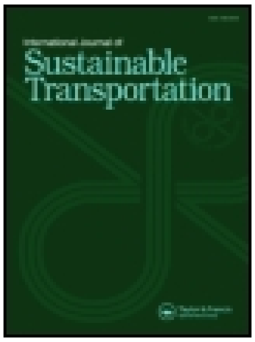


International Journal of Sustainable Transportation

Taylor & Francis
Taylor & Francis Group

ISSN: 1556-8318 (Print) 1556-8334 (Online) Journal homepage: http://www.tandfonline.com/loi/ujst20

Commuting Efficiency Gains: Assessing Different Transport Policies with New Indicators

Jiangping Zhou, Enda Murphy & Ying Long

INTERNATIONAL JOURNAL OF SUSTAINABLE TRANSPORTATION
https://doi.org/10.1080/15568318.2018.1510562

RESEARCH-ARTICLE

Commuting Efficiency Gains: Assessing Different Transport Policies with New Indicators

Jiangping Zhou[a], Enda Murphy[b], and Ying Long[c]

[a]Department of Urban Planning and Design, Faculty of Architecture, The University of Hong Kong, Hong Kong SAR, China; [b]School of Architecture, Planning and Environmental Policy, University College Dublin, Dublin, Ireland; [c]School of Architecture and Hang Lung Center for Real Estate, Tsinghua University, Beijing, China

ABSTRACT

This paper outlines new indicators for evaluating the probable impacts of introducing different land use/transport policies on the commuting efficiency of a city. It uses Beijing as a case study to describe how smartcard data can be used to derive a large number (*n* = 216,884, 9% of the population) of bus commuters' workplace and residential locations. Using existing excess commuting indicators and new commuting efficiency gain indicators established to assess policy options, it exemplifies how to assess impacts of different policies on bus commuting efficiency gains. The case study indicates policies that directly target bus commuters (such as BRT) bring greater commuting efficiency gains when compared to other policies such as car usage restrictions and stringent travel demand management measures in the downtown area.

Key Notations: T_{min}: the theoretical average minimum commute; T_{pol}: the probable average commute after a policy is adopted; T_{act}: the average actual or observed commute; T_{rand}: the average random commute; T_{max}: the theoretical average maximum commute; M_d: the existing travel cost matrix for all the existing TAZ pairs; M_g: the new travel cost matrix for all the existing TAZ pairs where a policy is adopted; M_t: the commuting trip matrix between the existing TAZ pairs; M_p: the new commuting trip matrix for a policy scenario

ARTICLE HISTORY
Received 16 August 2017
Revised 2 August 2018
Accepted 6 August 2018

KEYWORDS
Smartcard data; gravity model; policy scenario; excess commuting; bus

1. Introduction

In the United States, many cities and city-regions exhibit a severe separation of the location of employment and housing opportunities. In the academic literature this is referred to as a jobs-housing imbalance, or more formally, as a "geographic mismatch in the location of jobs and housing" (Cervero, 1991, p.10). In theory, living in closer proximity to the workplace provides the opportunity for reducing the average commuting distance; indeed, shorter distances between home and work might also promote more sustainable modes of travel such as bus, rail, walking, and cycling and thereby contribute to reducing car dependency, congestion, pollution, and resultant greenhouse gas (GHG) emissions. As a result, a number of academics and policymakers have advocated improving the jobs-housing balance as a public policy goal (e.g., Atlanta Regional Commission, 2012; California Planning Roundtable, 2008; Cervero, 1991; Weitz, 2003). In California, cities and counties that are able to demonstrate an increased housing supply relative to existing jobs within their respective jurisdictions are eligible to apply for a jobs-housing balance incentive grant from the state government.

In the academic literature, there have been competing ideas surrounding the jobs-housing balance concept, its potential usefulness and measurement (see for instance, Kanaroglou et al., 2015; Niedzielski et al., 2013; Suzuki & Lee, 2012). Within this context, it is the excess commuting (EC) framework that has provided the most suitable approach for evaluating the jobs-housing balance and wider commuting efficiency in city-regions. Within this framework, all jobs, workers and housing units are assumed to be homogeneous within a city-region while workers are permitted to switch jobs and/or housing at no cost. For any city-region a theoretical minimum commute (T_{min}) exists if workers, on average, travel to the closest possible workplace, where the travel is quantified by some measure of separation (e.g. time or distance). T_{min} captures the optimal jobs-housing balance that is permitted by the existing spatial distribution of homes and workplaces. However, the actual or observed commute (T_{act}) for all workers is invariably greater than T_{min}; this difference between T_{act} and T_{min} has been referred to as "wasteful commuting" or "excess commuting" in the literature (Hamilton & Röell, 1982; Ma & Banister, 2006a, b; White, 1988). The notion that commuting is excessive or wasteful is, of course, relative to a situation where individuals are assumed to behave optimally in terms of their home and workplace choices and the arrangement of the transport network.

CONTACT Jiangping Zhou zhoujp@hku.hk Department of Urban Planning and Design, Faculty of Architecture, The University of Hong Kong, Hong Kong SAR, China; Ying Long ylong@tsinghua.edu.cn School of Architecture, Tsinghua University, Beijing 100084, China.

In overall terms, the EC framework has been useful for understanding and measuring both jobs-housing balance and commuting efficiency in city-regions. While there has been a steady flow of literature on EC since its emergence as a separate concept in the 1980s, there continues to be gaps in terms of our understanding and use of the framework. For example, primarily due to data availability issues, the literature rarely disaggregates EC by mode of travel despite some notable exceptions (see Horner & Mefford, 2007; Murphy, 2009; and Murphy & Killen, 2011). A further gap is in relation to the application of the framework to developing countries where there has been comparatively little research completed to date although this has improved considerably in recent years (e.g., see Zhang et al., 2017; Zhou & Long, 2014; Zhou et al., 2014a,b, 2017). In addition, there have been few applications of the framework to assess various land use/transport policy scenarios in cities or regions where public transit usage is still salient. The current paper is an attempt to address some of these shortcomings.

As a whole, this study utilizes a revised EC framework to study the commuting patterns of bus riders in developing countries under different policy scenarios. The analysis builds on the previous work of Zhou et al. (2014a) and Zhou and Long (2014) but extends their work by demonstrating how the same set of smartcard data can be utilized to assess transport policy scenarios using the City of Beijing as an illustrative empirical study. More specifically, the paper contributes to existing EC studies in three dimensions. First, it provides an adapted application of the EC framework in the Chinese context, where bus commuting is still popular for a high percentage of workers. Second, it is one of a few studies utilizing smartcard data in an attempt to understand commuting patterns in major world cities like Beijing where different transport policies such as car use restrictions and stringent travel demand management measures have been adopted and could be emulated elsewhere. Third, the paper is the first of its kind to develop additional EC indicators for the evaluation of transport policy options and utilize them in a case study.

2. Excess Commuting and Associated Measures

Hamilton & Röell (1982) pioneered the EC concept. He used the monocentric model to assess the difference between T_{min} and T_{act}. T_{min} is necessitated by the physical separation between where people live and work. The difference between that and the observed commute represented "wasteful" or EC in that it was not necessitated by distribution of home and workplace land use functions. Specifically, EC can be quantified as follows:

$$\mathrm{EC} = \left(1 - \frac{T_{min}}{T_{act}}\right) \times 100 \tag{1}$$

White (1988) extended Hamilton's work by utilizing the transportation problem of linear programing (TPLP) to estimate T_{min}. The TPLP is more appropriate in that it uses the existing distribution of homes and workplaces in a zonal configuration of the city for the calculation of T_{min}. In the literature, T_{min} is calculated as follows:

$$\text{Min:}\quad Z = \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n} CijXij \tag{2}$$

$$\text{s.t.}\quad \sum_{i=1}^{n} Xij = Dj \quad \forall j = 1, \ldots, m \tag{3}$$

$$\sum_{j=1}^{m} Xij = Oi \quad \forall i = 1, \ldots, n \tag{4}$$

$$Xij \geq 0 \quad \forall i, j, \tag{5}$$

where

m = number of origins;
n = number of destinations;
O_i = commuting trips beginning at zone i;
D_j = commuting trips destined for zone j;
C_{ij} = travel cost from zone i to zone j;
X_{ij} = number of trips from zone i to zone j;
N = total number of commuters.

In subsequent work Horner (2002) demonstrated how the inverse of the minimization problem (i.e., the maximum commute (T_{max})) could be considered as the upper limit of a city's commuting range and as an indicator of the level of jobs-housing imbalance. In mathematical terms, the objective function of T_{max} is the inverse of (1) and has identical model constraints:

$$\text{Max:}\quad Z = \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n} -CijXij, \tag{6}$$

T_{max} also allows for an additional way in which to measure commuting efficiency - capacity utilization (C_u) - which provides a gauge of how much of the available commuting range has been consumed (Horner, 2002). Specifically, C_u is calculated as:

$$C_u = \frac{T_{act} - T_{min}}{T_{max} - T_{min}}. \tag{7}$$

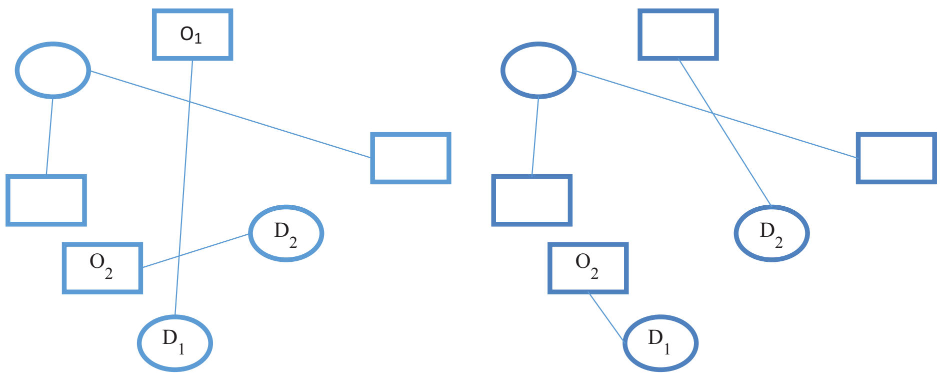
When taken together, C_u and EC provide dual measurements of the city's commuting efficiency.

The framework has also been extended in other ways. Following Hamilton's original work Charron (2007) argued that the random commute (T_{rand}) is a better benchmark for analyzing commuting efficiency because it was a more realistic gauge of the commuting possibilities available to workers. He calculates a slightly biased T_{rand} as follows:

$$T_{rand} = \frac{1}{N^2}\sum_i \sum_j O_i D_j C_{ij} \tag{8}$$

Subsequent work by Murphy and Killen (2011) utilized a more reliable T_{rand} based on a Monte Carlo method. Based on this new T_{rand}, they proposed two additional measures of commuting efficiency - commuting economy (C_e) and normalized commuting economy (NC_e) given by the following equations:

$$C_e = \left(1 - \frac{T_{act}}{T_{rand}}\right) \times 100 \tag{9}$$



(a) Commuting Pattern w/o T_{pol} (b) Commuting Pattern with T_{pol}

Figure 1. Commuting patterns in a hypothetical example.

$$NC_e = \left(\frac{T_{rand} - T_{act}}{T_{rand} - T_{min}}\right) \times 100. \tag{10}$$

C_e demonstrates the extent to which actual behavior is reacting to the cost of consuming separation – typically positively - that exists between residences and workplaces in the urban region. NC_e normalizes the extent to which observed commuting behavior deviates from random behavior within a framework where T_{rand} is taken as the upper bound of commute inefficiency. In overall terms, the foregoing measures represent the main innovations in the EC literature hitherto since its emergence in the 1980s.

3. Excess Commuting and Policy Scenarios

Based on the existing EC indicators outlined above, we propose a new indicator, the commute of policy relevance (T_{pol}), or more specifically, the likely average commute after a policy is adopted, which can be used to assess the impacts of implementing a specific transport/land use policy on commuting outcomes. Like T_{min}, T_{act}, T_{rand} and T_{max}, T_{pol} can be measured in terms of time, distance or even a monetary cost. When the associated commuting trip distribution is known or forecasted T_{pol} is calculated as follows:

$$T_{pol} = \frac{1}{N}\sum\nolimits_{i=1}^{n}\sum\nolimits_{j=1}^{m} C_{ij}X_{ij}^{pol}, \tag{11}$$

where X_{ij}^{pol} is the number of commuting trips between two zonal units when a policy is introduced while the remaining notation is the same as those in Equations (2) to (10). T_{pol} is a measure of the change in commuting cost (C_{ij}) resulting from the introduction of a specific transport/land use policy.

In order to demonstrate the usefulness of T_{pol} consider the following. Assume Person X lives at place O and works at place D. There are two distinct bus routes (R_1 and R_2) between O and D. Initially, the bus fare is the same regardless of which route person X chooses but person X takes R_1 because it is a shorter commute. Given that R_1 is a shorter route than R_2, the number of people using the route increases and it becomes overcrowded. The relevant transport authority responds by raising the fare of R_1 and reducing the fare for R_2. As a result, person X and many other bus commuters who are sensitive to the monetary cost associated with the route begin to utilize R_2. Indeed, it may also be the case that some bus commuters swap their workplaces and residences in an attempt to minimize their commuting costs. If it were possible to simulate the route choices that person X and other bus commuters would make in advance, we could better evaluate the impacts of fare changes on the resultant commuting patterns and make informed decisions in order to mitigate any negative impacts associated with a particular policy option. The route choices of individuals would, of course, also have "an intrinsically behavioural interpretation" as described by Murphy and Killen (2011: p.1261) and this would be of particular interest to policy-makers and scholars.

T_{pol} can be simulated via a doubly constrained gravity model which expands on the work of Yang and Ferreira (2008). In such a model, X_{ij}, the number of workers who reside in zone i and work in zone j can be estimated using the following formula:

$$X_{ij} = A_iB_jO_iD_j \exp(-\mu C_{ij}), \tag{12}$$

where

μ is a preference parameter;

A_i and B_j are estimated adjustment factors based on the origin and destination constraints. Other notation remains the same as Equations (1) to (11).

Following Yang and Ferreira (2008), T_{min} is achieved when μ approaches infinity. T_{act} is achieved when an empirical value μ^* is introduced into equation (11), which produces estimates of X_{ij} that best match the actual commuting trips between Zones i and j. Building on the work of Yang and Ferreira (2008), we further explain how μ is related to T_{pol}. It is important to note that T_{pol} does not change μ^*; rather, it introduces certain policy interventions to alter either the route choice of commuters, their choice of workplace/residence and/or their costs of travel (e.g., time or distance). As a result, the impact of introducing such policies is to reduce the overall average commuting cost across the city.

Figure 1 provides a hypothetical example of how public policy intervention can reduce average commuting costs. In this example, a city has m residential zones ($O_{1,2,3\ldots,m}$) and n employment zones ($D_{1,2,3\ldots n}$). Some zones such as D_1, D_2, O_1 and O_2 have an equal number of residences and jobs. Before T_{pol} is introduced, all workers residing in O_1 commute to D_1 and all workers residing in Residence Zone O_2 commute to Employment Zone D_2 (see Figure 1a). As a public policy intervention, let us assume that the government subsidizes employers via financial incentives so that all employers in D_2 and D_1 agree to swap their locations. As a result, both workers in O_1 and O_2 do not have to change their preferred/current employment and residence but would nevertheless see a decrease in their commuting distance and time thanks to route choice changes (Figure 1b). For all workers in the city, they would see a decrease in average commuting distance and time. This has occurred in reality. In the US context, for instance, despite the increased size of metropolitan areas, commuters have still enjoyed relatively stable average commuting times. This is referred to as the so-called co-location phenomenon, where employers and employees undertake relocation and/or mode-shifting efforts to ensure that employees do not have to endure a prohibitively long commute (Kim, 2008). In addition to co-location efforts, the government could also undertake measures to stabilize local commuting times. It could, for instance, subsidize, charge or regulate employees so that their commuting trips are more dispersed over a longer period of time. This way, commuting employees would see less congestion along their commuting routes and ultimately lead to a reduction in travel cost for everyone. This is, to a large extent, what happened in the congestion pricing cases of Stockholm, London and Singapore (Armelius and Hultkrantz, 2006; Goh, 2002; Litman, 2011).

The introduction of T_{pol} thus enriches the existing EC framework, enabling us to better connect the relatively abstract EC framework/indicators to real-world policy-making and assessment. Using T_{pol} as a new input for Equations (9) and (10), we propose two extra indicators for EC studies: absolute commuting efficiency change for policy scenarios ($C_{\times}^{pol}$) and absolute normalized commuting efficiency change for policy scenarios ($NC_{\times}^{pol}$), where

$$C_{\times}^{pol} = |\left(1 - \frac{T_{pol}}{T_{act}}\right)| \times 100 \qquad (13)$$

$$NC_{\times}^{pol} = |\left(\frac{T_{rand} - T_{pol}}{T_{rand} - T_{min}}\right)| \times 100, \qquad (14)$$

where "*" can be either "g", which indicates that the gravity model is applied or "ng" which indicates that the gravity is not applied.

Both $C_{\times}^{pol}$ and $NC_{\times}^{pol}$ provide a new gauge of how a policy may change the C_e of the study area relative to a do-nothing scenario. Intuitively, it would be expected that T_{pol} is smaller than both T_{rand} and T_{act}; however, there could be cases whereby T_{pol} turns out to be larger. One example might be where a city shuts down a shorter route between two nodes and forces commuters onto a much longer route. In reality, where T_{pol} is equal to or larger than T_{rand} or T_{act} a policy has no or even negative impacts on commuting and should not, therefore, be adopted or should be adapted as time progresses. In this study, we have introduced two scenarios for two types of T_{pol}: the short-term and long-term. For the short-term, we assume that commuters would not change their residences and workplaces and for the long-term, commuters would. Our simulated commuting trip distribution where the gravity model is applied generates T_{pol}'s (measured by distance) that are always larger than T_{act} in the policy scenarios under consideration. This is because our input for the gravity model to forecast trip distribution was a matrix that reflected the cost of travel between any pairs of O_i and D_j decreases by 0 to 20% when a new policy such as car use restriction was introduced. When the cost of travel by distance decreases, commuters tend to live relatively further away from their workplace; however, their average commuting time could still remain relatively stable (Levinson & Kumar, 1994; Levinson & Wu, 2005). In the ensuing sections we show such long-term impacts of a new policy as well the short-term ones, where the gravity model was not applied. In the latter case, commuters do not change their route choice and/or residential or workplace locations despite the changes in the cost of travel. Corresponding T_{pol}'s represent the short-term impacts of a new policy. As expected, in the short-term, T_{pol} is always smaller than T_{act} (See Table 6).

Table 1. Mode share of Beijing residents.

Mode	2008 (%)	2010 (%)
Bus only	28.8	28.9
Subway	8.0	10.0
Taxi	7.4	7.1
Car	33.6	34.0
Bike and walking	20.3	18.1
Company shuttle	1.9	1.9
Total	100	100

Sources: BTRC (2012).

4. Empirical Study in Beijing

4.1. The Site

As of 2014, Beijing has over 20 million residents and covers an area of 16,410 square kilometers. Beijing Public Transportation Company (BPTC), a state-owned company provides public bus services in Beijing, has 28,343 buses on 948 bus routes with a total service length of 187,500 kilometers as of 2011. In 2011 alone, these buses produced vehicle kilometers traveled of 1.7 billion with a total of 4.9 billion passengers being transported. Table 1 shows the modal split for local residents in 2008 and 2010. It is evident that bus commuting is a very important and popular mode of transport in the city. In Beijing, 28.9% of residents reported that their mode of travel was bus only in 2008 when the city had 8.2 million workers at that time (Beijing Municipal Bureau of Statistics, 2010).

Since 2005, over 90% of bus riders in Beijing have swiped an anonymous smartcard when boarding and alighting (for suburb routes and long distance routes in the central city) or when boarding (for inner-city routes) to pay for their fare (Long et al., 2012). The high rate of smartcard usage among bus riders is largely because of the subsidy the government gives to riders who pay their bus fare with a smartcard. Those riders enjoyed 60% discounts on any routes in the local bus system in 2011. Smartcards can also be used by commuters to pay for other services such as taxis, electricity and waste bills. It should be mentioned that the fare policy of public transit in Beijing was updated in 2015 and since then all bus lines require both swipe on and swipe off for payment. However, our data/files were for 2008 and thus still represents the travel behavior of bus commuters prior to 2015.

In 2008, BICP partitioned Beijing into 1,118 TAZs. On average, each TAZ is about 14 square kilometers. For the inner city, TAZs are much smaller than the global average, as shown in Figure 2. The average size of TAZs in Beijing's core is about 3 square kilometers, which is comparable to or even smaller than that of the TAZs or sub-divisions used in most existing EC studies. For instance, in Small and Song (1992), the 3,341-square-kiliometer Southern California Region was divided into 706 TAZs, where each TAZ is about 5 square kilometers. In Murphy (2009) and Murphy and Killen (2011), the Greater Dublin Region consists of 463 sub-divisions and covers 6,982 square kilometers, where each sub-division is about 15 square kilometers.

4.2. The Data

To show the usefulness of the above-mentioned new indicators, we conduct empirical studies of Beijing, where we have access to local smartcard data, which capture a high percentage (over 90%) of local bus riders. Based on this data from Beijing Institute of City Planning (BICP), we can derive a large number of homes, workplaces and/or both of local workers. The period covered by the data was between 7 and 13 April, 2008. They thus represent a situation where the Beijing Olympic Games had not yet occurred. In addition to the smartcard data, we acquired the ShapeFiles (an open GIS data format) of local traffic analysis zones (TAZs) and the road network produced by BICP for its 2008 local travel demand model. All the above data/files enabled us to run all necessary models presented in this study.

When a cardholder in Beijing uses their smartcard to pay for bus services, the card reader installed on the bus automatically generates the following information:

a. bus trip origin and/or destination stop[1];
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 were granted access to a full week's historical data from the server administrator, which contains 77,976,010 bus trips associated with 8,549,072 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 this process for seven days. In order to determine a person's workplace the criteria were established based on a decision-tree method and local household travel survey (see Long and Thill [2015] for more details). Thus, a cardholder's workplace was identified if it met the following criteria:

a. The card type is not a student card;
b. $H_j >= 6$ hours, where H_j is the duration that a cardholder stays at place j, which is associated with all bus stops within 500 meters of one another;
c. $j <> 1$, which means that j is not the first place in a weekday that the server records.

The place where a cardholder visited most frequently in five weekdays is defined as the workplace of the cardholder in this study. Based on the above, we identified 1,113,913 workplaces associated with 8,549,072 distinct cardholders.

Similarly, a cardholder's home was identified if they met the following conditions:

a. The cardholder has an identified workplace;
b. The card type is not a student card;
c. $T_h >= 6$ hours, where T_h is the duration that a cardholder stays at place h, which is associated with all bus stops within 500 meters of one another;
d. $F_h >= F_j$, where F_h is the first and the 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, 3,778,673 homes associated with 8,549,072 distinct cardholders were identified.

To further validate the deduced workplaces and homes, we checked our derived information with local household travel survey data, as well as a parcel-level land use map (for more details, see Long & Thill (2015)). In addition, to ensure that we singled out commuters solely by bus, we only selected cardholders that had continuous bus swipes. That is, our study does not consider commuters who are multimodal. Subway swipe information was available but the authors were not granted access to this data. The other thing that should be noted here is that for most cardholders, we could only derive their home or workplace. We could only detect a much smaller number of cardholders' homes and workplaces simultaneously.

Similar to existing studies such as Frost et al. (1998), Horner (2002), Murphy (2009) and Murphy and Killen (2011), bus trips originating from and destined for locations outside the study boundary were excluded in this study. In the end, we successfully validated and identified both homes and workplaces of 216,844 cardholders out of a total of 8,549,072. Thus, if we assume workers' mode of travel is the same as that of residents, there were 2.4 million workers who commuted by bus only (see Table 1 and associated text

[1]For those inner-city routes, the card holder is only required to swipe his/her card when boarding the bus but not when getting off the bus. In this case, we deduced the bus origin and destination using the same card's two swipes on a weekday.

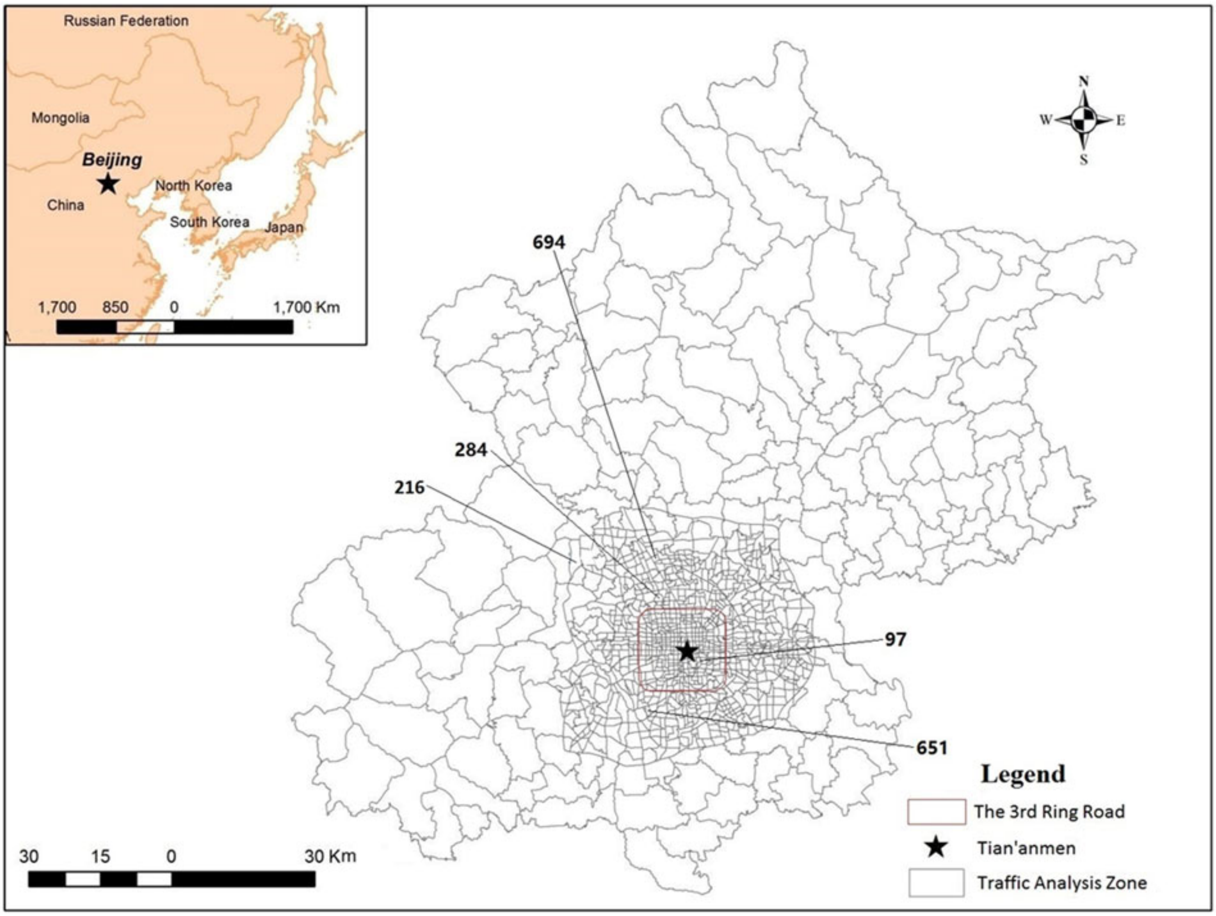


Figure 2. 1,118 TAZs in Beijing.

Table 2. Descriptive statistics of Homes and Workplaces by TAZ.

	Count*	Sum	Min.	Max.	Mean	Std. Dev.
Homes	729	216,844	1	2,880	297	346
Workplaces	752	216,844	1	2,340	288	368

*Altogether, there are 767 distinct TAZs that have at least one home or workplace.

above). Our bus rider/worker samples ($n = 216,844$) derived from local smartcard data thus represent approximately 9% of local workers.

We then geocoded and aggregated commuters' home and workplace data by TAZ. Descriptive statistics associated with commuters' homes and workplaces are given in Table 2.

Figure 3 presents the top 500 TAZ pairs that have the most bus commuters, with the local job (derived from smartcard data) distribution as the background. The graphic shows the tendency for radial trip making around employment subcenters beyond the core of the city.

Based on the home and workplace distribution information, we constructed two matrices: one is for the commuting trips (i.e., Origin-Destination flow information) (M_t) and another for journey-to-work travel cost between OD pairs (M_d). In contrast with the straight-line distances used in existing studies, in M_d, we used network distances between TAZ centroids. The network distances were calculated in TransCAD 5.0. TAZ ShapeFiles and road network information from BICP are used as inputs. For trips within the same TAZ, we assumed that the commuting distance equals R_i, where,

$$R_i = \sqrt{\frac{A_i}{\pi}}. \tag{15}$$

A_i is the approximated area of each TAZ_i, as utilized in previous studies such as Frost et al. (1998), Horner (2002), Murphy (2009) and Murphy and Killen (2011).

4.3. Policy Scenarios for Beijing

In theory, we can devise as many as policy scenarios as we want and calculate corresponding T_{pol}'s for them. But some policy scenarios are always of greater interest to policy-makers than others. In the case of Beijing, we explored three policy scenarios and corresponding T_{pol}'s to assess how bus commuters' EC indicators would compare to a base do-nothing scenario. The following policy scenarios were considered and all are based on revised distance-based trip matrices.

Policy Scenario 1: Beijing authorities restrict car usage on the basis of the last digit of a car's number plate on weekdays. This results in a 0%–20% decrease in travel costs between TAZs. In reality, Beijing has enforced such restrictions since 2008 and about 20% of all the private cars are not permitted on weekdays in the city.

Policy Scenario 2: Beijing adopts comprehensive travel demand management (TDM) measures resulting in a 0%–20% decrease in traffic flow and travel costs for trips to

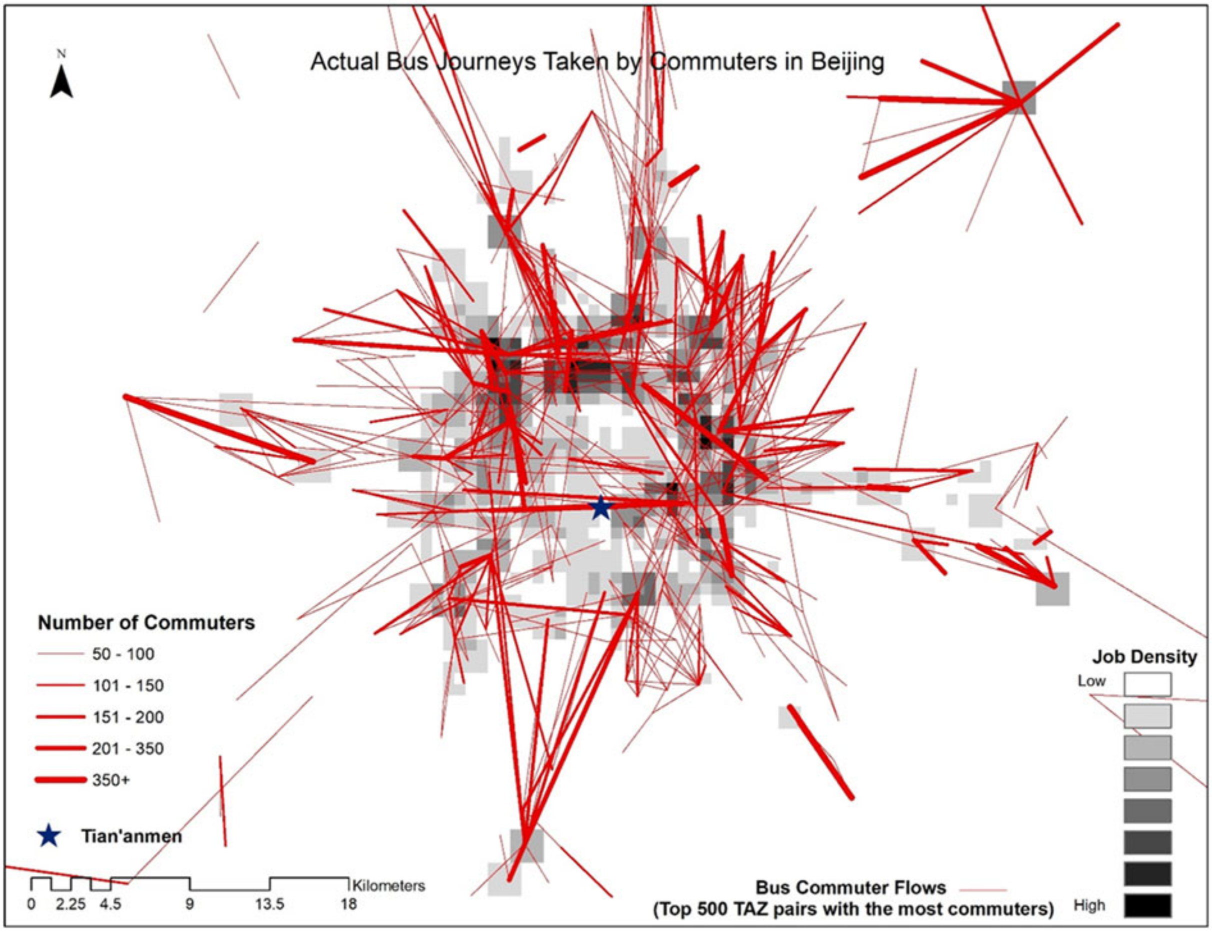


Figure 3. The TAZ Pairs with the Most Bus Commuters.

and from TAZs within the 3rd ring road of Beijing. In reality, Beijing authorities have already started adopting a series of TDM measures since 2011. They include annual quotas and lotteries for new vehicle registrations, increased parking prices in the inner city and prioritizing public transportation projects (Gu et al., 2017).

Policy Scenario 3*:* In light of existing large volumes of bus riders to several employment centers (TAZs 97, 216, 284, 651 and 694—see Figure 3 for their respective locations on the local map) where there are more than 2,000 incoming bus commuters per day, Beijing now operates bus rapid transit (BRT) to and from these centers and consolidates services of certain existing bus routes. As a result, all bus trips to and from these top employment centers see a reduction of travel cost between 0% and 20%. In reality, Beijing did introduce two BRT routes since the year 2000. Some commuters from the south and the north have greatly reduced their commuting times to the central city as a result.

For each policy scenario, we created a new distance-based travel cost matrix (M_g) for all the existing TAZ pairs, which has an existing travel cost matrix: M_d. Ideally, both M_g and M_d should be measured in time. But because we were not granted sufficient input data, we were not able to measure M_g and M_d in time. Nevertheless, given that travel time and travel distance are roughly positively correlated, our ensuing results do still somehow reflect the ideal scenario.

For M_g, the cost of commuting between all or certain TAZ pairs results in a random (0%–20%) decrease in commuting costs. In this study, we assume that the cost of commuting is the only utility that is of concern to commuters and that it has a positive linear correlation with the network distance between the centroids of any two TAZs. More details regarding how we assembled new M_g for each policy scenario and how each policy scenario is related to bus commuting is summarized in Table 3. What should be noted is that we made simple assumptions (e.g., all relevant commuting cost decreases by a random factor) about what would happen to commuting costs when a policy is introduced. This is largely because we do not have access to important or historical local data before and after Beijing enforced car usage restrictions based on plate number. Therefore, our scenario analysis is illustrative rather than evaluative in nature. Thus, it provides a proof of concept for the use of the EC framework for analyzing various policy scenarios and should be evaluated on this basis rather than on the specific scenario assumptions being adopted.

With the new M_g's and the known numbers of homes ("trip productions") and jobs ("trip attractions") by TAZ, we used a gravity model to generate new commuting trip matrices (M_p's) for Policy Scenarios 1 to 3, which would be necessary as input into the various modeling exercises of the EC framework that would allow us to obtain values for T_{pol}. The gravity model is the most widely used trip distribution

Table 3. Policy scenarios and bus commuting.

Policy scenario	Cost matrix characteristics (As compared to the baseline scenario)	Relationship to bus commuting	Notes
1: Car restriction	Travel costs (M_g) between any two TAZs decrease by a random factor of 0%–20%	On average, bus commuters would see a decrease in commuting cost as long as they travel within Beijing	Trip distribution of bus commuters would change across the city
2: TDM	Travel costs (M_g) to and from any TAZs within the 3rd ring road (See Figure 3) decrease by a random factor of 0%–20%	On average, bus commuters to and from the TAZs within the 3rd ring road would see a decrease in commuting cost	Only trip distribution of bus commuters to and from TAZs within the 3rd ring road would change
3: BRT	Travel costs (M_g) to and from any of the top five TAZs which have the most employment decrease by a random factor of 0%–20%	On average, bus commuters to and from the top employment centers would see a decrease in commuting cost	Only trip distribution of bus commuters along certain corridors would change

model (Caliper Cooperation, 2015) and can be automatically run in modern commercial software packages like TransCAD, PTV and Cube once M_g, trip productions, trip attractions and parameters of the gravity model are known. Mathematically, a typical gravity model can be expressed as:

$$T_{ij} = K_i K_j O_i D_j f(C_{ij}) \tag{16}$$

$$K_i = \frac{1}{\sum_j k_j D_j f(C_{ij})} \tag{17}$$

$$K_j = \frac{1}{\sum_j k_i O_i f(C_{ij})}, \tag{18}$$

where
$K_{i,j}$ are balancing factors to be solved interactively;
f(x) is the impedance function, which takes the form of

$$f(C_{ij}) = K \times C_{ij}{}^{a} \times \exp\left(b \times C_{ij}\right); \tag{19}$$

all other notation follows from previous equations.

In a doubly constrained gravity model, for any i, the total number of trips from i predicted by the model always (mechanically, for any parameter values) equals the real total number of trips from zone i. Similarly, the total number of trips to zone j predicted by the model equals the real total number of trips to j, for any j. In our case study of Beijing, we applied a doubly constrained model for Beijing because the typical EC framework assumes a fixed distribution of residences and workplaces. The values of K, a and b in the impedance function were calibrated based on the derived bus commuting trip matrix M_t, which contains 216,844 commuting trips. Table 4 presents the resultant calibrated K, a and b values (coefficients).

The technical procedures regarding how to estimate the values of K, a and b can be found in Sen and Pruthi (1983). With the presence of commercial software packages, those procedures can be automatically implemented once M_t or M_p are known. In our case, we used MATLAB to implement the technical procedures and TransCAD to validate the means of K, a and b that MATLAB generated. The calibrated K, a and b allow us to simulate the bus commuting trips with good approximation in terms of trip distance distribution (Figure 4). Our results indicate that the actual commuting pattern in Beijing based on our 216,844 sample largely follows a conventional gravity model. The gravity model was able to simulate the bus commuting trips' distance distribution with a small margin of error. Larger errors tend to occur when the trip distance is shorter than 8 km. But even the largest margin of error is still acceptable: for trips between 0 and 2 km, there are 5.5% and 4.9% of all the actual/simulated trips, respectively. In other words, the largest error of margin is only 0.6%.

Table 4. Coefficients for the Gravity Model.

	K	a	b
Coefficient (Calibrated)	0.011	1.01111	−1.1216

5. Results

Table 5 presents results for the baseline scenario as well as a selection of results identified from previous EC studies to add additional context to the Beijing case. It is notable that the results outlined for T_{min}, T_{act}, T_{max}, EC and C_u are different from earlier work by Zhou et al (2014) due to the use of bus/road network distances over air/desired distance as a proxy for commuting cost. The other key difference with the results emerging from this study is that T_{rand} has been calculated whereas it was not for the previous study. The introduction of T_{rand} has allowed for the calculation of Beijing bus commuters' C_e and NC_e, the first in the literature for the city. This enables a comparison of bus commuters' C_e between Beijing and a western city (Dublin, Ireland), which is also the first of its kind in the literature. The comparison shows that Beijing's bus commuters have a lower C_e and NC_e – 6% - to their public transport counterparts in Dublin for 2001 (28 and 40% respectively). This indicates bus commuters in the two cities differ significantly in their actual commuting behavior in terms of how they react to the cost of separation between homes and workplaces. Bus commuters in Dublin are further away from behaving as random commuters than their counterparts in Beijing, i.e., where the cost of commuting distance between land uses is considered irrelevant. Between Dublin and Beijing, we have observed divergence instead of convergence in bus/transit commuting/behaviours (c.f., Zhong et al., 2015, 2016). These results imply that Beijing's bus network as of 2008 was still organized inefficiently relative to the existing distribution of homes and workplaces in the city. One caveat, however, is that NC_e and C_e can be affected by city size and thus the differences in NC_e and C_e between

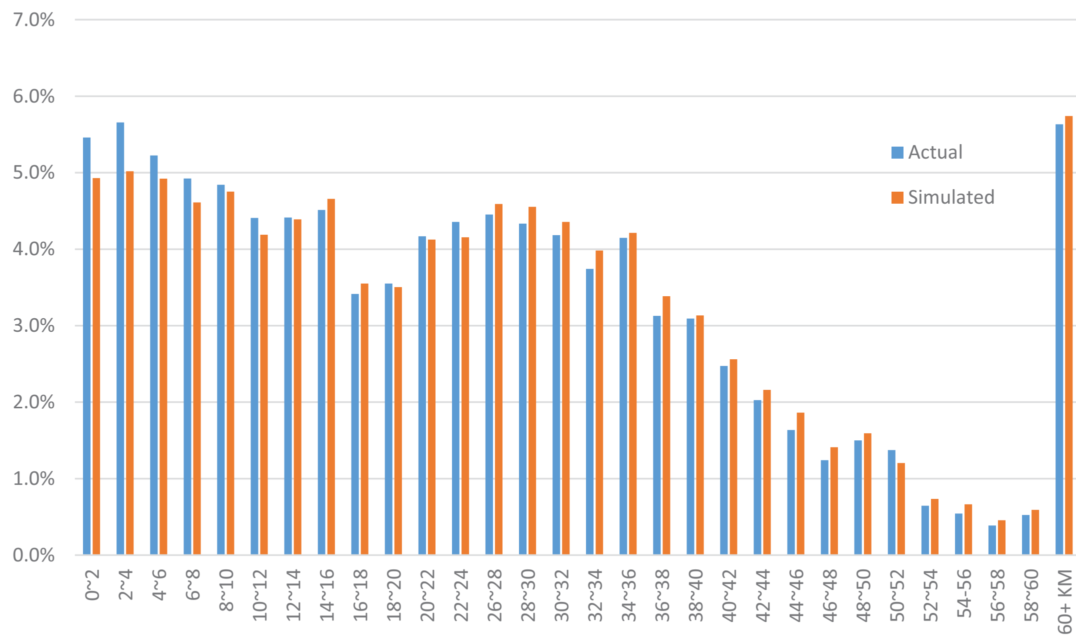


Figure 4. Distance distribution of the actual and simulated bus commuting.

Table 5. Different EC indicators of Beijing and Dublin.

Study/Survey	Mode, year	Sample size	T_{min}	T_{act}	T_{rand}	T_{max}	EC	C_u	C_e	NC_e
			km				%			
Current study (Network distance)	Bus, 2008	216,844 (Commuting only)	1.9	22.7 (23.5#)	24.1	40.7	92	54	6	6
Current study (Straight-line distance)			2.5	8.1	11.6	24.7	69	25	29	37
Murphy and Killen (2011)	Public transport*, 2001 (Dublin)	Not reported	2.8	6.5	9.0	11.6	60	58	28	40

*Public transport system in Dublin consists of bus and light rail and bus trips account for the bulk share of the public transport trips.
#T_{act} based on the gravity model.

Table 6. EC Indicators for Policy Scenarios.

	T_{pol} (Gravity model not applied, short-term impacts)	T_{pol} (Gravity model applied, long-term impacts)	T_{min}	T_{act} (All from the baseline scenario)	T_{rand}	T_{max}	Gravity model not applied: C^{pol}_{ng}	Gravity model not applied: NC^{pol}_{ng}	Gravity model applied: C^{pol}_{g}	Gravity model applied: NC^{pol}_{g}
Unit			km					%		
1: Car Restriction	16.7	23.6	1.9	22.7	18.3	30.0	26	10	4	32
2: TDM	22.5	31.3	1.8		25.3	47.4	0.8	12	38	26
3. BRT	22.6	31.6	1.9		25.5	48.7	0.4	12	39	26

Beijing and Dublin may not as big as we had observed (Kanaroglou et al., 2015).

In addition to the EC indicators for the baseline scenario, we also calculated EC indicators for policy scenarios proposed in the methodology section (Table 6). When calculating the indicators, we have two sets of T_{pol}'s and of C^{pol} and NC^{Pol}, respectively. One applies the gravity model and the other does not. We can regard them as the long-term and short-term impacts of different policies on EC.

Based on the results in Table 6, we found that all the policy scenarios considered could lead to significant commuting efficiency gains (for bus commuters) over the short term implying city-wide reductions in vehicle kilometers traveled, environmental emissions as well as improved fuel economies. This is unsurprising: it is quite typical for a policy to be introduced on the basis of anticipated positive benefits for the traveling public. What is interesting is that the policy scenarios that require larger scale efforts do not necessarily result in more commuting efficiency gains than other scenarios. The TDM scheme (Policy Scenario 2), for instance, which involves pricing/restricting car commuters in the city, actually has similar impacts on commuting efficiency gains as the introduction of BRT (Policy Scenario 3) in the short term (measured by NC^{pol}_{ng}). In the long term, assuming that commuters switch route choices and/or workplaces/residences because of lower travel cost introduced by different polices, it can be seen that Policy Scenarios 1 to 3 would all eventually suffer from C_e loss.

The results also show that under all policy scenarios C_e improves in the short term. The results for C^{pol}_{ng} demonstrate that commuters are 26, 0.8 and 0.4% gains in commuting economy under Policy Scenarios 1 to 3 respectively compared to the baseline scenario. However, when the data is normalized relative to the available commuting capacity of the city (assuming T_{rand} as the upper limit on commuting behavior) it can be seen that Policy Scenario 1 is less effective than Policy Scenarios 2 and 3. This indicates that larger-scale efforts do not always increase C_e/efficiency as one would expect.

As a whole, Table 6 and the corresponding policy scenario design and quantification also exemplifies that the introduction of new indicators such as T_{pol}, $NC_{\times}^{pol}$and $C_{\times}^{pol}$ allow us to better connect the EC framework to real world policymaking and policy evaluation. These new indicators enable us to quantify the probable impacts of different policies on commuting as compared to the baseline scenario. They also allow us to compare those impacts and decide which policy should be prioritized. Of course, the results of different policy scenarios presented in Table 6 are based on some simplifications of the real world and considerable assumptions about commuting cost. If we want to put the above new indicators to better use so that they are more relevant to real-world decision-making or policy assessment, we need to collect more input data so as to design better policy scenarios and obtain more convincing results about them than those presented in this illustrative study.

6. Discussion and Conclusions

In this paper, smartcard data, a type of 'big data', has been used to derive the residential and workplaces of bus commuters, thus enabling the retrieval of a much larger sample than would typically be the case in traditional EC studies. This paper has, above all, proposed new and transferrable indicators in light of local public policies, which helps to connect the excess commuting framework to daily and monthly policy evaluations and decision-making. It maintains some assumptions of the framework such as the total number of jobs and residences will be fixed within zonal units but assumes that policy changes will result in travel cost and trip distribution changes between zonal units. This is largely consistent with reality: it is very difficult to alter specific land use patterns once they are there; but it is very possible to introduce measures that result in travel-cost changes between origins and destinations. For the case of Beijing, we outline that introducing certain transport-policy options can improve the C_e of the city.

The paper has compared results emerging from the excess-commuting framework from Beijing with Dublin specifically in terms of bus C_e. It has also outlined and demonstrated the application of new C_e measures for transport policy scenarios. Quite a few concrete findings have emerged. Our results show that smartcard data can be used to generate most if not all of the input data necessary for the development of indicators for bus commuters within the excess-commuting framework.

They also show that the introduction of additional indicators such as T_{pol}, $C_{\times}^{pol}$and $NC_{\times}^{pol}$ can facilitate the development and evaluation of policy scenarios that inform policymakers about the potential impacts of a range of options on C_e/efficiency. Indeed, our results demonstrate that policies directly targeting bus commuters (such as BRT) tend to bring comparable benefits to bus commuters like other larger-scale interventions (e.g., parking restrictions for all commuters to and from the downtown). However, we must admit that the specification of absolute values of these indicators means that although we can judge the effects of a policy change, we cannot evaluate its direction.

In addition, our study can further inspire other researchers and decision-makers in at least three other aspects associated with smartcard data usage. First, like Beijing, many other cities have introduced a similar smartcard into their public transportation system. For instance, Chicago has the "Chicago Card Plus", Los Angeles has the "Tap card", Washington D.C. has the "SmarTrip" card and Atlanta has the "Breeze" card. All these cards could generate similar raw smartcard data that we used in the case of Beijing. Such data, when appropriately processed and validated, could provide an alternative form of input data to support studies of commuting patterns and their determinants/impacts. In theory, these data can enable temporal studies across hours, days, weeks, months and years, and are much more efficient and economical to acquire than traditional data such as censuses and surveys. Such data has the potential to induce a paradigm shift in public transport planning, operation and related plan and policy evaluations (c.f., Batty 2013).

Second, is linking smartcard data to other conventional data. Most smartcards do not store personal information such as home address, income, gender and age. However, they can still be used to generate home and workplace information with quite refined geographic resolutions, for instance, at the bus stop level. This means that the smartcard data can be aggregated later to larger units of analysis such as zip code, TAZ, or district. In this paper, we aggregated relevant data by TAZ. There are many variables such as the number of workers, residences and mean travel time available at the TAZ level from local transport planning entities. We did attempt but did not succeed in gaining access to these variables. For other cities where transport-planning-related data availability and access are not a sensitive issue, smartcard data can be aggregated to other units of analysis, say census block or tract. At those levels, there are often many publicly available variables, which would enable researchers to do more research. In the US, for instance, one can access age, sex, race, income, education, housing, mode choice, travel time, etc. information at the census tract level through the US Census' website. If all the above data are not available, one can still design conventional surveys to supplement smartcard data. Combing smartcard data with these conventional/extra data would engender more relevant and useful studies of bus/transit commuters, which could inspire more informed plans and policies.

Third, smartcard data can be used to help policy or plan evaluations. In this study, we devised three different policy scenarios, estimated M_p's, the most probable trip distributions associated with them and calculated corresponding T_p's and C_e. Our policy scenarios are by no means perfect as they are illustrative in nature and we did not have all the desired input data to estimate M_p's. But our procedures, models and methodologies are replicable and transferrable. They highlight how researchers might do a better job in other cities where there are more or better input data. With such data, we are confident that they could complete more excess-commuting studies with policy relevance, even by simply replicating our efforts described above using local data as input.

Finally, research in this area can be enhanced in three aspects in the future. First, smartcard data need to be better linked to customized household travel survey data, particularly for bus commuters. This would give us better knowledge of local bus commuters travel behavior, housing choice and their determinants. Of course, there are barriers to be overcome in this regard because of the tradition among local agencies to hoard most data they have from researchers in the Chinese context (Zhou and Wang, 2014). Second, combining our data with that of other modes, particularly automobiles, would be highly beneficial. The absence of reliable data for automobiles has prevented us from getting a fuller picture of local commuting patterns across different modes of travel. The trip distribution of auto commuters in different policy scenarios can significantly differ from that of bus commuters. But they mutually influence each other in terms of trip generation, travel time, route and/or mode choices. Third, it would also be beneficial to integrate bus data with rail data (e.g. subway data). Local agencies have this data but due to the privacy/security concerns mentioned above they were unwilling to share it. It is likely that adding extra information about subway users would make smartcard data processing, derivation of homes and residences and corresponding validation more complicated and challenging. But this extra information would enable us to better understand the EC of more commuters, in particular, those commuters who primarily or solely use subway, those who use both bus and subway, or those who walk/bike and use subway. This could be extremely important for a metropolis like Beijing where the subway system is growing rapidly together with an increase in the share of multimodal commuters whose primary mode is subway.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Jiangping Zhou http://orcid.org/0000-0002-1623-5002

References

Armelius, H., & Hultkrantz, L. (2006). The politico-economic link between public transport and road pricing: An ex-ante study of the Stockholm road-pricing trial. *Transport Policy*, 13(2), 162–172.

Atlanta Regional Commission. (2012). *Jobs-housing balance*. Atlanta, GA: Atlanta Regional Commission.

Batty, M. (2013). Big data, smart cities and city planning. *Dialogues in Human Geography*, 3(3), 274–279.

Beijing Municipal Bureau of Statistics. (2010). *The second economic census of Beijing*. Beijing, China: Beijing Municipal Bureau of Statistics.

Beijing Transportation Research Center (BTRC). (2012). 2011 annual report of Beijing's transportation development (In Chinese), Beijing, China. Available at: http://www.bjtrc.org.cn/JGJS.aspx?id =5.2&Menu=GZCG. Accessed November 2015.

Caliper Cooperation. (2015). *TransCAD 7.0 User Guide*. Newton, MA: Caliper Cooperation.

California Planning Roundtable. (2008). *Deconstructing jobs-housing balance*. Sacramento, CA: California Planning Roundtable.

Cervero, R. (1991). Jobs-housing balance as public policy. *Urban Land*, *10*, 14.

Charron, M. (2007). From excess commuting to commuting possibilities: more extension to the concept of excess commuting. *Environment and Planning A*, 39(5), 1238–1254.

Frost, M., Linneker, B., & Spence, N. (1998). Excess or wasteful commuting in a selection of British cities. *Transportation Research A*, 32(7), 529–538.

Gu, Y., Deakin, E., & Long, Y. (2017). The effects of driving restrictions on travel behavior evidence from Beijing. *Journal of Urban Economics*, 102, 106–122.

Goh, M. (2002). Congestion management and electronic road pricing in Singapore. *Journal of Transport Geography*, 10(1), 29–38.

Hamilton, B. W., & Röell, A. (1982). Wasteful commuting. *Journal of Political Economy*, 90(5), 1035–1051.

Horner, M. W. (2002). Extensions to the concept of excess commuting. *Environment and Planning A*, 34(3), 543–566.

Horner, M. W., & Mefford, J. N. (2007). Investigating urban spatial mismatch using job-housing indicators to model home-work separation. *Environment and Planning A*, 39(6), 1420–1440.

Kanaroglou, P. S., Higgins, C. D., & Chowdhury, T. A. (2015). Excess commuting: A critical review and comparative analysis of concepts, indices, and policy implications. *Journal of Transport Geography*, 44, 13–23.

Kim, C. (2008). Commuting time stability: A test of a co-location hypothesis. *Transportation Research A*, 42(3), 524–544.

Levinson, D., & Kumar, A. (1994). The rational locator: Why travel time have remained stable? *Journal of American Planning Association*, 60(3), 319–332.

Levinson, D., & Wu, Y. (2005). The rational locator: Re-examined. *Transportation*, 32(2), 187–202.

Litman, T. (2011). London Congestion Pricing Implications for Other Cities. Available at: https://www.vtpi.org/london.pdf, Accessed 07 January, 2018.

Long, Y., & Thill, J.-C. (2015). Combining smart card data and household travel survey to analyze jobs–housing relationships in Beijing. *Computers, Environment and Urban Systems*, 53, 19–35.

Long, Y., Zhang, Y., & Cui, C. (2012). Analysing jobs-housing relationship and commuting pattern of Beijing using bus smartcard data (In Chinese). *Acta Geographica Sinica*, 67, 1339–1352.

Ma, K. R., & Banister, D. (2006). Excess commuting: a critical review. *Transport Reviews*, 26(6), 749–767.

Ma, K. R., & Banister, D. (2006). Extended excess commuting: A measure of the jobs-housing imbalance in Seoul. *Urban Studies*, 43(11), 2099–2113.

Murphy, E., & Killen, J. E. (2011). Commuting economy: An alternative approach for assessing regional commuting efficiency. *Urban Studies*, 48(6), 1255–1272.

Murphy, E. (2009). Excess commuting and modal choice. *Transportation Research A*, 43(8), 735–743.

Niedzielski, M. A., Horner, M. W., & Xiao, N. C. (2013). Analyzing scale independence in jobs-housing and commute efficiency metrics. *Transportation Research A*, 58, 129–143.

Sen, A., & Pruthi, R. K. (1983). Least Squared calibration of the gravity model when interzonal flows are unknown. *Environment and Planning A*, 15(11), 1545–1550.

Small, K. A., & Song, S. (1992). Wasteful commuting: A resolution. *Journal of Political Economy*, 100(4), 888–898.

Suzuki, T., & Lee, S. (2012). Jobs-housing imbalance, spatial correlation, and excess commuting. *Transportation Research A*, 46(2), 322–336.

Yang, J., & Ferreira, J. R. (2008). Choices vs. Choice sets: A commuting spectrum method for representing job–housing possibilities. *Environment and Planning B: Planning and Design*, 35(2), 364–378.

Weitz, J. (2003). *Jobs-housing Balance*. Chicago, IL: American Planning Association.

White, M. J. (1988). Urban commuting journeys are not "wasteful". *Journal of Political Economy*, 96(5), 1097–1110.

Zhang, P., Zhou, J., & Zhang, T. (2017). Quantifying and visualizing jobs-housing balance with big data: A case study of Shanghai. *Cities*, 66, 10–22.

Zhong, C., Manley, E., Arisona, S. M., Batty, M., & Schmitt, G. (2015). Measuring variability of mobility patterns from multiday smart-card data. *Journal of Computational Science*, 9, 125–130.

Zhong, C., Batty, M., Manley, E., Wang, J. Q., Wang, Z. J., Chen, F., & Schmitt, G. (2016). Variability in regularity: mining temporal mobility patterns in London, Singapore and Beijing using smart-card data. *PLoS One*, 11(2), e0149222.

Zhou, J., & Long, Y. (2014). Jobs-housing balance of bus commuters in Beijing. *Transportation Research Record: Journal of the Transportation Research Board*, 2418(1), 1–10.

Zhou, J., Murphy, E., & Long, Y. (2014). Commuting efficiency in the Beijing metropolitan area: an exploration combining smartcard and travel survey data. *Journal of Transport Geography*, 41, 175–183.

Zhou, J., Zhang, C., Chen, X., Huang, W., & Yu, P. (2014). Has the legacy of Danwei persisted in transformations? the jobs-housing balance and commuting efficiency in Xi'an. *Journal of Transport Geography*, 40, 64–76.

Zhou, J., & Wang, Y. (2014). Mobile-Source GHG Modeling Institutions and Capacities in China. *China City Planning Review*, 23(2), 14–23.

Zhou, J., Wang, Y., Cao, G., & Wang, S. (2017). Jobs-housing balance and development zones in China: A case study of Suzhou Industry Park. *Urban Geography*, 38(3), 363–380.