Integrating road carrying capacity and traffic congestion into the excess commuting framework: the case of Los Angeles

Abstract: The excess commuting (EC) framework has advanced a series of metrics through which a city or a region's jobs-housing balance and commuting efficiency can be measured. This study seeks to add to the conceptual development and extension of the EC framework. Specifically, it considers the carrying capacity (of links) and related congestion issues in the EC framework and demonstrates that by overlooking these characteristics has important implications for EC metrics. Drawing on an empirical case study, it shows that when carrying capacity and traffic congestion are accounted for, the actual commute is longer than otherwise. EC tends to be higher than its counterparts in previous excess commuting studies. The findings suggest that future EC studies should take account of the carrying capacity and congestion in determining EC metrics. Moreover, high-quality connections (preferably via public transport) between jobs and housing allied with sufficient carrying capacity of popular links/routes for commuters are crucial preconditions for cities and regions to harvest the full benefits of jobs-housing balance policies targeted at the reduction of the average commute distance and vehicle miles travelled.

Keywords: Excess commuting; Jobs-housing balance; Los Angeles; Assignment modelling; Traffic Congestion

1. Introduction

Since the publication of Hamilton's (1982) seminal work, the excess commuting (EC) framework has received much scholarly attention. The EC has been utilised to derive a suite of benchmarks and indices for the analyses of the, jobs-housing balance, spatial mismatch, and commuting efficiency in urban regions (see for example, Giuliano and Small, 1993; Horner, 2007; Horner and Mefford, 2007; Kanaroglou et al., 2015; Scott et al., 1997; Small and Song, 1992). Within the EC framework, urban form is typically considered to be a pattern of disaggregated spatial units/zones each containing a fixed number of jobs and/or residences. Trips between zones are associated with a cost which can be measured either in terms of travel time, distance, or even some form of monetary cost. Assuming all commuters are homogeneous in their skills and preferences and all jobs and residences are exchangeable and offer the same utility to commuters, the minimum commute (T_{min}) -- that is necessitated by the fixed distribution of jobs and housing -- in a city or region occurs when commuters collectively coordinate their housing and job choices in a manner which minimises commuting cost; the result being that commuters will, on average, live to their closest possible workplace (Murphy, 2016; Zhou and Long, 2015). The extent to which the observed commute (T_{act}) exceeds T_{min} is considered to be EC (White, 1988). It is, of course, unrealistic for any city-region to have T_{act} equivalent to T_{min}; nevertheless, comparing T_{act} and T_{min} can be an important benchmark for measuring the balance/imbalance of jobs and housing in the region or the efficiency of commuting more generally. It is from this conceptual base that the EC framework has been extended to include four benchmarks including T_{min}, T_{act}, random (T_{rand}) and maximum (T_{max}) commutes as well as four indices – EC, capacity utilisation (C_u,), commuting economy (C_e), and normalised commuting economy (NC_e) (Kanaroglou et al., 2015). These benchmarks and indices are discussed in detail in the next section.

This study seeks to add to the conceptual development and extension of the EC framework. Specifically, it proposes a way to account for the carrying capacity of links and link congestion as an integral component of estimating the EC benchmarks. Further, we demonstrated through an empirical case study that overlooking road carrying capacity and traffic congestion can introduce bias and inaccuracies in EC indices which have the potential to negatively influence policy decisions. Drawing on the six-county Los Angeles region as the case study context, this paper demonstrates that ignoring carrying capacity and traffic congestion in assignment methods (such as the transportation problem of linear programming) leads to: (1) distorted estimates of EC indices and benchmarks; and (2) underestimates of the average minimum commute distance (T_{min}). In addition, it shows that when optimising the jobs-housing balance and corresponding commuting flows, decision-makers should ensure that popular and/or shortest-distance corridors/routes between locales where both jobs and housing units are in high density should be provisioned with sufficient carrying capacity. If the issue of carrying capacity is ignored, commuting efficiency will be substantially worse than is estimated using the traditional transportation problem of linear programming (TPLP) approach where the road carrying capacity is assumed to be unlimited.

Bearing the issues of road carrying capacity and traffic congestion in mind, this study makes two contributions to existing literature: First, it shows both theoretically and empirically that carrying capacity of alternative routes between worksites and residences should be considered in both jobs-housing balance studies and policies based on the EC framework, and; second, it quantitatively demonstrates (via a study of a region known to be notorious for traffic congestion) the extent of bias that can result in EC benchmarks and indices when the carrying capacity and congestion of a link or route are ignored. In this study, we focus on the carrying capacity impacts on the results emerging from EC studies. Thus, it is possible to reflect upon, and indeed reconsider, the results from existing studies within this context.

Related literature

The EC framework: an overview

Following Kanaroglou et al. (2015), four EC benchmarks and four EC indices that are derived from the benchmarks together make up the EC framework. Contemporary EC studies typically utilise the TPLP for the calculation of T_{min}. TPLP was not proposed by EC scholars rather it was first introduced into EC studies by White (1988).

Once T_{min} is calculated, it can be compared with T_{act} to determine the level of commuting that is excessive (EC). EC is a measure of the deviation of an urban area's observed average travel cost from a theoretical minimum average travel cost as follows:

$$\mathrm{EC} = \frac{T_{act} - T_{min}}{T_{act}} \qquad (1)$$

Within the EC framework, Horner (2002) determined that T_{max} is the upper limit of a city's commuting range and which can be used to measure jobs-housing imbalance within cities. T_{max} is calculated using the inverse of the objective function of the TPLP in (1) where the objective is to maximise rather than minimise the average commuting costs between residences and workplaces.

T_{max} also allows for an additional way to measure commuting efficiency -- C_u -- which provides a gauge of how much of a city's available commuting range has been consumed (Horner, 2002). If we consider $T_{max} - T_{min}$ as a finite scale determined by the distribution of jobs and housing, then the relative location of the observed average travel cost, T_{act}, on the scale indicates the amount of available travel resources being used by collective urban commuting. Specifically, C_u is calculated as:

$$C_u = \frac{T_{act} - T_{min}}{T_{max} - T_{min}} \qquad (2)$$

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

In further developments, Murphy and Killen (2011) argue that T_{rand} is a better benchmark of the upper boundary of commuting possibilities than T_{max} given that it is a more realistic gauge of commuting behaviour where commuting cost is considered irrelevant to the physical separation of jobs and housing. Similar to Murphy and Killen (2011), Yang and Ferreira (2008) also questioned the appropriateness of using T_{max} to measure the upper limit of a city's commuting range and to represent jobs-housing possibilities. Both studies recommended the use of proportionally matched commuting (PMC) as an alternative to T_{max} to measure the upper limit of a city's commuting range and represent jobs-housing possibilities. Kanaroglou et al. (2015) recently demonstrated that PMC is mathematically equivalent to T_{rand} and note that the two can be used interchangeably.

In a 2011 study, Murphy and Killen (2011) utilised T_{rand} to propose two additional measures of commuting efficiency: C_e and NC_e given by the following equations:

$$C_e = (1 - \frac{T_{act}}{T_{rand}})*100 \quad (3)$$

$$NC_e = \left(\frac{T_{rand} - T_{act}}{T_{rand} - T_{min}}\right)*100 \quad (4)$$

C_e gives the extent to which T_{act} is positive or negative T_{rand} i.e. the degree to which behaviour as expressed by T_{act} is becoming either more or less random. This represents the extent to which individuals are economising on commuting costs. NC_e represents the extent to which T_{act} is below T_{rand} relative to the theoretical extent to which this could happen as determined by the geography of land-use i.e. T_{min}. Thus, this is considered to be a normalised commuting economy indicator. These metrics collectively represent the key conceptual advancements in the development of the EC framework.

Conceptual development: traffic congestion and the EC framework

Several studies have comprehensively reviewed existing scholarship within the EC framework (see for example, Ma and Banister, 2006; Layman and Horner, 2010; Barr et al., 2010; Kanaroglou et al., 2015). However, there has been no study which has explicitly considered carrying capacity of alternative routes/links and resulting congestion within the EC framework.

Table 1: EC studies where carrying capacity/congestion issues are considered

Study	**Cost measurement**	**Traffic congestion considered?**		**Specific implications for this study**
		T_{act}	T_{min}	
White, 1998	Reported time	Yes	No	Represents the first study to employ the transportation problem method.
Small and Song, 1992	Peak-hour time and network distance based on local travel demand models	Yes	No	Comparable EC of L.A. for this study
Giuliano and Small, 1993				
Merriman, et al., 1995	Estimated time based on transit schedules plus reported time	Yes	No	Complexity in constructing a commuting time matrix
Scott et al., 1997	Peak-hour time based on local travel demand models	Yes	Yes	T_{min} measured by time can be shorter when traffic congestion is considered

Manning, 2003	Reported distance from census	Yes	No	How a worker classification influences EC
Ma and Banister, 2006b	Reported time and network-based distance	Yes	No	Qualitative and quantitative jobs-housing balance concepts and how they are related to EC
Horner, 2010	Variable travel times	Yes	Yes	T_{act} is more likely to be affected by variable travel times.
Murphy, 2009	Peak-hour time based and distance on local travel demand models	Yes	No	EC by mode of travel
Murphy and Killen, 2011				Monte Carlo method to derive T_{rand}; C_e; NC_e
Hu and Wang., 2015	Estimated distance and time based on surveys	Yes	No	Reporting errors in travel time from surveys; How to mitigate the miscalculation of EC

To avoid repeating reviews from other scholars, here we review and summarise existing studies with specific attention to whether or not they accounted for traffic congestion when estimating EC benchmarks (Table 1). An exhaustive search revealed 11 studies that indirectly dealt with congestion and one piece of work (Scott el al., 1997) that explicitly considered the impacts of traffic congestion on T_{min} and T_{act}. Scott and colleagues showed that when traffic congestion is considered for both T_{min} and T_{act} measured in terms of journey time, the EC for their study area (Hamilton, Canada) was 3 percent higher than if traffic congestion is considered for T_{act} alone. Their explanation for the difference was that (a) there were more short trips in the case of optimum assignment under T_{min} compared with T_{act}; and (b) the modal split in the case of T_{min} was different to that of T_{act}. However, the study did not consider how traffic congestion influences T_{min} as measured by travel distance, rather they solely considered travel time. While Horner's (2010) study does not directly focus on traffic congestion, it does consider travel time between origins and destinations as something uncertain and examines how changing travel times between specific origins and destinations affects the values of T_{act}, T_{min} and EC. He finds that T_{act} is more likely to be affected by uncertain travel times than T_{min}. However, the study does not directly model how available road capacity and traffic congestion might impact the route choice of commuters. Indeed, neither Scott et al. (1997) or Horner (2010) note that the optimum assignment pattern (T_{min}) determined using the TPLP, there will likely be some routes where the assigned flows far exceed the available carrying capacity along certain links, which forces some commuters to choose alternative routes. In reality, therefore, there may be routes congested under the optimum assignment pattern and in such cases commuters may choose routes that are longer in distance to reduce commuting time. In reality therefore, the average minimum distance of commuters may be longer than predicted using the TPLP.

To illustrate the carrying capacity/congestion issue further, Figure 1 provides a hypothetical example. It is well-known that congestion extends commuting time; but it can also extend commuting distance, which is not always evident to commuters. Consider Figure 1 as an illustration.

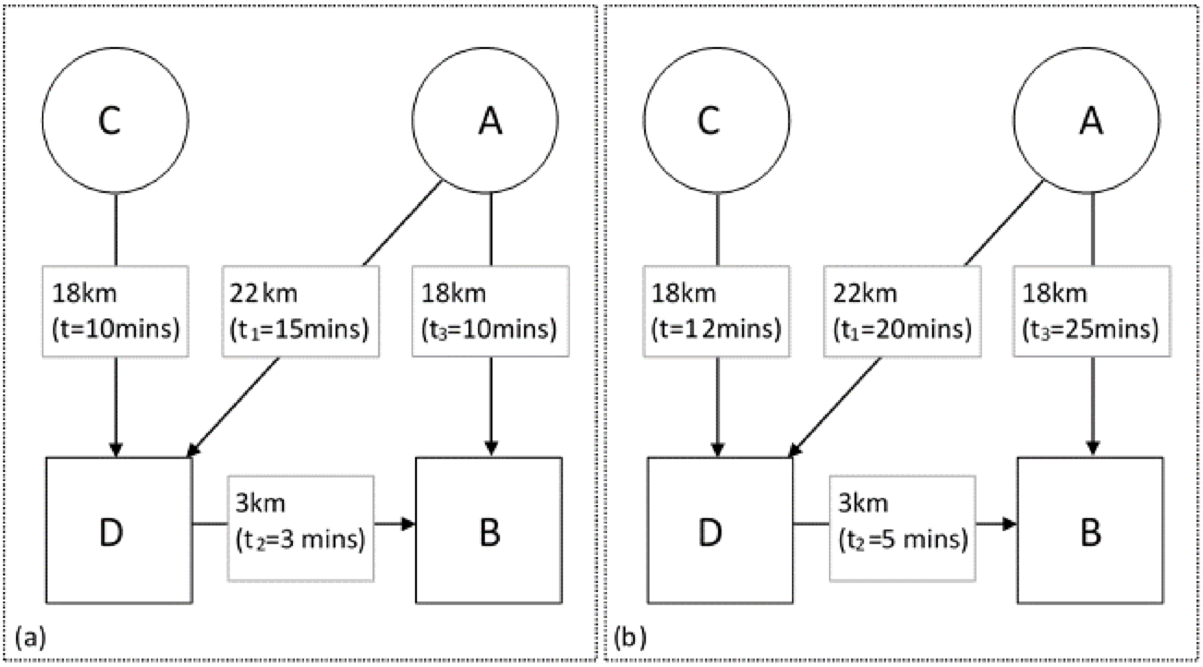

Figure 1. Schematic depicting the impact of traffic congestion on commuting distance

Figure 1(a) outlines a scenario where traffic congestion is absent and commuters from A and C travel to their respective workplaces D and B in ten minutes on the shortest paths (CD or AB); this equates to 18 kilometres. The route ADB is not considered by commuters originating at A given its longer commuting time. Figure 1(b) outlines a scenario where traffic congestion exists between A and B resulting from capacity constraints along the route. Under the assumption that rational commuters seek to minimise travel time, commuters from A would then use routes AB and ADB to travel to B; both routes have equal travel times (25 minutes). However, for commuters using the route ADB, their commute distance is now 25 kilometres instead of 18 kilometres. In other words, commuters would choose the shortest-time route instead of the short-distance route. Thus, even when all jobs and housing are near to one another, there may still be issues of insufficient road capacity along links connecting those origins and destinations.

From the foregoing illustration and a review of the relevant literature, it is evident that little research has been undertaken that considers the impacts of available carrying capacity and resulting traffic congestion on the route choice of commuters and the impact of these issues on the resulting EC benchmarks and indices. This motivates the current study within the Los Angeles region to demonstrate how these issues can be integrated into, and more broadly considered within EC studies.

Methodology

Study area and input data

Our study area is the six-country Los Angeles region (Figure 1). This particular study area has been studied quite extensively by other scholars interested in urban form and commuting for two core reasons. First, it is the second-largest metropolitan area in the US: as of 2006 the area covers a land area of 258,178 square kilometres accommodating 22.6 million residents.

Second, it is notorious for urban sprawl with a number of studies pointing to issues of jobs-housing imbalance, high EC levels, and environmental sustainability issues, including high energy use from private transport and high levels of air pollution (see for example Giuliano, 1991; Giuliano and Small, 1993; Small and Song, 1992; Giuliano et al., 2007, 2012).

The Southern California Association of Governments (SCAG), which is the local transportation planning agency, divides the LA region into 4,191 traffic analysis zones (TAZs), including 82 "virtual TAZs", which are important egresses or ingresses in the region such as airports, railyards, highway exits/entrances and seaports. Unlike actual TAZs, there are no internal trips within virtual TAZs as they are only considered as nodes. The 4,102 TAZs (not including the visual TAZs) have an average area of 62.78 square kilometres. Located in the core of the region (Los Angeles and Orange Counties), the TAZs are much smaller than those in the east, that include large expanses of desert. The core also contains over 80% of the region's jobs and residences (Figure 2). Indeed, Niedzielski et al. (2013) analysed the relationship between Modifiable Areal Unit Problem (MAUP) effects and indicators in the EC framework. Their analysis and others (see also Horner and Murray, 2002) shows that EC and CU indicators vary substantially with the size of the zonal unit being utilised for analysis. However, it is notable that values of C_e and NC_e display much less variability with the level of zonal aggregation (Niedzielski et al., 2013)

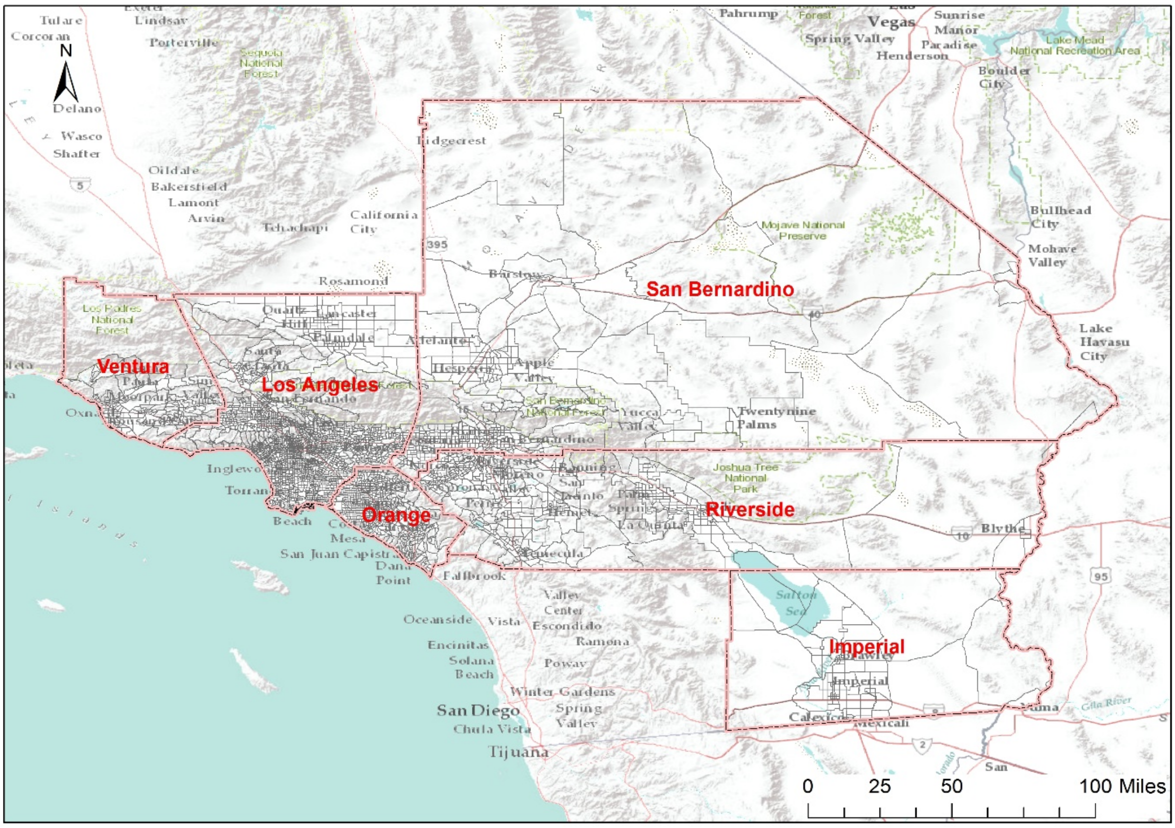

Figure 1: The six-county LA region

Figure 2: Spatial distribution of employment and residential density (per kilometre) by TAZ across the LA region

The data utilised for this study was taken from SCAG's 2003 travel demand model, which was validated and released in 2007 (SCAG, 2007). Note that the dataset is entirely automobile-based and does not consider rail-based commuting, which accounts for less than one percent of travel in the region (SCAG, 2007). The dataset contains the following information for the morning peak period (06:00 to 09:00):

- OD vehicle flows for solo drivers;
- OD vehicle flows for car poolers (including drivers);
- OD vehicle flows for light-duty truckers;
- OD vehicle flows for medium-duty truckers;
- OD vehicle flows for heavy-duty truckers;
- The GIS polyline files for the road network (comprising a total length of 32,726 centreline kilometres);
- The GIS polygon files for local TAZs.

Table 2 shows the descriptive statistics for each of the categories of vehicle flows contained within the SCAG dataset. Results reveal that the overwhelming majority of trips are associated with lone car drivers followed closely by car poolers. It is notable that only seven percent of the workers who drive alone have their workplace and residence in the same TAZ. Thus, the vast majority of workers must travel to another TAZ to reach their workplace destination. As was the case for other relevant studies (see for example, Small and Song, 1992; Giuliano and Small, 1993; Murphy, 2009; Murphy and Killen, 2011), this study treats all solo-drivers in the SCAG model as workers/commuters.

Table 2: Vehicle flows in the LA region (proportions in parentheses)

Type of commuter	**Total vehicle count**	**Vehicle count by area**	
		Within TAZs	**Between TAZs**
Solo drivers	5,041,306	354,462 (7)	4,686,844 (93)
Car poolers	2,112,983	255,031 (12)	1,857,952 (88)
Light-duty truckers	75,131	10,001 (13)	65,130 (87)
Medium-duty truckers	44,497	1,733 (4)	42,764 (96)
Heavy-duty truckers	54,485	1,191 (2)	53,294 (98)
Total	7,328,402	622,418 (8)	6,705,984 (92)

Other travellers (car poolers and various truckers) in the SCAG dataset were considered as background traffic: their total numbers do not change but each time the route choice of solo-drivers changes, their route choice also changes. Together, solo-drivers and other travellers interact and contribute to the final equilibrium of the road traffic, that is, the user equilibrium, system equilibrium, all-or-nothing assignment, multi-modal multi-class assignment or stochastic assignment that travel demand modellers often refer to (Caliper, 2015). In this study, we combine the user equilibrium and multi-modal multi-class assignment to best account for traffic congestion on the road network given that we have five classes of travellers' OD flow matrices, which were either not utilised or were unavailable in previous studies.

Calculation approaches for EC benchmarks

The first element of this study involves the calculation of T_{min}, T_{max}, T_{rand} and T_{act}. For T_{min}, the TPLP was used to determine the assignment of trips from origin to destination that minimised average commuting cost. The objective function and constraints of the TPLP are given by:

$$\text{Min: } Z = \frac{1}{N} \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} X_{ij} \qquad (5)$$

$$\text{s.t.} \quad \sum_{i=1}^{n} X_{ij} = D_j \qquad \forall j = 1, \dots, m \qquad (6)$$

$$\sum_{j=1}^{m} X_{ij} = O_i \qquad \forall i = 1, \dots, n \qquad (7)$$

$$X_{ij} \geq 0 \qquad \forall i, j \qquad (8)$$

where, m = number of origins; n = number of destinations; O_i = trips beginning at zone i; D_j = 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, and N = total number of trips. The objective function (2) minimises average transport costs. Constraint (3) ensures that trip demand at each destination zone is satisfied while constraint (4) limits the number of trips leaving each origin zone to the number of trips originating there. Constraint (5) restricts the decision variables, X_{ij}, to non-negative values. It should be noted that travel costs, c_{ij}, may be expressed in terms of any measure of zonal separation, for example travel distance, travel time or indeed a generalised cost measure. For this study, Euclidean distance (ED) and route network distance (ND) were both used as measures of zonal separation.

T_{max} was also determined using the TPLP where the objective function is the inverse of the minimisation problem discussed previously (5) and is given by:

$$\text{Max } Z = \frac{1}{N} \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} X_{ij} \qquad (9)$$

The approach for calculating T_{rand} utilised the Markov Chain Monte Carlo (MCMC) hit-and-run algorithm outlined by Murphy and Killen (2011, p.1261-62). It is given by the following:

$$N_T = N! \qquad (10)$$

$$\text{s.t.} \quad \sum_{i=1}^{n} X_{ij} = D_j \qquad \forall i = 1, \dots, n \qquad (11)$$

$$\sum_{j=1}^{m} Xij = Oi \qquad \forall j = 1,...,m \qquad (12)$$

$$Xij \geq 0 \qquad \forall i, j \qquad (13)$$

Where: N_T is the number of possible commuting configurations in a city; N! is the factorial of the total number of trips in the urban area. Constraints (11-13) are identical to those of the transportation problem and they limit commuting possibilities to those supported by the fixed distribution of jobs and residences. Because it takes days to complete a random traffic assignment for the LA region, only ten assignments were run to derive the values of T_{rand} quoted in the current paper. Finally, T_{act} was calculated from observed trip data and associated commuting costs.

Assignment and travel demand modelling

Travel demand modellers have developed and used different traffic assignment models to consider the impacts of traffic congestion on travel behaviour. These models include but are not limited to the following: the user equilibrium, system equilibrium, all-or-nothing assignment, multi-modal multi-class assignment, and stochastic assignment. The aforementioned models can now be implemented more conveniently in practice using specialist software packages such as TransCAD, EMME, PTV and Cube. Previous related studies use two different approaches to derive different EC benchmarks: the all-or-nothing assignment (i.e. the TPLP) and the user equilibrium multi-modal multi-class assignment. The former effectively assumes that road capacity is always sufficient and thus congestion is non-existent on the route network. In this study, we follow this approach to generate the base results where congestion is not considered. Next, the user equilibrium multi-modal multi-class assignment is used where congestion is considered because the approach allows considerations of road capacity and traffic congestion.

Based on Bechmann et al. (1956), the mathematical formulation of the user equilibrium approach is to find an optimal solution to the following problem:

$$\text{Min} \qquad \sum_{a \in A} \int_0^{f_a} l_a(x)dx$$

$$\text{s.t.} \qquad \sum_{p \ni a} f_P = f_a \qquad \text{for all } a \in A, \qquad (14)$$

$$\sum_{P \in P_k} f_p = d_k \qquad \text{for all } k \in K \qquad (15)$$

$$f_a \leqslant C_a \qquad \text{for all } a \in A \qquad (16)$$

$$f_P \geqslant 0 \qquad \text{for all } p \in P \qquad (17),$$

Where:

a is a link of the set of links A between an origin and a destination(OD);

f_a is the traffic passing a;

$l_a(x)$ is the cost function, which can be used to calculate the cost of travel for a when there is x amount of traffic;

f_p is the amount of traffic using path (route) p, which contains a;

d_k is the rate of traffic from a point to another point in OD pairs of K;

C_a is the feasible flow capacity for a;

P are all possible paths (routes) between the origin-destination pairs of K.

Usually, the user equilibrium approach only considers one class of vehicle and thus applying it requires the conversion of all vehicles into a common unit of analysis such as passenger car equivalent (Caliper, 2015).

All-or-nothing assignment does not require condition (16) to be met and assumes that all traffic between an OD pair in K is always assigned to the least cost route and, crucially, that that can accommodate unlimited traffic volumes. Multi-modal multi-class assignment has a similar mathematical formalisation to the user equilibrium approach but adopts a $l_a(x)$ that is more complex and that accounts for interactions of different classes of vehicles on roads and does not assume that routes can accommodate unlimited volumes of traffic. It takes the following form:

$$l^m = \sum_{b \in B} \{VOT^m VDF(T_b, C_b, \sum_m PEC_m X_{b,} \ldots) + FT_b^m\} + \sum_{i \in M^m} MT_m^i$$
(18),

Where:

l^m is the generalized cost between origin and destination for mode m;

m is mode of travel;

b is a link in a network;

B is the set of links on the shortest path between origin and destination for mode m;

VOT^m is the value of time of mode m;

T_b is free flow travel time on b;

C_b is the capacity on b;

FT_b^m is fixed toll on link b for mode m;

M^m is set of nodes based tool section between origin and destination for mode m;

MT_m^i is toll value for section i, mode m;

VDF is the volume delay function;

X_b is total volume on link b, $\sum_m PEC_m X_b$;

X_b^m is flow of type m on link b;

PCE^m is passenger car equivalent for mode m (Caliper, 2015).

Mathematically, there are different ways to find the optimal solution to a multi-modal multi-class assignment problem. All-or-nothing assignment and the user equilibrium are simply two such options. In this study, we use both options.

Specifically, we used the Los Angeles data described previously as the input data. For the parameters of the volume delay function, which is required when implementing different traffic assignment models, we adopted the Bureau of Public Roads function as follows:

$S(X_b)=T_b*(1+0.15*(X_b/C_b)^4)$ (19)

Where:

$S(X_b)$ is the average travel time for a traveller/vehicle on link b;

All other notation is the same as that outlined in equation (18).

To convert different vehicle classes into PCEs, we used the weighted average conversion factors specified in SCAG (2007) (Table 3).

Table 3: Converting different vehicle classes into PCEs

Vehicle class	Car/SUV	Light-duty trucks	Medium-duty trucks	Heavy-duty trucks
Conversion factors	1.0	2.9	3.6	4.8

Because trips inside TAZs are always treated separately in traffic assignment processes, this study only considers commuting trips between TAZs when calculating EC benchmarks which is consistent with previous studies such as Small and Song (1992) and Giuliano and Small (1993) for the same region. Thus, this differs from some existing studies, e.g., Frost et al., (1998), which considers the distance of all trips within a TAZ as:

$D_i=\sqrt{A_i}/\pi$ (20)

Where:

D_i is the distance for all trips within a TAZ i;

A_i is the area of the TAZ i.

The formula assumes that the intrazonal commuting distance travelled by individuals assigned to a workplace within in the same zone as they live is equivalent to the radius of the circle which approximates the area of a TAZ. However, there has not been a consistent way to estimate the time duration of commuting trips inside TAZs. Typically, these trips measured by time are not considered when calculating T_{act}, T_{min}, T_{max}, and T_{rand}. In travel demand modelling, commuting trips inside TAZs are usually assumed to occur on "virtual links" between the centroid of the TAZ and several egresses/ingresses of the lowest-level of links/roads of the local transport network in question. These "virtual links" are called "centroid connectors" by travel demand modellers. Most of the time these "centroid connectors" are so short that travel times/distances on them can be ignored and overall traffic assignment results' (including EC benchmarks) are not significantly affected. In the case of Los Angeles, for instance, if Equation (20) is adopted to derive the average of D_i for all TAZs, the average is only 1.6 miles while the average distance between any two TAZs is 37.5 miles.

Results

EC in LA

Table 4 presents estimates for the four EC benchmarks (T_{min}, T_{act}, T_{rand} and T_{max}) as well as the four indices (EC, C_u, C_e and NC_e) using the all-or-nothing assignment approach alongside previous EC studies (see Small and Song 1992 and Giuliano and Small, 1993) in terms of travel distance and time. Comparing the results of this study with others for the same region reveals a number of interesting trends. Table 4 also outlines the results of the modelling exercise that explore the impact of road carrying capacity and traffic congestion on EC benchmarks and indices.

First, the values of T_{min}, T_{act} and EC have changed over time. It can be seen that our results report a higher T_{act} for LA workers relative to that revealed in previous studies. Our study calculated a commuting average time of 36.0 minutes compared to 22.1 and 23.0 minutes for the Small and Song (1993) and Giuliano and Small (1993) studies respectively – equating to a value between 57 to 63 percent higher. In terms of distance, T_{act} is more than double (20.1 miles) than that reported in Small and Song's (1992) study (10.0 miles) in the 1980s. This implies that behavioural characteristics of LA commuters have evolved in such a way that they have moved more towards choosing origins and destinations in a manner which trades off distance for (in relative terms) improvements in commuting time.

Turning to the results for T_{min}, they show that the average commuting time is also higher than those reported in previous studies (Small and Song, 1992; Giuliano and Small, 1993), however the difference is only marginally. Whereas the difference for T_{act} was in the region of 57 to 63 percent, the difference for T_{min} is only 6 to 17 percent. For distance, the T_{min} value for our study is substantially lower at 0.9 miles compared to 3.1 miles , that was reported by Small and Song (1992) . Overall, our results for T_{min} imply substantial declines in travel distance but increases in travel time values for LA commuters since the 1990s. This suggests that jobs and housing arrangements have become more physically juxtaposed than in the 1990s; yet, has not led to proportional declines in the value of T_{min} or T_{act} (for time or distance). In this regard, the values of T_{rand}, measured by time, are particularly noteworthy because they suggest that actual commuting behaviour is, on average, only 2 minutes away (or 7 percent as indicated by the NC_e values) from what would be expected if commuter's behaved randomly. Put another way, in terms of journey time the relative results for T_{rand} and T_{act} suggest that commuters are collectively quite close to behaving in a manner where the cost of separation between origins and destinations is irrelevant to them. However, when the input cost is distance, this trend does not hold, indicating that commuting behaviour is much closer to what would be expected under random conditions for time rather than distance. Using distance, commuters are 36 percent away from behaving randomly indicating that distance minimisation is more important for LA commuters than time when choosing residences and workplaces. In this regard, the results for C_e in LA are by and broadly comparable to those for Dublin, Ireland (Murphy and Killen, 2011) and similar to those found by Schleith et al (2016) for Columbus, Milwaukee, and Portland in the United States.

In sum, the foregoing results suggest that the jobs-housing balance hypothesis does not hold in the case of LA – reducing the average distance between jobs and housing does not lead to expected benefits in observed commuting time/distance efficiency. In theory, workers in LA on average have more jobs in proximity as time progresses while in reality they are either unable to find a job that matches their expectations or skills nearby or willing to make trade-off between job proximity and other amenities such as better schools and safer neighbourhoods. On the one hand, this is somewhat consistent with Giuliano (1991); however, on the other, it in part highlights the importance of the match between jobs and workers' expectations or skills (Benner and Karner, 2016; Stoker and Ewing, 2014). Of course, in a metropolis like LA, labour division has become so sophisticated, the sprawl of certain jobs is so severe, and housing prices are so high in certain areas that it may also be difficult for many to find an appropriate job near his/her residence, or vice versa.

Table 4: EC estimates for LA

	T_{min}	T_{act}	T_{rand}	T_{max}	EC (%)	C_u (%)	C_e (%)	NC_e (%)
Traffic congestion not considered (TPLP, i.e., the all-or-nothing assignment)								
Results for this study								
Time (mins)	8.9	36.0	38.0	92.0	75	33	5	7
Trips considered*	1,942,191	4,686,844	NA	5,039,202	NA	NA	NA	NA
Dist. (miles)	0.9	20.1	30.9	47.6	95	40	35	36
Trips considered*	4,992,857	4,686,844	NA	5,039,202	NA	NA	NA	NA
*Small and Song's (1992) Study***								
Time (mins)	7.6	22.1	NA	NA	66	NA	NA	NA
Dist. (miles)	3.1	10. 0	NA	NA	69	NA	NA	NA
*Giuliano and Small's (1993) Study****								
Time (mins)	8.4	23.0	NA	NA	63	NA	NA	NA
Dist. (miles)	NA	NA	NA	NA	66	NA	NA	NA
Traffic congestion considered (Initial T_{min} /T_{max} flow distribution based on distance), this study								
Time (mins)	2.2	81.6	104.3	225.2	97	55	22	22
Trips considered*	4,992,857	4,686,844	NA	5,039,202				
Dist. (miles)	1.0	25.2	46.9	75.5	96	32	46	47
Trips considered*	4,992,857	4,686,844	NA	5,039,202				
Traffic congestion considered (Initial T_{min} /T_{max} flow distribution based on time), this study								
Time (mins)	4.3	81.6	104.3	216.3	94	36	22	23
Trips considered*	1,942,191	4,686,844	NA	5,039,202				
Dist. (miles)	2.4	25.2	46.9	75.4	90	31	46	49
Trips considered*	1,942,191	4,686,844	NA	5,039,202				

*As we have no way to consider traffic congestion on centroid connectors, which are "virtual links" within TAZs, only trips between TAZs are considered. But even if those trips on centroid connectors were considered, they should make T_{min}, T_{max} T_{rand} and T_{act} consistently smaller as centroid connectors are usually shorter than most of the travel distances between any two TAZs in the local transport network.

** Their study area had 706 zones covering 1,289 square miles.

*** Their study area had 1,146 zones and has 10.6 million residents and 4.6 million jobs; they did not provide information on the land area.

Impact of traffic congestion on commuting efficiency

Table 4 shows estimates of the four EC benchmarks (T_{min}, $T_{act,}$ $T_{rand,}$ and T_{max}) when traffic congestion (or more specifically, different road segment's carrying capacity in relation to traffic volume) is considered in the assignment model set against the results where congestion is not considered. To derive these estimates, the model allows for two possible inputs regarding the optimised flow distribution in the final equilibrium of traffic flows for the region (i.e. for deriving T_{min}): one is based on travel distance while the other is based on travel time (i.e. where the travel cost matrix (C_{ij}) specified in Equation (5) is based on distance or time). The matrix based on travel time considers traffic congestion along the route network while the matrix for distance does not. The hypothesis was that traditional EC studies where distance is typically minimised overestimate the extent of distance-based commuting efficiency because they do not consider the fact that individual's trade off time for distance when making commuting decisions. In other words, the optimum allocation of individuals from home to work is likely to be associated with longer distance minimum commutes on average if congestion is considered in the assignment process.

The results are interesting because they reveal the impact of congestion in the optimum assignment process. Table 4 shows that when distance is minimised T_{min} is 1.0 miles on average whereas when time is minimised T_{min} is 2.4 miles – more than twice that when distance is considered in isolation. This suggests that if we overlook road capacity and/or congestion, we could mistakenly assume that workers will always take the shortest path between their home and workplace. Our modelling results indicate that when some routes become congested, it can lead to significant increases in the average distance travelled both in terms of T_{min}/T_{max} but also in terms of observed commuting distance (T_{act}). In relation to the latter, it can be seen that T_{act} is much higher when congestion is considered indicating that congestion also leads to significant increases in observed travel distance and time costs. Specifically, T_{act} for distance increased from 20.1 to 25.2 miles (25 percent) while T_{act} for time increased from 36.0 to 81.6 minutes (143 per cent) when congestion is considered. This elucidates the fact that observed commuting time is much more likely to suffer during times of route congestion and as such commuters tend to place a premium on the minimisation of time over distance during periods of congestion. Bearing this in mind, the foregoing results imply that individuals trade off distance for improvements in time when making route choices between home and work.

The results also suggest that traditional EC studies that rely solely on distance as the inter-zonal measure of commuting costs are likely significantly underestimating the true extent of T_{min} and T_{max}. Moreover, the results for EC in Table 4 show that when congestion is considered, it is considerably higher (between 94 and 97 percent) when time is used as the inter-zonal separation measure versus when no congestion is considered (75 percent). However, it is also notable that the EC measures for distance are similar when congestion is considered and not.

Turning next to the values for C_u, C_e and NC_e some interesting results also emerge. LA's commuting efficiency (as measured by C_u) and commuting economy (indicated by C_e and NC_e) when congestion is not considered is considerably higher if the input trip cost is measured by distance rather than time. In relation to C_u this suggests that commuters are consuming a greater proportion of available commuting distance capacity than time capacity while for C_e and NC_e it suggests that LA's commuters are much closer to behaving as random commuters when commuting distance is utilised as the input versus travel time. It is also notable that when congestion is considered and where the input cost is measured by distance, commuters are much further away from behaving as random commuters (see values for C_e and NC_e in Table 5) when compared with the C_e and NC_e results presented thereby implying more efficient behaviour than we might have concluded by assessing only the values for when congestion is not considered. Ultimately, the results show that considering congestion in the assignment process allows us to draw quite different behavioural interpretations regarding the nature of the efficiency of a city regions collective commuting efficiency than if congestion was not considered.

Discussion and conclusions

The EC framework has evolved considerably over the last three decades to incorporate a range of new benchmarks, indices, disaggregated analysis by type of commuter, as well as application to specific local contexts. However, previous studies have rarely considered the impacts of road carrying capacity and traffic congestion on EC benchmarks and indices. As Higgins (2017) notes, congestion does increase impedance and unpredictability of travel times and, as a result, it can reduce an individual's perceived control over the duration of their commute. Using empirical data and modelling approaches, this study demonstrates the extent to which EC benchmarks and indices change when traffic congestion is taken into account in the assignment modelling process. The empirical results from LA indicate that traditional EC studies that rely solely on distance as the inter-zonal measure of commuting cost are likely to underestimate, in particular, the true extent of $T_{min,}$ T_{act}, and T_{max} as well as the indices that rely on those values. The traditional calculation of T_{min} and T_{max} using the TPLP is based on an all-or-nothing assignment approach that assumes routes in the optimum assignment have unlimited carrying capacity. Quite clearly, this approach is unrealistic and our study utilised a modelling approach that accounts for changes in behaviour when a route reaches its carrying capacity and becomes congested. In overall terms, the results here suggest that future EC studies should take account of the carrying capacity/congestion in determining EC benchmarks and indices.

In the case of LA, the results generated in this study suggest that LA's commuting efficiency has changed since the 1990s in line with its land use arrangements. The T_{min} results suggest that residences and workplaces have become more inter-mixed in LA over time. While this has resulted in a greater jobs-housing balance, it has not led to reductions in average observed commuting distances (T_{act}). This has notable implications for policies aimed at improving jobs-housing balances in cities because it suggests that greater jobs-housing balance alone does not lead to reductions in the average length of commutes. It seems that improvements in jobs-housing balance has resulted in commuters trading off longer distance commutes for shorter time commutes which is the antithesis to improving environmental sustainability in the region. Our modelling results indicate that: (1) workers may use an alternative route in order to avoid recurrent traffic congestion which is considerably longer (in distance) than the shortest path available; (2) while jobs-housing balance policies should be encouraged, their desired benefits in terms of reducing the overall distance of commuting or vehicle miles travelled (VMT) can only be achieved when there is no traffic congestion that forces commuters to take a longer route to minimise commute time rather than commute distance. Thus, in addition to achieving a balance between housing affordability, earnings and mixed land use (see Benner and Karner, 2016; Stoker and Ewing, 2014), our study implies that high-quality connections (preferably via public transport) between jobs and housing along with sufficient carrying capacity of the related infrastructure are crucial ingredients for cities and regions if they are to harvest the full benefits of jobs-housing balance policies for reducing commuting distance and VMT.

It is important to note the limitations of the current study and, related, the avenues it presents for future research. First, when considering the potential route choice of commuters, our model assumes that travel time (a proxy of congestion) is the only determinant for decision-making. It does not account for a range of other factors that might impact upon route choice decision-making such as available modal choice, access to real time travel information concerning alternative routes, trip-chaining needs, and indifference to commute time because of path-dependence or daily habits among many other factors as well as MAUP effects. Commuters, for instance, may stick to a familiar route so long as the perceived level of congestion and/or additional commute time does not exceed a certain threshold. In this regard, the introduction of a stochastic assignment approach may be useful for considering this dimension in future work. Second, is the issue of modal shift among commuters. Because of the lack of stated preference data about commuters' mode choice, this study assumes that commuters who drive would always drive regardless of the level of congestion or availability and quality of alternative modes. This may be true in the context of the LA region, where driving is the preferred mode of travel and often the most feasible/convenient for the majority of commuters. However, in other situational contexts where public transit services are more competitive, modal shift is likely to play a more important role and this should be incorporated into the modelling framework. It is likely that the modelling results would have utility for better evidence-based decisions that point towards how improvements to public transit and/or adjustments to land uses could enhance public transit uptake. In other words, this study's practical relevance could be enhanced by considering modal shift in future work.

References

Barr, S, Fraszczyk, A and Mulley, C (2010) Excess travelling—what does it mean? New definition and a case study of excess commuters in Tyne and Wear, UK. European Transport Research Review 2 (2): 69–83.

Beckmann, M, McCuire, CB and Winsten, CB (1956) Studies in the economics of transportation. New Haven, CT: Yale University Press.

Benner, C and Karner, A (2016) Low-wage jobs-housing fit: identifying locations of affordable housing shortages, Urban Geography, 37(6): 883-903

Caliper (2015) Travel demand modelling with TransCAD 7.0. Newton, MA: Caliper Cooperation.

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

Giuliano, G, Reddearn, C, Agarwal, A and He, S (2012) Network accessibility and employment centres. Urban Studies 49(1): 77-95.

Giuliano, G (1991) Is jobs-housing balance a transportation issue? Transportation Research Record 1935: 305-312.

Giuliano, G, Reddearn, C, Agarwal, A, Li, C and Zhuang, D (2007) Employment concentrations in Los Angeles, 1980-2000. Environment and Planning A 39: 2935-2957

Giuliano, G and Small, K (1993) Is the journey to work explained by urban structure? Urban Studies 30(9): 1485–1500.

Hamilton, BW (1982) Wasteful commuting. Journal of Political Economy 90 (5): 1035–1051.

Horner, MW (2010) Exploring the sensitivity of jobs-housing statistics to imperfect travel time information. Environment and Planning B: Planning and Design 37(2): 367-375.

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

Horner, MW (2007) ‘Optimal’ accessibility landscapes? Development of a new methodology for simulating and assessing jobs–housing relationships in urban Regions. Urban Studies 45(8): 1583-1602.

Horner, MW and Mefford, JN (2007) Investigating urban spatial mismatch using jobs–housing indicators to model home-work separation. Environment and Planning A 39: 1420–1440

Horner, MW and Murray, AT (2002) Excess Commuting and the Modifiable Areal Unit Problem, Urban Studies, 39, 131-139.

Higgins, CD, Sweet, MN and Kanaroglou, PS (2017) All minutes are not equal: travel time and the effects of congestion on commute satisfaction in Canadian cities. Transportation: 1-20.

Kanaroglou, PS, Higgins, CD and Chowdhury, TA (2015) Excess commuting: A critical review and comparative analysis of concepts, indices, and policy implications. Journal of Transport Geography 44: 13-23.

Hu, Y and Wang, F (2015) Decomposing excess commuting: A Monte Carlo simulation approach. Journal of Transport Geography 44: 43-52.

Ma, KR and Banister, D (2006a) Excess commuting: A critical review. Transport Review 26 (6):749–767.

Ma, KR and Banister, D (2006b) Extended excess commuting: A measure of the jobs-housing imbalance in Seoul. Urban Studies 43(11): 2099-2113.

Merriman, D, Ohkawara, T and Suzuki, T (1995) Excess commuting in the Tokyo Metropolitan Area: Measurement and policy simulations. Urban Studies 32(1): 69-85

Manning, A (2003) The real thin theory: Monopsony in modern labour markets, Labour Economics 10: 105–131.

Murphy, E (2016) Excess commuting and commuting economy: peak and off-peak variation in travel efficiency measures, In: Cohen, S. and Yannis, G (eds). Traffic Management. New Jersey, USA: Wiley. , pp.251-266.

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

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

Niedzielski, MA, Horner, MW, Xiao, N (2013) Analyzing scale independence in jobs–housing and commute efficiency metrics, Transportation Research A, 58, 129–143.

Schleith, D, Widener, M and Kim, C (2016) An examination of the jobs-housing balance of different categories of workers across 26 metropolitan regions. Journal of Transport Geography, 57: 145-160.

Scott, DM, Kanaroglou, PS and Anderson, WP (1997) Impacts of commuting efficiency oncongestion and emissions: Case of the Hamilton CMA, Canada. Transportation Research Part D 2: 245–257.

Small, KA and Song, S (1992) Wasteful commuting: A resolution. Journal of Political Economics 100(4): 888-898.

Southern California Association of Governments (SCAG) (2007) The state of the region 2007. Los Angeles: SCAG.

Stoker, P and Ewing, R (2014) Job–worker balance and income match in the United States, Housing Policy Debate 24(2): 485-497

White, MJ (1988) Urban commuting journeys are not wasteful. Journal of Political Economics 96(5):1097–1110.

Yang, J and Ferreira, JR (2008) Choices versus choice sets: A commuting spectrum method for representing job-housing possibilities. Environment and Planning B 35(2): 364–378.

Zhou, J and Long, Y (2016) Losers and Pareto Optimality in optimizing commuting patterns. Urban Studies 53(12): 2511–2529.