Bus Commuters’ Jobs-housing balance in Beijing:
An exploration using large-scale synthesized smart card data

By
Jiangping Zhou*
Assistant Professors
Department of Community and Regional Planning
Iowa State University
Ames, IA, 50011
zjp@iastate.edu
Phone: 515-294-5470

and

Ying Long
Beijing Urban Planning and Design Institute
Beijing 100045, China

Word count: 5,749+1,750 (3 figures + 4 tables) =7,499

July 10, 2013

*Corresponding author.
ABSTRACT

Existing jobs-housing studies have rarely (a) used smart card data provided by the public transportation agency and (b) focused on commuters of bus mode. In this study, massive smart card data were used to estimate 216,844 bus commuters’ workplace and residence locations in Beijing. This enables a jobs-housing study of bus commuters in the metropolis with a much larger sample size than most existing studies. This study indicates that Beijing’s bus commuters (a) have a shorter actual required commuting (ARC) and minimum required commuting (MRC) than commuters in four other auto-dependent western cities with comparable population and/or land use sizes; (b) have a longer ARC and MRC than commuters of all modes in Guangzhou, another metropolis in South China that is half of the size of Beijing. Using local expert consultations, field surveys and information provided by on-line housing search engines to supplement smart card data, this study has established five land-use prototypes of jobs-housing imbalance and has proposed countermeasures to address the imbalance.
INTRODUCTION

Car dependence, traffic congestion, long commute and associated air pollution and Greenhouse Gas (GHG) emissions have become notable phenomena that characterize many populous cities. Planners, policy analysts and public agencies have advocated or even incentivized jobs-housing balance as a way to reduce peak-period travel and optimize commutes, in particular, commutes by driving alone (1-6). Not surprisingly, different factors contribute to jobs-housing balance or imbalance among different social groups in different locales. In the United States (US), for instance, suburbanization of jobs, housing segregation, inefficient public transportation services, race and “automobile mismatch” may all more or less contribute to jobs-housing imbalance (or spatial mismatch) for the low-income minority or new immigrants (7-10). In China, the disappearance of working unit (“Danwei”) compounds and the introduction of the commercial housing market led to increasing jobs-housing imbalance of workers in large cities such as Beijing and Guangzhou (11-15).

There has been a large body of literature on jobs-housing balance and related topics such as excessive commute and commute efficiency. Horner (16) and Ma and Banister (17), for instance, have provided a good review of such literature in the western context, respectively. Many existing studies of jobs-housing balance have dealt with total commuting flow and have treated workers and job/workplace as homogeneous (18). This is especially true when we look at two seminal manuscripts (19, 20) and articles that attempt to extend them in different contexts (4, 21-23). To increase policy relevance of their research, an increased number of authors have paid attention to worker, workplace and/or employer heterogeneity. Crane (24), Gordon et al. (25) and Kim (26), for instance, all show that employees can be owners or renters and this status affects their commuting distance, that is, jobs-housing balance. Though not directly dealing with jobs-housing, Kwan (27) indicate that individual and household activity schedules and time-budgets are critically important too in terms of influencing one’s activity sphere. Giuliano and Small (28)
study jobs-housing balance among workers of different occupations, finding that service workers have the lowest average commutes.

However, few authors have looked at jobs-housing balance of commuters by mode, that is, mode choice heterogeneity. This may be because of three reasons. **First**, conducting and processing surveys to get reliable information about individual-level jobs-housing balance and mode choice has not been cheap and has become increasingly expensive. In the US, for instance, the cost to get the information was $195/household in 1995 and was $411 in 2001¹. **Second**, driving is the dominant commute mode in most developed countries and thus comparatively more attention has been given to jobs-housing balance of driving commuters in these countries, where most relevant studies have been conducted. **Third**, despite that we know commuters of different mode choices could face different degree of jobs-housing imbalance (e.g., see (7)), little has been done on causes, implications and cures of the differences. In developing countries such as China, nevertheless, the study of jobs-housing balance of bus commuters is still quite relevant. There are a significant percentage of residents or commuters by bus and there are millions of them, even in leading cities there. In Beijing, for instance, over five millions or 28% of the residents still commuted by bus as of 2010².

Given the above, there is a need to separately study the jobs-housing balance of bus commuters in developing countries. Using smart card data processed by Beijing Institute of City Planning, this manuscript/study investigates the character of jobs-housing balance of bus commuters in Beijing. Adopting the concepts of excessive commute, minimum/maximum required commutes and commute efficiency, this study estimates the amount of excessive commute of these commuters and how efficient

---


² Based on 2011 Annual Report of Beijing’s Transportation Development (In Chinese, internally circulated report), Beijing Transportation Research Center.
their commute is. Also, this study compares its findings with those in existing studies whenever possible in order to expand our knowledge about jobs-housing balance and relevant topics. This study differs from existing ones and possibly contributes to the research on jobs-housing in these dimensions: (a) It specifically deals with bus commuters, a subgroup of the general commuter that has rarely been studied so far; (b) It uses commuter data that are of much greater sample size than those in most existing studies--there are 216,884 samples in this study; (c) It explores a new way to synthesize widely available smart card data in the public transportation sector and other supplementary data and information publicly available to establish land-use prototypes of jobs-housing imbalance so as to propose better countermeasures.

This manuscript is organized as follows. Next section (Section 2) is a survey of relevant literature. Section 3 provides an overview of the case city and describes how the 216,884 samples were extracted from the smart card data. Section 4 empirically estimates excessive commute, minimum/maximum required commutes and commute efficiency of the samples. Section 5 discusses the findings based on the empirical evidence and proposes prototypes of areas suffering severe jobs-housing imbalance in Beijing. Section 6 concludes and discusses directions of further studies.

RELEVANT LITERATURE

Jobs-housing balance describes the relationship of the total numbers of jobs and housing units within a given geographical area or within a given travel distance or travel time (3, 4, 29, 30). Under the strong assumptions that all jobs, workers and housing units are homogeneous and workers can switch jobs or housing at little or no cost, authors show that there exists a Minimum Required Commute (MRC) for all
workers (19, 20). But the Actual Required Commute (ARC) for all workers is always bigger than MRC. The differences between MRC and ARC are called wasteful commute (19) or excessive commute (17, 18, 23, 31-41).

Based on the concepts of ARC and MRC, Horner (23) proposes a way to evaluate the commute efficiency of different cities. He points out that in addition to MRC, there exists the maximum required commute (MaxRC). Employing the following two ratios, one can best compare commute efficiency or jobs/housing balance across cities:

\[
E = \left(\frac{ARC - MRC}{ARC}\right) \times 100 \quad (1) 
\]

\[
C_u = \left(\frac{ARC - MRC}{MaxRC - ARC}\right) \times 100 \quad (2), 
\]

where \(E\) shows how much excess commute there is or how good jobs/housing balance is in a city, \(C_u\) reflects how much “commute potential” of a city has been used. The smaller \(C_u\) is, the more efficient the commute or the better jobs/housing balance is.

Using the 1990 US census data, Horner (23) compares the commute efficiency of 26 US cities. Like most existing studies highlighted in (17), Horner’s comparisons do not differentiate commuters by mode choice. Thus, one cannot conclude from his studies that whether bus commuters’ jobs/housing balance is better or worse than other commuters. But when mode choice of commuters is considered, there tends to be more relevant insights. Levinson (42), for instance, finds that accessibility to jobs and housing has a negative relationship with distance, and that transit commuters appear to have higher accessibility than automobile users. In another study, Cervero (43) finds that low density, single use and non-integrated features of many US suburban employment centers (SECs) contribute to auto dependence. The emergence of SECs
with high densities, rich mixtures of land use and nearby affordable housing (that is, better jobs-housing balance) would increase public transit efficiency and usage as well as mitigate congestion over the long run. Cervero and Duncan (44) and Kockelman (45) use micro-data from travel diaries, regressing individual or household vehicle miles traveled (VMT) on land use measures (including jobs-housing balance metrics) plus variables that controlled for household income and survey respondent gender, age, and ethnicity among other demographic characteristics. Their studies are more sophisticated than studies that used data aggregated to census tracts or other geographic areas (e.g. 23, 30, 46). They show how jobs-housing imbalance could partially contribute to auto dependence and higher VMT. Largely built on the above work and many others, Horner and Mefford (7) develop a conceptual approach synthesizing research on spatial mismatch and jobs–housing balance, including disaggregation studies of the phenomena by mode of transportation, within the broader excess commute framework. Using residential and workplace location data from Atlanta, they show that (a) minority’s home-work alternatives (after controlling mode of travel) are more spatially constrained; (b) race and mode choice should be simultaneously considered when studying jobs-housing balance. In a recent study, Murphy (38) specifically studies the extent of mode-based excess commuting in Dublin, Ireland. He finds that excess commute is considerably greater for drivers. In other words, there is greater commute inefficiency and worse jobs-housing balance associated with driving.

In addition to the above studies related to jobs-housing balance in the western context, there are also parallel studies in the Chinese context. Based on travel survey data of a small sample of workers (n ≤ 750) in Beijing, Wang and Chai (11) and Zhao et al. (14) find that:

(a) The traditional employer-provided housing system in China contributed to a better jobs-housing balance and a shorter commute;
(b) The marketization of housing supply in China decreased jobs-housing balance and lengthened workers’ commute.

Largely due to urban expansion and suburbanization of affordable housing, all commuters and commuters from affordable housing projects also see an increase in their commuting distance in selected cities such as Beijing (15, 47-51). But again, few studies about jobs-housing balance in the Chinese context have focused on bus commuters. One of the few such studies is (47), which focuses on jobs-housing balance in two affordable housing projects in Beijing and found that bus commuters’ commuting time was greater than that of auto commuters.

THE CASE CITY AND THE DATA

As the capital of China, Beijing now has over 20 million residents and is becoming one of the most populous cities in the world. The land size of Beijing metropolitan area is 16,410 km², making it one of the biggest in China. Beijing Public Transportation Company, a state-owned company provides public bus services in the metropolitan area. As of 2011, the company has 28,343 buses on 948 bus routes with a total length of 187,500 km. In 2011 alone, these buses produced vehicle kilometers traveled of 1.7 billion and transported total passenger of 4.9 billion. As of 2005, over 90% of bus riders in Beijing swiped an anonymous smart card when boarding and alighting (for suburb routes) or when boarding (for inner-city route) to pay for their fare. The swipes automatically generate the following information of cardholders:

(a) Bus trip origin and/or destination information;

(b) Boarding and/or alighting time;

---

(c) Unique card number and card type (student card at a discount vs. regular card);

This information (a) to (c) is instantly sent to and is stored at a central server. For this study, we requested a whole week’s data from the server administrator. The data eventually provided to us contain 77,976,010 bus trips of 8,549,072 non-distinct cardholder records between April 7 and April 13, 2008.

To identify a cardholder’s workplace, we queried one-day data on a MS SQL Server and repeated the work for seven days based on these rules:

(a) The card type is not a student card;
(b) \( D_j \geq 6 \) hours, where \( D_k \) is the duration that a cardholder stays at place \( j \), which is associated with all bus stops within 500 meters of one another;
(c) \( j \neq 1 \), which means that \( j \) is not the first place in a weekday that the server records.

The most frequent identified workplace of a cardholder in seven days will be defined as the final workplace of the cardholder in this study.

Similarly, we deduced from the data queries that a place would be a cardholder’s home if the data meet these conditions:

(a) The cardholder has an identified workplace;
(b) The card type is not a student card;
(c) \( D_h \geq 6 \) hours, where \( D_k \) is the duration that a cardholder stays at place \( h \), which is associated with all bus stops within 500 meters of one another;
(d) \( F_h \geq F_j \) where \( F_h \) is the first most frequent place a cardholder starts a bus trip of a day within the week, \( F_j \) is the trip frequency to or from \( j \) that the cardholder has.
Based on the above rules or conditions, we successfully found that there were 216,844 distinct cardholders/workers commuting by bus in Beijing⁴. We then geocoded and aggregated cardholders’ home and workplace by traffic analysis zone (TAZ). In Beijing, following conventional rules about defining TAZ boundaries, local planning agency divides the metropolitan area into 1,118 TAZs. Figure 1 is a map of these TAZs and Tian’ an Men Square, the agreed-upon center of Beijing.

In the end, for the 216,844 cardholders/workers, they reside in 729 distinct TAZs and work at 752 distinct TAZs. Based on this information, we constructed the matrices of the commuting trips ($M_t$) as well as journey-to-work distances ($M_d$) in support of this study. To be consistent with the distances used in existing studies so that we can make apple-to-apple comparisons later on, the distances in $M_d$ were linear distances between centroids of TAZs for trips between two different TAZs. For trips within the same

---

⁴ Technical details about how we obtained the data can be found in Long Y, Thill J-C, 2012, “Combining smart card data, household travel survey and land use pattern for identifying housing-jobs relationships in Beijing” Working Paper, Beijing Institute of City Planning.
TAZ, the distance equals $R_i$, where $R_i = \sqrt{A_i/\pi}$, $A_i$ is the area of TAZ$_i$. A portion of $M_t$ and $M_d$ is visualized in Table 1.

**TABLE 1: A Portion of $M_t$ and $M_d$**

(a) $M_t$

<table>
<thead>
<tr>
<th>TAZ$_h$</th>
<th>TAZ$_j$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>120</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>80</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>...</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>...</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>...</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>729</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

(b) $M_d$

<table>
<thead>
<tr>
<th>TAZ$_h$</th>
<th>TAZ$_j$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.47</td>
<td>1.53</td>
<td>1.87</td>
<td>2.48</td>
<td>2.92</td>
<td>...</td>
</tr>
<tr>
<td>2</td>
<td>1.53</td>
<td>0.79</td>
<td>0.95</td>
<td>1.85</td>
<td>1.92</td>
<td>...</td>
</tr>
<tr>
<td>3</td>
<td>1.87</td>
<td>0.95</td>
<td>0.43</td>
<td>0.22</td>
<td>1.08</td>
<td>...</td>
</tr>
<tr>
<td>4</td>
<td>2.92</td>
<td>1.92</td>
<td>1.08</td>
<td>0.66</td>
<td>1.02</td>
<td>...</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>729</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

**EMPIRICAL ANALYSIS**

**ARC of the commuters**

We used the following formula to get the ARC of the 216,844 bus commuters in Beijing:

$$ARC = \sum_{h} \sum_{j} \frac{t_{hj} \cdot d_{hj}}{T},$$

where

- $t_{hj}$ shows the number of commuters from TAZ$_h$ (home) to TAZ$_j$ (workplace);
- $d_{hj}$ is the distance between TAZ$_h$ and TAZ$_j$;
- $T$ is the total number of commuters.
MRC and MaxRC of the commuters

For the MRC and MaxRC calculations, we employed the standard linear programming method. The calculations equal to solving the optimization problems as follows,

\[
\min \frac{1}{T} \left( \sum_{h} \sum_{j} t_{hj} \times d_{hj} \right) \text{ or } \max \frac{1}{T} \left( \sum_{h} \sum_{j} t_{hj} \times d_{hj} \right)
\]

Subject to

\[
\sum_{j} t_{hj} = O_{h},
\]

\[
\sum_{h} t_{hj} = D_{j}, \text{ and}
\]

\[
t_{hj} \geq 0, \text{ where}
\]

\[O_{h}\] is the total number of commuters living in TAZ\(_{h}\),

\[O_{j}\] is the total number of commuters working at TAZ\(_{j}\).

E and C\(_{u}\)

After getting ARC, MRC and MaxRC, E and C\(_{u}\) were calculated using equations (1) and (2).

Results

Table 2 below presents the ARC, MRC and MaxRC, E and C\(_{u}\) of Beijing bus commuters as well as relevant results we identified from selected existing studies.


<table>
<thead>
<tr>
<th>Study/Survey</th>
<th>Mode, year</th>
<th>Sample size (City, if not Beijing)</th>
<th>ARC (km)</th>
<th>ARC (min)</th>
<th>MRC (km)</th>
<th>MaxRC (km)</th>
<th>E</th>
<th>C_u</th>
</tr>
</thead>
<tbody>
<tr>
<td>This study</td>
<td>Bus, 2008</td>
<td>216,844</td>
<td>8.1</td>
<td>36.0****</td>
<td>3.0</td>
<td>37.3</td>
<td>64</td>
<td>17</td>
</tr>
<tr>
<td>The 2005 Survey*</td>
<td>Bus, 2005</td>
<td>6,651</td>
<td>8.4</td>
<td>40.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>The 2010 Survey*</td>
<td>Bus, 2010</td>
<td>9,778</td>
<td>-</td>
<td>60.3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Liu and Wang (52)</td>
<td>Bus, 2007</td>
<td>307</td>
<td>-</td>
<td>46.3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Wang and Chai (11)</td>
<td>Bus, 2001</td>
<td>227</td>
<td>-</td>
<td>55.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Zhao et al. (14)</td>
<td>Bus and rail rapid system, 2001</td>
<td>220</td>
<td>-</td>
<td>52.4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Liu et al. (15)</td>
<td>All, 2001</td>
<td>1,500(Guangzhou)</td>
<td>4.5</td>
<td>-</td>
<td>1.9</td>
<td>-</td>
<td>58</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>All, 2005</td>
<td>1,500(Guangzhou)</td>
<td>5.0</td>
<td>-</td>
<td>2.7</td>
<td>-</td>
<td>44</td>
<td>-</td>
</tr>
<tr>
<td>Horner (23)</td>
<td>All, 1990</td>
<td>n/a (26 US cities)</td>
<td>6.7~16.8</td>
<td>-</td>
<td>3.0</td>
<td>10.1~14.4</td>
<td>48~67</td>
<td>17</td>
</tr>
<tr>
<td>Horner (7)</td>
<td>Non-Hispanic white, 1990 - driving alone</td>
<td>n/a (Atlanta)</td>
<td>-</td>
<td>-</td>
<td>9.8</td>
<td>43.6</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Non-Hispanic white, 1990-bus</td>
<td>-</td>
<td>-</td>
<td>5.9</td>
<td>27.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Non-Hispanic black, 1990-driving alone</td>
<td>-</td>
<td>-</td>
<td>7.7</td>
<td>30.2</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Non-Hispanic black, 1990-bus</td>
<td>-</td>
<td>-</td>
<td>6.0</td>
<td>18.8</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Non-Hispanic, other, 1990-driving alone</td>
<td>-</td>
<td>-</td>
<td>6.2</td>
<td>35.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Non-Hispanic, other, 1990-bus</td>
<td>-</td>
<td>-</td>
<td>5.8</td>
<td>24.7</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Hispanic white, 1990-driving alone</td>
<td>-</td>
<td>-</td>
<td>7.2</td>
<td>35.3</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Hispanic white</td>
<td>-</td>
<td>-</td>
<td>6.2</td>
<td>21.7</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cropper et al. (53)</td>
<td>All, n/a (Baltimore)</td>
<td>15.4</td>
<td>-</td>
<td>6.5</td>
<td>-</td>
<td>58</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Kim (34)</td>
<td>All, 1991</td>
<td>782** (Los Angeles)</td>
<td>24.6</td>
<td>-</td>
<td>16.5</td>
<td>-</td>
<td>33</td>
<td>-</td>
</tr>
<tr>
<td>Frost et al. (22)</td>
<td>All, 1991</td>
<td>n/a (London)</td>
<td>13.3***</td>
<td>-</td>
<td>10.8***</td>
<td>-</td>
<td>19</td>
<td>-</td>
</tr>
<tr>
<td>Bulning et al. (31)</td>
<td>All, 1990</td>
<td>61,453*** (Toronto)</td>
<td>11.4</td>
<td>-</td>
<td>4.0</td>
<td>-</td>
<td>65</td>
<td>-</td>
</tr>
</tbody>
</table>

*Citywide travel survey conducted by Beijing Municipal Government.
**Households, not workers.
***Including inward commuters from outside the limits of the City of London.
****Based on the smart card data, in-bus time only.

It can be seen from Table 2 that:
First, thanks to the processed smart card data, this study was able to have the biggest sample size among all similar studies identified. Bus commuters in Beijing have a shorter ARC and/or MRC than commuters of all modes in four other cities (Los Angeles, Atlanta, London and Toronto) of comparable land size and/or population size. What’s a little surprising is that, the ARC/MRC of bus commuters in Beijing is larger than those of all commuters in Guangzhou, another populous city in South China. But given that the land size of Guangzhou (about 7,400 km²) is only about half of that of Beijing, this is probably understandable.

Second, if we disregard the time and sample selection issues, in-vehicle commuting time based on the processed smart card data is smaller than total commuting time from travel surveys of bus commuters in Beijing across the six studies/surveys highlighted in the table. This indicates that (a) smart card data do provide accurate estimates of in-vehicle commuting time, and (b) the in-vehicle commuting time could be used to cross-check the reliability of time given by travel surveys. For instance, the latter should at least be greater than the former since most would have to spend time finishing home-bus stop and bus stop-job trips.

Third, the percentage of excess/wasteful commute for bus commuters in Beijing (64%) is notably higher than most the other cities. The percentage of excess/wasteful commute for bus commuters in Beijing is comparable to that for all commuters in Philadelphia (67%, reported in (23)) and in Toronto (65%). Domestically, the percentage of excess/wasteful commute that housing mismatch is larger than those of commuters of all modes in Guangzhou. All the above comparisons between Beijing and other cities, as a whole, indicate that bus commuters in Beijing have rather noticeable jobs-housing imbalance.
Fourth, the $C_u$ value of Beijing bus commuters is the lowest among all values presented in the table. This indicates that the jobs-housing imbalance of bus commuters in Beijing could potentially get worse. The commuting time differences between two local surveys in 2005 and in 2010 have notably “encroached” this potential.

DISCUSSIONS

Commute Efficiency for Bus Commuters

The above analyses have showed that bus commuters in Beijing face quite significant amount of excess/wasteful commute. Two natural questions one could have after getting this finding are:

(a) What caused it?

(b) What we can do about it?

With only the smart card data, it is impossible for us to provide convincing answers to these questions. But extra processing of the smart card data could still offer some clues. Based on $M_d$ and $M_i$, two matrices about trips and commuting distances between TAZs, specifically, we mapped out mean commuting distances (MCDs) by TAZs within about 30 kilometers from Tian’an Men Square (Figure 2).
FIGURE 2: Mean Commuting Distances by TAZ in Beijing*

In the figure, TAZs where MCD is smaller or equals to the metropolis’ ARC distance (8.1 km) is shown in snowfield-ice polygons. Origin TAZs where MCD is bigger than 8.1 km are shown in graduated black colors, the darker the colors, the longer distances are. If we assume that TAZs where the distance is smaller or equals to 8.1 km have achieved jobs-housing balance then these balanced TAZs are quite evenly distributed across the space and so as the imbalanced TAZs. This indicates that unlike the conventional perception among local residents, one’s residential location’s distance to the center (DtC) (Tian’an Men Square) has little to do with commuting distance/time. To validate, we investigated the relationship between MCDs by TAZ (M=8.1, SD=3.8) and DtC by TAZ (M=20.4, SD=15.8) using Pearson correlation coefficient. Analyses were performed to ensure no violation of * TAZs with no-data are not mapped.
assumptions of normality and linearity. The results indicate that there was an only weak positive

correlation between MCD and DtC ($r=0.04, p=0.25, n=729$). Based on the above analysis, we concluded

that space (or more specifically DtC) does not play a significant role in determining bus commuters’ jobs-

housing balance or imbalance.

The areas (TAZs) with the most severe jobs-housing imbalance, however, tend to follow these spatial

patterns:

(a) Clustering around in areas where is about 7 km northwest to Tian’an Men Square;
(b) Scattering near to each other in the west, southwest or southwest to Tian’an Men Square;
(c) Scattering around a half-circle with its center as the Tian’an Men Square and with a radius of 30

km (See Figure 2, the dash lines show the half-circle).

To better investigate the physical environment and community characteristics of the above areas, experts

(n=10) with profound local knowledge were consulted and field surveys were made after Figure 2 was

obtained. Based on the consultation and field surveys, we identified five land-use prototypes that could

have severe jobs-housing imbalance of bus commuters in Beijing. Table 3 summarizes the characteristics

of these prototypes and provides spatial index regarding how to find real-world examples in Figure 2. To

help one to locate concrete examples of these prototypes in Beijing, landmarks and directions are also

offered in Table 3.
**TABLE 3:** Land-use Prototypes of Extreme Jobs-housing Imbalance of Bus Commuters

<table>
<thead>
<tr>
<th>Prototype</th>
<th>Characteristics</th>
<th>Landmarks and Directions</th>
<th>Spatial Index in Figure 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>University campuses, hotels and old Danwei compounds left with mostly apartment buildings</td>
<td>Minzu and Jiaotong Universities and areas in between; Beijing Technology and Business University and Capital Normal University (east campus) and areas adjacent to them; Beijing University of Aeronautics and Astronautics and Beijing University of Science and Technology and adjacent areas</td>
<td>A (Areas around and areas to its northeast)</td>
</tr>
<tr>
<td>2</td>
<td>Parks with luxury hotels, high-end apartments, specialized research institutes, hospitals and some mixed-use residential areas</td>
<td>Areas north to Yu Yuan Tan Park; Area adjacent to Tian Tan Park in the east</td>
<td>A (South to A, the darkest area); The U-shaped area south to the star</td>
</tr>
<tr>
<td>3</td>
<td>Traditional Hu’ tong with old, cheap, small, shared and underserviced rental housing units</td>
<td>Areas in between Qian Men Da Jie and Zhu Shi Kou Da Jie</td>
<td>B</td>
</tr>
<tr>
<td>4</td>
<td>Residential areas with mixed-age housing units adjacent to freeway interchanges or arterials, railways within the fifth ring road</td>
<td>Areas near Yong Ding Men and Nan Sha Wo Bridges</td>
<td>C</td>
</tr>
<tr>
<td>5</td>
<td>Low density, developing areas with relatively cheap housing units in the suburb</td>
<td>Areas adjacent to the sixth ring road and Jingshi Freeway interchange; Areas adjacent to Yan Chun Railway Station</td>
<td>D</td>
</tr>
</tbody>
</table>

To visualize the areas highlighted in Table 3, we used [www.baidu.com](http://www.baidu.com), a leading search engine providing map services in China, to generate bird’s eye photos and detailed maps (see Figure 3).
*Within the circle are the areas north to Yu Yuan Tan Park, the blue lake in the lower left. Here, there is the General Hospital of Navy and Danwei Compounds of the Hospital and “Han Tian”, Astronautics Science and Technological Research Institute of China. To its north are Beijing Technology and Business University and Capital Normal University (east campus) and areas adjacent to them.

*Within the circle are traditional Hu’tong in between Qian Men Da Jie and Zhu Shi Kou Da Jie.
(c) Prototype 4*

*Within the circle the bride crossing the river is Yong Ding Men Bridge. To the bridge’s southeast and southwest there co-exist new real estate development (the high-rise buildings) emerged in the past two decades or so and low-rise old apartments built prior to the former.

(d) Prototype 5*

*Within the circle the bigger interchange is the sixth ring road and Jingshi Freeway interchange. Around the interchange, there are low-density developing areas with cheap housing units.

**FIGURE 3:** Visualization of the Prototypes

At this point, due to the fact that we could not identify any specific bus commuters directly from the smart card data and conduct surveys of, or interviews with them, we still do not know what exactly caused severe jobs-housing imbalance in Beijing. Publicly available disaggregate housing, travel and
employment data about Beijing are very limited. The ones we could find offer no clues. Thus we turned
to other ways. We conducted internal discussions among a small group of experienced local urban
planners and analysts (n=10) and referred to (11,12, 47, 49-50), which generated speculated reasons and
empirical evidences summarized in Table 4.

**TABLE 4: Speculated Reasons Causing Severe Jobs-housing Imbalance**

<table>
<thead>
<tr>
<th>Prototype</th>
<th>Speculations</th>
<th>Empirical evidences*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Housing side: Some fresh university graduates still live in a cheap, shared apartment near their respective alma mater but work in another locale; other fresh graduates also prefer to live in or around a university campus, where they could efficiently find a roommate to reduce rent, cheap food and services Job side: universities often do not offer a variety of jobs like CBDs and thus residents living in or around have to find a job in another locale</td>
<td>Many local housing on-line search engines offer tailored functions for users to efficiently locate housing near university campuses; A room’s rent in a shared apartment near a university can be as cheap as a couple hundred RMB per month (see: <a href="http://bj.ganji.com/xxx">http://bj.ganji.com/xxx</a>)</td>
</tr>
<tr>
<td>2</td>
<td>Housing side: Some renters managed to live an affordable apartment or room in such a premium area Job side: Specialized research institutes, luxury hotels and hospitals offer only limited types of job opportunities such as researcher, nurse, doctor and front desk clerk and thus local residents want other opportunities have to work elsewhere; some workers simply made trade-offs between residential amenities and commuting distance</td>
<td>Funan Residential Area, an old community which is adjacent to Yu Yuan Tan Park, for instance, offers housing prices that are comparable to areas nearby despite its premium location, based on <a href="http://www.anjuke.com">www.anjuke.com</a>.</td>
</tr>
<tr>
<td>3</td>
<td>Housing side: Cheap rental housing units relatively ample in Hu’tong Job side: Few non-service-sector and well-paid job opportunities available in Hu’tong</td>
<td>Except hotels and restaurants, very few employers are located in Hu’tong, based on <a href="http://www.baidu.com">www.baidu.com</a>.</td>
</tr>
<tr>
<td>4</td>
<td>Housing side: Housing units adjacent to freeway interchanges or arterials, railways are often relatively cheaper due to noise and pollution from through traffic but this means affordability to some workers Job side: Residential areas predominantly built for housing provide few job opportunities for local residents</td>
<td>Average second-hand apartment price to the southeast of Yong Dong Men Bridge, for instance, is around RMB 26,000/square meter according to <a href="http://www">www</a>. <a href="http://beijing.anjuke.com">http://beijing.anjuke.com</a> as of July 2012, which is cheaper than the average elsewhere around the inner city</td>
</tr>
<tr>
<td>5</td>
<td>Housing side: Much cheaper housing prices than in the inner city Job side: low-density developing areas in suburb offer few job opportunities of any kind</td>
<td>New housing pricing near the sixth ring road and Jingsh Freeway, for instance, is around RMB 8,000-13,000/square meter, which is much cheaper than anywhere in the inner city</td>
</tr>
</tbody>
</table>

*Few public data were available and so we turned to leading private on-line housing or generic searching engines to get evidences.
Based on reasons listed in Table 4, we now could better answer Question (b) posed above: “What we can do about bus commuters’ excess commute or jobs/housing imbalance in Beijing?”.

First, we have to understand better why commuters chose to live in the prototype areas. In prototype area I, for instance, we speculated that fresh graduates prefer to live in or around university campuses. But we do not know exactly why they do so despite that we guessed that the primary reason is efficiency in finding a roommate, housing/services affordability and/or familiarity with local communities (for those living in or around their alma mater).

Second, we probably need to provide access to pleasant parks and other public infrastructure such as quality schools and hospitals across the space so that commuters do not need to sacrifice commuting distance for such access. In prototype II, we speculate that some commuters have done that. There are also empirical evidences that support our speculation in (54).

Third, we should consider affordable housing options near employment centers so that employees do not have to live in areas that are cheap but far away from their workplace. We know this is not easy and most large-scale affordable housing projects (such as Hui Long Guan and Tian Tong Yuan with decades of history) completed in Beijing are located in the suburb, where few established employment centers were located. Even today, according to (54), except Wang Jing and its hosting district\(^5\) provide employment opportunities for 75 percent of local residents, all other large-scale affordable housing projects or new housing projects in Beijing’s suburb and their respective hosting districts could only do so for 30 percent of local residents. But we could think about hundreds and even thousands of small-scale affordable

---

\(^5\) District is an important administrative unit in China. In Beijing’s inner city, for instance, there are eight districts. Each district has its own government and is responsible for providing a wide range of services like a large city in the US.
housing projects by private parties. We could incentivize them to build, own and/or operate affordable
apartments or units in job-rich and housing-deficient districts or areas.

Fourth, we could move jobs closer to employees. There are now lots of “affordable” housing projects in
the suburbs of Beijing. We could selectively encourage small employers to establish their business in the
project areas by offering tax deduction and low-interest business loans. Similarly, we could provide
residents in the suburb who self-employ the same treatment and on-the-job trainings. Another option is to
encourage telecommuting, which makes commuting and jobs-housing balance much less relevant than
ever. Given the relatively low set-up costs, this may work best for Prototype V areas in Beijing. To make
this happen, however, we need to work on both ends: the employee end and the employer end.

Potentials of Smart Card Data

Most existing studies of jobs-housing balance and related topics have relied on self-reported travel survey
data. Due to the cost to administrate a large-scale travel survey and to process relevant responses, the
sample size of the survey is often quite small. In Table 1, for instance, most of the other studies about
Beijing have only a couple hundred of samples. Despite that these many samples could reasonably
represent the commuter population in terms of their travel behaviors and socio-demographic
characteristics they often cannot well represent global phenomena such as jobs-housing balance. In the
case of Beijing, assuming that we got workplace and housing information 6,000 samples from travel
surveys and these samples are evenly distributed across all TAZs, there would be only four or five
samples for each TAZ as there are 1,118 TAZs. No matter how hard we try, deducing any useful insights
into metropolitan-scale jobs-housing balance based on four or five samples at each TAZ is still
technically challenging. In comparison, smart card data can efficiently and economically provide
workplace and housing information of samples that are forty or even fifty times bigger. This facilitates studies of metropolitan-scale jobs-housing balance and significantly increases the representation and reliability of related analyses. Therefore, smart card data have provided an alternative way for analysts to study jobs-housing balance and related topics at the metropolitan scale.

But there are still enhancements that can be made to increase the value of smart card data, which contain no socio-economic information about cardholders/commuters. One proposed enhancement is to link smart card data to resident card data for all cardholders who authorize the data administrator to do so. In China and many other countries, resident card data are universally available and they at least provide information such as age, gender, marriage status and current home address of a resident. The other proposed enhancement is to ask anyone who purchases a smart card to voluntarily fill in an anonymous form that collects minimum amount of socio-economic information about him or her of interest to policy analysts. In return, she or he can get some bonus monetary value on the smart card or can reclaim a new smart card for free if she or he loses the original card, which is not offered currently.

CONCLUSIONS

In this study, massive smart card data were used to get 216,844 bus commuters’ workplace and residential locations. This enables a jobs-housing study of bus commuters in a metropolis with a much larger sample size than most existing studies. This study indicates that bus commuters in Beijing (a) have a shorter ARC and MRC than commuters in four other auto-dependent western cities with comparable population and/or land use sizes; (b) have a longer ARC and MRC than commuters of all modes in Guangzhou, another metropolis which is half of the size of Beijing. Overall, Beijing’s bus commuters face severe excess commute (E=64%) but still only use a relatively small portion of the commute potential (Cu=17%). As the
city expands and if nothing is done about correcting their jobs-housing imbalance, Beijing’s bus
commuters could encounter even more severe excess commute. Bus commuting time differences between
two local travel surveys in 2005 and in 2010 have partially reflected this trend.

To help reverse the trend, this study has supplemented the smart card data with local expert knowledge,
field surveys and with information provided by on-line housing searching engines so as to better
investigate bus commuting imbalance in Beijing and ways to correct it. This study has
established five land-use prototypes of jobs-housing imbalance. It has used local experts to speculate
reasons for the imbalance of each prototype. It also proposed solutions to the imbalance based on the
prototypes and speculated reasons. By doing so, the study has expanded the relatively ‘static’
analysis of the status quo (i.e., one-time quantification of a city’s jobs-housing balance, commuting
efficiency, excess commute, etc.) of a city into one of relevant dynamics. Understanding the dynamics
enables us to propose more relevant solutions to the problems identified. To other researchers who want
to use smart card data to conduct more studies, this study may have provided a good example and a
generic roadmap regarding how to enhance the value of smart card data with complementary data or
knowledge.

Lastly, despite that the above merits of this study, it can still be improved in at least three aspects in the
future. **First,** it needs to be linked to local household travel surveys, particularly the responses given by
bus commuters. This would allow one to know better these commuters’ job-, housing- and transportation-
related decisions and concerns. This would also enable one to better validate or invalidate existing expert
speculations about them. Of course, there are barriers to be overcome as there has been a tradition among
local agencies to hoarding survey responses and share them within the agencies. **Second,** it can collect
data to study jobs-housing balance and related issues among auto and bicycling commuters too. Now, the
study does not make any comparisons between them and bus commuters. Thus one does not know whether problems identified in this study are unique to bus commuters. **Third**, this study would be greatly enhanced if smart card data that include rail rapid system swipe information would be used. The smart card data administrator has such information but for some reasons they were unwilling to provide it for this study. Of course, adding this extra information would make smart card data processing and validation more complicated and challenging. However, with this extra information would enable one to better understand jobs-housing balance of more commuters, in particular, those commuters who primarily or solely use rail rapid system or those who use both bus and rail rapid system. This could be extremely important for a metropolis like Beijing that sees a growing rail rapid system and an increased share of rail rapid system commuters.

**REFERENCES**


