Article:

Xiang, L, Gould, M and Stillwell, J orcid.org/0000-0002-8803-2001 (2022) What determines pupils' travel distance to school in China? A multilevel analysis of educational access in Beijing. Urban Studies, 59 (5). pp. 1048-1067. ISSN 0042-0980

https://doi.org/10.1177/0042098021996227

What determines pupils' travel distance to school in China? A multilevel analysis of educational access in Beijing

Abstract

Whilst access to school is one crucial aspect of education equality, there is a lack of research on factors that influence the distance that pupils travel to school. Previous studies have failed to reveal the relationship between pupils' socio-spatial characteristics and travel distance. This paper uncovers the multilevel structure, ignored hitherto, that underpins the determinants of pupils' travel distance. Using detailed travel survey data for Beijing, and an appropriate multilevel modelling approach, this research reveals that contextual variation remains having taken account of compositional (individual-level) variables; and that contextual factors, i.e. school density and neighbourhood context, are more influential when compared to individual-level factors except for education stage and housing type. The policy implications include improved planning for schools in comparatively deprived areas, increased provision of affordable housing, and enhanced education opportunities for migrant children.

Keywords

travel distance, multilevel modelling, access, socio-spatial inequalities, context effects

Introduction

Education equality is valued by numerous leading worldwide development agencies (OECD, 2016; UNESCO, 2015; United Nations, 2015). Equality in education matters not only because of its intrinsic worth but also because it is essential for realising other human rights, offering the chance to counter existing inequalities (e.g., economic, wellbeing). Spatial access to educational facilities (e.g., travel-to-school distance) is an important aspect of education equality (Levin, 1976; Farrell, 1999). Different travel distances provide students with different amounts of disposable time. Mark Witham (1997) found that the time children spend commuting ('lost opportunity time') led to serious erosion of their study time. Longer commutes result in greater physical and mental weariness (Talen, 2001; Hamnett and Butler, 2011) and children living far from school also have less opportunity to be involved in after-school activities and to get other additional benefits (Pacione, 1989). Consequently, spatial access to education facilities can have an important effect on children receiving education.

So, what determines a pupil's spatial access to school? From a social justice (Brighouse, 2000; Israel and Frenkel, 2017) and policy-making perspective, it is important to explore how social and spatial factors influence pupils' travel distance, whether long commuting distances to school are problematic for specific individuals or groups with certain characteristics, and whether long commutes are more pronounced in particular areas. There is a scarcity of small-area research on education inequality in China and beyond. Some initial research by Emily Talen (2001) suggested the relationship between pupils' socioeconomic traits (i.e., 'who they are') and travel distance was neither conclusive, nor consistent across different counties. Talen's so-called 'unpatterned inequality' is due to the neglected influence of socio-spatial context (i.e., 'where they live') on pupils' travel distance and also the hierarchical nature of the education system. Pupils are nested within the geographical areas (e.g. sub-districts) in which they live and; pupils' residential location has a crucial impact on their access to available educational opportunities and resources (e.g., school places, school quality, enrolment policy) (Gordon and Monastiriotis, 2007; Hamnett and Butler, 2011). Accordingly, the level and nature of education that pupils can access is not only influenced by who they are - their personal characteristics - but also by where they live.

To reduce inequalities and promote educational equality by making sure all residents are treated equally in terms of school access and resources regardless of their geographical location and potential mobility (e.g., socioeconomic characteristics, disability), it is essential that the complex socio-spatial multilevel structures that create these inequalities be examined more closely. Thus, a more flexible statistical methodology, multilevel modelling, which acknowledges the structure of hierarchical data (Goldstein and Silver, 1989; Owen *et al.*, 2016), is adopted in this research. Using individual observations from two comprehensive and geographically identifiable pupil travel surveys, we assess the influence of contextual factors (spatial factors) in determining pupils' travel distance in Beijing, simultaneously taking account of possible confounding pupils' individual level characteristics (e.g., hukou, household income, parents' qualifications) and thereby taking account of the composition of different areas.

The remainder of the paper is organised as follows; the next section reviews factors influencing travel distance that inform the multilevel modelling. Thereafter, the data and the distribution of education resources in Beijing are described before the modelling strategy and specifications of model estimation are outlined. The results, discussion and policy implications are then presented before concluding.

Individual and contextual determinants of educational commuting

Geographical location, and socio-economic and demographic composition of pupils' residential areas are key dimensions that determine the distance of travel to school. Previous research concerning geographical location has focused on the distribution of educational resources or pupils' accessibility to these (Talen, 2001; Dadashpoor *et al.*, 2016). However, locational discrimination in education occurs due to biases imposed on certain social groups linked to their geographical locations. People living in similar neighbourhoods tend to share similar social status, have a similar habitus and potentially enjoy similar chances of accessing education resources (Israel and Frenkel, 2017). Spatial inequality is therefore normally related to and/or derived from certain social inequalities, with social factors such as social class and ethnicity shaping spatial or locational discrimination in terms of access to education resources (Soja, 2009).

As socio-spatial processes are normally complex and associated with various factors, analysis may be more effective when combining different social variables rather than just employing one key explanatory variable (Williams and Wang, 2014). For example, rapid urbanisation and the specific institutional background in China (e.g., socialist market economic system and hukou system[1]) have generated new forms of socio-spatial differentiation and segregation, which are based on multiple factors like income, occupation and residential registration status (Gu *et al.*, 2006; Xiang *et al.*, 2018). However, there is limited research that considers multi-dimensional and socio-spatial differentiation. Xiang *et al.* (2018) used geodemographic classification

[1] Hukou is the Chinese system whereby people are legally required to register as residents of a particular area. This system enables population movement control in China. Residents registered with local communities, with local hukou in urban or rural areas, are granted priority access to local services, such as public housing and education, while those migrating from rural areas to the cities, with a non-local hukou, have limited access to the same services and are disadvantaged in receiving education in urban areas.

techniques to create the first education-related area typology for Beijing (China). This classification is employed in this paper to characterise the socio-spatial context for pupils' access to schools. Moreover, the impact of pupils' individual demographic characteristics (e.g., age, sex, ethnicity), and social economic background (e.g., poverty level, home ownership, car ownership, family income) on their educational mobility and access are analysed (Talen, 2001; Hamnett and Butler, 2011; Logan *et al.*, 2012; De la Fuente *et al.*, 2013; Williams and Wang, 2014).

We suggest using single-level generalised models based solely on individuals' characteristics neglects the contextual neighbourhood effects and the hierarchical structure of the education system (Goldstein and Silver, 1989) and limits conclusions. A single-level model assumes that all the observations are independently and identically distributed, indicating that individuals are expected to be independent of one another and come from an unstructured random sample (Goldstein, 2011). However, there is a hierarchical structure in both our research context and data where individual pupils are nested within residential neighbourhoods (i.e., sub-districts); and individuals in the same neighbourhood tend to be more alike than individuals chosen randomly from the whole population (Wodtke *et al.*, 2011; Xiang *et al.*, 2018).

Although contextual differences can be modelled through a fixed-effects approach by including dummy neighbourhood variables in a single-level model, this limits further inference of contextual characteristics (level-2 or neighbourhood level variables), which may be of substantial importance in explaining variations between neighbourhoods. All neighbourhood-level characteristics will be confounded with the fixed effects of each neighbourhood, making the identification of the influence of each neighbourhood-level variable mathematically impossible (Fielding, 2004). Additionally, modelling neighbourhoods as fixed effects can be inefficient when the number of neighbourhoods is large and may result in poor estimates (Jones and Bullen, 1994). We propose multilevel modelling pupil travel distance using a random part expansion, treating neighbourhoods as a separate level and assuming these higher level units are normally distributed.

In summary, the modelled relationships are allowed to vary both between individuals and also from place to place. The advantage of using multilevel modelling is that it works by specifying models at macro (e.g., neighbourhood) and micro (e.g., individual) levels which are then combined into an overall hierarchical model. This allows us to explore both how individual travel distance varies by type of pupil, how it varies across neighbourhoods, and also investigate neighbourhood covariates. By doing so, we properly consider how ‘geography matters’.

Data and the distribution of education resources in Beijing

Two datasets, the Fifth Comprehensive Household Travel Survey (CHTS) and the Pupil Travel Survey (PTS) are used in our analyses. Both were launched by the Beijing Municipal Commission of Transport in 2014. The Fifth CHTS is a household survey and 0.52% sample of the Beijing population across all districts and counties. The surveyed areas of focus are within Central Beijing; with samples for the main new towns/urban fringe areas also collected. Given that the purpose of this research, only school commuter trips of 6-18-year-old respondents are selected. The CHTS data is hierarchically structured with 5,373 pupils (level-1 units) nested within the 251 sub-districts (level-2 units).

In contrast to countries with greater freedom choice of school (e.g. UK), a ‘nearby enrolment policy’ was stipulated by the Compulsory Education Law of China (SCNPC, 2006). Accordingly, school-age children are required to enrol in schools near their current residence. Nearby enrolment in Beijing is implemented by adopting school districts, the catchment areas for enrolment into primary and secondary schools. However, those pupils wishing to enrol in a nearby public school must show proof of local hukou registration or parent’s property ownership certificate. Only those families who have local hukou registration or sufficient economic capital to buy a property are guaranteed a place for their children under the nearby enrolment policy; whilst migrants without local hukou who are in rental housing cannot enrol their children in nearby schools unless these are undersubscribed. Boarding schools and private schools are not restricted by this policy. Given only a small percentage of children choose to enrol in these (4.5% primary pupils chose non-government schools in 2019-20 according to Beijing Education Development Statistics, 2019-20), the schools discussed in this paper are all public schools.

The input variables are selected based on related literature, correlation analysis and the specific institutional background of Beijing and China. Most of the predictors are categorical variables recoded using contrast coding to create a series of dummy variables. The base categories used in models for each variable are denoted in bold italics in Table 1.

Table 1 CHTS used in multilevel modelling

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

To avoid biased sampling in the CHTS and confirm the characteristics and patterns revealed, we have also used the PTS, a larger dataset with 460,306 valid records after data cleaning, covering more than half the pupils (62.5% of all primary pupils and 53.5% of all secondary pupils). Whilst this has fewer variables than the CHTS, we focus on those that are common to both datasets.

The median travel distances of primary, middle and high school pupils contained in the CHTS are around 2, 4 and 6 kilometres respectively. The travel distance distribution is more variable for high school pupils. According to the 2015 Statistical Yearbook of Beijing, the number of primary school pupils increased most (170%) between 2005-2014, while the number of pupils attending middle school stayed almost the same and high school pupils decreased by 60%. A strict school enrolment policy means that in-migrants are not allowed to attend public high school in Beijing. Thus, most high school pupils have Beijing hukou except for those who enrol in private high schools; the proportion of school-age pupils whose hukou is outside Beijing decreases from primary to middle to high school (23%, 15% and 6% respectively according to the CHTS).

In terms of median travel distance for pupils living in different geodemographic clusters in Central Beijing, the travel distance for 'In-migrant children living in deprived suburbs', is the longest (Figure 1) followed by those living in the 'Urban-rural transition belt'. Additionally, the travel distance for 'Long settled Beijing society', and 'Employees of the state-owned enterprises and white collar workers' are shorter than the median distances for the other three clusters.

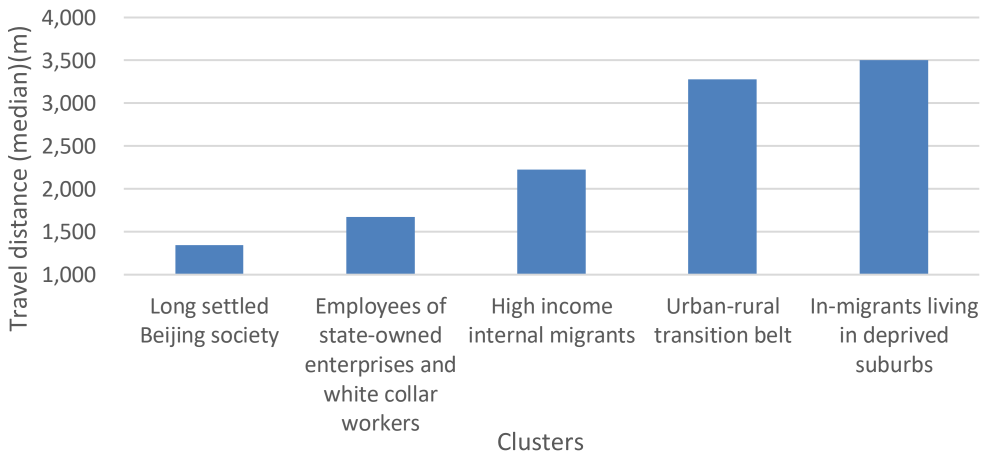
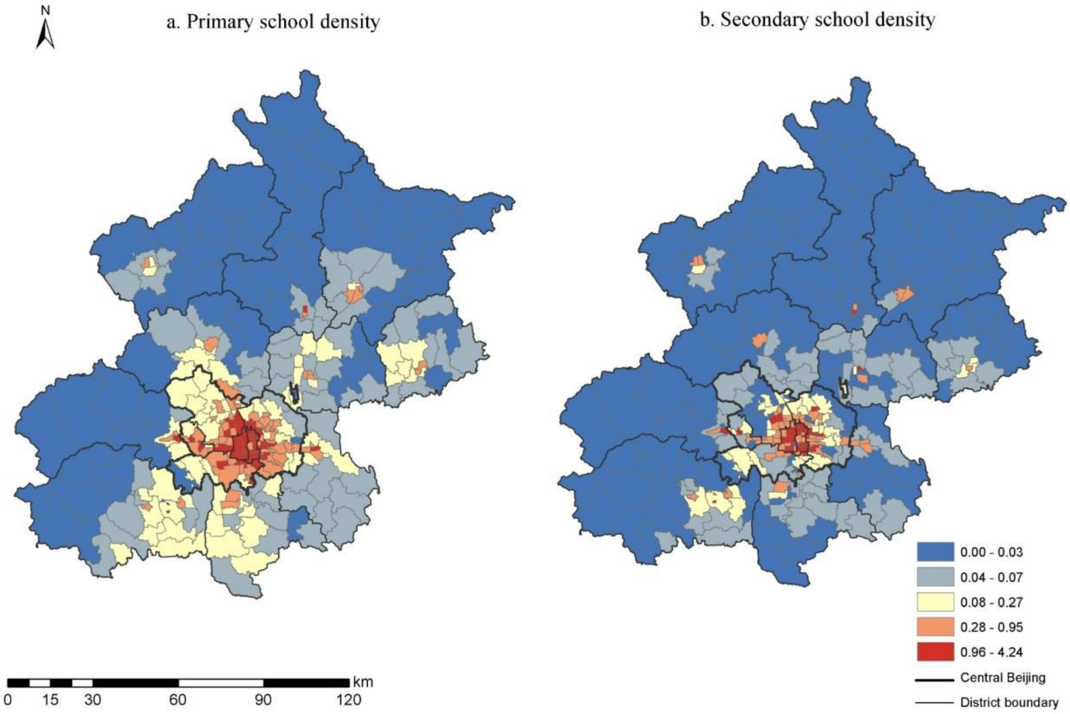


Figure 1. Median travel distance of pupils from different geodemographic clusters in Central Beijing (CHTS)

The distribution of education resources in Beijing is very unequal. Figure 2 shows the general pattern in 2014 using school density calculated by dividing the number of schools located in each sub-district by the area of each sub-district (schools per km^2), and decreases considerably from inner to outer areas, although the distributions of both types of school are similar.



Data source: National Administration of Surveying, Mapping and Geo-information.

Figure 2. School density quintile distribution of primary and secondary schools by sub-district in Beijing, 2014

Figure 3 provides boxplots of school density by the geodemographic classifications. The 'Long settled Beijing society' and 'Employees of state-owned enterprises and white collar workers' have the lowest travel distance have the highest school density. In contrast, 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. The 'Outside Central Beijing' cluster is not 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, is already as high as that of the 'Long settled Beijing society' and 'Employees of state-owned enterprises and white collar workers'.

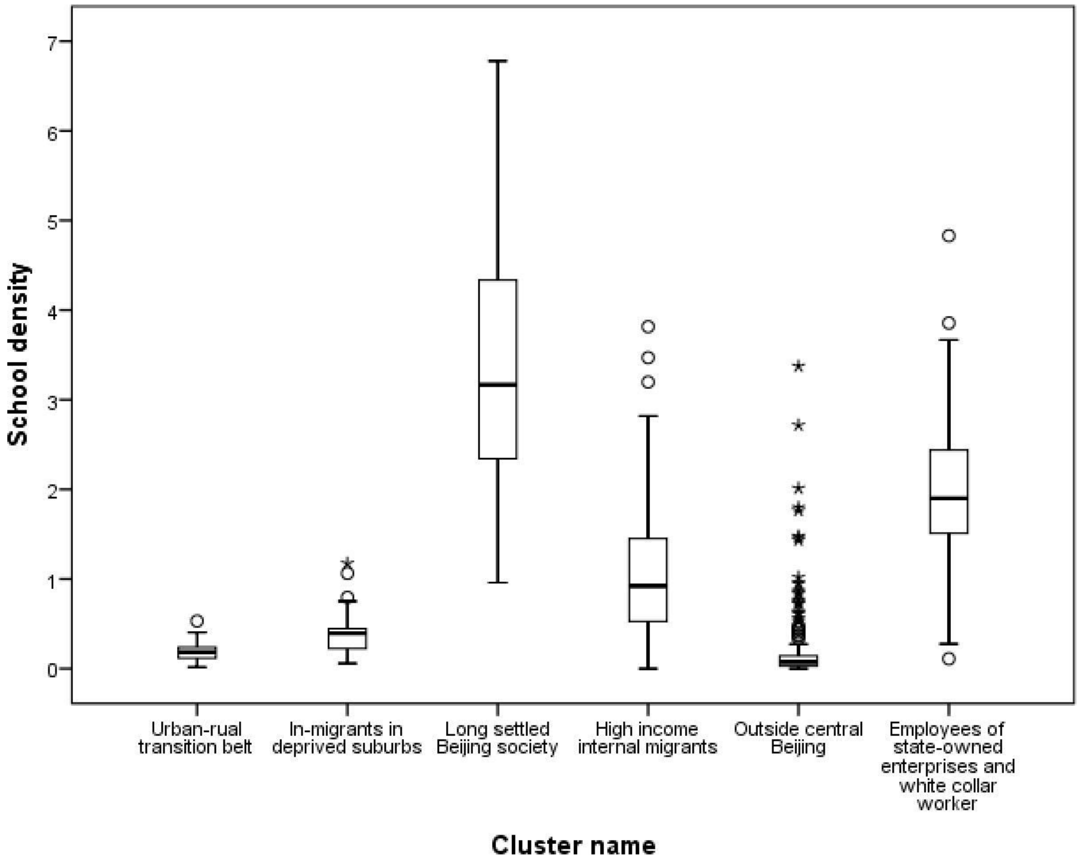


Figure 3. Boxplot of school density and geodemographic cluster of residence

The CHTS survey suggests that more than 60% of pupil households own a car and those households with higher income and higher qualifications are more likely to own cars. Purchased public housing (20.2%) and rented housing (16.3%) are the two most important housing types except for commercial housing in Central Beijing, whilst self-built housing is the other most important housing type in areas outside Central Beijing.

The larger PTS dataset validates the reliability of the smaller CHTS dataset; whilst the latter has more detailed individual information than the former. Comparisons between the two datasets suggest the results acquired from the CHTS-based multilevel models are reasonably robust.

Modelling multilevel variations in pupils' travel distance

Pupil travel distance from home to school is the response variable, which is positively skewed. Thus, it needs log-transformation before modelling (Shuttleworth and Gould, 2010). The explanatory variables included in multilevel model are listed in Table 1 and are based on the descriptive analysis and previous research. All the micro-level characteristics are selected from the CHTS, including hukou status, income, housing type, travel mode and education stage. These individual-level data are supplemented by two contextual indicators derived from other datasets. The education-related geodemographic classification for Central Beijing (Xiang *et al.*, 2018) summarises the characteristics of residents in specific clusters and provides a variable option for operationalising origin (residence) context. The sub-districts of Central Beijing are classified into five 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. In this research, sub-districts outside Central Beijing are treated as a sixth area-type. School density is the other context variable that is used.

In terms of the modelling strategy, to confirm that there is strong evidence of between sub-district differences in logged travel distance, a null model (no predictors) with random intercepts specified (model 1) has been fitted first to compare the level-2 between sub-districts variance with the level-1 between individuals variance (Table 2). Thereafter, compositional differences and context effects are considered by introducing the level-1 and level-2 covariates into the fixed part of the model. Following this, if there remains large variance at a higher level (and statistically significant), this is regarded as evidence of contextual effects. Including the micro-level covariates in models reveals their influence on variations in pupils' travel distance can also help identify geographical variations that are a consequence of compositional differences. Two contextual variables, geodemographic classification and school density within each sub-district, are included in Models 2 and 3 respectively. To uncover variability in distances travelled for different groups, the level-1 heterogeneity for the individual level predictors, hukou, education stage, income and housing, are included in models 4-7 (Table 3). All the above models are estimated in MLwiN, a software package dedicated to the implementation of a wide range of multilevel models (Steele, 2008, Fielding, 2010). The Iterative Generalised Least Squares (IGLS) algorithm, a fast and reliable algorithm to achieve maximum likelihood estimators, is used (Goldstein, 2011). Wald-tests are used to assess the statistical significance of fixed- and random-part model estimates, while the Likelihood Ratio Test (LRT) has been employed to judge overall model fit.

Table 2 Multilevel models with covariates and random intercepts (Models 1-3)

	Model 1		**Model 2**		**Model 3**	
Parameter	Coeff	S.E.	Coeff	S.E.	Coeff	S.E.
Fix 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 variance: between sub-district	0.053	0.006	0.037	0.004	0.033	0.004
Level-1 variance: between-individual	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*log likelihood)	5357.709		4458.464		4438.925	

* Denotes statistical significance ($p<0.05$)

Table 3 Results from fitted multilevel models (models 4-7)

	Model 4	**S.E.**	**Model 5**	**S.E.**	**Model 6**	**S.E.**	**Model 7**	**S.E.**
Parameter	Coeff	S.E.	Coeff	S.E.	Coeff	S.E.	Coeff	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; 'Employees 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	

* Denotes statistical significance ($p<0.05$)

Model results

Variance component model

Model 1 is a two-level null variance component model for pupils nested within their sub-district of residence (Table 2). Although the majority of variation is still found between individuals, there is also strong evidence of contextual variation. Level-2 sub-district variation accounts for 26.8%[2] of the total variance in pupil travel distance. As the micro-level predictors have not been included in this model, one cannot conclude that there are substantial place-based variations, until micro-level compositional predictors are included, and before contextual covariates are introduced.

Random intercept models with covariates

The individual-level and sub-district level context covariates are included in models 2 and 3. This leads to substantial model improvement, as evidenced by a large decrease in model deviance (around 900) compared with model 1. The inclusion of level-1 and level-2 covariates results in a reduction in level-1 and level-2 variances. Both models share the 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.

All the covariates are incorporated in the fixed-part of the model as additive effects. In model 2, the transformed estimate for the constant suggests a mean distance travel to school of 1.4km[3] for the base category pupils (whose hukou is within a district, with middle income parents, living in commercial housing, own a car, attend a primary school, and live in sub-district labelled as 'Employees of state-owned enterprises and white-collar workers'). For the individual-level predictors, the pupils' education stage

---

2 The variance partition coefficient (VPC) is calculated as :

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

and thus for model 1 is: 0.053 / (0.053+0.145) = 0.268.

3 The antilog of the constant, 3.156: $10^{3.156} \approx 1432$ (metres).

provides the largest effects with high school pupils travelling a 1.76km[4] and middle school pupils travelling 0.57km further than primary school pupils, holding the other covariates constant. Two types of housing also are associated with pupil travel distance: those 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 pupil commute distance; those with hukou outside their district of residence but within Beijing travelling an extra 0.31km when other effects are controlled for and is consistent with results reported in Xiang *et al.* (2018). However, the effect of having hukou outside Beijing is not statistically significant. Regarding 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.

Other than education stage, the largest differential effects on travel distance are found for the contextual variables. Pupils who live in clusters labelled 'Urban-rural transition belt', 'In-migrants in deprived suburbs', 'Outside Central Beijing' and 'High income internal migrants' travel an extra 1.7km, 1.2km, 0.96km and 0.4km respectively compared to pupils in the base category, 'Employees of state-owned enterprises and white collar workers'. The mean distance travelled by pupils in the 'Long settled Beijing society' cluster is not statistically different from pupils in the base category.

In model 3, school density replaces the geodemographic clusters as the contextual variable and displays very similar estimation results. The estimates for the fixed part and random part are very similar to that of model 2. Pupil travel distance decreases by 0.3km for a unit increase of school density, holding all the other variables constant. The deviance of model 3 is smaller than that of model 2, and the better fit for model 3 is probably due to the coarse definition of the cluster, 'Outside Central Beijing', in model 2, which has been explained in the previous section (see Figure 3). However, the inclusion of geodemographic groups in the model has the advantage of revealing the relationship between the pupil travel distance and the socio-spatial

[4] Obtained by using the antilog of the sum of the coefficient of high school , 0.348, and the constant, 3.15, to minus the antilog of the constant: $10^{3.156+0.348} - 3.156$: $10^{3.156}$ ≈1759 (metres).

characteristics of their residential areas directly, while school density can only reflect spatial inequalities.

When the level-2 sub-district scale residuals are mapped (Figure 4), it is apparent that pupils living in the urban fringes having a tendency to travel longer distances than those living in inner-city areas. Also, the travel distance of pupils who reside in new towns outside Central Beijing is shorter than surrounding areas. The pattern of the unobserved contextual effects is, to some extent, consistent with the distribution of school resources, although school density has been controlled in the model. This suggests there are some omitted variables, e.g., school size and residential density, due to the limitation of data availability.

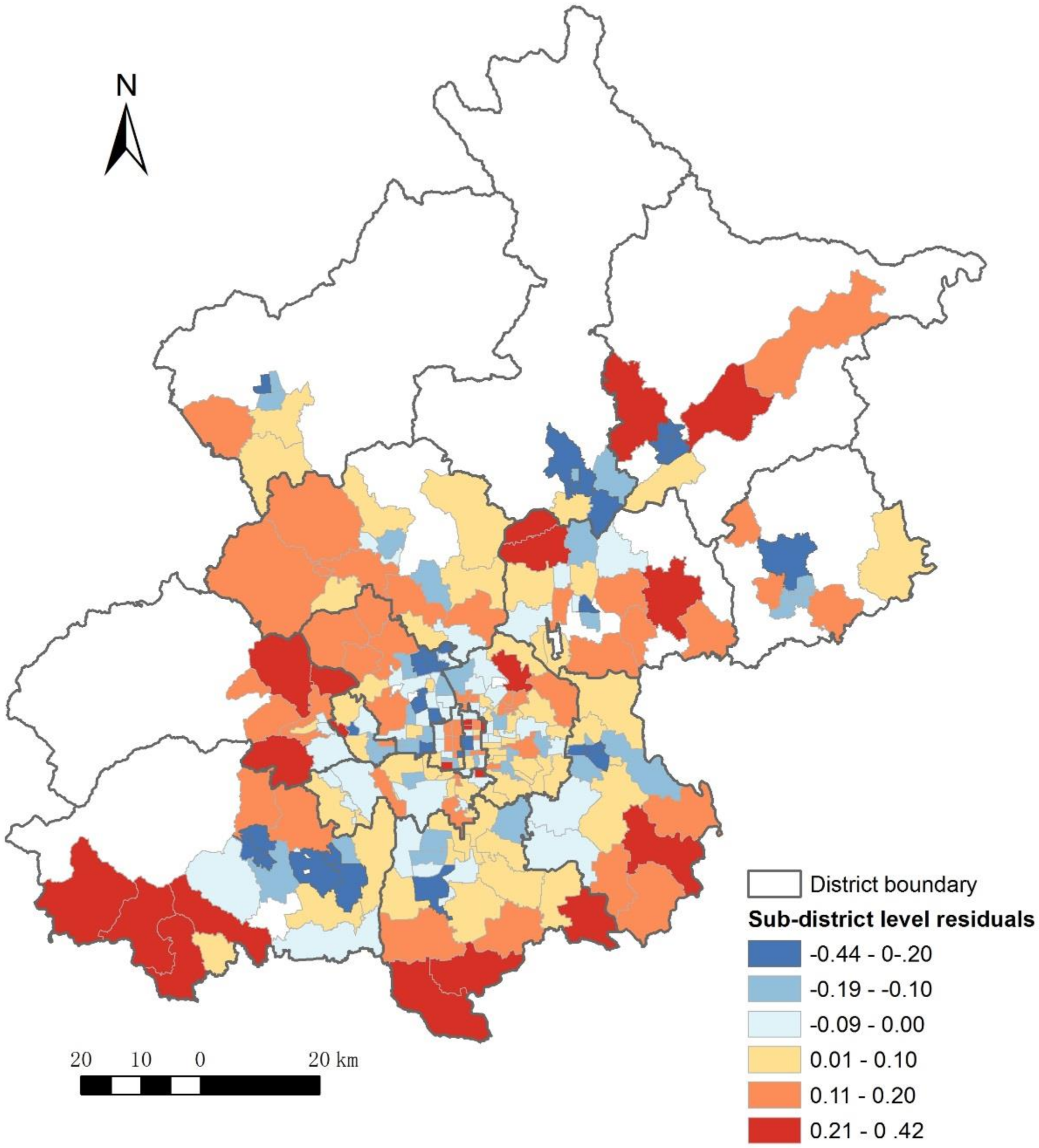


Figure 4. Estimated sub-district level random effects from model 3

A further investigation of pupil heterogeneity effects

The differential variability between pupils with different characteristics is explored by fitting a series of models in which the level-1 equal variance assumption is relaxed (Table 3). Models 4-7 permit level-1 variance to be associated with hukou status, education stage, income and housing type while containing the same set of fixed part variables. These four models all showed improved model fit by employing the LRT test. 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 to occur 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 travel distance. It is probably not only because high school pupils are old enough to travel further but also because they are enrolled across the districts within Beijing and thus have more flexibility in their school choices. In model 6, the pupils from high-income households have the highest variability, whilst pupils in low-income households have the lowest variation. Finally, 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.

Discussion and policy implications

Spatial inequalities of education resources

With the help of detailed descriptive analysis and multilevel modelling, the influence of micro-level covariates and macro-level context effects on travel-to-school distance have been identified. Some implications of these results in terms of access to education are discussed in this section. Figure 2 revealed a very uneven distribution of school places in Beijing. 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 Figure 2, school density decreases from inner areas to outer areas in Central Beijing and decreases from some new towns to the outer areas outside Central Beijing; correspondingly, the travel distance increases from inner areas to outer areas and increases from some new towns to outer areas outside Central Beijing. Unlike education provision in countries experiencing counterurbanisation, like the UK (Hamnett and Butler, 2011), Beijing is still in the process of rapid urbanisation. This together with comparatively lagging education planning means that education provision is insufficient in outer areas.

Influence of social-spatial contextual characteristics

If long travel distance is associated with disadvantaged social groups (e.g., non-local or low-income residents), it may increase their financial burden, decrease the opportunities of pupils receiving education, (re)produce disadvantages and result in a vicious cycle. As the inclusion of school density can only reveal spatial equalities, accordingly, 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 advantaged groups, such as those living in ‘Long settled Beijing society’ and ‘Employees of state-owned enterprises and white-collar workers’ travel shorter distances. Figure 3 showed 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 inequality of education in terms of travel distance has revealed that areas in which disadvantaged groups are concentrated tend to have lower school density (or insufficient educational resources) and force pupils to travel further, after controlling in multilevel models for the individual-level variations. Compared to the socially advantaged groups, the disadvantaged in these areas are more vulnerable to travelling longer travel distances. Thus, intervention measures, such as the construction of more schools to reduce the distance of travel, could improve education provision and social well-being in the educationally disadvantaged areas. The results of the models also show that geographical context has an important effect on pupils’ travel distance, and that macro-level covariates have bigger effects than most of the individual level covariates except for education stage.

Influence of individual-level characteristics

The following discussion relating to the influence of individual-level covariates in variations of pupil travel distance is based on model results that control for spatial context effects.

Hukou

The model results show the influence of the effects of hukou status, which is a unique and critical factor of China's institutional structure. 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. One reason for this result is probably due to the coarse division of the hukou in the CHTS, where the hukou status, hukou within a district, contains two types of hukou (hukou within a sub-district and hukou outside a sub-district but within a district). The school choice and travel distance for pupils with the two types of hukou are different. The travel distance of pupils whose hukou is within a sub-district may be shorter than that of non-local pupils, and the latter will be assigned to any school throughout the whole district. However, the pupils who have hukou outside a sub-district but within a district also tend to travel long distances within the range of district (Xiang *et al.*, 2018).

Additionally, although pupils who do not have Beijing hukou are normally assigned to undersubscribed schools with spare places within the whole district, most of them live in rented accommodation (which is more flexible when it comes to moving); and will probably move to residences close to assigned schools to reduce travel costs. Moreover, the accommodation which is close to unpopular schools tends 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 existing evidence to demonstrate the commute distances of in-migrant children are long, the pupils are indeed deprived of chances of access to high-quality education. In contrast, the pupils whose hukou is outside a district but within a municipality are proven to travel slightly longer distances than pupils with 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 are registered rather than their current residence, which leads them to commute longer distances and across district boundaries.

Income and housing

The coefficients for income indicate that pupils from the high-income group commute further than those from the middle-income group, while pupils from the low-income group or the group without cars 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 middle-class households are able and willing to travel longer distances to assess 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. Hamnett and Butler (2013) suggest that advantaged parents in east London not only seek out quality schools but also avoid schools with a higher proportion of pupils from disadvantaged backgrounds. Similarly, the high-income group in Beijing also do 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 school quality.

Our results suggest that pupils who live in affordable housing and also tend to be in the low-income group, are also shown to travel longer distances. This is likely to be due to spatial mismatches between economically affordable housing and public facilities (e.g., schools), which also has been pointed out by others (Yang *et al.*, 2014; Chen *et al.*, 2015). Economically affordable housing, including the Economic and Comfortable Housing and Capped-Price Housing, is subsidised housing targeted at low and medium-income households with local hukou (local resident registration). 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 access to public services (Curley, 2005).

In terms of education, accessibility to schools can strongly influence their opportunities to receive education. Compared to higher socioeconomic groups, access to education facilities could impact more on the education of low socioeconomic groups since 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 improving school supply should be integrated into affordable housing policy design. A well-designed affordable housing programme can help low-income households to compensate for their disadvantages with better access to education facilities.

The distance travelled by pupils living in self-built housing in the countryside is also longer, with very small individual-level variations or heterogeneity compared to other housing types. It reflects the universally 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 the city.

Conclusions

This paper has used detailed Chinese microdata to identify, for the first time, the relationship between pupils' access to school and their socioeconomic status by controlling for the varied context effects of the sub-districts where pupils live. It has rejected the notion of 'unpatterned-inequality' (Talen, 2001) and revealed a consistent relationship between an individual's socioeconomic status and travel distance. The results reveal that, at the individual level, the slightly longer distances are travelled by those with advantaged socioeconomic status, such as higher family income, hukou outside a district but within a municipality, and car availability, with parents seeking better education opportunities for their children.

However, this does not mean that socio-economically disadvantaged groups are more advantaged in term of access to schools. Firstly, those living in economically affordable housing or self-built housing in the countryside, who are normally socially disadvantaged, also are more likely to travel longer distances. This may be due to the relatively sparse distribution of residents and inadequate school education provision around them. Secondly, the pupils with disadvantaged socioeconomic status, like those with low income, no car and without local hukou, tend to travel shorter distances because their families have less mobility and social capital (i.e. 'exchange of information') (Sampson, 2003) to inform school choice and are consequently sent to nearby schools irrespective of educational quality and disadvantaged.

This research has confirmed the importance of context through the use school density and a geodemographic classification, on pupil travel distance, and shown that macro-level covariates have more influence in general compared to most micro-level covariates, except for education stage and housing type. The results have demonstrated socio-spatial education inequalities in school commuting. The statistically significant contextual effects of the geodemographic clusters on pupils' access to school indicate that pupils in areas of potential educational disadvantage, such as in the 'Urban-rural transition belt' or areas with 'In-migrants in deprived suburbs', also tend to commute longer distance due to lower school density. Therefore, the socio-economically disadvantaged groups living in these areas also tend to be disadvantaged in term of travel distance.

This research has avoided not only the ecological fallacy (Robinson, 1950) but also the atomistic (or individualistic) fallacy (Subramanian *et al.*, 2009) through its use of multilevel modelling. It has revealed the influence of contextual effects without ignoring the compositional variations (Pickett and Pearl, 2001). Understanding how different individual-level or geographical factors influence pupil travel distance is crucial for addressing spatial inequalities in terms of access to education resources and introducing proper interventions to promote education justice.

Finally, there are some important policy implications relating to pupils' access to school. Insufficient education provision and low school density in outer areas are due to the lack of effective school planning and rapid urban expansion. Also, the education requirements of migrant children have not been properly considered by school planners in local government. 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 areas of concentration, such as in the 'Urban-rural transition belt' and 'In-migrants in deprived suburbs' areas. Furthermore, for the low-income groups, the effects of the construction of affordable housing should not only focus on making housing economically viable for those people but 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.

References

Ball SJ, 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, 52-78.

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, 31-48.

Brighouse H (2000) *School Choice and Social Justice,* Oxford: Oxford University Press.

Chen MX, Zhang WZ and Lu DD (2015) Examining spatial pattern and location choice of affordable housing in Beijing, China: Developing a workable assessment framework. *Urban Studies,* 52, 1846-1863.

Curley AM (2005) Theories of urban poverty and implications for public housing policy. *Journal of Sociology and Social Welfare,* 32, 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, 159-172.

De La Fuente H, Rojas C, Salado MJ, Carrasco JA and Neutens T (2013) Socio-spatial inequality in education facilities in the Concepción Metropolitan Area (Chile). *Current Urban Studies,* 1, 117-129.

Farrell JP (1999) Changing conceptions of equality of education: Forty years of comparative evidence. *Comparative Education: The Dialectic of the Global and the Local*, 149-178.

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, 3-9.

Fielding A (2010) Module 8, Multilevel Modelling in Practice: Research Questions, Data Preparation and Analysis. *LEMMA VLE, University of Bristol, Centre for Multilevel Modelling. Retrieved December,* 16, 2014.

Goldstein H (2011) *Multilevel Statistical Models,* Chichester, West Sussex: Wiley.

Goldstein H and Silver R (1989) Multilevel and multivariate models in survey analysis. In: Skinner CJ, Holt D and Smith TMF (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, 1203-1228.

Gu C, Chan RCK, 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, 249-310.

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, 479-500.

Hamnett C and Butler T (2013) Distance, education and inequality. *Comparative Education,* 49, 317-330.

Israel E and Frenkel A (2017) Social justice and spatial inequality: Toward a conceptual framework. *Progress in Human Geography,* 42, 647-665.

Levin HM (1976) Educational opportunity and social inequality in Western Europe. *Social Problems,* 24, 148-172.

Logan JR, Minca E and Adar S (2012) The geography of inequality why separate means unequal in American public schools. *Sociology of Education,* 85, 287-301.

Ming HH (2014) *The Education of Migrant Children and China's Future: The Urban Left Behind,* Abingdon, Oxon: Routledge.

OECD (2016) Education at A Glance 2016: OECD Indicators. Paris: OECD Publishing.

Owen G, Harris R and Jones K (2016) Under examination: Multilevel models, geography and health research. *Progress in Human Geography,* 40, 394-412.

Pacione M (1989) Access to urban services-the case of secondary schools in Glasgow. *The Scottish Geographical Magazine,* 105, 12-18.

Pickett KE and Pearl M (2001) Multilevel analyses of neighbourhood socioeconomic context and health outcomes: A critical review. *Journal of Epidemiology & Community Health,* 55, 111-122.

Robinson WS (1950) Ecological correlations and the behavior of individuals. *American Sociological Review,* 15, 351-357.

Sampson RJ (2003) The neighborhood context of well-being. *Perspectives in biology and medicine,* 46, S53-S64.

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, 1221-1238.

Small ML and Newman K (2001) Urban poverty after the truly disadvantaged: The rediscovery of the family, the neighborhood, and culture. *Annual Review of Sociology,* 27, 23-45.

Soja E (2009) The city and spatial justice. *Justice Spatiale/Spatial Justice,* 1, 1-5.

Standing Committee of the National People's Congress (2006) Compulsory Education Law of PRC. Article 12.

Steele F (2008) Introduction to multilevel modelling-MLwiN practicals. *LEMMA VLE Module 5*, 1-48.

Subramanian SV, Jones K, Kaddour A and Krieger N (2009) Revisiting Robinson: The perils of individualistic and ecologic fallacy. *International Journal of Epidemiology,* 38, 342-360.

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, 465-486.

UNESCO (2015) Education 2030 Incheon Declaration and Framework for Action.

United Nations (2015) *Goal 4: Ensure Inclusive and Quality Education for All and Promote Lifelong Learning* [Online]. Available: http://www.un.org/sustainabledevelopment/education/ [Accessed 2015].

Waters JL (2017) Education unbound? Enlivening debates with a mobilities perspective on learning. *Progress in Human Geography,* 41, 279-298.

Williams S and Wang F (2014) Disparities in accessibility of public high schools, in metropolitan Baton Rouge, Louisiana 1990–2010. *Urban Geography,* 35, 1066-1083.

Witham M (1997) The Economics of (Not) Closing Small Rural Schools. *Paper presented at Ph.D. Symposium for Candidates and Supervisors: A Focus on Rural Issues, Townsville, Queensland, Australia, July 1997.*

Wodtke GT, Harding DJ 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, 713-736.

Wu Q, Zhang X and Waley P (2016) Jiaoyufication: When gentrification goes to school in the Chinese inner city. *Urban Studies,* 53, 3510-3526.

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, 59-70.

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, 521-540.