Is “Attending Nearby School” Near? An Analysis of Travel-to-School Distances of Primary Students in Beijing Using Smart Card Data

Cong Liao 1,2 and Teqi Dai 3,*

1 Institute of Remote Sensing and Geographical Information Systems, Peking University, No. 5 Yiheyuan Road, Haidian District, Beijing 100871, China; liaocong233@pku.edu.cn
2 Beijing Key Laboratory of Spatial Information Integration & Its Applications, Beijing 100871, China
3 Beijing Key Laboratory for Remote Sensing of Environment and Digital City, School of Geography, Faculty of Geographical Science, Beijing Normal University, No. 19, Xinjiekouwai Street, Haidian District, Beijing 100875, China
* Correspondence: daiteqi@bnu.edu.cn

Abstract: The distance between home and school is crucial for children’s mobility and education equity. Compared with choice-based enrollment systems, much less attention has been given to the commuting distance to school in proximity-based systems, as if the institutional arrangement of assigning children to nearby schools can avoid the problem of long commuting distances. Using student-type smart card data, this study explored the spatial characteristics of the commuting distance to primary schools by public transport and the residence-school spatial pattern under the proximity-based system in Beijing. The relationships between long school commutes and house price/age were investigated under the context of school gentrification. For the identified primary student users, fewer than 35% of the students travelled fewer than 3 km to school, while more than 80% of students travelled long distances greater than 5 km, which indicated that the policy of “attending nearby school” did not guarantee a shorter commuting distance to school. Long distances to school greater than 5 km correlate negatively with a lower average house price/building age and fewer students. This finding verified the assumptions from China’s school gentrification that people might buy older school-district houses but live far from the school district for a new house. These findings provide a complementary view of previous survey studies and reveal the actual commuting distance by public transport for a group of primary students in a proximity-based enrollment system.

Keywords: educational equality; house price; proximity-based system; commuting modal; children’s mobility

1. Introduction

Access to education is one important aspect for individual development and social sustainability. Generally, travel to school is a daily activity for children to access education. Analyzing the travel characteristics of students is an important prerequisite for understanding children’s mobility and further space intervention in educational inequality. The distance between home and school is a crucial factor for a wide range of topics on children’s mobility and travel characteristics. The distance between home and school is a main correlating factor in acquiring the admission qualification of primary schools, and in influencing travel mode choice and has been associated with equal educational opportunity [1–3], academic achievement [4,5], active school transport and health benefits [6–8]. Moreover, it affects a city’s spatial and functional structure and contributes to urban problems such as transport congestion [9]. Therefore, the distance between home and school becomes a key element for policy making in educational equity, children mobility and urban planning.

A common trend of increasing distance has been observed in many Western developed countries, such as the USA, England, Canada, Australia, Sweden and Ireland [10–14]. This
trend has often been placed in the context of liberal educational reforms pertaining to school choice in many Western countries since the late of 1980s. Most studies of travel-to-school distances assume that the distance between home and school increases as the price of more school choices towards choice-based systems [13,15]. The social status of families and other factors of travel-to-school distance have been investigated and have rich implications for school choice, student health or educational equity [16–19].

In contrast, the distance from home to school under the proximity-based enrollment system has been under studied, although equity issues under the nearby enrollment policy have received an increasing amount of attention in recent years. Existing studies have explored the complex issues of the nearby enrollment policy on educational equity, such as the unbalanced distribution of education resources or social stratification from parents’ bidding on residence properties for high-quality schools, focusing on education gentrification [20–22], school district distortion [23], or education-led spatial disparities [24–26]. Few studies have discussed school distance under a proximity-based enrollment system. One reason may be that the institutional arrangement of assigning children to nearby schools is automatically regarded as strictly implemented and could effectively avoid the problem of long commuting distance. However, this situation may not hold in reality [27,28]. The study found that the ratio of long school commuting distances under the proximity-based enrollment system in China was rather high, approximately 13% to 27% across Beijing for distances greater than 5 km [29]. This finding was echoed by Yang et al. and Xiang et al. [30,31]. Zhang et al. revealed that long school commuting distances greater than 6 km in Beijing were mostly travelled by public transport [32]. Compared with the situation of travelling by car or school bus, this group is worthy of further research and would well fit the advantage of spatial analysis with geospatial big data of public transport.

With the rapid development of information and communications technology (ICT), the individual-based geospatial big data collected by various location-aware devices, such as mobile phones, subway smart cards and social media check-ins, have been employed to offer observations of human mobility, although there is not a clear and widely accepted definition [33,34]. Big data provide good opportunities to understand the characteristics of commuting distances between home and school. Compared with the low frequencies and vast workload of survey data, big data can broaden the knowledge on commuting and gain insights that are previously unattainable through traditional datasets [35,36]. There is a rapid growth of literature in conducting empirical studies on human mobility patterns and their dynamics characteristics by using geo-tagged big data [37,38]. Most papers have focused on the mismatch of work commuting patterns [39–41]. Certain papers have discussed specific groups of people, such as extreme travelers and elderly people [42–44]. Although there was a “spatial turn” in education [45], studies of spatial analyses in education research are still limited [46]. Moreover, few studies have applied geospatial big data to analyze the characteristics of the school commuting behaviors of basic education student groups. To the best of our knowledge, in social sensing research by geo-tagged big data, only one paper addressed university students [47] and only one addressed children’s school commuting behaviors [48]. Both papers did not relate their studies to educational policy or equity.

The group of students with long commuting distance to school means the law of “attending nearby school” may not be complied with. We prefer to focus this research with a different assumption than previous research, that is, the nearby enrollment policy is not a panacea to ensure nearby enrollment, and at least two groups should be distinguished. The first group comprises those who comply with the policy, while the second group includes those who do not comply with the policy. The latter group might use very different spatial strategies to obtain educational resources, and existing studies have revealed that the proportion of this group is not small, at least in Beijing [29–31]. However, this group was largely disregarded in previous studies. Both the approaches of survey and big data can be utilized in this study. Considering that an interviewee may provide a school-district address instead of their actual residential address, this study uses smart card data (SCD).
Certain limitations, such as a lack of personal information or possible errors from the algorithm, may arise. However, big data may reflect the overall situation, and public transportation reflects the majority of students with long commuting distances [32].

This study aims to use student-type smart card data to explore the spatial characteristics of long school commuting by public transport in Beijing and to explore the relationships between school commuting and residence in the context of school gentrification.

The remainder of this article is organized as follows. We first describe our methodology, followed by the results analysis. The discussion section presents the implications for policy making, and the final section presents the conclusions.

2. Methodology
2.1. Study Area

As the capital of China, Beijing has a special political symbolism. We expect that Beijing could be representative of Chinese cities in educational institutional arrangements [49]. More than 86.5% of Beijing’s population was classified as urban in 2015, most of which lived in six urbanized areas of Beijing. As shown in Figure 1, the six urbanized districts include Dongcheng District, Xicheng District, Chaoyang District, Fengtai District, Shijingshan District and Haidian District. Dongcheng District and Xicheng District are regarded as core districts of Beijing. The study spans an area of 1369 square kilometers, inhabited by 12.8 million people, which is subdivided into 103 subdistricts and 5429 residential quarters as of 2015. There were 550 primary schools within six urbanized areas of Beijing in 2015, with a total of 521,277 students. Among them, the core districts have 144 primary schools and 115,939 students.

Beijing has a convenient public transport system, and the transit price is quite low due to large subsidies by the government. Approximately 780 bus lines pass through the case study area and more than 3000 bus stations are within the study area. In terms of subway system, there are 17 subway lines and 256 subway stations within the study area.

![Figure 1. Overview of the study area.](image)

2.2. Data Collection and Preprocessing

This work was supported by data of bus smart card records and subway card records of student type during workdays of one week from 13–17 June 2016, obtained from Beijing Public Transport Group. The dataset contains 6,080,225 and 879,089 records of bus and subway cards with card swiping times of pick-ups and drop-offs, respectively. Data
preprocessing was conducted to remove irrelevant attributes or invalid records. After data preprocessing, the sample of each record contained 6 attributes, as listed in Table 1. The student-type card is used in this research, indicated by number 18 or 19. Student cards can only be used by students at primary schools, secondary schools and universities in Beijing. Student cards can provide a special transit price discount of 2.5%.

Table 1. Typical sample record of SCD.

<table>
<thead>
<tr>
<th>Card ID</th>
<th>Swiping Time</th>
<th>Line Num.</th>
<th>Bus Stop to Get On</th>
<th>Bus Stop to Get Off</th>
<th>Card Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>19151591</td>
<td>20160614071535</td>
<td>331</td>
<td>12</td>
<td>18</td>
<td>18</td>
</tr>
</tbody>
</table>

The housing price/age data are extracted from the LIANJIA platform (https://bj.lianjia.com/, accessed on 30 January 2019), a widely employed housing trading platform in China. The housing price data include approximately 89,742 second-hand house transaction records of the study area in 2018.

Based on the preprocessed data, several steps were completed to investigate the residence-school spatial pattern, as shown in Figure 2.

Figure 2. Workflow of the study.

2.3. Identifying Primary Student Travelers

The first step of this study is to identify travelers who are primary students using public transport. The records of student-type SCD are extracted. In Beijing, primary school students are just one category of students who are eligible to use student cards. Other types of school students, such as college students and high school students, can also own student cards. Therefore, distinguishing different travel characteristics between primary school students and other types of students is critical.

In existing studies, temporal information, spatial information and frequency in travel behavior are commonly employed to describe the travel regularity and identify of the specific groups [50]. In this research, we use this kind of information to identify primary students, focusing on the following special features of Beijing primary schools. In Beijing, primary schools start at 8 o’clock in the morning and end at 3 o’clock in the afternoon. The end time of primary school is different from that of other types of schools. According to this special regular pattern of Beijing primary schools, we set the following four rules to distinguish primary students from other types of students:

(1) The card type is a student card, which excludes the travel records of nonstudent cardholders;
(2) The swiping time is between 5:00 and 8:00 AM or 15:00–18:00 PM, which excludes travel behaviors of students who do not commute to school in the morning or afternoon bell times in Beijing;
(3) Cards are swiped two times a day and three days a week, which considers the periodicity of primary students’ commuting to reduce the interference of other non-primary students;
(4) There is a primary school near the alighting station, which expresses that students’ travel destinations are to primary schools.

Though students in Beijing go to school five days a week, we set the threshold as three, considering too high a threshold may miss some students who might use private cars or
taxis in certain conditions. This threshold identifies regular school commuting and sets these student-card users as primary students.

Based on the above rules, there are 35,847 primary student cardholders on weekday commutes to primary schools by public transport. The ratio of primary students who commute to schools by public transport to all primary students is 6.88%. The ratio is close to the public transport travel data (approximately 9.4%) in the Pupil Travel Survey by the Beijing Municipal Commission of Transport in 2014.

After a card record is identified as the primary student type, the residence and corresponding school is allocated. The most frequently used daily boarding stations are identified as proxies of residence by examining the maximum frequency of all boarding stations during the week. As distance is one of the most important factors in inferring trip purposes [51,52], the distances from stations to schools are utilized to identify the residence and the school. In this research, following previous studies [47,53], the boarding station is set as the students’ residence address and the nearest school to alighting stations as students’ destination.

2.4. Defining Long-Distance Schooling Commuters

There are no distinct and consistent specifications for the distance threshold of “nearby”. Although the policy of “attending nearby school” is obliged by the Compulsory Education Law of China, there is no official definition for “nearby”. We address four key distance thresholds in Figure 2. The first two distance thresholds pertain to the standard of “near”. The first threshold is the average shortest Euclidean distance from all residential locations to schools. The second threshold of 2 km or 3 km is obtained from existing papers [14,15,54]. The other two distance thresholds pertain to the standard of “far”. The third distance threshold was set to 5 km, which is regarded as too far to commute for primary school pupils [55,56]. In this study, a distance of 5 km was selected as the threshold distance for long distance. To reveal extreme commuting, a distance of 10 km was also investigated.

2.5. Relating Long-Distance Commuting to House Price/Age

This study examines how the distribution of long commuting distances to school may correlate with the spatial variation in average housing price and building age based on the context of school gentrification in Beijing. According to the catchment zone policy for school enrollment, many well-off middle-class parents would seek to purchase house property in the catchment zone of a good school to obtain a school place for their child [20]. The competition for these houses within the school catchment areas of leading statement schools has led to a serious inflated price for school district houses [57]. Due to the short-term occupancy (generally nine years of elementary and junior middle schooling), these middle-class parents generally have no incentive to refurbish those old buildings before they sell them to the new gentrifying families [21,22]. Wu et al. referred to this variant form of gentrification as “Jiaoyufication” in the inner city [58], i.e., a middle-class makeover of inner-city districts connected with schools. Inflated prices and old buildings are two typical characteristics of “Jiaoyufication”. The place dependence on schools drives gentrifying families to move home in Beijing [30]. Consistent with the above discussion, we accordingly hypothesize that a long commuting distance to school leads to a lower average house price and newer average building age.

The spatial interpolation method of the inverse distance weighted (IDW) technique is selected to calculate the house price data at locations without transaction records [59]. The search radius used in the interpolation method is set to 1 km. The method transforms the housing locations into a raster system, so that the relationship between housing price/age and distribution of students can be assessed at the same scale. Figure 3 illustrates the spatial data of housing price/age of the study area. The grid values are extracted to facilitate the correlation analysis.
Figure 3. Spatial data of housing price/age; (a) housing price and (b) housing age.

3. Results

Figure 4 illustrates the distribution and cumulative proportions of primary student commuters with public transport by travel distances with a step of 1 km. The curve of distribution proportions showed an upwards trend first and then a downwards trend. The dividing point between the proportion of rising and falling is 5 km. Before 5 km, the number of students who use public transport increases with an increase in the commuting distance. Between 5 and 10 km, a high proportion of students is generally maintained, approximately 6%. The total number of students between 5 km and 10 km accounts for nearly 30% of students using public transport. At distances greater than 10 km, the number of students decreases with an increase in the commuting distance. The average shortest Euclidean distance from all residential locations to schools is 0.37 km. Approximately 96% of students who travel by public transport traveled over this threshold distance. The proportion of students with commuting distances within 2 km or 3 km is about 4.69% and 8.91%, respectively. This finding suggests the disadvantage of public transport in the modal competition for short commuting distance. The proportion of students with commuting distances greater than 5 km is approximately 81.19%. For extreme commuters, the corresponding ratio of students with distances greater than 10 km is approximately 51.89%. Overall, according to the various definitions of “near” or “far”, the policy of “attending nearby school” was poorly implemented in the study area.

Figure 5 illustrates the kernel density distribution of students with commuting distances greater than 5 km by their boarding and alighting points. Using the Jenks method in ArcGIS, the density values are divided into five grades. This method can minimize the differences within each type and maximize the differences among each type. For the boarding points, there are several high-density regions around the core areas of Beijing. A relatively high-density band is formed between the second ring road and the fourth ring road. Due to the small number of students near to the peripheral areas, the overall density in these areas is also small. For the alighting points, the high-density areas are more concentrated. Compared with the density by boarding points, the density value of alighting points at the high level is greater. The core areas are important destinations for students with commuting distances greater than 5 km. In addition to the core areas, the east and northwest of the core areas are two other important high-density areas. On the whole, compared with the students’ residences, the area covered by students’ destination is more compact.
Figure 4. Distribution of commuters by distance.

Figure 5. Kernel density of students’ boarding (a) and alighting points (b) with commuting distance greater than 5 km.

Figure 6 illustrates the relationship between the average house price and commuting distance to school that is greater than 5 km. All the data were segmented and aggregated at a step of 1 km. The abscissa represents the average distance to school. The ordinate represents the average house price within this distance interval. The size of the ball indicates the number of students in this range. There was an overall decreasing trend of house prices as the distance increased to values greater than 5 km. Average house prices decrease with average distance. The goodness of fit of the linear function is 0.88. This finding implies a similar tradeoff between school commuting and average housing price, as shown in the classic Alonso model on working, commuting and housing.
The results indicate that most students who travel by public transport do not attend their nearby schools. Only 34.89% of students attend a school located fewer than 3 km from home, while 55.76% of them attend a school more than 5 km from home. For students with commuting distances greater than 5 km, several high-density regions of boarding points around the core areas of Beijing appear. Compared with the density of boarding points, the density value of alighting points is more concentrated in the core areas. The commuting distance is significantly affected by housing prices and housing age. The analysis of the
relationship between average house price and average commuting distance that is greater than 5 km and the number of students in these disadvantaged groups illustrates that higher house prices correlate negatively with long school distance. Existing Western literature also reveals that a rather high ratio of parents would send their children outside the nearby school district [27,28]. For this study, we restate another issue regarding different spatial strategies to obtain educational resources and the ambiguous and distorted district zoning [23].

Although the policy of “attending nearby school” has legal status in China, the results of the case study by public transport indicate that the policy was not well implemented. Under this proximity-based system, the efficiency of attending school is a priority. The eligibility to attend school is determined by school district, which is generally delineated according to the distance rule. A recent study by Xiang et al. explored the influencing factors on school distance under a proximity-based enrollment system, but the group of students with long school commuting distances were not analyzed separately [31]. The actual commuting distance to school is more complicated. Besides the unclear division rules of school district boundaries [23], this study revealed that inconsistencies exist between the actual residential address and the eligible registration address for enrollment. Many students may not actually live in these registered home addresses, as some school district housings are usually old, small and not suitable for living [22]. This study illustrated that actual commuting distance to school can be better captured by individuals’ trajectory information of geo-tagged big data.

A long commuting distance has equality applications. The existing research reveals that geography has an important role in education [2,16]. The travel distance to schools, especially good schools, is related to students’ equal access to opportunities [1]. Affordable housing prices and building age vary by space, and these two factors have a significant influence on a long commuting distance to school. In many cases, the ability of students to attend a good school is closely related to the social and cultural capital of their families [26,30]. A negative effect is discussed in the studies related to “capitalization of school district” and “Jiaoyufication” [57,58]. Institutional factors, such as registration status, complicate this issue [23]. To promote the equal allocation of educational opportunities in a proximity-based system, improvements in the spatial assignment mechanism towards equality are needed [60].

5. Conclusions

By using student-type SCD and other geo-tagged big data on housing prices and primary schools, this research investigates the residence-school spatial pattern of students in Beijing, focusing on the effectiveness of the proximity-based allocation system on commuting distance to school from an equity perspective. This research highlighted not only the importance of using SCD for transport analysis, but also social equity and educational policy making. The study contributes to the present literature on the geography of education and geo-tagged big data.

It was found that about 81.19% of students who commuted to school by public transport traveled a long distance of more than 5 km. Above the threshold distance of 5 km, the number of primary students decreased with the distance to school. In addition, there was a highly negative correlation between the house price/building age and distance to school. The analysis can offer information on spatial distribution of long-school-commuting students and may raise concerns on the effectiveness of the policy of “attending nearby school”, as well as children’s health, traffic congestion and other sustainability issues. In light of the close association to lower house prices far from school and old buildings near school, this study provides evidence for “Jiaoyufication” and has implications on complex location decisions for school commuting.

The analysis showed a comprehensive picture of the distance to school of all primary students by public transport, therefore providing a sound basis for further survey research. Future work could be extended in the following directions: First, this study only analyses
the travel-to-school distance by public transport. An integrated analysis that combines other modes, such as walking and private cars will be valuable for revealing the system characteristics of students’ commuting distance in the future. Second, students’ attribute information, such as gender and household structure, can be further enriched by home interview surveys in addition to housing price/age. More attribute information better describes students’ choices of attending school. Third, due to the lack of data to evaluate the quality of schools, this study cannot assess who is attending good schools. Fourth, dynamic analysis of the residence-school spatial pattern is necessary, and the development of computational tools to conduct periodical analysis and prognosis is interesting [61,62].

Author Contributions: C.L.: conceptualization, methodology, writing—original draft; T.D.: writing—review and editing, supervision, project administration and funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References
Sustainability 2022, 14, 4344


30. Yang, Z.; Hao, P.; Wu, D. Children’s education or parents’ employment: How do people choose their place of residence in Beijing. Cities 2019, 93, 197–205. [CrossRef]


38. Egu, O.; Bonnel, P. Investigating day-to-day variability of transit usage on a multimonth scale with smart card data. A case study in Lyon. Travel Behav. Soc. 2020, 19, 112–123. [CrossRef]


41. Manley, E.; Zhong, C.; Batty, M. Spatiotemporal variation in travel regularity through transit user profiling. Transportation 2018, 45, 703–732. [CrossRef]


48. Mei, D.; Xiu, C.; Feng, X.; Wei, Y. Study of the school—Residence spatial relationship and the characteristics of travel to school distance in Shenyang. Sustainability 2019, 11, 4432. [CrossRef]

49. Song, Y.; Ding, C.; Knaap, G. Envisioning Beijing 2020 through sketches of urban scenarios. Habitat Int. 2006, 30, 1018–1034. [CrossRef]


60. Liao, C.; Scheuer, B.; Dai, T.; Tian, Y. Optimizing the spatial assignment of schools to reduce both inequality of educational opportunity and potential opposition rate through introducing random mechanism into proximity-based system. *Socio-Econ. Plan. Sci.* **2020**, *72*, 100893. [CrossRef]
