012	0.141	0.012
High school	0.348	0.013	0.349	0.014	0.35	0.013	0.348	0.013
Sub-district variable								
Urban-rural transition belt	0.338	0.074	0.334	0.075	0.331	0.075	0.342	0.076

In-migrants in deprived suburbs	0.264	0.056	0.262	0.057	0.263	0.057	0.268	0.058
Long settled Beijing society	-0.058	0.057	-0.056	0.057	-0.061	0.057	-0.061	0.059
High income internal migrants	0.118	0.05	0.11	0.05	0.109	0.05	0.113	0.052
Outside Central Beijing	0.224	0.04	0.23	0.04	0.222	0.04	0.237	0.041
Random Part								
Level-2: Between sub-district of home variance	0.037	0.004	0.038	0.004	0.038	0.004	0.041	0.004
Level-1: Between-individual variance								
Variance：Hukou within a district	0.121	0.003						
Variance：Hukou outside a district but within a municipality	0.148	0.013						
Variance：Hukou outside a municipality	0.129	0.006						
Variance：primary school			0.113	0.003				
Variance：middle school			0.119	0.005				
Variance：high school			0.155	0.007				
Low income (Below 5,000)					0.098	0.004		
Middle income (5,000 - 20,000)					0.131	0.003		
High income (Above 20,000)					0.155	0.015		
Variance：Commercial housing							0.135	0.004
Variance：Purchased public housing							0.138	0.008
Variance：Rented housing (private)							0.124	0.007
Variance：Social rental housing							0.112	0.014
Variance：Economically affordable housing							0.156	0.014
Variance：Self-built housing in the countryside							0.074	0.004
Variance：Other housing types							0.112	0.008
Units: Sub-district of home	251		251		251		251	
Units: Individual (pupils)	5373		5373		5373		5373	
Estimation:	IGLS		IGLS		IGLS		IGLS	
Deviance (-2*loglikelihood)	4451.808		4416.908		4416.467		4357.895	

It worth noting that some other potentially important micro-level predictors (e.g. the relationship with the head of household, the highest qualification and occupation) had been included in preliminary models to attempts to explain the possible remaining variations at individual level. However, these variables have not been included in the final model as their inclusion was neither statistically significant nor important.

7.5 Discussions and policy implications

With the help of detailed descriptive analysis and multilevel modelling, not only the influence of microlevel covariates and the context effects of macrolevel variables have been identified, but some implications relating to the social and spatial disparities of pupils in terms of access to education also have been revealed and will be discussed in this section.

7.5.1 Spatial mismatch of education provision and consumption

The descriptive analysis in Section 7.2 revealed a very uneven distribution of school places in Beijing. In addition, the available school places were not matched with the school age pupils in each district, as indicated by the sufficiency index shown in Figures 7.2 and 7.3, which reveal a significant spatial mismatch between education provision and consumption. This has led pupils to commute long distances in some places in Beijing, sometimes across district boundaries in search of educational opportunities. The districts with sufficient (spare) education resources are the areas with net inflows of pupils, while the districts where there are insufficient education resources are the districts with high net outflows of pupils. The pupils living in the areas lacking enough resources have to commute long distances. If the long travel distance is attached to disadvantaged social groups (e.g. non-local or low income residents), it may increase their financial burden, decrease the opportunities of pupils receiving education, produce or reproduce disadvantages and result in a vicious cycle.

7.5.2 The influence of individual level characteristics

7.5.2.1 Hukou

The multilevel models show that the influence of the joint effects of varied hukou status, which are a unique and critical factor under China's specific institutional background, are statistically significant. Surprisingly, the travel distance of pupils whose hukou is outside Beijing is not significantly different from that of the base category pupils whose hukou is within a district. This contradicts with our assumption that in-migrants are likely to have longer commute distances. One reason for this result is probably due to the coarse division of the hukou, where the hukou statuses (hukou within a sub-district and hukou outside a sub-district but within a district) are combined as hukou within a district. The travel distance of pupils whose hukou is within a sub-district may be shorter than that of non-local pupils, as the pupils with non-local hukou will be assigned to a school within the range of the whole district. However, the pupils who have hukou outside a sub-district but within a district also have the possibility to commute long distances (discussed in Chapter 6). They also may commute across sub-district boundaries and enrol in schools within the range of the district as non-local pupils. It is probably why the travel distance of in-migrant pupils is not statistically significantly longer than that for those with combined hukou status, i.e. hukou within a district.

In addition, although the pupils who do not have Beijing hukou will normally be assigned to schools with spare places available to them (undersubscribed schools) within range of the whole district, most of them live in rented accommodation (which is more flexible when it comes to move), thus, they will probably move to residences close to assigned schools to reduce their travel costs. Moreover, the accommodation which is next to unpopular schools will not be very expensive and will tend to be affordable for in-migrant families. The above assumptions not only provide a possible explanation for the shorter commute distances for in-migrant children; they also suggest a location adjusting process for in-migrants' accommodation led by their children's enrolment. As there is a high proportion of non-local pupils in unpopular schools (Bi and Zhang, 2016), their accommodation adjustment may lead to the emergence of some in-migrant concentration areas near unpopular schools. As a consequence, the educational institutions (e.g. enrolment policy) will impact on the socio-spatial transformations within the urban area and lead to residential segregation based on school segregation.

Although there is not enough proof to demonstrate the commute distances of in-migrant children are long, they are indeed deprived of chances of access to high quality education. In contrast, the pupils whose hukou outside a district but within a municipality are proven to travel significantly longer distances than the other hukou statuses according to the model results. Their longer commute distances are more likely to be due to their school choices based on the separation of their current residence and the place of hukou registration. The quality schools they prefer are located in sub-districts where they registered rather than their current residence, which will lead them to commute longer distance and across district boundaries. This is also validated by the breakdown of the nearby enrolment rate (enrol in the same district) by different hukou statuses in Section 7.3.1.2, which indicates that only 61% pupils with hukou outside a district but within a municipality enrol in the same district (hukou within a district 83%; hukou outside Beijing 89%).

7.5.2.2 Income and housing

The coefficients for income shown in the models in Section 7.4.3 indicate that the high income group commutes further than the middle income group, while the low income group or the group without a car travel shorter distances. This is consistent with the conclusions from existing research in western countries (Ball *et al.*, 1995; Waters, 2017) and China (Bi and Zhang, 2016; Wu *et al.*, 2016). Ball *et al.*(1995) pointed out that the middle-class households are able and willing to travel longer distances to access better educational opportunities compared to working-class families. Bi and Zhang (2016) also found the most ‘powerful’ natives or non-natives tend to actively choose quality schools without necessarily living within the catchment area. According to the school choice pattern depicted in the above studies, the high income groups are more likely to own a car and have enough capital to make a school choice for their children, thus tending to travel longer to enrol in a better school. Hamnett and Butler (2013) suggest that advantaged parents not only seek out quality schools but also avoid schools with a high proportion of pupils from disadvantaged backgrounds. Similarly, the high income group in Beijing also not want their children to enrol in schools with a high proportion of migrant children (Ming, 2014). In contrast, the low income group tend to send their children to nearby schools irrespective of the school quality. However, the pupils who live in affordable housing and also tend to be in the low income group, are also shown to travel longer distances in the models that have been calculated.

The models have demonstrated that the pupils who live in economically affordable housing tend to travel further and the result is statistically significant. This is likely due to the significant spatial mismatches between economically affordable housing and public facilities (e.g. schools), which also has been pointed out by other scholars (Yang *et al.*, 2014; Chen *et al.*, 2015). The economically affordable housing, including the Economic and Comfortable Housing (ECH) and Capped-Price Housing (CPH), is subsidised housing targeted at low- and medium-income households with local hukou (local resident registration) (Yang *et al.*, 2014). Increasingly, studies denote that the poverty of households is not only due to the absence of financial resources or accommodation, but also occurs because of the lack of access to urban resources and the ability to function effectively in society (Sen, 2001; Small and Newman, 2001). Thus, a housing system should not only solve the habitation problems for the low- and medium-income households, but also should foster their well-being and social status by improving their conjunction with public services (Curley, 2005).

In terms of education, accessibility to schools can strongly influence their opportunities to receive education. A well-designed affordable housing programme can help low income households to compensate for their disadvantages with better access to education facilities. Compared to higher socioeconomic groups, access to education facilities could impact more on the education of low socioeconomic groups due to the fact that they are more sensitive or vulnerable to the cost and their mobility barriers (e.g. not owning cars) *(Yang et al.*, 2014). Thus, the long travel distance of pupils living in economically affordable housing revealed by the models might put low income groups in an even worse situation and lead to further economic disparities. Therefore, government action on prompting school supply should be integrated into affordable housing policy design.

The distance travelled by pupils living in self-built housing in the countryside is also longer, with a very small individual level variation or heterogeneity compared to other housing types. It reflects the universal poor education provision for suburban areas or rural areas in Beijing, an issue that should also attract more attention from the government in terms of education provision given the rapid expansion of urban areas in Beijing.

7.5.3 Socio-spatial education inequality

All of the above discussions related to the influence of individual level covariates in variations of pupils' travel distance are conducted based on controlling for the spatial context effects. With the help of multilevel modelling, the dependent geographical (neighbourhood) context effects have been explored with the additional benefit of being able to identify pupils from areas where pupils tend to commute longer distance by including the effects of macro-level variables and allowing for the unexplained macro-level variance. The cluster-level variables, such as school density and geodemographic classification of each sub-district, are called contextual variables, thus their effects on response are referred to as contextual effects. The results of the models show that there are very significant geographical context effects on pupils' travel distance, and the effects of macro level covariates outperform most of the individual level covariates except for education stage.

As shown in Figure 7.1, the spatial variations of school density are very dramatic, indicating the distribution of school places is very uneven in Beijing. School density was thus included in multilevel models to illustrate the effect of school density on pupils' travel distance and explore whether low school density will lead to longer travel distances. The coefficients of school density validated our assumption that the sub-districts with high density would lead to low commute distances and vice versa. According to the school density distribution shown in Figure 7.1, school density decreases from inner areas to outer areas in Central Beijing and decreases from some new towns to outer areas outside Central Beijing; correspondingly, the travel distance increases from inner areas to outer areas in Central Beijing and increases from some new towns to outer areas outside Central Beijing.

As the inclusion of school density can only reveal spatial equalities, the other sub-district level variable, the geodemographic classification of each sub-district, has been included in the models to associate geographical spaces with the socio-economic characteristics of their compositional population. The results show that those pupils that live in socially disadvantaged clusters, such as in areas classified as 'Urban-rural transition belt' and 'In-migrants in deprived suburbs', tend to travel longer distances, while the socially educational advantaged groups, such as those living in 'Long settled Beijing society' and 'Employees of state-owned enterprises and white collar workers' travel shorter distances. The boxplot of the school density distribution for sub-districts by the bespoke clusters shows that the sub-districts in disadvantaged clusters, such as 'Urban-rural transition belt' and 'In-migrants in deprived suburbs', also have lower school density. Therefore, the longer travel distances in these areas are very likely to occur because of the low school density in these areas. With the help of the clusters, the socio-spatial inequity of education in terms of travel distance has revealed that the concentration areas of the disadvantaged groups, i.e. the areas with higher percentage of disadvantaged groups, tend to have lower school density (or insufficient educational resources) and force pupils to travel longer, after controlling for the individual level variations. Compared to socially advantaged groups, these disadvantaged groups in these areas are more vulnerable to travelling longer travel distances. Thus, intervention measures, such as the construction of more schools in the socially disadvantaged places to reduce the distance of travel, can be designed and introduced to improve the education provision and social well-being in educationally disadvantaged areas.

7.6 Conclusions

Using detailed microdata, CHTS, and employing multilevel modelling, this chapter reveals how household socio-economic characteristics influence pupils' travel distance; how spatial context, such as school density and social composition of sub-district, affect pupils' commute distances; and how these factors lead to social and spatial education inequities. Given the suitable data in CHTS, multilevel models are able to analyse the influence of macro level context variables, while keeping individual level characteristics as the object of study (Owen *et al.*, 2016). Compared to the 'unpatterned-inequality' (Talen, 2001) described in previous research, one of the most important contributions of this research is that the results of the analysis have identified, for the first time the relationship between pupils' access to school and their socioeconomic status, when controlling for the varied context effects of the sub-districts where pupils live. Meanwhile, the geography or context effects have been introduced in the models without ignoring the effects of individuals, which is one of the advantages of employing multilevel modelling (Pickett and Pearl, 2001). In this way, the analysis avoids not only the ecological fallacy (Robinson, 1950) but also the atomistic (or individualistic) fallacy (Subramanian *et al.*, 2009).

At the individual level, the longer distances are travelled by those with advantaged socio-economic status, such as higher income, hukou outside a district but within a municipality, and car availability, seeking for better education opportunities for their children. The people living in economically affordable housing or self-built housing in the countryside, who are normally socially disadvantaged, however, also are more likely to travel longer distances, which is due to the inadequate school education provision around them. The pupils in disadvantaged socio-economic status, like those with low income and no car, tend to travel shorter distances, which indicates that they are not deprived in terms of travel distance, however, they are more likely have less access to high quality education. Similarly, the influence of hukou outside Beijing on travel distance is not statistically significant, which might be due to the coarse definition of the base category or because of their more flexible housing mobility. However, there is no doubt that the non-local pupils are also deprived of high quality education. In addition, their accommodation search process probably leads to the concentration of non-local households around some unpopular schools and results in residential and school segregation in these areas.

The modelling results also confirmed the importance of contextual effects of school density and geodemographic classification, on pupils' travel distance, which is conditional on the individual level predictors. School density represents the education provision in pupils' sub-district of residence, while the geodemographic clusters summarise the principal features of the residents within each sub-district based on a set of their socio-economic characteristics. The model reveals that school density negatively influences pupils' travel distance such that lower school density tends to lead to longer travel distance. School density is very unevenly distributed in Beijing so that the density decreases significantly from inner areas and some new towns of Beijing to outer areas. Thus, the model reflects a very dramatic spatial inequality on pupils' access to school due to the uneven school density distribution. Unlike the education provision under counter urbanisation in Western countries like the UK (Hamnett and Butler, 2011), Beijing is still in the process of rapid urbanisation. Rapid urban expansion and comparatively lagging education planning mean that education provision is insufficient in outer areas. In addition, the characteristics of sub-districts of residence play important roles in conditioning the pupils' desire and ability for enrolment and school choice. The significant context effects of the geodemographic clusters on pupils' access to school, travel distance, indicate that the pupils in areas of potential educational disadvantage, such as in the 'Urban-rural transition belt' or in areas with 'In-migrants in deprived suburbs', also tend to commute longer distance with lower school density. As there is a higher proportion of socio-economic disadvantaged groups living in these areas, which are more vulnerable to the cost of longer travel distance, the inclusion of clusters in the model revealed very dramatic social and spatial education inequalities in Beijing.

Understanding how different individual level factors or geographical factors influence pupils' travel distance is crucial for addressing spatial inequalities in terms of access to education resources and introducing proper interventions to promote education justice. Insufficient education provision and low school density in outer areas are due to the lack of effective school planning and rapid urban expansion. In addition, the education requirements of migrant children have not been properly considered by school planners in local government. In order to bring the current education provision and demand into equilibrium, school planning in outer areas should be undertaken proactively. Moreover, the education demand of migrant children should be considered in education planning, especially where there are perceived concentration areas, such as in the 'Urban-rural transition belt' and 'In-migrants in deprived suburbs' areas. Furthermore, for the low income groups, the effects of construction of affordable housing should not only focus on making housing economically viable for those people but the household well-being, like education opportunities, should not be ignored. Attention should also be directed towards creating liveable neighbourhoods for low income urban groups to promote social equality and social mobility.

Chapter 8
Conclusions and Future Recommendations

8.1 Introduction

This thesis has addressed a number of research questions (see Section 1.2) that have been formulated in relation to the geographies of educational inequality in China, a much under-researched area of study. The identification of these questions provided the rationale for setting the following two broad research aims:

- to accurately measure existing regional and rural-urban education inequalities and quantify how they have changed over time so as to assess the extent to which the policies have been successful in reducing regional education inequalities, and
- to reveal the education inequalities existing within urban areas by exploring the influence of residents' socio-economic characteristics and areas' specific socio-economic composition on pupils' access to education.

The thesis has successfully addressed these two aims by establishing a series of seven objectives that form the basis for analysing education inequalities in China at different geographical scales by using appropriate quantitative methods. Section 8.2 contains summaries of the extent to which each objective has been achieved and the key findings emerging from the empirical analyses undertaken. The findings of this research have implications for the actions that Central or local governments might take to alleviate problems and improve opportunities for different sections of the population living in different types of environment to benefit from education. These implications and recommendations for policy and practice are discussed in Section 8.3, whilst the limitations of the research are considered in Section 8.4 and the potential for future studies are presented in Section 8.5.

8.2 Research summary

The objectives formulated and presented in Chapter 1 will now be revisited and discussed in turn in this subsection to demonstrate how each has been achieved.

8.2.1 Objective I

Objective I - To review the concept and discussion of education equality; give the definition of education equality used in this research; explore and review key research literature on education inequality from different geographical scales; introduce the geographies, specific institutional background and education system in China; build a research framework for this research based on the given theoretical and empirical context.

Through the reviews of concepts and policy in Chapters 2, 4, 5, 6 and 7, the thesis provides an essential theoretical, empirical and institutional context for analysing education inequalities from different geographical scales in China. Given that education can be equal and unequal in many different ways and pervasive, it was necessary to define education equality in this research, so as to establish a firm basis on which to measure and regulate education inequalities. The concept of education equality has been considered from various perspectives, and the main differences among various egalitarian theories revolve around two core questions: ‘equality for whom’, which indicates the target individuals and groups for comparison; and ‘equality of what’, which refers to the specific themes or focal variables for measurement. Spatial education inequalities are also pervasive forms of education equality and usually reflect other forms of education inequality, like education inequalities among different social groups. Moreover, spatial inequalities in education can also create or exacerbate education inequalities, thus great attention should be paid to spatial education inequalities.

A three-dimensional framework was proposed to define education equality in this study, incorporating ‘equality of where’ as one important dimension. Compared to the previous two-dimensional (non-spatial) understanding of education equality, the given framework in this research, defining education equality from the dimensions, where, whom and what, is more comprehensive. In addition, it is also flexible for researchers to define or analyse education equality according to their specific applications. The 'equality of where' refers to the geographical scales and target areas for comparison, therefore education inequalities can be investigated at national, regional and local scales, etc. Since inequality at different geographical scales may be shaped by varied social practices and have their own underlying theoretical base, different determinants are involved. The framework also emphasises the 'vertical' links between various scales of geography, which can reveal the hierarchical or multilevel structure of phenomenon under investigation. Accordingly, this research provides an open ended and relational framework to assist scholars to define, evaluate and analyse education inequalities at different spatial scales.

Empirical research in geography demands the selection of an appropriate scale (or scales) with which to examine and analyse a particular spatial phenomenon. This research has investigated education inequalities in China from three scales, provincial, local and individual scales. At the provincial scale, the thesis has demonstrated a new metric to measure regional education inequality and evaluate the effectiveness of regional education policies.

With the current rapid urbanisation in China, there are increasingly more migrants moving from rural areas or less developed areas to big cities in search of better job opportunities and better welfare, including high quality education. Inequalities between local residents and non-local residents emerge and interwoven with inequalities among local residents with varied socio-economic characteristics. Due to the absence of direct educational data relating to small areas, it is apparent that there is very little research about education inequalities within urban areas. As Beijing has experienced rapid urbanisation and a large influx of in-migrants, it is selected as the case study area to examine education inequalities at the local scale.

In western countries, the influence of 'Where they live', which refers to the geographical location and social composition of the residential neighbourhood, on pupils' access to schools is frequently with respect to the residential segregation (e.g. ethnicity and social class). However, due to the large number of in-migrants and the restrictions of the hukou system, education related socio-spatial differentiation in China’s cities, like Beijing, are more complicated and tend to be formed by multiple factors (e.g. income, occupation and hukou status). This research has investigated education inequality in terms of social-spatial access in 136 sub-districts in Central Beijing by using a new data exploration method, geodemographics, and available census data.

Moreover, given the lack of research to explore how socio-spatial structures (where they live) and individual characteristics (who they are) jointly affect one’s access to education resources, multilevel modelling has been used as a suitable modelling framework to deal with such complexities.

8.2.2 Objective II

Objective II - To review the methodological approaches and critically introduce corresponding data used in analysing educational inequality in China.

The review of various methodological methods and dataset used in this research was essential to provide necessary context and justification for the quantitative analysis undertaken in the thesis. In order to evaluate the effectiveness of policies targeted at reducing regional/provincial educational inequalities in China since 2005, a new multidimensional index, the Index of Regional Education Advantage (IREA), was introduced underpinned by Amartya Sen’s capability approach. Compared to previous narrowly defined measurement approaches, which focus on attainment or resources, the proposed index is more comprehensive in terms of the dimensions that are included. The measures based only on actual attainment ("functioning") or input resources are unable to capture issues of context and past history in enabling people to exercise their true abilities. Thus education equality has been conceptualised and evaluated beyond the ‘basic’ measures, and the capability approach has been applied to provide a theoretical basis for the measurement of educational disparities.

The IREA is not only concerned with the allocation of resources (provision) but also considers the ability of areas to convert resources into capabilities (enrolment and attainment). It provides a more comprehensive measurement method which can capture the real regional education inequality and it also can be directly linked to the analysis of individual capability.

Compared to previous monitoring indices, such as the indices of Education for All (EFA) of UN, the IREA index has been given a theoretical justification. In addition, it is a more suitable measure for providing feedback and incentives to policymakers compared to using education attainment (AYS) on its own. This analysis and evaluation framework was used to evaluate real regional education advantaged and disadvantaged areas in Chapter 4 and filled the current theoretical void. It also provides a reference for measuring regional education inequalities in other countries or contexts.

This thesis has used data from various sources to undertake analyses at different spatial scales but within the limits imposed by availability. All the data sources are characterised by varying degrees of accuracy, coverage and detail and have varied weaknesses and strengths. Administrative sources of data have an advantage in their timeliness, comprehensiveness, reliability and complete regional scale geographic coverage. Thus the data adopted to construct the IREA index are all obtained from the Educational Statistical Yearbook and Educational Finance Statistical Yearbook of China. According to the proposed capability approach-based framework, 17 education-related variables across three dimensions (enrolment, attainment and provision) are integrated into a single score for each provincial level unit to quantify the extent of educational inequality between different provinces in China. However, only a limited number of attributes were available for rural and urban areas within each provincial level unit and so the comparison of rural-urban education inequalities was limited accordingly.

At the local scale, given the absence of direct data on education relating to small areas, geodemographics has been used to yield new ideas and insights about potential inequalities in education within urban areas. Geodemographic classification is able to assign a category to similar neighbourhoods according to the characteristics of their residents and thus provides a framework for understanding the characteristics and behaviour of people who live in these neighbourhoods. It is thus usually regarded as a multidimensional technique to measure socio-economic differentiation. As the level of children’s access to education services can be strongly influenced by the geographical location and the social composition of the residential neighbourhood, the potential socio-spatial education inequality in terms of access to education can be investigated by the bespoke geodemographic classification. Within urban areas, the decennial census data are the most reliable and comprehensive data source and the sub-district is the smallest geographical unit in census data, thus geographical inequalities within Central Beijing have been examined at the sub-district level by using geodemographics to simplify a range of education-related attributes into distinct spatial clusters. The application of geodemographics in this research provides a point of reference for other cities, where educational research is also limited by the availability of data for small areas.

The third method, multilevel modelling, has been used in this research to analyse the factors that influence the distance that pupils travel to school in Beijing. The education system has a clustered or hierarchical structure with pupils nested within the sub-district where they live. As discussed in the literature review section, the pupils' access to school is not only influenced by 'who they are', the pupils' socio-economic characteristics, but also 'where they live', the context effects. Multilevel modelling is a suitable method to deal with such complexities as it can be applied to various hierarchical datasets and allows the identification of heterogeneity between pupils (micro level) but also heterogeneity between sub-districts (macro level) (Bullen *et al.*, 1997). In addition, as the approach can analyse contextual effects while keeping the micro scale of individuals as objects of study, it also avoids both the ecological and atomistic fallacies. As the fifth Comprehensive Household Travel Survey (CHTS) collected information about pupils' origins (sub-districts) and socio-economic and demographic background for the whole Beijing, it was used to conduct multilevel analysis in Chapter 7. As sample size in CHTS is relatively small, the other dataset, the Pupil Travel Survey (PTS), which is large samples covering more than half of all the pupils in Central Beijing, but fewer variables and smaller geographical coverage than the CHTS, has been used to confirm the characteristics and patterns revealed by the CHTS.

8.2.3 Objective III

Objective III - To accurately measure regional education inequalities in China and evaluate the effectiveness of policies targeted at reducing these inequalities by introducing a new multidimensional index.

In Chapter 4, Objective III of the thesis has been achieved by introducing a new multidimensional index, the Index of Regional Education Advantage (IREA), which evaluated the effectiveness of policies targeted at reducing regional/provincial educational inequalities in China since 2005. IREA scores for 2004, 2009 and 2014 were calculated in this research to enable observation of aggregate regional educational inequalities and their evolution in ten years.

The patterns of the IREA and its component facets display a different 'way of looking' at education inequalities in China. Education in the north-east areas of China appears better than that in the south-west areas. This pattern is different from the area division used by the Central Government as the base of policy implementation. Therefore, it is not reasonable for the Government to implement education policies, especially the fiscal policies, based on dividing China into eastern, central and western regions, which will decrease the effectiveness of these policies in reducing education inequalities. These policies should be formulated by considering the variations within regions and their different development trajectories.

Similarly, in previous analysis (Section 4.4.2.3), it has been shown that the increase of per-pupil education expenditure (PPEE) is much quicker than that of per capita GDP (PCGDP) (Figure 4.6) in the last ten years, due to the 'three increases' policy specified in previous education related laws. As early as 1995, The Education Law of the PRC had already stipulated the policy of 'three increases':

> *"Article 55 Educational appropriations of the People's Governments at different levels shall be listed as a separate item of the financial budget according to the principle of consistency of business power and financial power.*
>
> *The **increase** of financial allocation to education by the People's Government at different levels shall be higher than the growth of frequent income of the finance. The People's Government shall make the average per capita education appropriation for all pupils in the school **increase** progressively and ensure that the teacher's salary and the per capita public fund for pupils **grow** gradually."* (Standing Committee of the National People's Congress, 1995, Article 55)

The Compulsory Law of Education in 2006 also had emphasised the 'three increases' policy:

> *"The percentage of increase in government funds allotted for compulsory education by the State Council and the local People's Governments at various levels shall be higher than the percentage of increase in regular government revenues, in order to ensure the gradual increase in the average amount of funds for compulsory education per pupil in school, in the salaries of the teaching staff and in the average amount of funds per pupil for public use."* (Standing Committee of the National People's Congress, 2006f, Article 42)

The 'three increases' policy has had a very positive influence on education development and ensured that enough funds are available for the improvement of teaching staff salaries and school facilities in less developed areas. However, these policies have not considered the specific conditions in each area, which may cause wastage of funding on the construction of educational facilities in developed areas where there are already advanced educational facilities. In order to meet the requirement of these laws, some developed areas have had to upgrade their facilities or rebuild some buildings, which are still in very good condition, to spend up the money allocated for education. These unified policies will decrease the efficiency of the Central Government's input, and high likelihood will lead to a waste of education funds.

Moreover, the IREA has also been decomposed into the scores for different facets to reveal the detailed variations and rationale behind the aggregate regional education disparities. For example, despite the increased resources that have been allocated, the attainment falls short in some regions like Tibet. This is because of Tibet's historically poor education foundation, low enrolment and attainment, which limits its ability to translate education resources into capability. In addition, Tibet is also limited by its physical and social context. Thus, continuing educational investment should be input to these areas to gradually improve their weak education foundation. In addition, the enrolment score used in the IREA reveals that the spatial variations of the enrolment rate for high school are still substantial compared to that of compulsory education. The extremely low enrolment rate for high school in some less developed areas, like Tibet, leads to inadequate high quality labour and then influences economic development in these areas. However, most of the policies aimed at reducing regional education inequalities only paid attention to compulsory education. Since the funding for high school education is still only guaranteed by local government, the development of high school education is directly linked to their uneven economic development. Therefore, in order to effectively reduce regional education disparities, high school education also requires more policy attention from the Central Government and the financial burden on local authorities and families should be reduced.

Furthermore, the enrolment score in the IREA also highlighted the education inequalities for in-migrant children, who normally cannot receive high school education in megacities, like Beijing. As providing equal education to child migrants will greatly increase the financial burden of urban authorities, they usually display a negative attitude towards it. This leads to quality education not being available for these children and the high proportion of 'left behind' children. Although, the policy, notification of further improvement of the funding guarantee system for urban and rural compulsory education (NUR), was implemented in 2015 to encourage local governments in urban areas to offer equal educational opportunities to migrant children by unifying the share of Central Government's support via rural-urban education expenditure, local government still need to cover the fund for teacher's salaries and all the fund for high school education. As the funds for high school education is still not sponsored by Central Government, this leads to many migrant children dropping out after compulsory education or being left behind in rural areas and not receiving consistent education in urban areas.

Furthermore, a region's education capability plays an important role in the development of other capabilities and is fundamental to other different capabilities. The discussion of regional education inequality can provide a reference to the research on other regional capability sets and can fit into a wider context to facilitate research on regional development.

Within a wider context, our evidence from China in terms of education inequalities not only sheds new light on the debate over the impact of fiscal decentralisation and centralisation on educational resources redistribution but also contributes to understanding the role of institutions in determining equal educational distributions among regions and different groups of people. The discussions of policy practice in China that aims at reducing regional education inequality provide an alternative model specification, offer a point of reference for research in other regional capability sets and can fit into different contexts to facilitate research on regional development.

8.2.4 Objective IV

Objective IV - To detect the education inequalities between rural and urban areas by comparing rural-urban education and their variations across space.

In order to achieve Objective IV, Chapter 5 has examined the educational inequalities between rural and urban areas and their spatial variations and temporal evolution. As Central Government policies have attempted to reduce regional education inequality but also narrow the education gap between rural and urban areas, rural and urban education inequality in China has therefore been the focus of interest. Given not all the variables in the IREA have separate data for rural and urban areas respectively (e.g. enrolment rate), the IREA has not been applied to conduct rural-urban comparisons. The variables related to attainment and provision have been compared between rural and urban areas and the results reveal that the overall rural-urban gap in terms of per-pupil educational expenditure (PPEE), teacher-pupil ratios and teacher qualifications, which are provision related variables, have significantly narrowed between 2004 and 2014. However, the overall rural-urban attainment gap has not been reduced but marginally increased from 2000 to 2010. According to the capability-based analysis framework, the previous comparatively low education performance (education foundation), and the disadvantageous natural physical and social environment (e.g. parents' cultural perspective) in rural areas has led to low conversion rates in rural areas. Thus, even given the same education provision, rural areas will probably still achieve lower education attainment than urban areas. Currently, the education provision gap between rural and urban areas is narrowing, but urban areas still tend to have better education provision, especially for middle and high school education, and therefore the existing policy will not lead the gap between rural and urban areas to keep increasing but the gap will still exist and will be unlikely to narrow under current policies.

The above conclusion was drawn from the temporal evolution of the overall rural-urban educational gap. However, there are significant spatial variations in terms of rural-urban educational gap. Some provinces in the north part of China already have higher PPEE in rural areas for both primary and middle school education. Due to the lack of attainment data for each provincial level unit, the change of educational attainment for each province cannot be revealed. However, we can still speculate that the rural-urban educational gap in these areas could be decreased in the ten years with more investment in rural education.

Moreover, the educational advantage in terms of educational attainment is still closely related to advantage in economic development. The north-east areas, that have higher a proportion of the urban population and a more developed economy, not only have higher education attainment in rural areas, but have a smaller rural-urban educational gap in 2010. In contrast, the south-west areas, which have lower urbanisation rates, have lower educational attainment in both rural and urban areas and have larger rural-urban attainment gaps. Furthermore, these results further validated the pattern identified by IREA index. In addition, the results also revealed that the spatial variation in the rural and urban PPEE gap for middle schools is different from that for primary schools in both 2004 and 2014. For example, the PPEE of primary school education for rural areas in some south-west undeveloped provinces is already higher than urban areas, but that of middle school education is still lower. It indicates that even within the same area, education development at different education stages can be different, thus, education related policy should take the education condition of each education stage of each area into consideration. The analysis of Chapter 5 provides more evidence that it is unreasonable to implement the education related policies based on the Government’s area partitioning.

In addition to the above reasons, the School Consolidation strategy has led to very substantial negative effects on rural education and decreases in the efficiency of financial support from Central Government. It worth noting that, in 2014, the teacher-pupil ratio for primary and middle education in most provincial units was higher in rural areas, while that for high schools in most provinces was slightly lower in rural areas (compared to urban areas). The higher teacher-pupil ratio for primary and middle school education in rural areas is likely to be due to sharply decreased pupil numbers in rural areas during rapid urbanisation. The lower teacher-pupil ratio for high school education indicates that a large volume of migrant children go back to rural areas to receive high school education due to the enrolment discrimination policies in urban areas.

8.2.5 Objective V

Objective V - To recognise potentially educational advantage and disadvantage areas in terms of access to compulsory education by creating bespoke geodemographic classifications which is associated with residents' demographic and socio-economic composition, like housing type, parental occupation, qualifications and, in particular, registration status in Central Beijing.

Chapter 6 has explored the use of geodemographics as a means of assessing potential inequality in access to compulsory education. The thesis argues that area classification allows consideration of the influence of multi-dimensional socio-spatial structure on pupils' school choice. Variables from the 2010 Census have been used to create a sub-district classification of Central Beijing to identify disadvantaged and advantaged groups for access to schools. Recognition of areas that are most appropriate for additional investment can help improve the efficiency and equity of the allocation of resources to schools. This study also provides new knowledge about the evolution of urban social space, explaining why the implementation of 'nearby enrolment' policy can be both inefficient and inequitable.

Specifically, five education related clusters were created and labelled based on their distinctive and principal variables. The classification offers some crucial education-related implications and reveals potential inequalities in access to schools for those residents of each type of cluster, which is associated with nearby enrolment policy, housing type, parental occupation, qualifications and registration status. The great influence of 'hukou' on the distances that pupils travel to school and the quality of schools that pupils can attend has been discussed in detail. Hukou status has mediated the influence of income on pupils' access to school. The local residents are likely to be more advantaged compared to migrants with same or even higher incomes. The migrants in cluster labelled 'In-migrants living in deprived suburbs' are expected to be the most educationally disadvantaged groups in Central Beijing. Compared to local low-income groups living in the areas labelled 'Long settled Beijing society' and 'Urban rural transition belt', these migrants are far more disadvantaged due to their non-local hukou. They are more likely to travel a long distance to enrol in unpopular schools. Moreover, the migrants who cannot provide required documents do not have the right to enrol their children in schools in Beijing at all.

Except for identifying potential education inequalities within each type of area, this research also demonstrates that hukou and property ownership based enrolment for compulsory education can make 'nearby enrolment' policy inefficient. The migrant children are firstly excluded from consideration by the 'nearby enrolment' policy due to their non-local hukou. Then, some high income groups in 'High income internal migrants' areas also will not enrol in nearby schools. As they not only want to live in outer areas where there is a better living environment but also want their children can receive a better education in the places of their hukou registration (High quality schools are located in crowded inner areas of Beijing). Accordingly, the hukou and the property-based enrolment policy further strengthens the high income groups' advantages comparing to the actual residence-based enrolment policy in western countries.

In addition, due to the educational gap between rural-urban areas, people living in areas labelled 'Urban rural transition belt' are also likely to be disadvantaged in terms of education.

This chapter identifies for the first time the potential socio-spatial education inequality in Beijing using a geodemographic classification and created the first geodemographic classification in the public sector in China. The results have been published in Xiang *et al.* (2018). Within a wider context, this work also offers new theoretical perspectives for school enrolment policy related research and will be a reference for the cities where educational research is restricted by the limited data for small areas.

8.2.6 Objective VI

Objective VI - To explore the influence of context factors (e.g. geodemographic classifications) and confounding individual level predicators (e.g. hukou, household income, and parents' qualifications) on pupils' travel distance in Beijing by adopting multilevel modelling.

Employing multilevel modelling using detailed microdata, Chapter 7 reveals the influence of contextual factors and confounding individual level predictors on pupils' travel distance to school in Beijing and the underlying social and spatial education inequities within urban areas. Thus, Objective VI is fulfilled in this chapter. Before the multilevel modelling, descriptive analysis of the CHTS was conducted to give an overview of the data. Except for the individual-level characteristics, like hukou, income and housing type, obtained from the CHTS, the contextual variables representing the ecological information about sub-districts from other datasets were also included in the model. Contextual effects were derived by summarising the characteristics of individuals in specific areas. In this research, the geodemographic classification in Central Beijing, defined in Chapter 6, and the school density of each sub-district, were included in models as contextual variables.

The detailed descriptive analysis and multilevel modelling enabled the influence of micro-level covariates and the contextual effects of macro-level variables to be identified, and some implications relating to the social and spatial disparities of pupils in terms of access to education were also revealed. Firstly, it is clear that there is a very significant spatial mismatch between education provision (supply) and consumption (demand) within Beijing, which has led pupils who live in areas with insufficient education provision to have to commute long distances, across district boundaries in many cases, to receive better education opportunities. Secondly, the mean travel distance of in-migrant pupils is not significantly longer than that for those with combined hukou status, i.e. hukou within a district. This is different from our finding in Chapter 6 that in-migrant children would commute further and might be due to the coarse division of the hukou status, where hukou within a sub-district and hukou outside a sub-district but within a district are combined as hukou within a district. It might also be because of the accommodation search process of in-migrant families. Although there is not enough proof to demonstrate in-migrant children commute further, they are certainly deprived of chances to access to educational resources of high quality. In contrast, the pupils whose hukou are outside the district of residence but within this municipality are proven to travel significantly longer distances than those pupils with other hukou statuses according to the model results. This result validates the analysis in Chapter 6 that the people in 'High income internal migrants' cluster tend to commute longer distance due to their choice for high quality schools based on the separation of the place of hukou registration and their current residence.

Similarly, the high income group commutes further than the middle income group, while the low income group or the group whose household without a car travels shorter distances. It is because the high income groups are more capable to make school choice and not sensitive to the cost of longer travel distance. In contrast, the children of the low income background tend to enrol in nearby schools irrespective of the school quality, as they normally don't have enough capital to do school choice. However, the models show that the pupils who live in affordable housing and who also tend to be in the low income group are also shown to travel further. This is due to the fact that school supply has not been fully integrated into the planning of affordable housing programmes and most of the affordable housing is normally located in the urban fringes where there is lack of education provision. As people of low socioeconomic are more sensitive or vulnerable to cost and mobility barriers, long travel distance of pupils living in economically affordable housing might put these low income groups in an even worse situation and lead to further economic disparity.

The model results also confirm the importance of context effects on pupils' travel distance and show that macro level covariates have more significant influence compared to the most of the micro level covariates except for education stage. As the geodemographic clusters associate geographical spaces with the distinct features of the compositional population, their coefficients can reflect socio-spatial education inequalities. The model results validate our finding in Chapter 6 that the pupils living in educational advantaged clusters (e.g. ‘Urban-rural transition belt’, ‘In-migrants in deprived suburbs’) travel longer, while those living in educational advantaged clusters (e.g. ‘Long settled Beijing society’, ‘Employees of state-owned enterprises and white collar workers’) travel shorter distance. Since school density reflects educational provision in each sub-district, the coefficients of school density reveal that the sub-districts with more sufficient education provision would have shorter school commute and *vice versa*. Moreover, the boxplot of school density and geodemographic cluster found that the long commute distance of those living in disadvantaged clusters, such as ‘Urban-rural transition belt’ and ‘In-migrants in deprived suburbs’, is very likely to be due to the low school density in these areas. Accordingly, these clusters are not only socially disadvantaged (demand), but also are more likely to have insufficient educational provision (supply).

8.2.7 Objective VII

Objective VII - To summarise the fulfilled aforementioned objectives to answer the main research questions; to review the success of the research project and then give advice of improvement and possible future research

Whilst education inequality is a topic of great importance and broad interest, it has become apparent that there is relatively little education research in China which has paid attention to spatial equality of education, i.e. research into education inequalities from a geographical perspective. Therefore, through the use of related data from different sources, appropriate methods, this project has studied education inequalities in China at different geographical scales and has successfully achieved the research objectives, met the overall project aims and answered the research questions set out in Chapter 1.

To conclude, this thesis proposed a three-dimensional framework to measure and analyse education inequalities; it assessed and monitored the effects of current policies in China, which aimed at reducing education inequalities, by accurately measuring existing regional inequalities and their change over time; it also revealed the education inequalities within urban areas in terms of social and spatial access to education for the first time by using geodemographics and a multilevel approach based on the limited data available. This research has enriched our theoretical and empirical knowledge about education inequality by considering its spatial characteristics. A comprehensive and flexible three-dimensional framework is proposed for defining and analysing education equality, which extended the more traditional two-dimensional non-spatial framework and can help users to define the concepts of education equality according to their specific research context. Through the empirical research on China, it has provided some unique insights into regional, rural-urban education inequalities and education within urban areas in China and contributed to the substantive literature (Xiang *et al.*, 2018).

8.3 Recommendations for policy

Education inequalities vary spatially and temporally and thus it is not reasonable to set policies without considering the variations within regions and their different development trajectories. This will decrease the effectiveness of these policies in reducing the provincial level and rural-urban educational disparities and cause a waste of education funds to some extent. The optimum allocation of resources and funds should be based on the specific development level of each area and each education stage.

With continuing urbanisation, how to ensure that urban authorities provide equal education for migrants and the host population is a major problem confronting the Chinese Government. The education inequality for migrants is related to the education inequalities at different scales.

More policies should be implemented by the Central Government to reduce urban governments' fiscal pressure of providing equal education to migrant children. Although some policies have already employed to solve the above issues, such as NUR, there are still many significant limitations of these policies that need to be addressed. Firstly, the policy should not be implemented based on the problematic regional division. In addition, the local government of the main target areas of migrants, eastern areas of China, have to cover a higher proportion of education expenditure. Secondly, only compulsory education of migrant children is sponsored by the Central

Government. As the important bridge between compulsory education and higher education, high school education of child migrants also should be supported by the Central Government. Moreover, teachers' salaries should be partly guaranteed by the Central Government.

In order to ensure the migrant children can enjoy equal education rights, the following measures should be considered by urban governments. Firstly, the migrants with rented properties in urban areas also should be guaranteed by nearby enrolment policy as those with local hukou or property ownership. Secondly, the migrant children's education demand should be properly considered by educational planning in urban areas. The educational provision for child migrants and the educational attainment of these child migrants should be included in the performance evaluation of urban educational sectors. Thirdly, the school density distribution for sub-districts by the geodemographic clusters (Figure 7.19) shows that the sub-districts in disadvantaged clusters (e.g. 'Urban-rural transition belt' and 'In-migrants in deprived suburbs') also have lower school density. Thus, intervention measures, for example, more schools, should be designed and introduced in the educationally disadvantaged areas to improve the education provision and social well-being within these areas. Furthermore, the private schools which provide education specifically for migrant children should get more support, in terms of funding, qualified teachers and facilities. In conclusion, ensuring education equality for migrant children at different geographical scales needs effective collaboration among different levels of government.

Moreover, high school education should acquire more policy attention. Owing to the popularisation of compulsory education, the spatial and rural-urban variations of primary and middle school are already very small. At present, due to the rapid urbanisation and transformation of industrial structure in China, there is increasingly more demand for high quality labour. Accordingly, high school education should be gradually included into the compulsory education system and a standard share of government expenditure for high school education should be specified by policies. The financial burden on families and local government should be reduced and more efforts by the Central Government should be used to support the senior secondary education development in less developed areas.

Furthermore, the construction of affordable housing should not only focus on making housing economically viable for low income groups, but school supply should also be integrated into the affordable housing design to create more liveable neighbourhoods. For the already built affordable housing neighbourhoods, more accessible schools and more accessible transportation system should be promoted. On the other hand, decreasing the privilege of Beijing (local) hukou and providing similar quality education services citywide are keys to reducing the long commute distances and stopping high income groups or the residents with 'Hukou outside a district but within a municipality' 'gaming the system'.

8.4 Limitations of the research

It is acknowledged there are some limitations to this research. On the one hand, there are due to the drawbacks of applied methods, such as the ecological fallacy (Burrows *et al.*, 2005; Graham, 2005; Leventhal, 2016) and the Modifiable Areal Unit Problem (MAUP) (Openshaw, 1994), and issues of using a composite index and a geodemographic classification. There is a degree of variability within each area, like each provincial level unit and sub-district. In addition, the accuracy of geodemographic clusters will gradually degrade over time due to its reliance on the decennial census data (Leventhal, 2016);

On the other hand, this research has been limited to the available data. For example, all the determinants related to pupils' travel distance cannot be fully included in the multilevel models due to the limitations in the available variables; the comprehensive rural-urban education comparison has been restricted by the lack of separate data for rural and urban education; there is no separate data for hukou within a sub-district and hukou outside a sub-district but within a district in CHTS; Furthermore, the education inequalities research within urban areas also are retrained because of lacking data related to school performance.

8.5 Potential for future research

First, this study provides a three-dimensional framework for education equality analysis, which can be applied to various geographical scales and contexts. This thesis focused on regional education inequalities and education inequalities within urban areas, at the macro-scale and micro-scale respectively. However, education inequalities and variations among counties and municipalities within each provincial have not been revealed in this research, which might be useful for the provincial level government to set improvement policies. Thus, a mesoscale study might be conducted in future to form a complete map of education inequalities in China. Moreover, the framework also can be used in other countries, which may enable some comparative study of spatial education inequalities to be undertaken.

Second, the multilevel models constructed in this research only focus on the influence of pupils' socio-economic characteristics and population composition of the sub-districts where they live (origin). As there is an apparent cross-classified structure where each pupil is naturally nested in one sub-district of residence (origin) and at the same time belongs to one school sub-district (destination), cross-classified multilevel modelling can be specified in the next stage to simultaneously consider the context effects of where pupils live and study. This approach might provide more insights into how individual factors affect pupils' travel distance and how they vary between origins or destinations. Moreover, there is no direct education quality related data included in the above analysis within urban areas due to the data limitations. Thus, in future research, other datasets that can be used to evaluate education quality variations, which should be explored and included to make conclusions more convincing.

Finally, more in-depth and comprehensive qualitative research should be conducted in some sub-districts to investigate the detailed variations and disparities within each cluster (e.g. access to schools and attitude to education) and inform on the influences of areas on a disadvantaged group's education. When new census data are available, an education related classification could be reconstructed to thereby monitor changes in the effects of education related policy on education inequalities.

List of References

Abbas, J.,Ojo, A. and Orange, S. 2009. Geodemographics-a tool for health intelligence? *Public Health.* **123**(1), pp.35-39.

Adam, A.,Delis, M.D. and Kammas, P. 2014. Fiscal decentralization and public sector efficiency: Evidence from OECD countries. *Economics of Governance.* **15**(1), pp.17-49.

Alderman, H.,Orazem, P.F. and Paterno, E.M. 2001. School quality, school cost, and the public/private school choices of low-income households in Pakistan. *Journal of Human Resources.* pp.304-326.

Allen, R. 2007. Allocating pupils to their nearest secondary school: The consequences for social and ability stratification. *Urban Studies.* **44**(4), pp.751-770.

Amelsvoort, H.W.C.G. and Scheerens, J. 1997. Policy issues surrounding processes of centralization and decentralization in European education systems. *Educational Research and Evaluation.* **3**(4), pp.340-363.

Amin, A. 2002. Spatialities of globalisation. *Environment and planning A: Economy and Space.* **34**(3), pp.385-399.

Ashby, D.I. and Longley, P.A. 2005. Geocomputation, geodemographics and resource allocation for local policing. *Transactions in GIS.* **9**(1), pp.53-72.

Baden, B.M.,Noonan, D.S. and Turaga, R.M.R. 2007. Scales of justice: Is there a geographic bias in environmental equity analysis? *Journal of Environmental Planning and Management.* **50**(2), pp.163-185.

Baker, J.,Lynch, K.,Cantillon, S. and Walsh, J. 2016a. Dimensions of equality: A framework for theory and action. In: Baker, J.,Lynch, K.,Cantillon, S. and Walsh, J. (eds). *Equality: From Theory to Action.* Hampshire; New York: Springer, pp.21-46.

Baker, J.,Lynch, K.,Cantillon, S. and Walsh, J. 2016b. Equality and education. In: Baker, J.,Lynch, K.,Cantillon, S. and Walsh, J. (eds). *Equality: From Theory to Action.* Hampshire; New York: Springer, pp.140-168.

Baker, J.,Lynch, K.,Cantillon, S. and Walsh, J. 2016c. New challenges to an unequal world. In: Baker, J.,Lynch, K.,Cantillon, S. and Walsh, J. (eds). *Equality: From Theory to Action.* Hampshire; New York: Springer, pp.3-20.

Ball, S.J. 2003. *Class Strategies and The Education Market: The middle Classes and Social Advantage.* London: Routledge.

Ball, S.J.,Bowe, R. and Gewirtz, S. 1995. Circuits of schooling: A sociological exploration of parental choice of school in social class contexts. *The Sociological Review.* **43**(1), pp.52-78.

Bandura, R. 2008. A survey of composite indices measuring country performance: 2008 update. *New York: United Nations Development Programme, Office of Development Studies (UNDP/ODS Working Paper).*

Bardhan, P. 2002. Decentralization of governance and development. *The Journal of Economic Perspectives.* **16**(4), pp.185-205.

Barro, R.J. and Lee, J.-W. 1993. International comparisons of educational attainment. *Journal of Monetary Economics.* **32**(3), pp.363-394.

Baschieri, A. and Falkingham, J. 2009. Staying in school: Assessing the role of access, availability, and economic opportunities-the case of Tajikistan. *Population, Space and Place.* **15**(3), pp.205-224.
Becker, W.E. and Powers, J.R. 2001. Student performance, attrition, and class size given missing student data. *Economics of Education Review.* **20**(4), pp.377-388.
Beijing Municipal Commission of Education. 2015. *2014-2015 School Year Beijing Education Development Statistics.* [Online]. Available from: http://zfxxgk.beijing.gov.cn/columns/63/3/565666.html
Beijing Statistical Bureau. 2015. *Beijing Statistical Yearbook.* Beijing: China Statistical Press.
Bi, B. and Zhang, X. 2016. “Attending nearby schools” in Central Beijing: Influencing factors and the policy distortion. *International Review for Spatial Planning and Sustainable Development.* **4**(1), pp.31-48.
Bilton, T. 2002. *Introductory Sociology.* Basingstoke: Palgrave.
Boterman, W.R. 2013. Dealing with diversity: Middle-class family households and the issue of 'black' and 'white' schools in Amsterdam. *Urban Studies.* **50**(6), pp.1130-1147.
Bradford, M. 1990. Education, attainment and the geography of choice. *Geography.* pp.3-16.
Breen, R. and Jonsson, J.O. 2005. Inequality of opportunity in comparative perspective: Recent research on educational attainment and social mobility. *Annual Review of Sociology.* **31**(1), pp.223-243.
Brighouse, H. 2000. *School Choice and Social Justice.* Oxford: Oxford University Press.
Brighouse, H.,Howe, K.R.,Tooley, J. and Haydon, G. 2010. *Educational Equality.* 2nd ed. London: Continuum.
Bullen, N.,Jones, K. and Duncan, C. 1997. Modelling complexity: Analysing between-individual and between-place variation-a multilevel tutorial. *Environment and Planning A.* **29**(4), pp.585-609.
Burrows, R.,Ellison, N.,Woods, B. and Joseph Rowntree, F. 2005. *Neighbourhoods on the Net: The Nature and Impact of Internet-based Neighbourhood Information Systems.* Bristol: Policy Press.
Burt, M.E. and Park, N. 2009. Education inequality in the Republic of Korea: Measurement and causes. In: Holsinger, D.B. and Jacob, W.J. (eds). *Inequality in Education: Comparative and International Perspectives.* Hong Kong: Springer Science & Business Media, pp.261-289.
Butler, T. and Hamnett, C. 2007. The geography of education: Introduction. *Urban Studies.* **44**(7), pp.1161-1174.
Butler, T.,Hamnett, C. and Ramsden, M.J. 2013. Gentrification, education and exclusionary displacement in East London. *International Journal of Urban and Regional Research.* **37**(2), pp.556-575.
Cai, F. 2007. Rural urban income gap and critical point of institutional change. *Economic Change and Restructuring.* **40**(1), pp.189-206.
Chan, K.W. 1994. *Cities with Invisible Walls: Reinterpreting Urbanization in Post-1949 China.* Oxford; Hong Kong: Oxford University Press.
Chang, H.-S. and Liao, C.-H. 2011. Exploring an integrated method for measuring the relative spatial equity in public facilities in the context of urban parks. *Cities.* **28**(5), pp.361-371.

Charrad, M.,Ghazzali, N.,Boiteau, V. and Niknafs, A. 2014. Nbclust: An R package for determining the relevant number of clusters in a data set. *Journal of Statistical Software.* **61**(6), pp.1-36.

Chen, M.,Liu, W. and Tao, X. 2013. Evolution and assessment on China's urbanization 1960-2010: Under-urbanization or over-urbanization? *Habitat International.* **38**, pp.25-33.

Chen, M.X.,Zhang, W.Z. and Lu, D.D. 2015. Examining spatial pattern and location choice of affordable housing in Beijing, China: Developing a workable assessment framework. *Urban Studies.* **52**(10), pp.1846-1863.

Cheshire, P.C. and Gordon, I.R. 1998. Territorial competition: Some lessons for policy. *The Annals of Regional Science.* **32**(3), pp.321-346.

Cheung, H.Y. and Chan, A.W.H. 2008. Relationships amongst cultural dimensions, educational expenditure and class size of different nations. *International Journal of Educational Development.* **28**(6), pp.698-707.

Christmann, E.P. 2005. The effects of social control on Chicago's public schools. *The Educational Forum.* **69**(1), pp.92-93.

Connelly, R. and Zheng, Z. 2007a. Educational access for China's post-cultural revolution generation: Enrolment patterns in 1990. In: Park, A. and Hannum, E. (eds). *Education and Reform in China.* London: Routledge, pp.64-80.

Connelly, R. and Zheng, Z. 2007b. Enrollment and graduation patterns as China's reforms deepen, 1990-2000. In: Park, A. and Hannum, E. (eds). *Education and Reform in China.* London: Routledge, pp.81-92.

Courgeau, D. 2003. *Methodology and Epistemology of Multilevel Analysis: Approaches from Different Social Sciences.* Dordrecht: Kluwer Academic Publishers.

Cowan, R. 2005. *The Dictionary of Ubanism.* Tisbury: Streetwise Press.

Craglia, M.,Haining, R. and Signoretta, P. 2001. Modelling high-intensity crime areas in English cities. *Urban Studies.* **38**(11), pp.1921-1941.

Croxford, L. and Raffe, D. 2014. Social class, ethnicity and access to higher education in the four countries of the UK: 1996-2010. *International Journal of Lifelong Education.* **33**(1), pp.77-95.

Curley, A.M. 2005. Theories of urban poverty and implications for public housing policy. *Journal of Sociology and Social Welfare.* **32**(2), pp.97-119.

Dadashpoor, H.,Rostami, F. and Alizadeh, B. 2016. Is inequality in the distribution of urban facilities inequitable? Exploring a method for identifying spatial inequity in an Iranian city. *Cities.* **52**, pp.159-172.

Darling-Hammond, L. 2007. *The Flat Earth and Education: How America's Commitment to Equity will Determine our Future.* Washington: American Educational Research Association.

Daun, H. 2009. 15. Religious education and Islam in Europe. In: Holsinger, D.B. and Jacob, W.J. (eds). *Inequality in Education: Comparative and International Perspectives.* Hong Kong: Springer, pp.348-368.

Davis-Kean, P.E. 2005. The influence of parent education and family income on child achievement: The indirect role of parental expectations and the home environment. *Journal of Family Psychology.* **19**(2), pp.294-304.

De la Fuente, H.,Rojas, C.,Salado, M.J.,Carrasco, J.A. and Neutens, T. 2013. Socio-spatial inequality in education facilities in the Concepción Metropolitan Area (Chile). *Current Urban Studies.* **1**(04), pp.117-129.

De Vries, M.S. 2000. The rise and fall of decentralization: A comparative analysis of arguments and practices in European countries. *European Journal of Political Research.* **38**(2), pp.193-224.

Decancq, K. and Lugo, M.A. 2013. Weights in multidimensional indices of wellbeing: An overview. *Econometric Reviews.* **32**(1), pp.7-34.

Deer, C. 2009. 14. Different paths, similar effects: Persistent inequalities and their sources in European higher education. In: Holsinger, D.B. and Jacob, W.J. (eds). *Inequality in Education: Comparative and International Perspectives.* Hong Kong: Springer Science & Business Media, pp.324-347.

Dennett, A. and Stillwell, J. 2011. A new area classification for understanding internal migration in Britain. *Population Trends.* (145), pp.146-171.

Desai, M. and Shah, A. 1988. An econometric approach to the measurement of poverty. *Oxford Economic Papers.* **40**(3), pp.505-522.

Deutsch, J. and Silber, J. 2005. Measuring multidimensional poverty: An emprical comparision of various approaches. *Review of Income and Wealth.* **51**(1), pp.145-174.

Dickerson, A. and McIntosh, S. 2013. The impact of distance to nearest education institution on the post-compulsory education participation decision. *Urban Studies.* **50**(4), pp.742-758.

Diette, T.M. 2012. The whiter the better? Racial composition and access to school resources for black students. *The Review of Black Political Economy.* **39**(3), pp.321-334.

Ding, W. and Lehret, S. 2007. Incentives and quality of teachers and schools. In: Park, A. and Hannum, E. (eds). *Education and Reform in China.* London: Routledge, pp.191-204.

Dworkin, R. 2002. *Sovereign Virtue: The Theory and Practice of Equality.* Cambridge, Massachusetts; London, Enland: Harvard University Press.

Dworkin, R. 2018. What is equality? Part 2: Equality of resources. *The Notion of Equality.* New York: Routledge, pp.143-205.

Easton, S. and Ferrari, E. 2015. Children's travel to school-the interaction of individual, neighbourhood and school factors. *Transport Policy.* **44**, pp.9-18.

Eliason, S.R. 1993. *Maximum Likelihood Estimation: Logic and Practice.* London; Newbury Park: Sage.

Erikson, R.,Goldthorpe, J.H.,Jackson, M.,Yaish, M. and Cox, D.R. 2005. On class differentials in educational attainment. *Proceedings of the National Academy of Sciences of the United States of America.* **102**(27), pp.9730-9733.

Fang, Y.P.,Logan, J.R. and Pal, A. 2015. Emerging socio-spatial pattern of Chinese cities: The case of Beijing in 2006. *Habitat International.* **47**, pp.103-112.

Farrell, J.P. 1999. Changing conceptions of equality of education: Forty years of comparative evidence. *Comparative Education: The Dialectic of the Global and the Local.* pp.149-178.

Fekjær, S.N. and Birkelund, G.E. 2007. Does the ethnic composition of upper secondary schools influence educational achievement and attainment? A multilevel analysis of the Norwegian case. *European Sociological Review.* **23**(3), pp.309-323.

Fielding, A. 2004. The role of the Hausman test and whether higher level effects should be treated as random or fixed. *Multilevel Modelling Newsletter.* **16**(2), pp.3-9.

Fox, M. 1996. Rural school transportation as a daily constraint in students' Lives. *Rural Educator.* **17**(2), pp.22-27.

Fu, Q. and Ren, Q. 2010. Educational inequality under China's rural-urban divide: The hukou system and return to education. *Environment and Planning A.* **42**(3), pp.592-610.

Fu, Q.,Zhu, Y.S. and Ren, Q. 2015. The downside of marketization: A multilevel analysis of housing tenure and types in reform-era urban China. *Social Science Research.* **49**, pp.126-140.

Gale, C.,Singleton, A. and Longley, P. 2017. Creating a new open geodemographic classification of the UK using 2011 census data. In: Stillwell, J. ed. *The Routledge Handbook of Census Resources, Methods and Applications: Unlocking the UK 2011 Census.* London: Routledge, pp.213-229.

Garner, C.L. and Raudenbush, S.W. 1991. Neighborhood effects on educational attainment: A multilevel analysis. *Sociology of Education.* **64**(4), pp.251-262.

Goesling, B. 2001. Changing income inequalities within and between nations: New evidence. *American Sociological Review.* **66**(5), pp.745-761.

Goldstein, H. 1986. Multilevel mixed linear model analysis using iterative generalized least squares. *Biometrika.* **73**(1), pp.43-56.

Goldstein, H. 1987. *Multilevel Models in Educational and Social Research.* London: Griffin.

Goldstein, H. 1989. Restricted unbiased iterative generalized least-squares estimation. *Biometrika.* **76**(3), pp.622-623.

Goldstein, H.,Browne, W. and Rasbash, J. 2002. Partitioning variation in multilevel models. *Understanding Statistics.* **1**(4), pp.223-231.

Goldstein, H. and Silver, R. 1989. Multilevel and multivariate models in survey analysis. In: Skinner, C.J.,Holt, D. and Smith, T.M.F. (eds). *Analysis of Complex Surveys.* Chichester, UK: John Wiley & Sons.

Gordon, I. and Monastiriotis, V. 2007. Education, location, education: A spatial analysis of English secondary school public examination results. *Urban Studies.* **44**(7), pp.1203-1228.

Gould, M.I. and Fieldhouse, E. 1997. Using the 1991 Census SAR in a multilevel analysis of male unemployment. *Environment and Planning A.* **29**(4), pp.611-628.

Graham, S.D.N. 2005. Software-sorted geographies. *Progress in Human Geography.* **29**(5), pp.562-580.

Green, A.,Preston, J. and Sabates, R. 2003. Education, equality and social cohesion: A distributional approach. *Compare: A Journal of Comparative and International Education.* **33**(4), pp.453-470.

Gu, C.,Chan, R.C.K.,Liu, J. and Kesteloot, C. 2006. Beijing's socio-spatial restructuring: Immigration and social transformation in the epoch of national economic reformation. *Progress in Planning.* **66**(4), pp.249-310.

Gu, C.,Wang, F. and Liu, G. 2005. The structure of social space in Beijing in 1998: A socialist city in transition. *Urban Geography.* **26**(2), pp.167-192.

Gutiérrez, J.,Kitchin, R. and Thrift, N. 2009. Transport and accessibility. *International Encyclopedia of Human Geography.* pp.410-417.

Gyatso, S.,Postiglione, G.A. and Jiao, B. 2005. Education in rural Tibet: Development, problems and adaptations. *China: An International Journal.* **3**(1), pp.1-23.

Hallett, R.E. and Venegas, K.M. 2011. Is increased access enough? Advanced placement courses, quality, and success in low-income urban schools. *Journal for the Education of the Gifted.* **34**(3), pp.468-487.

Halsey, A.H.,Heath, A.F. and Ridge, J.M. 1980. *Origins and Destinations: Family, Class, and Education in Modern Britain.* Oxford: Clarendon Press.

Hamnett, C. and Butler, T. 2011. 'Geography matters': The role distance plays in reproducing educational inequality in East London. *Transactions of the Institute of British Geographers.* **36**(4), pp.479-500.

Hamnett, C. and Butler, T. 2013. Distance, education and inequality. *Comparative Education.* **49**(3), pp.317-330.

Hannum, E. 2003. Poverty and basic education in rural China: Villages, households, and girls' and boys' enrollment. *Comparative Education Review.* **47**(2), pp.141-159.

Hannum, E.,Park, A. and Cheng, K.-M. 2007. Introducation: Market reforms and educational opportunity in China. In: Park, A. and Hannum, E. (eds). *Education and Reform in China.* London: Routledge, pp.1-24.

Harabet, V. 2014. Equal access to education as a factor of social protection of a person: Historical and economical aspect. *Koncept : Scientific and Methodological E-magazine.* **1**, pp.691-695.

Hare, R.M. 1952. *The Language of Morals.* Oxford: Oxford Paperbacks.

Harris, R. 2012. Local indices of segregation with application to social segregation between London's secondary schools, 2003–08/09. *Environment and Planning A.* **44**(3), pp.669-687.

Harris, R. 2013. Geographies of transition and the separation of lower and higher attaining pupils in the move from primary to secondary school in London. *Transactions of the Institute of British Geographers.* **38**(2), pp.254-266.

Harris, R. 2016. Measuring segregation as a spatial optimisation problem, revisited: a case study of London, 1991-2011. *International Journal of Geographical Information Science.* **30**(3), pp.474-493.

Harris, R. 2017. Measuring the scales of segregation: Looking at the residential separation of white British and other schoolchildren in England using a multilevel index of dissimilarity. *Transactions of the Institute of British Geographers.* **42**(3), pp.432-444.

Harris, R. and Feng, Y. 2016. Putting the geography into geodemographics: Using multilevel modelling to improve neighbourhood targeting-a case study of Asian pupils in London. *Journal of Marketing Analytics.* **4**(2-3), pp.93-107.

Harris, R. and Johnston, R. 2008. Primary schools, markets and choice: Studying polarization and the core catchment areas of schools. *Applied Spatial Analysis and Policy.* **1**(1), pp.59-84.

Harris, R.,Johnston, R. and Burgess, S. 2007. Neighborhoods, ethnicity and school choice: Developing a statistical framework for geodemographic analysis. *Population Research and Policy Review.* **26**(5/6), pp.553-579.

Harris, R.,Sleight, P. and Webber, R. 2005. *Geodemographics, GIS, and Neighbourhood Targeting.* Chichester: Wiley.

He, S.,Chung, C.K.L.,Bayrak, M.M. and Wang, W. 2018. Administrative boundary changes and regional inequality in provincial China. *Applied Spatial Analysis and Policy.* **11**(1), pp.103-120.

Helander, E. 1993. *Prejudice and Dignity: An Introduction to Community-based Rehabilitation.* New York: United Nations Development Programme.

Herrero, C.,Martínez, R. and Villar, A. 2012. A newer human development index. *Journal of Human Development and Capabilities.* **13**(2), pp.247-268.

Herz, B.K. and Sperling, G.B. 2004. *What Works in Girls' Education: Evidence and Policies from the Developing World.* New York: Council on Foreign Relations Press

Holloway, S.L.,Hubbard, P.,Jöns, H. and Pimlott-Wilson, H. 2010. Geographies of education and the significance of children, youth and families. *Progress in Human Geography.* **34**(5), pp.583-600.

Holsinger, D.B. and Jacob, W.J. 2009. *Inequality in Education: Comparative and International Perspectives.* Hong Kong: Springer Science & Business Media.

Holt, L. 2007. Children's sociospatial (re)production of disability within primary school playgrounds. *Environment and Planning D: Society and Space.* **25**(5), pp.783-802.

Holt, L.,Bowlby, S. and Lea, J. 2019. Disability, special educational needs, class, capitals, and segregation in schools: A population geography perspective. *Population, Space and Place.* pe2229.

Howitt, R. 1998. Scale as relation: Musical metaphors of geographical scale. *Area.* **30**(1), pp.49-58.

Hox, J.J. 2010. *Multilevel Analysis: Techniques and Applications.* 2nd ed. New York: Routledge.

Hu, W. 2012. Thirty years of education in China: Between change and no change. *Chinese Education & Society.* **45**(1), pp.84-94.

Huang, Z.,Wang, T. and Li, X. 2015. The political dynamics of educational changes in China. *Policy Futures in Education* **14**(1), pp.24-41.

Hulton, L. and Furlong, D. 2001. *Gender Equality in Education: A Select Annotated Bibliography.* Brighton: Bridge.

Hung, C.-C. and Cheng, S.Y. 2009. 12. Access and equity: Who are the students at Taiwan's top universities? In: Holsinger, D.B. and Jacob, W.J. (eds). *Inequality in Education: Comparative and International Perspectives.* Hong Kong: Springer Science & Business Media, pp.290-306.

Húsen, T. 1974. *The Learning Society.* London: Methuen.

Illich, I. 2002. *Deschooling Society.* London: Marion Boyars.

Israel, E. and Frenkel, A. 2017. Social justice and spatial inequality: Toward a conceptual framework. *Progress in Human Geography.* **42**(5), pp.647-665.

Jacob, W.J. and Holsinger, D.B. 2009. 1. Inequality in education: A critical analysis. In: Holsinger, D.B. and Jacob, W.J. (eds). *Inequality in Education: Comparative and International Perspectives.* Hong Kong: Springer Science & Business Media, pp.1-33.

Jiang, Y.H. and Shen, J.F. 2013. Weighting for what? A comparison of two weighting methods for measuring urban competitiveness. *Habitat International.* **38**, pp.167-174.

Johnston, R.,Burgess, S.,Wilson, D. and Harris, R. 2006. School and residential ethnic segregation: An analysis of variations across England's local education authorities. *Regional studies.* **40**(9), pp.973-990.

Johnston, R.J. 2000. *The Dictionary of Human Geography.* Malden, MA: Blackwell Publishers.

Jones, K. 1991. Specifying and estimating multi-level models for geographical research. *Transactions of the Institute of British Geographers.* pp.148-159.

Jones, K. and Bullen, N. 1994. Contextual models of urban house prices: A comparison of fixed-coefficient and random-coefficient models developed by expansion. *Economic Geography.* **70**(3), pp.252-272.

Jones, K. and Duncan, C. 1996. People and places: The multilevel model as a general framework for the quantitative analysis of geographical data. In: Longley, P. and Batty, M. (eds). *Spatial Analysis: Modelling in A GIS Environment.* Cambridge: John Wiley &Son, pp.79-104.

Jones, K. and Gould, M. 2014. Multilevel modeling: Concepts and applications In: *Belfast NILS User Group Workshop 3rd & 4th Nov. 2014*.

Jones, K. and Subramanian, S. 2011. *Developing multilevel models for analysing contextuality, heterogeneity and change using MLwiN* Bristol: Centre for Multilevel Modeling, University of Bristol

Jones, M. 2009. Phase space: Geography, relational thinking, and beyond. *Progress in Human Geography.* **33**(4), pp.487-506.

Jorda, V. and Alonso, J.M. 2017. New estimates on educational attainment using a continuous approach (1970-2010). *World Development.* **90**, pp.281-293.

Kennedy, P. 1979. *A Guide to Econometrics.* Oxford: Martin Robertson.

Koo, J. and Lee, H. 2015. Regional capability and regional disparity: A conceptual framework and applications within Korea. *International Review of Public Administration.* **20**(2), pp.121-135.

Kyriacou, A.P.,Muinelo-Gallo, L. and Roca-Sagalés, O. 2017. Regional inequalities, fiscal decentralization and government quality. *Regional Studies.* **51**(6), pp.945-957.

Law, W. and Pan, S. 2009. Legislation and equality in basic education for all in China. *Interchange: A Quarterly Review of Education.* **40**(4), pp.337-372.

Leckie, G. and Bell, A. 2013. Cross-classified multilevel models-MLwiN practical. *LEMMA VLE Module 12.* pp.1-60.

Leventhal, B. 2016. *Geodemographics for Marketers: Using Location Analysis for Research and Marketing.* London: Kogan Page.

Levin, H.M. 1976. Educational opportunity and social inequality in Western Europe. *Social Problems.* **24**(2), pp.148-172.
Li, B. 2005. Urban social change in transitional China: A perspective of social exclusion and vulnerability. *Journal of Contingencies and Crisis Management.* **13**(2), pp.54-65.
Li, L. and Liu, H.M. 2014. Primary school availability and middle school education in rural China. *Labour Economics.* **28**, pp.24-40.
Li, N. and Placier, P. 2015. Migrant and non-migrant families in Chengdu, China: Segregated lives, segregated schools. *Social Sciences.* **4**(2), pp.339-360.
Li, Y. and Wei, Y.H.D. 2010. The spatial-temporal hierarchy of regional inequality of China. *Applied Geography.* **30**(3), pp.303-316.
Lin, D.,Zhang, L.,Mathur, A.,Stanton, B.,Fang, X.,Xiong, Q. and Li, X. 2010. Schooling of migrant children in China: Perspectives of school teachers. *Vulnerable Children and Youth Studies.* **5**(1), pp.79-87.
Lin, J.-J.,Huang, Y.-C. and Ho, C.-L. 2014. School accessibility and academic achievement in a rural area of Taiwan. *Children's Geographies.* **12**(2), pp.232-248.
Lipman, P. 2005. Metropolitan regions-new geographies of inequality in education: The Chicago metroregion case. *Globalisation, Societies and Education.* **3**(2), pp.141-163.
Liu, J. and Jacob, W.J. 2013. From access to quality: Migrant children's education in urban China. *Educational Research for Policy and Practice.* **12**(3), p177.
Logan, J.R. and Burdick-Will, J. 2016. School segregation, charter schools, and access to quality education. *Journal of Urban Affairs.* **38**(3), pp.323-343.
Logan, J.R.,Minca, E. and Adar, S. 2012. The geography of inequality why separate means unequal in American public schools. *Sociology of Education.* **85**(3), pp.287-301.
Lopez-Acevedo, G. 2009. 18. Mexico: Evolution of education and inequality in the last two decades. In: Holsinger, D.B. and Jacob, W.J. (eds). *Inequality in Education: Comparative and International Perspectives.* Hong Kong: Springer Science & Business Media, pp.417-443.
Ma, J.,Mitchell, G. and Heppenstall, A. 2015. Exploring transport carbon futures using population microsimulation and travel diaries: Beijing to 2030. *Transportation Research Part D-Transport and Environment.* **37**, pp.108-122.
MacKinnon, D. 2011. Reconstructing scale: Towards a new scalar politics. *Progress in Human Geography.* **35**(1), pp.21-36.
Marcuse, P. 2009. Spatial justice: Derivative but causal of social injustice. *Justice Spatiale/Spatial Justice.* **1**, pp.1-6.
Martin, R. 2008. National growth versus spatial equality? A cautionary note on the new ‘trade-off’ thinking in regional policy discourse. *Regional Science Policy & Practice.* **1**(1), pp.3-13.
McDonald, N.C. 2010. School siting: Contested visions of the community school. *Journal of the American Planning Association.* **76**(2), pp.184-198.
Mei, H.,Jiang, Q.,Xiang, Y. and Song, X. 2015. School consolidation: Whither China's rural education? *Asian Social Work and Policy Review.* **9**(2), pp.138-150.

Milligan, G.W. and Cooper, M.C. 1985. An examination of procedures for determining the number of clusters in a data set. *Psychometrika.* **50**(2), pp.159-179.

Minas, R.,Wright, S. and van Berkel, R. 2012. Decentralization and centralization: Governing the activation of social assistance recipients in Europe. *International Journal of Sociology and Social Policy.* **32**(5/6), pp.286-298.

Ming, H.H. 2014. *The Education of Migrant Children and China's Future: The Urban Left Behind.* Abingdon, Oxon: Routledge.

Ministry of Education of PRC. 1992. *The Compulsory-education Implementing Detailed-rules of PRC. Article 26* [Online]. [Accessed 14 March]. Available from: http://old.moe.gov.cn/publicfiles/business/htmlfiles/moe/moe_620/200409/3177.html

Ministry of Education of PRC. 2005. *Statistical Report on the Implementation of the National Education Funds in 2004.* [Online]. [Accessed December 31st]. Available from: http://www.moe.edu.cn/srcsite/A05/s3040/200512/t20051230_78264.html

Ministry of Education of PRC. 2006. *The Answer of the Compulsory Education Law (3)* [Online]. [Accessed 4 Sep]. Available from: http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/moe_1294/200610/18176.html

Ministry of Education of PRC. 2011. *Statistical Report on the Implementation of the National Education Funds in 2010.* [Online]. [Accessed December 31st]. Available from: http://www.moe.edu.cn/srcsite/A05/s3040/201112/t20111223_128871.html

Ministry of Education of PRC. 2015a. *The Fund Guarantee System of China are More Perfect* [Online]. [Accessed November 20th]. Available from: http://www.moe.edu.cn/jyb_xwfb/moe_2082/zl_2015n/2015_zl60/201511/t20151123_220433.html

Ministry of Education of PRC. 2015b. *Statistical Indicator System of China's Educational Monitoring and Evaluation.* [Online]. [Accessed August 10th]. Available from: http://www.moe.edu.cn/srcsite/A03/s182/201509/t20150907_206014.html

Ministry of Education of PRC. 2016. *National Education Development in 2015.* [Online]. [Accessed December 19th]. Available from: http://www.gov.cn/xinwen/2016-12/19/content_5149958.htm#1

Ministry of Education of the People's Republic of China. 2015a. *China Education Monitoring and Evaluation Index System.* [Online]. [Accessed August 18th]. Available from: http://www.moe.edu.cn/srcsite/A03/s182/201509/t20150907_206014.html

Ministry of Education of the People's Republic of China. 2015b. *Net Enrolment Ratio of School-age Children in Primary Schools.* [Online]. [Accessed October 17th]. Available from: http://www.moe.edu.cn/s78/A03/moe_560/jytjsj_2015/2015_qg/201610/t20161012_284487.html

Ministry of Human Resources and Social Security of the People's Republic of China and Ministry of Education. 2015. *The Notice Issued by Ministry of Human Resources and Social Security of the People's Republic of China and Ministry of Education about " The Guidance of Deepening the Reform of Teachers' Professional Title System".* [Online]. [Accessed August 28th]. Available from: http://www.moe.edu.cn/jyb_xxgk/moe_1777/moe_1779/201509/t20150902_205165.html

Morris, J.M.,Dumble, P.L. and Wigan, M.R. 1979. *Accessibility Indicators for Transport Planning.* Vermont South, Vict: Australian Road Research Board.

Müller, S. 2011. Assessment of school closures in urban areas by simple accessibility measures. *Erdkunde.* pp.401-414.

National Bureau of Statistics of the PRC. 1999. *The Regulations about Urban Rural Division (Trial).* [Online]. [Accessed December 6th]. Available from: http://www.stats.gov.cn/tjsj/pcsj/rkpc/5rp/html/append7.htm

National Bureau of Statistics of the PRC. 2017. *2003-2006 National Per Capita Disposable Income Gini Coefficient.* [Online]. [Accessed Oct 10th]. Available from: http://www.stats.gov.cn/ztjc/zdtjgz/yblh/zysj/201710/t20171010_1540710.html

National People's Congress. 2004. *Constitution of PRC.*

Nielsen, I.,Nyland, B.,Nyland, C.,Smyth, R. and Zhang, M. 2006. Determinants of school attendance among migrant children: Survey evidence from China's Jiangsu province. *Pacific Economic Review.* **11**(4), pp.461-476.

Norman, P.D. and Darlington-Pollock, F. 2018. The changing geography of deprivation in Great Britain: Exploiting small area census data, 1971 to 2011. In: Stillwell, J. ed. *The Routledge Handbook of Census Resources, Methods and Applications: Unlocking the UK 2011 Census.* Abingdon, Oxon; New York: Routledge, pp.404-420.

OECD. 2008. *Handbook on Constructing Composite Indicators: Methodology and User Guide.* Paris: OECD.

OECD. 2012. *Equity and Quality in Education: Supporting Disadvantaged Students and Schools.* [Online]. Available from: https://www.oecd.org/education/school/50293148.pdf

OECD. 2016. *Education at A Glance 2016: OECD Indicators.* Paris: OECD Publishing.

Openshaw, S. 1994. *Developing Smart and Intelligent Target Marketing Systems.* School of Geography, University of Leeds.

Owen, G.,Harris, R. and Jones, K. 2016. Under examination: Multilevel models, geography and health research. *Progress in Human Geography.* **40**(3), pp.394-412.

Pacione, M. 1989. Access to urban services-the case of secondary schools in Glasgow. *The Scottish Geographical Magazine.* **105**(1), pp.12-18.

Pan, A.,Bosch, D. and Ma, H. 2017. Assessing water poverty in China using holistic and dynamic principal component analysis. *Social Indicators Research.* **130**(2), pp.537-561.

Paterson, L. and Goldstein, H. 1991. New statistical methods for analysing social structures: An introduction to multilevel models. *British Educational Research Journal.* **17**(4), pp.387-393.
Pearce, J.,Witten, K.,Hiscock, R. and Blakely, T. 2007. Are socially disadvantaged neighbourhoods deprived of health-related community resources? *International Journal of Epidemiology.* **36**(2), pp.348-355.
Pickett, K.E. and Pearl, M. 2001. Multilevel analyses of neighbourhood socioeconomic context and health outcomes: A critical review. *Journal of Epidemiology & Community Health.* **55**(2), pp.111-122.
Postiglione, Gerard A. 2009. Dislocated education: The case of Tibet. *Comparative Education Review.* **53**(4), pp.483-512.
Postiglione, G.A.,Jiao, B. and goldstein, M.C. 2011. Education in the Tibetan Autonomous Region: Policies and practices in rural and nomadic communities. In: Ryan, J. ed. *Education Reform in China: Changing Concepts, Contexts and Practces.* New York: Routledge, pp.92-110.
Prud'homme, R. 1995. The dangers of decentralization. *The World Bank Research Observer.* **10**(2), pp.201-220.
Qian, X. and Smyth, R. 2008. Measuring regional inequality of education in China: Widening coast-inland gap or widening rural-urban gap? *Journal of International Development.* **20**(2), pp.132-144.
Quinlan, J.R. 1986. Induction of decision trees. *Machine Learning.* **1**(1), pp.81-106.
Rangel, C. 2013. *Choosing success? Inequalities and opportunities in access to school choice in nine United States districts*. PhD thesis, University of Illinois at Urbana-Champaign.
Rao, J. and Ye, J.Z. 2016. From a virtuous cycle of rural-urban education to urban-oriented rural basic education in China: An explanation of the failure of China's rural school mapping adjustment policy. *Journal of Rural Studies.* **47**, pp.601-611.
Rawls, J. 1971. *A Theory of Justice.* Cambridge, Mass: Belknap Press of Harvard University Press.
Rickles, J.,Ong, P.M. and Houston, D. 2004. *School Integration and Residential Segregation in California: Challenges for Racial Equity.* [Online]. [Accessed May 1st]. Available from: https://escholarship.org/uc/item/75g9g9zp
Robeyns, I. 2005. The capability approach: A theoretical survey. *Journal of Human Development.* **6**(1), pp.93-117.
Robinson, W.S. 1950. Ecological correlations and the behavior of individuals. *American Sociological Review.* **15**(3), pp.351-357.
Rodríguez-Pose, A. and Ezcurra, R. 2009. Does decentralization matter for regional disparities? A cross-country analysis. *Journal of Economic Geography.* **10**(5), pp.619-644.
Saito, M. 2003. Amartya Sen's capability approach to education: A critical exploration. *Journal of Philosophy of Education.* **37**(1), pp.17-33.
Saltelli, A. 2007. Composite indicators between analysis and advocacy. *Social Indicators Research.* **81**(1), pp.65-77.
Satz, D. 2012. Unequal chances: Race, class and schooling. *Theory and Research in Education.* **10**(2), pp.155-170.
Schnell, I.,Diab, A.A.B. and Benenson, I. 2015. A global index for measuring socio-spatial segregation versus integration. *Applied Geography.* **58**, pp.179-188.

Schofield, J.W. 1989. Review of research on school desegregation's impact on elementary and secondary school students.
Sen, A. 1979. *Equality of what? Tanner Lecture on Human Values. .* Stanford University.
Sen, A. 1992. *Inequality Reexamined.* Oxford: Oxford University Press.
Sen, A. 2001. *Development as Freedom.* New York: Anchor Books.
Shuttleworth, I. and Gould, M. 2010. Distance between home and work: A multilevel analysis of individual workers, neighbourhoods, and employment sites in Northern Ireland. *Environment and Planning A.* **42**(5), pp.1221-1238.
Singleton, A.D. 2010. *Educational Opportunity: The Geography of Access to Higher Education.* Farnham: Ashgate.
Singleton, A.D. and Longley, P. 2015. The internal structure of Greater London: a comparison of national and regional geodemographic models. *Geo: Geography and Environment.* **2**(1), pp.69-87.
Singleton, A.D. and Longley, P.A. 2009. Creating open source geodemographics: Refining a national classification of census output areas for applications in higher education. *Papers in Regional Science.* **88**(3), pp.643-666.
Small, M.L. and Newman, K. 2001. Urban poverty after the truly disadvantaged: The rediscovery of the family, the neighborhood, and culture. *Annual Review of Sociology.* **27**, pp.23-45.
Soja, E. 2009. The city and spatial justice. *Justice Spatiale/Spatial Justice.* **1**(1), pp.1-5.
Soja, E.W. 2010. *Seeking spatial justice.* Minneapolis: University of Minnesota Press.
Song, Y.,Zenou, Y. and Ding, C. 2008. Let's not throw the baby out with the bath water: The role of urban villages in housing rural migrants in China. *Urban Studies.* **45**(2), pp.313-330.
Stafford, M. and Marmot, M. 2003. Neighbourhood deprivation and health: Does it affect us all equally? *International Journal of Epidemiology.* **32**(3), pp.357-366.
Standing Committee of the National People's Congress. 1995. The Education Law of the PRC. **Article 55**.
Standing Committee of the National People's Congress. 2006a. Compulsory Education Law of PRC. **Article 2**.
Standing Committee of the National People's Congress. 2006b. *Compulsory Education Law of PRC.*
Standing Committee of the National People's Congress. 2006c. Compulsory Education Law of PRC. **Article 12**.
Standing Committee of the National People's Congress. 2006d. Compulsory Education Law of PRC. **Article 42,Article 44, Article 45**.
Standing Committee of the National People's Congress. 2006e. Compulsory Education Law of PRC. **Article 22**.
Standing Committee of the National People's Congress. 2006f. Compulsory Education Law of PRC. **Article 42**.
Steele, F. 2008a. Introduction to multilevel modelling-concepts. *LEMMA VLE Module 5.* pp.1-45.
Steele, F. 2008b. Introduction to multilevel modelling-MLwiN practicals. *LEMMA VLE Module 5.* pp.1-48.

Steele, F.,Vignoles, A. and Jenkins, A. 2007. The effect of school resources on pupil attainment: A multilevel simultaneous equation modelling approach. *Journal of the Royal Statistical Society. Series A (Statistics in Society).* **170**(3), pp.801-824.

Subramanian, S.V.,Jones, K.,Kaddour, A. and Krieger, N. 2009. Revisiting Robinson: The perils of individualistic and ecologic fallacy. *International Journal of Epidemiology.* **38**(2), pp.342-360.

Sun, M. 2010. Education system reform in China after 1978: Some practical implications. *International Journal of Educational Management.* **24**(4), pp.314-329.

Sun, X. 2012. The national medium-and long-term educational reform and development guideline (2010-20). *Chinese Education & Society.* **45**(3), pp.9-22.

Talen, E. 2001. School, community, and spatial equity: An empirical investigation of access to elementary schools in West Virginia. *Annals of the Association of American Geographers.* **91**(3), pp.465-486.

Teese, R. and Polesel, J. 2003. *Undemocratic Schooling: Equity and Quality in Mass Secondary Education in Australia.* Carlton: Melbourne University Press.

Terzi, L. 2007. The capability to be educated. In: Walker, M. and Unterhalter, E. (eds). *Amartya Sen's Capability Approach and Social Justice in Education.* New York: Palgrave Macmillan, pp.25-44.

The Central People's Government of PRC. 2005. *Administrative Division.*

The Economist. 2016. *The Class Ceiling.* [Online]. [Accessed Jun 4th 2016]. Available from: http://www.economist.com/news/china/21699923-chinas-education-system-deeply-unfair-class-ceiling

The National People's Congress of PRC. 1958. *Household Register System of People's Republic of China.*

The state Council of China. 2001. *The Decision on the Reform and Development of Compulsory Education.* [Online]. [Accessed May 29th]. Available from: http://old.moe.gov.cn/publicfiles/business/htmlfiles/moe/moe_16/200105/132.html

The State Council of China. 2010. *National Medium- and Long-Term Educational Reform and Development Plan Outline* [Online]. [Accessed July 29th]. Available from: http://www.gov.cn/jrzg/2010-07/29/content_1667143.htm

The State Council of China. 2014. *To Further Promote the Reform of Household Registration System.* [Online]. [Accessed July 30th]. Available from: Http://www.gov.cn/zhengce/content/2014-07/30/content_8944.htm

Thomas, M.,Stillwell, J. and Gould, M. 2015. Modelling multilevel variations in distance moved between origins and destinations in England and Wales. *Environment and Planning A.* **47**(4), pp.996-1014.

Thomas, M.J.,Stillwell, J. and Gould, M. 2014. *Modelling individual and place variations in residential moves using commercial data and official statistics*. thesis, School of Geography, University of Leeds

Thomas, S.M. and Peng, W.J. 2010. Methods to evaluate educational quality and improvement in China. In: Ryan, J. ed. *Education Reform in China: Changing Concepts, Contexts and Practices.* London: Routledge.

Thomas, V. and Wang, Y. 2009. 2. Distribution of opportunities key to development. In: Holsinger, D.B. and Jacob, W.J. (eds). *Inequality in Education: Comparative and International Perspectives.* Hong Kong: Springer Science & Business Media, pp.34-58.

Tickle, M.,Moulding, G.,Milsom, K. and Blinkhorn, A. 2000. Primary dental care: Socioeconomic and geographical influences on primary dental care preferences in a population of young children. *British Dental Journal.* **188**(10), pp.559-562.

Tomintz, M.N.,Clarke, G.P. and Rigby, J.E. 2009. Planning the location of stop smoking services at the local level: A geographic analysis. *Journal of Smoking Cessation.* **4**(02), pp.61-73.

Tomul, E. 2009. Measuring regional inequality of education in Turkey: An evaluation by Gini index. *Procedia-Social and Behavioral Sciences.* **1**(1), pp.949-952.

Tonks, D. 1999. Access to UK higher education, 1991-98: Using geodemographics. *Widening Participation and Lifelong Learning.* **1**(2), pp.6-15.

Tonks, D.G. and Farr, M. 1995. Market segments for higher education: Using geodemographics. *Marketing Intelligence & Planning.* **13**(4), pp.24-33.

Tsang, M.C. 1996. Financial reform of basic education in China. *Economics of Education Review.* **15**(4), pp.423-444.

Tsang, M.C. 2000. Education and national development in China since 1949: Oscillating policies and enduring dilemmas. *China Review.* pp.579-618.

Tsou, K.-W.,Hung, Y.-T. and Chang, Y.-L. 2005. An accessibility-based integrated measure of relative spatial equity in urban public facilities. *Cities.* **22**(6), pp.424-435.

Tyler, W. 1977. *The Sociology of Educational Inequality.* London: Methuen.

UNESCO. 2006. Understandings of literacy. *Education for All Global Monitoring Report 2006.*

UNESCO. 2010. *Deprivation and Marginalization in Education (DME).* [Online]. Available from: http://www.unesco.org/fileadmin/MULTIMEDIA/HQ/ED/GMR/html/dme-1.html

UNESCO. 2015. Education 2030 Incheon Declaration and Framework for Action.

UNESCO. 2016. *Global Education Monitoring Report 2016: Education for People and Planet: Creating Sustainable Futures for All.* (978-92-3-100167-3). Paris: UNESCO Publishing.

Unger, J. 1984. Severing the links between school performance and careers: The experience of China's urban schools, 1968-1976. *Comparative Education.* **20**(1), pp.93-102.

United Nations. 2015. *Goal 4: Ensure Inclusive and Quality Education for All and Promote Lifelong Learning.* [Online]. Available from: http://www.un.org/sustainabledevelopment/education/

van Leeuwen, B. and Foldvari, P. 2014. Educational and income inequality in Europe, ca. 1870-2000. *Cliometrica.* **8**(3), pp.271-300.

Vaughan, R. 2007. Measuring capabilities: An example from girls' schooling. In: Walker, M. and Unterhalter, E. (eds). *Amartya Sen's Capability*

Approach and Social Justice in Education. New York: Palgrave Macmillan, pp.109-130.

Vickers, D. and Rees, P. 2007. Creating the UK national statistics 2001 output area classification. *Journal of the Royal Statistical Society. Series A (Statistics in Society).* **170**(2), pp.379-403.

Voas, D. and Williamson, p. 2001. The diversity of diversity: A critique of geodemographic classification. *Area.* **33**(1), pp.63-76.

Vyas, S. and Kumaranayake, L. 2006. Constructing socio-economic status indices: How to use principal components analysis. *Health Policy and Planning.* **21**(6), pp.459-468.

Walker, M. 2005. Amartya Sen's capability approach and education. *Educational Action Research.* **13**(1), pp.103-110.

Walker, M. and Unterhalter, E. 2007. The capability approach: Its potential for work in education. In: Walker, M. and Unterhalter, E. (eds). *Amartya Sen's Capability Approach and Social Justice in Education.* New York: Palgrave Macmillan, pp.1-18.

Wang, F. 2012a. Measurement, optimization, and impact of health care accessibility: A methodological review. *Annals of the Association of American Geographers.* **102**(5), pp.1104-1112.

Wang, H.,Guo, F. and Cheng, Z. 2015. Discrimination in migrant workers' welfare entitlements and benefits in urban labour market: Findings from a four-city study in China. *Population, Space and Place.* **21**(2), pp.124-139.

Wang, L. 2008. The marginality of migrant children in the urban Chinese educational system. *British Journal of Sociology of Education.* **29**(6), pp.691-703.

Wang, L. 2012b. Social exclusion and education inequality: Towards an integrated analytical framework for the urban-rural divide in China. *British Journal of Sociology of Education.* **33**(3), pp.409-430.

Wang, W. 2014. Decomposing inequality in compulsory education finance in China: 1998-2008. *Public Finance and Management.* **14**(4), pp.437-458.

Wang, W. and Zhao, Z. 2014. Spatial decomposition of funding inequality in China's basic education: A four-level Theil index analysis. *Public Finance and Management.* **14**(4), p416.

Wang, W. and Zhao, Z.J. 2012. Rural taxation reforms and reforms and compulsory education finance in China. *Journal of Public Budgeting, Accounting & Financial Management.* **24**(1), p136.

Wang, Y. and Yao, Y. 2003. Sources of China's economic growth 1952-1999: Incorporating human capital accumulation. *China Economic Review.* **14**(1), pp.32-52.

Waters, J.L. 2017. Education unbound? Enlivening debates with a mobilities perspective on learning. *Progress in Human Geography.* **41**(3), pp.279-298.

Whitty, G. 1997. Creating quasi-markets in education: A review of recent research on parental choice and school autonomy in three countries. *Review of Research in Education.* **22**, pp.3-47.

Wilkinson, R.G. and Pickett, K. 2010. *The Spirit Level: Why Equality Is Better for Everyone.* London: Penguin Books.

Williams, S. and Wang, F. 2014. Disparities in accessibility of public high schools, in metropolitan Baton Rouge, Louisiana 1990–2010. *Urban Geography.* **35**(7), pp.1066-1083.

Williams, S.E. 2012. *Accessibility to Public High Schools and School Performance in Metropolitan Baton Rouge, Louisiana 1990-2010*. thesis, Louisiana State University.

Williamson, T.,Ashby, D.I. and Webber, R. 2005. Young offenders, schools and the neighbourhood: A new approach to data-analysis for community policing. *Journal of Community & Applied Social Psychology.* **15**(3), pp.203-228.

Wodtke, G.T.,Harding, D.J. and Elwert, F. 2011. Neighborhood effects in temporal perspective: The impact of long-term exposure to concentrated disadvantage on high school graduation. *American Sociological Review.* **76**(5), pp.713-736.

Wu, B.M.,Birkin, M.H. and Rees, P.H. 2008. A spatial microsimulation model with student agents. *Computers, Environment and Urban Systems.* **32**(6), pp.440-453.

Wu, K.B. 2009. 4. Education and Inequality in the Developing World. In: Holsinger, D.B. and Jacob, W.J. (eds). *Inequality in Education: Comparative and International Perspectives.* Hong Kong: Springer Science & Business Media, pp.86-127.

Wu, Q.,Zhang, X. and Waley, P. 2016. Jiaoyufication: When gentrification goes to school in the Chinese inner city. *Urban Studies.* **53**(16), pp.3510-3526.

Wu, X. 2013. The school choice market in China: A case study. *Educational Research.* **55**(2), pp.195-218.

Wu, X. 2014. *School Choice in China: A Ddifferent Tale?* London: Routledge.

Wu, X. and Treiman, D.J. 2004. The household registration system and social stratification in China: 1955-1996. *Demography.* **41**(2), pp.363-384.

Xiang, L.,Stillwell, J.,Burns, L.,Heppenstall, A. and Norman, P. 2018. A geodemographic classification of sub-districts to identify education inequality in Central Beijing. *Computers, Environment and Urban Systems.* **70**, pp.59-70.

Xie, P. 2015. *Annual Report on the Development of China's Investment(2015).* Beijing: Social Sciences Academic Press(China).

Xie, Y. and Zhou, X. 2014. Income inequality in today's China. *Proceedings of the National Academy of Sciences.* **111**(19), pp.6928-6933.

Xinhuanet. 2015. *Comprehensive explanation of entrance policy in core areas of Beijing* [Online]. [Accessed 14th April]. Available from: http://www.bj.xinhuanet.com/fw/2015-04/14/c_1114961631.htm

Yang, Z.,Yi, C.,Zhang, W. and Zhang, C. 2014. Affordability of housing and accessibility of public services: Evaluation of housing programs in Beijing. *Journal of Housing and the Built Environment.* **29**(3), pp.521-540.

Yanqing, D. 2012. The problems with access to compulsory education in China and the effects of the policy of direct subsidies to students. *Chinese Education & Society.* **45**(1), pp.13-21.

Young, I.M. 2001. Equality of whom? Social groups and judgments of injustice. *Journal of Political Philosophy.* **9**(1), pp.1-18.

Zhang, D.,Li, X. and Xue, J. 2015. Education inequality between rural and urban areas of the People's Republic of China, migrants' children education, and some implications. *Asian Development Review.* **32**(1), pp.196-224.

Zhang, G.,Zhao, Y.,Lei, J.,Zhao, Y. and Qiu, W. 2012. Policy changes and educational reforms in China: Decentralization and marketization. *On the Horizon.* **20**(4), pp.313-323.

Zhang, H. 2011. School desegregation and white flight revisited: A spatial analysis from a metropolitan perspective. *Urban Geography.* **32**(8), pp.1208-1226.

Zhang, X. and Kanbur, R. 2009. Spatial inequality in education and health care in China. In: Fan, S.,Kanbur, S.M.R. and Zhang, X. (eds). *Regional Inequality in China: Trends, Explanations, and Policy Responses.* London: Rouledge.

Zhao, D. and Parolin, B. 2014. Merged or unmerged school? School preferences in the context of school mapping restructure in rural China. *The Asia-Pacific Education Researcher.* **23**(3), pp.547-563.

Zhao, P. and Howden-Chapman, P. 2010. Social inequalities in mobility: The impact of the hukou system on migrants' job accessibility and commuting costs in Beijing. *International Development Planning Review.* **32**(3), pp.363-384.

Zhao, P.,Lü, B. and de Roo, G. 2010. Urban expansion and transportation: The impact of urban form on commuting patterns on the city fringe of Beijing. *Environment and Planning A.* **42**(10), pp.2467-2486.

Zheng, S.,Hu, W. and Wang, R. 2015. How much is a good school worth in Beijing? Identifying price premium with paired resale and rental data. *The Journal of Real Estate Finance and Economics.* pp.1-16.

Zhu, M. and Peyrache, A. 2017. The quality and efficiency of public service delivery in the UK and China. *Regional Studies.* **51**(2), pp.285-296.