SIMPLIFIED MODEL OF LOCAL TRANSIT SERVICES

Giovanni Circella
Urban Land Use and Transportation Center
University of California, Davis
One Shields Avenue
Davis, CA 95616, USA
Phone: 1 530 752 1331
Fax: 1 530 752 6572
e-mail: gcircella@ucdavis.edu

Kevin John Stefan
HBA Specto, Inc
Calgary, T2P 3P4, AB, Canada
Phone 1 403 232 1060
Fax: 1 403 232 1060
e-mail: kjs@hbaspecto.com

John Douglas Hunt
Dept. of Civil Engineering
University of Calgary
2500 University Drive
NW Calgary, AB, Canada
Phone: 1 403 220 8793
Fax: 1 403 286 7206
e-mail: jdhunt@ucalgary.ca

Alan Thomas Brownlee
HBA Specto, Inc
Calgary, T2P 3P4, AB, Canada
Phone 1 403 232 1060
Fax: 1 403 232 1060
e-mail: atb@hbaspecto.com

Michael McCoy
Urban Land Use and Transportation Center
University of California, Davis
One Shields Avenue
Davis, CA 95616, USA
Phone: 1 530 754 9171
Fax: 1 530 752 3350
e-mail: mcmccoy@ucdavis.edu

TRB 90th Annual Meeting, January 23-27, 2011
Submission date: August 1, 2010
Revised version submitted on November 15, 2010
Word count: 7,404 words (249 words abstract) + 6 figures or tables
8,904 words total

Abstract

The California Statewide Travel Demand Modeling (CSTDM) Framework is a comprehensive model system designed and developed for use in transportation policy analysis and forecasting. It includes representation of all major components of both long and short distance transportation covering the entire state. A novel hybrid system is used to represent the full range of rail and bus transit services that are available. Rail and fixed busway services – including all long-distance rail, commuter rail, LRT and busway services – are represented in the standard manner, using explicit node and link networks, and the relevant in-vehicle and out-of-vehicle service characteristics for journeys are determined as standard skims of these networks. On-street bus services are not represented using explicit networks; rather, the relevant in-vehicle and out-of-vehicle service characteristics for journeys are determined using functions of other transportation network variables, land use descriptors and relevant policy indicators. These functions are simplified econometric models that have been estimated using observations of transit service obtained from Transit Data Feeds available on the Google platform. The network and simplified components are integrated in order to allow transit paths with both rail and on-street bus components to be considered by the various travel choice models included in the Framework. This hybrid system provides a suitable representation of transit for an area of such size, facilitating consideration of transit service policies, while also obviating the need for extensive transit coding, which would be a daunting task for such an area. As such, it is a very practical modeling approach.

Keywords: Transit Systems, Statewide Travel Demand Model, In-Vehicle Time, Waiting Time, Urban Density, Google Transit Data Feeds

1. Introduction

Public transportation is an important component of the transportation system. It offers affordable scheduled services and a valid alternative to driving for all classes of users, including those that, for economic conditions or physical inabilities, do not have access to private vehicles. The development of efficient public transportation services is also an important tool to promote accessibility of regions and cities and to promote the economic development and easy access to workplaces and central business districts.

The realistic representation of transit services is an important task in transportation modeling, and is required for the correct estimation of travelers' choice behavior and transit demand. This task may represent a very demanding activity, in particular in those areas that, for vast geographic extension or density of the network, present a complex transit network and a large number of transit operators and lines.

This paper discusses the development of a simplified representation of transit services in the California Statewide Travel Demand Modeling (CSTDM) Framework. The proposed methodology implies the use of a hybrid approach to model the two main components of the transit system: railways and buses. In the proposed approach, all rail-based services (including long-distance intercity railways, commuter railways, subways and light-rail systems) are explicitly coded in the public transportation network. However, bus services, which account for the vast majority of local transit services, are represented through the use of a simplified methodology. This simplified approach reduces the number of input variables used in the model, and the explicit coding activities required for the representation of bus transit services. The methodology is based on the estimation of econometric models that express local transit attributes (*in-vehicle* and *out-of-vehicle* times) as functions of other transportation and land use variables.

The methodology is based on the theoretical assumption that local transit attributes are influenced by other characteristics of transportation and land use. Therefore, it is possible to establish appropriate functional forms that express these attributes as functions of other transportation variables, as average speed and congested travel times on the network, and of the land use characteristics of the neighborhoods that are served by transit services.

The estimation of the econometric models for the local transit attributes is based on observed data for transit travel time collected from internet sources (Transit Data Feeds from the Google platform). Additional data on the land use and the transportation system are obtained from other available sources to create the required datasets for the model estimation. The methodology is developed for different time periods to account for the variability of service during the day.

The remainder of the paper is organized as follows. Section 2 presents the theoretical background of the research, and discusses the relationships between the development of transit systems and the urban form and other sociodemographic variables. Section 3 presents the California Statewide Travel Demand Modeling (CSTDM) Framework and the objectives that have supported the development of the simplified methodology for the representation of local transit. The simplified methodology for the representation of local transit services is the object of Section 4. Section 5 presents the estimation of the final model for local transit. Considerations on the use of the proposed approach in the CSTDM Framework are the object of Section 6. Finally, Section 7 presents the conclusions on the use of the proposed simplified approach in a statewide model.

2. The Characteristics of the Transit System and the Built Environment

Several studies have investigated the relationships among the characteristics of transit systems, land use features and travel behavior. Significant experience is available in literature on the analysis of the relationships among the demand for transit services and the characteristics of the land use and the built environment. Urban density, in particular, is an important determinant for transit use (Cervero and Kockelman, 1997; Badoe and Miller, 2000): neighborhoods with higher density are usually associated with higher use of transit, although the increase in demand for transit is sometimes limited (Cervero, 1994).

From the comparison of the transit demand at international level, several variables beside land use characteristics are found to affect ridership. Transit demand is particularly affected by the amount of transit services provided (*transit supply*), the area of coverage of transit services and the fare levels (Kain and Liu, 1999; Taylor *et al*., 2009; Litman, 2007). Many other characteristics of the transportation network and of the local geographic context are significant, such as auto ownership (Cervero, 2007; Paulley *et al*., 2006) and public transportation policies. Gas price has traditionally not been considered determinant in the definition of transportation (and, specifically, of transit) demand, which is usually considered inelastic (Wang and Skinner, 1984). Recent studies, however, have found an increasing role of gas price in affecting mode share and the demand for transit services, in part as a response to the sharp increase in gas price in the years 2007-2008, which might have contributed to modify travelers' behavior towards the use of transit (Lane, 2009; Maley and Weinberger, 2009).

In the majority of most developed countries, transit services have lost significant ridership during the last few decades. This phenomenon is the result of many contemporary changes in the society, in the priorities of policies in planning, as well as in consumers' habits and attitudes (Dittmar *et al*., 2004; Crawford, 2000). In the US, public transportation has not received enough support through policies and adequate funding. In the same years, cities were quickly reshaped in dependence of an extensive use of cars and other private vehicles. However, public transportation still plays an important role in higher density and more historical cities (Dunphy and Fisher, 1996). Other urban areas that significantly expanded in the last 40 years generally present lower-density and more sparse urban form, with high separation of land uses, which do not incentivize the use of public transportation. From the comparison with other countries, differences in gas pricing, policies and funding for transit, and urban density account for the large differences in transit patronage observed, for instance, between the US and Canada (Schimek, 1996) and the U.K.(Giuliano and Dargay, 2005).

The distance from the transit stations (*accessibility* to transit services) is an important determinant of transit patronage (Cervero, 1994). This confirms the important role of urban density: usually, better transportation services are found in more compact urban areas. In particular, the number of transit users is affected by the number of employees and workplaces in the central business district (CBD), more than by residential density (Hendrickson, 1985). Following the study from Newman and Kenworthy (1989), many studies have focused on the relationships among land use variables and transportation demand. Kenworthy and Laube (1996) proved that the mode share for transit significantly increases with an increase in the urban density and with the population of the urban area. The results were confirmed both in the most advanced and in developing countries.

The experience in literature on the relationships between transit supply and land use and sociodemographic variables is more limited. Few studies have discussed the topic, and limited findings are available: for instance, in an international comparison of the public transportation in 45 European cities, Albalate and Bel (2010) found that better transit services are usually associated with higher GDP. Moreover, transit supply is usually better in National Capitals. Additional regional patterns are observed, and better services are usually found in more historical regions with higher density. The role of subsidies and policies to support public transportation is also found to be important, although the efficiency of the subsidies highly depends on the size of the local transit network and on the way the subsidies are provided (Karlaftis and McCarthy, 1998). In addition, better transit supply is usually associated with a better integration of services with other transportation solutions, and with its coordination with pedestrian and bike networks (Buehler, 2009).

Despite the reported experiences, still additional investigation is needed in this field. The topic is of particular interest for transportation modelers, to support the estimation of functional relationships that link the provision of transit services to more easily measurable parameters of the transportation system and land use patterns. A similar approach is followed in the development of the simplified methodology for the representation of local transit in the California Statewide Travel Demand Modeling (CSTDM) Framework.

3. Public Transportation in the California Statewide Travel Demand Modeling Framework

The California Statewide Travel Demand Modeling (CSTDM) Framework is a comprehensive modeling framework designed for forecasting travel demand in the State of California. The program was launched in 2009 to support transportation planning and the evaluation of policy packages for transportation. The CSTDM Framework includes all major components of long and short distance travel demand. It adopts a tour-based microsimulation approach and is based on the analysis of 5191 Travel Analysis Zones (TAZ). The CSTDM includes five main models: the short distance personal travel model, the short distance commercial vehicle model, the long distance personal travel model, the long distance commercial vehicle model and the external trip model. All relevant means of transportation for both long distance and short distance trips throughout the State are included in the model: Single Occupancy Vehicle (SOV), High Occupancy Vehicle with 2 passengers (HOV2), High Occupancy Vehicle with 3 or more passengers (HOV3), Public Transportation (Bus and Rail), Airlines (for long distance trips), School Bus, Bike and Pedestrian. The model simulates travel demand for all trip purposes in the average weekday during the regular work and school season. Four times of the day are explicitly simulated in the model: AM Peak (from 6:00AM to 10:00AM), Midday (10:00AM to 3:00PM), PM Peak (from 3:00PM to 7:00PM) and Off-Peak (rest of the day).

The road network in the CSTDM Framework is coded in the Citilabs CUBE software package. The road network was developed for the years 2000 (calibration scenario) and 2008 (validation scenario), and it includes all road links that are relevant for a Statewide travel demand model. Additional information is included regarding the location of HOV lanes and dedicated ramps, bridges and tolls, and the access to all major transit and airport terminals.

Public transportation is represented by the local transit (rail and bus) services, the long-distance intercity railways and the air network. Different approaches were developed for the representation of the public transportation networks, depending on the different needs and the level of relevance of the public transportation services for either one or more of the five component models. The air network, which is only used in the long distance personal travel model, was coded explicitly through the definition of the airports that provide intra-state commercial air services, and the characteristics of the services offered on each route (travel time, headways, average fares, and reliability).

All railway services were explicitly coded in the public transportation network. The rail network is used in both the short distance and the long distance personal travel model. It accounts for a limited number of intercity and commuter railways, light rail and subway systems, which provide scheduled passenger services on fixed routes throughout the state.

A simplified methodology was developed for the representation of the remaining bus public transportation services. Bus services, which include hundreds of lines operated and managed by several different transit agencies and operators, were represented through an innovative numeric approach, which allowed a robust representation of the local transit services, without extensive coding requirements.

3.1. The Rail Network

Limited railway services are currently offered in California. In 2008, only three intercity railways were operated in the state. Besides, the rail network also includes four commuter railways and some light rail and subway systems. In 2009, there were 39 lines of rail transit in California, with average weekday ridership of approximately 1.12 million passengers.

The rail network is a relatively high cost, long-term capital system. Only limited modifications are introduced in railway operations over the years. This is due to the large investments required by rail projects, the fixed routes (railway tracks) and high sunk costs associated with this means of transportation.

In consideration of the characteristics described above, and the importance that the rail network has in providing a reliable alternative for both long distance and short distance trips in a statewide model, it was decided to code all railway lines explicitly in the development of the CSTDM public transportation network.

3.2. Local Transit Services

Many bus lines are currently operated in California, with level of service that varies from the frequent mass services in densely populated urban areas to sparse, low frequency service among more remote locations in rural counties.

In 2008, bus services were provided through more than 50 local transit operators in California, with over 1500 local bus routes. These services are an important component of the transportation system, and many times are the only alternative available to reach a destination for many users that do not have access to a private vehicle.

Bus services, on average, provide a limited contribution to the total transit ridership, if compared to the large amount of service provided (number of operated lines): according to ridership data from 2009, 191 bus routes were in service in the Los Angeles Metro bus system, with a total daily ridership of 1.18 million. Similarly, the three major Bay Area agencies combine for daily bus ridership of 0.84 million across 262 bus routes.

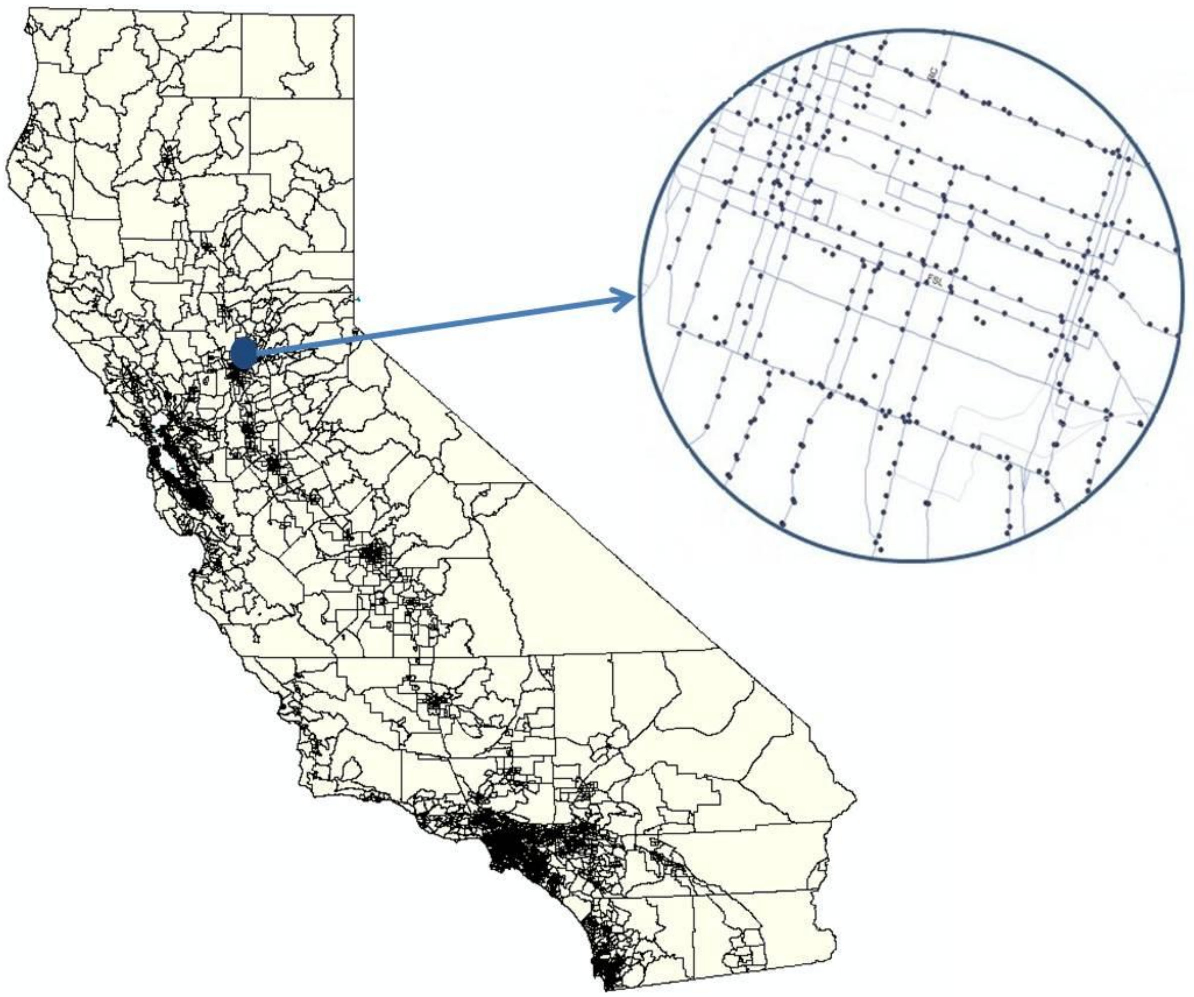

FIGURE 1 Distribution of local transit services in California. The right part of the figure shows an enlarged map of the local transit lines and bus stops in downtown Sacramento

The right part of Figure 1 provides an example of the extensive bus network that is operated in the downtown Sacramento, the State Capital. Several different lines provide transit services on a rather dense network in the area, each one counting dozens of bus stops in which users may access/egress the service. The bus routes shown in the example are only a very limited subset of the whole bus network system operated in California (left part of Figure 1).

The explicit coding of all bus lines and transit stops represents a time-intensive task that would require the allocation of a considerable amount of resources, which is not justified by the purposes of a statewide travel demand model. Furthermore, the characteristics of local bus services change frequently to adapt to modifications in travel demand, changes in the land use and the urban form of cities, funding and subsidies for public transportation. For all the reasons above, the definition of a valid alternative to the explicit coding of the bus network, which also reduces the efforts required for updating the network, is a valuable solution to adopt in such a large-scale modeling framework. In the CSTDM framework, the use of a simplified methodology for the representation of local transit services is introduced for the representation of local transit services in the short distance (shorter than 100 miles) personal travel model. In the simplified methodology, the characteristics of local transit services are represented through the use of a limited number of key input variables.

4. The Simplified Model for Local Transit

The simplified methodology for the representation of local transit services was developed with the aim of providing a robust representation of local bus services without the explicit coding of all local bus routes (which would imply a level of details beyond the scope of a statewide model such as the CSTDM). The methodology is also designed to reduce the efforts required for updating the information collected and for defining scenarios of future development of public transportation.

The simplified methodology was applied for the representation of local bus services in the short distance personal travel model. The proposed approach is based on a numeric methodology for representation of local transit travel attributes, based on the assumption that public transportation characteristics are strongly correlated with relevant features of the road network (e.g. distances and average speed) and land use variables.

4.1. Assumptions

The methodology is based on the estimation of econometric models that express the local transit attributes (*in-vehicle* and *out-of-vehicle times*) as a function of other more easily measurable transportation variables and land use patterns. These transit attributes, together with the transit fare, are used in the CSTDM travel demand model to represent the local transit choice option available (in the areas served by bus services) for short distance trips (under 100 miles). Specific coefficients are used in the model for weighting these components of travel time separately, and to account for the value of time. The synthetic methodology is also used to compute the travel times, and accessibility measures, to access railway facilities with connecting bus services, thus providing a realistic representation of the multimodal trips involving the use of both rail and bus.

The model is developed for four different times of the day, consistently with the time periods used in the development of the CSTDM framework, to account for the variability of transit services during the day.

At the base of the development of the model is the definition of the *catchment areas* for local transit. The catchment area is a measure of the geographical accessibility to transit services. Each of the 5191 TAZs in the modeling framework is assigned to a catchment area depending on the distance from the available transit lines in the area. TAZs that do not have access in a reasonable range to any local transit services are not included in any catchment area.

The local transit functions use four key inputs:

1. *Transfer areas*: the areas within which a person can travel (they include the possibilities of transfers among different operators in a region);
2. *Service areas*: the areas within which transit service is generally provided by a single operator (they are subdivisions of the larger transfer areas with multiple transit operators);
3. *Level of Service*: a single number representing the quantity of local bus service provided by the transit operator; and
4. *Fare*: a composite value, expressed in US dollars, indicating the typical fare paid by a customer.

The development of the model is based on the adoption of specific assumptions on the relationships among transit travel times and other relevant transportation and land use variables, which are presented in the following paragraphs.

4.2. Transfer and Service Areas

Transfer and Service Areas concur to the definition of the *catchment areas* for transit. The *transfer areas* measure the accessibility to transit services in the various regions. It is a measure of the portion of a region in which transit trips are possible using *any* of the operators that offer local transit services. A *service area* is a smaller region that is usually served by only one operator. In each service area, the level of service of transit is considered homogenous. 32 Transfer areas were defined in California. Multiple service areas are sometimes found in the largest transfer areas, as a result of the presence of multiple operators on a geographically large area. Service areas are generally indicated through the addition of a digit to the number of the transfer areas they belong to (for example, service areas 7.0, 7.1, 7.2 and 7.3 form the transfer area no. 7 in the "Sacramento region").

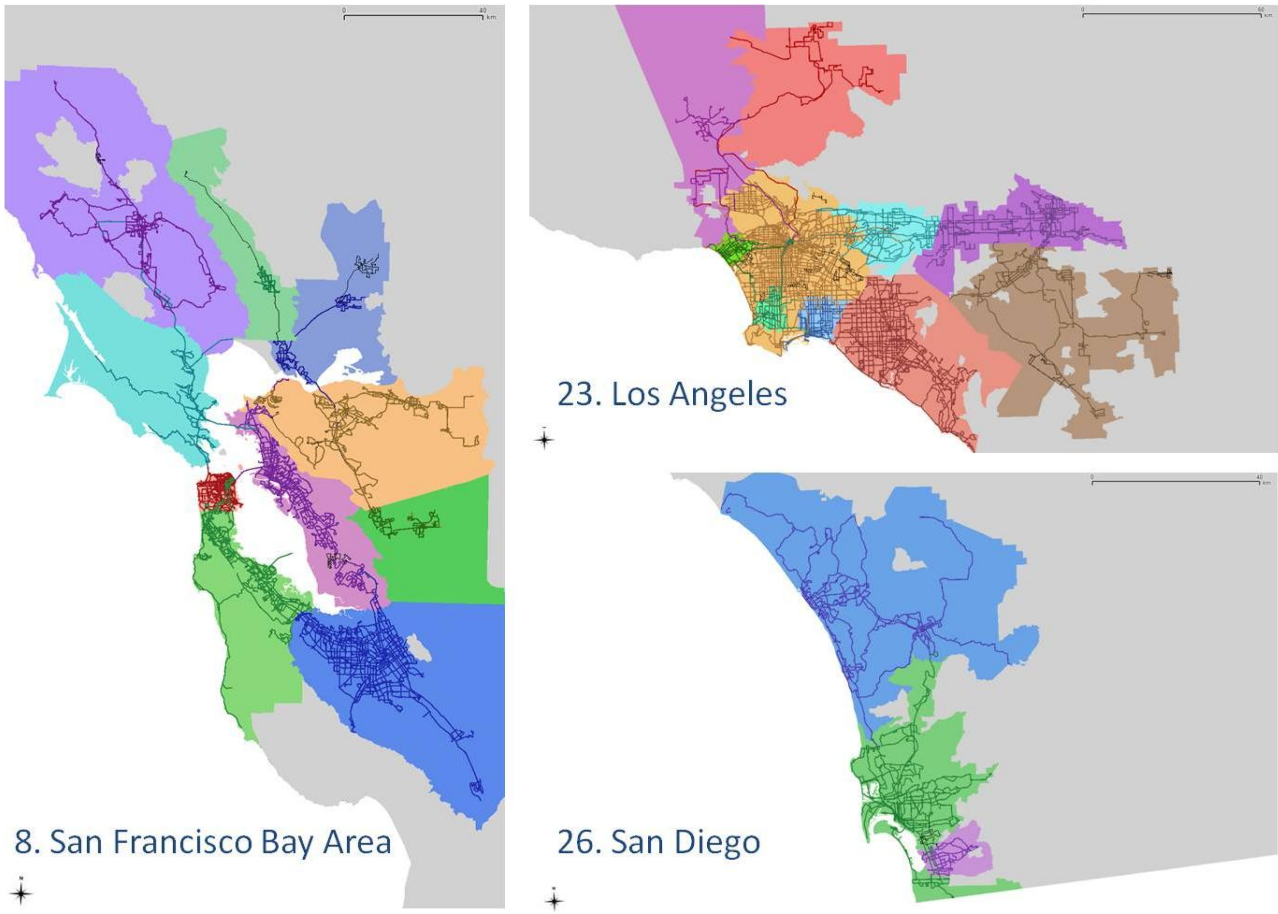

FIGURE 2 Local Transit *Catchment Areas* (Transfer and Service Areas) in three of the four major metropolitan areas of California: 8. San Francisco Bay Area; 23. Los Angeles; and 26. San Diego

Eventually, in more isolated urban areas and rural counties, transfer and service areas are identical. In these cases, transit trips are possible only among the TAZs of the same service area, and there are no possibilities for connecting trips that extend into other contiguous service areas. The four major urban areas (Los Angeles, San Francisco Bay Area, Sacramento and San Diego) are the only areas in which a single transfer area contains multiple service areas. In these regions, multiple operators can be used for creating longer trips that originate in one service area and are directed to a destination in a different service area. This two-level approach handles the major urban areas served by many operators (for instance, permitting local bus service from San Mateo into San Francisco). However, it prevents a traveler taking a local bus between e.g. Sacramento and San Francisco, even if the overall transit skim process does permit this transit trip by taking the AMTRAK Capitol Corridor train in addition to the local bus services.

A GIS shapefile of transit lines provided by the California Department of Transportation was used to develop the transfer and service areas. Each of the 5191 TAZs in the CSTDM framework is assigned to a catchment area, depending on the proximity from the TAZ centroid to the closest bus lines. The TAZ is assigned to the corresponding service and transfer area of the operator that runs the transit lines if such distance does not exceed 3 miles. If this test fails (some TAZ centroids are often located far from any bus line), but at least one bus line crosses the TAZ, then the TAZ is also included in the corresponding catchment area. Otherwise, the TAZ is not included in any catchment area (no transit services).

4.3. Level of Service and Fare

The Level of Service (LOS) is a single number that represents the quality of transit services provided by the local operator. The LOS variable is defined as the ratio of the *population served by a transit operator* (in its service area) divided by the *Annual Revenue Service Miles Provided* (in thousands). This variable is a measure of the *quantity* and *density* of the service provided (related to the population served by transit in the service area). For the way it is defined, the numeric value of the LOS decreases with an increase in the quantity of service provided by a transit operator. The LOS variable is computed using data from the National Transit Database of the Federal Transit Administration (FTA). In this measure, the amount of service is limited to that provided by bus and trolleybus, and does not include rail, which is modeled explicitly, nor demand responsive transit (which is not covered by the CSTDM).

Travel within a *Service Area* is determined by the *Level of Service* of the operator in that Service Area. Observed values for LOS range from 39.3 for San Francisco MUNI, to 484 in Thousand Oaks. In model operation, the value of LOS is capped at 200 (which affects Thousand Oaks, Santa Clarita and Gold Coast Transit), as an upper bound for the LOS associated with lower quality bus services. Several minor rural transit operators without data available were also assigned 200, and average LOS were computed for those service areas that are served by more than one operator. The weighted average LOS is 111, which is similar to the level provided by the Orange County Transportation Authority (OCTA), or the Santa Clara Valley Transportation Authority (VTA). This measure LOS has a number of beneficial properties:

- the value is a single number, which is easy to establish and understand;
- it relates to the actual transit provided, and is a policy input;
- by being based on population, a future "status quo" scenario with the same per capita service is an easy default option (by maintaining the LOS index constant); and
- the value, which is lower for better service, can easily be used directly in model estimation, and offers multiple possibilities for policy evaluation. For example, a doubling of service frequency or a doubling of service coverage would be represented by a halving of the LOS.

Fares for local bus transit were computed as the average single trip fare for each operator in a service area. When travel happens between two service areas that are included in the same transfer area, it requires the payment of the fare for both areas. A 5 minute penalty is incurred for transfer between services, and a weighted average of the Level of Service is used, with 2/3 of the weight on the poorer quality service to account for the increased difficulties for transfers among lines operated by low LOS operators.

4.4. In Vehicle Time

The model specification for the local transit functions is based on the assumption that local transit attributes are correlated with other transportation and land use variables used in the CSTDM framework. The *In-Vehicle Time (IVT)* is directly correlated with the travel time of the private vehicles that share the road, as an effect of the distance to the destination and of the traffic congestion on the network. Since HOV lanes and ramps (where available) can be used by buses in California, it is reasonable to expect that IVT is correlated with the congested travel time measured in the CSTDM for car users in high occupancy vehicles (HOV3 travel time).

Different model specifications were tested. In particular, the possibility of a quadratic relation between IVT and HOV3 congested travel time was suggested, as a way to allow possible non-linear effects of HOV3 auto travel time on the in-vehicle transit time: travel times are usually higher for buses than for cars, especially for shorter trips (also due to the frequent stops in the bus mode). The marginal effects of auto travel time tend to diminish for longer trips, also as an effect of the eventual availability of express bus services. IVT is also expected to depend on the LOS, which measures the quantity and density of service provided (as affected by investments in the transit system and other local conditions): as the LOS index increases, and less services per capita are provided, travel times are expected to increase as a result of longer detours and indirect routes needed to reach the desired destination.

4.5. Out of Vehicle Time

Similar assumptions to those introduced for the local transit function for IVT where used for the model specification for the *Out-of-vehicle Time (OVT).* This measure of time represents the sum of all components of "out of vehicle" time that are associated with a transit trip:

1. the *time to access* the transit stop from the origin of the trip (generally by walk);
2. the *waiting time* at the first stop;
3. (eventual) *transfer time(s)* in any intermediate stop(s); and
4. the *egression time* from the last bus stop to reach the final destination.

OVT is expected to depend on the distance on the road network between the origin and the destination of the trip: longer trips usually require higher OVT, because of the lower availability of direct services to a specific destination, and the higher waiting and transfer times required.

Road distances are measured on the HOV3 congested network (which may differ from the SOV distances depending on how common HOV lanes and ramps are in the network). Moreover, OVT is a function of the residential and employment densities of the TAZs of origin and destination, as a way to account for the general availability of better transit coverage in higher density areas. Similarly to the IVT function, OVT is expected to depend on the LOS: lower waiting and transfer times are required in those areas that have better level of service, as a result of more frequent transit services, and a more dense network (which increases the probability of easier transfer between lines, and higher availability of direct lines to the destination).

5. Estimation of the models

Two models are used for the Local Transit Functions. One represents *In-Vehicle Time (IVT)*, and the other represents the *Out-of-Vehicle Time (OVT)*. The estimation of these two models was carried out with observed data for travel times collected through the Google Transit Data Feed (database with transit service data optimized to provide travel information through the Google platform), merged with additional information on the transportation system, the land use patterns and sociodemographics available from other sources.

5.1. The Data

The econometric models for In-Vehicle Time (IVT) and Out (of) Vehicle Time (OVT) were estimated using observed data collected from internet sources. A large sample of transit travel time records was obtained through the databases stored by transit agencies on the Google platform and that generates the information for travel solutions by transit available on http://maps.google.com/.

Transit data available from the internet are still not often used in transportation research studies. However, the quality of the information of these sources has considerably increased in recent years: as transit agencies put considerable efforts in promoting their services through online platforms, the availability of reliable information from these sources has sharply increased. Besides, the standardization required by the adoption of the common platform and interface, defined by the provider of the internet services, allows easy collection of information for multiple operators in geographically separated areas through the same procedure, and with similar margins of error. For the data extraction, a python code, denominated Graphserver, was available. The code is a tool developed for transit agencies to test their data to post online. The code is a useful basis for the analysis of Google Transit Data Feeds from multiple operators, and was used in the process of collection of the data.

A total of 91,074 records were extracted with complete transit travel times for interzonal trips having origins and destinations in the centroids of the TAZs within the CSTDM region. Each record contained information on:

- the *time of the day* of the travel record;
- the *exact time of departure* from the origin of the trip;
- the *exact time of arrival* at the final destination;
- the *walking time* from the origin to the first transit stop;
- the *time* spent *on board* of the first bus;
- the *time* for the first *transfer* (*if any*);
- the *time* spent *on board* of the second bus (*if more than one busis required*);
- any additional *transfer times* and *in-vehicle times* for additional parts of the trip;
- the *walking time* from the last bus stop to the final destination.

Additional information was obtained through the integration of the data with the CSTDM TAZ system:

- the *transfer* and *service area* of the selected itinerary;
- the *TAZ of origin*;
- the *TAZ of destination*.

The data were collected from 29 different *service zones* across all four periods of time of the day. The data were merged with information available from other sources (e.g. Caltrans, US DOT, other components of the CSTDM framework) to create the required datasets for the estimation of the IVT and OVT functions. The additional information included HOV3 auto travel times and distances for each itinerary (estimated on the CSTDM road network), the LOS in each service area, and residential and employment densities for each TAZ.

5.2. In Vehicle Time

The IVT function is a linear regression that was estimated with 91,074 observations. Several alternative model specifications were tested, and a number of parameters tried as inputs. Four different functions were estimated for the four times of the day (AM Peak, Midday, PM Peak and Off-Peak). However, due to their similar trends and goodness of fit, they were combined in only two final models, respectively estimated for the Peak (6:00AM to 10:00AM and 3:00PM to 7:00PM) and the Off-Peak time (rest of the day). The sample sizes for the estimation of the final models were respectively 50,727 (Peak) and 40,347 (Off-Peak). Both models have quite good goodness of fit, with r-square respectively of 0.916 for the Peak model, and of 0.909 for the Off-Peak model. The estimated coefficients for the two models are reported in Table 1.

TABLE 1 In Vehicle Time (IVT) Functions for (a) Peak and (b) Off-Peak Time

a) *Peak Period Model (N=50727)*

Parameter	Estimated coefficient	Std. Error	t-statistic	p-value
HOV3_Time	**2.8921040**	.01478	195.69460	<0.001
HOV3_Time^2	**-.0174477**	.00040	-43.99996	<0.001
LOS*HOV3_Time	**.0057270**	.00011	52.59453	<0.001
R-Square	**0.915505**			
Std. Error of the Estimate	**14.44533**			

b) *Off-Peak Period Model (N=40347)*

Parameter	Estimated coefficient	Std. Error	t-statistic	p-value
HOV3_Time	**2.7813943**	.01734	160.42225	<0.001
HOV3_Time^2	**-.0029318**	.00055	-5.29485	<0.001
LOS*HOV3_Time	**.0046781**	.00013	35.90664	<0.001
R-Square	**0.909412**			
Std. Error of the Estimate	**13.63288**			

The estimated models describe *In Vehicle Time* as a function of the HOV3 auto travel time (both variables are measured in minutes). This represents the effect of network speed, connectivity and road geometry, as well as the effects of traffic congestion on bus travel times. The squared term provides attenuation for longer trips, which is likely due to the presence of limited service or express long haul service.

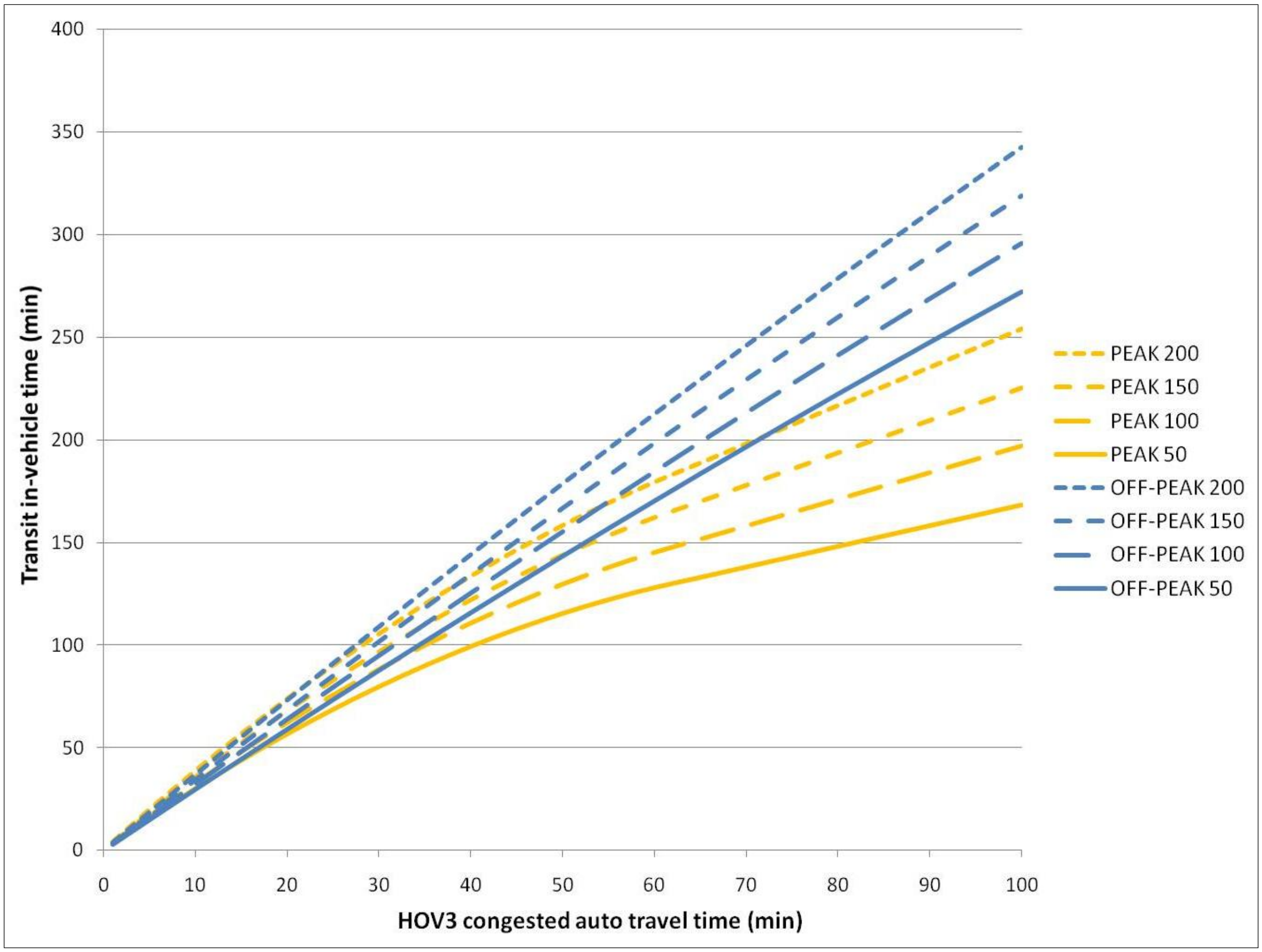

FIGURE 3 Transit In-Vehicle Time in dependence of HOV3 congested auto travel time for Peak and Off-Peak time, with different LOS Curves

The final term in the IVT models provides a policy sensitivity tool. As additional transit coverage is provided (better LOS), the in-vehicle time is reduced. This is due to both the provision of more direct lines (operators providing a low level of service typically provide very circuitous routes to ensure a minimum access to all residents), and also the increased likelihood of there being a route serving the specific OD pairs, rather than travelling to/from a transfer point.

5.3. Out of Vehicle Time

The OVT function is also a linear regression. The available observations for the estimation of this function were 88,730 observations, after excluding 2344 records with missing values for at least one of the variables used in the model. Similarly to the *in-vehicle time* model, four different models were estimated for the four times of the day. However, also in this case, two final models were estimated respectively for the Peak and for the Off-Peak time. Table 2 reports the estimated coefficients for the OVT equations. The sample sizes are respectively 49,263 for the Peak model and 39,467 for the Off-Peak model. Both models have quite good measures of goodness of fit, with r-square respectively of 0.840 for the Peak model and of 0.830 for the Off-Peak model. R-square values are lower than in the IVT models, probably because *out-of-vehicle* time has many more possible causes, including a sparse network leading to long walks, infrequent headways and variable transfer times.

TABLE 2 Out Vehicle Time (OVT) Functions for (a) Peak and (b) Off Peak Time

a) ***Peak Period Model (N=49263)***

Parameter	Estimated coefficient	Std. Error	t-statistic	p-value
SQ_LOS	**3.219780**	0.018315	175.79624	<0.001
LOS*HOV3_Dist	**0.006140**	9.9E-05	62.042392	<0.001
SQ_P2E_DENSITY	**-0.016737**	0.000669	-25.02035	<0.001
R-Square	**0.839672**			
Std. Error of the Estimate	**16.24132**			

b) ***Off-Peak Period Model (N=39467)***

Parameter	Estimated coefficient	Std. Error	t-statistic	p-value
SQ_LOS	**3.087907**	0.021532	143.4103	<0.001
LOS*HOV3_Dist	**0.007235**	0.000124	58.35179	<0.001
SQ_P2E_DENSITY	**-0.007630**	0.000745	-10.2447	<0.001
R-Square	**0.829657**			
Std. Error of the Estimate	**16.78981**			

This model is highly dependent on the level of service provided (through the squared root of the LOS): a better level of service can provide more closely spaced routes (and better coverage), lower headways and more efficient transfers (or direct service without a transfer). The dependence on the HOV3 distance is the result of the effects of the availability of fewer direct lines for longer trips (that therefore determine larger waiting and transfer times).

The third term uses the sum of the square roots of the P2E (sum of population and two times employment) densities for the TAZs of origin and of the destination. In this expression, the employment component is doubled to provide a more balanced representation between employment and population, consistently with the findings of the literature (Hendrickson, 1985) on the stronger relationships between employment density and the use of transit. This measure implies that denser areas will have lower *out-of-vehicle* transit times, which is due to operators typically focusing service on core activity nodes, as well as the reduced walking distances likely in denser areas. The origin and destination are considered separately, to account for the different conditions existing, for instance, between trips connecting more balanced OD pair (cross-town) and trips on radial routes from very sparse suburb to a dense downtown. Figure 4 provides some examples of the OVT curves in dependence of the LOS index, for Peak and Off-Peak time, for four trip types (diversified by the OD distance and the values of neighborhood densities at the trip origin and the destination).

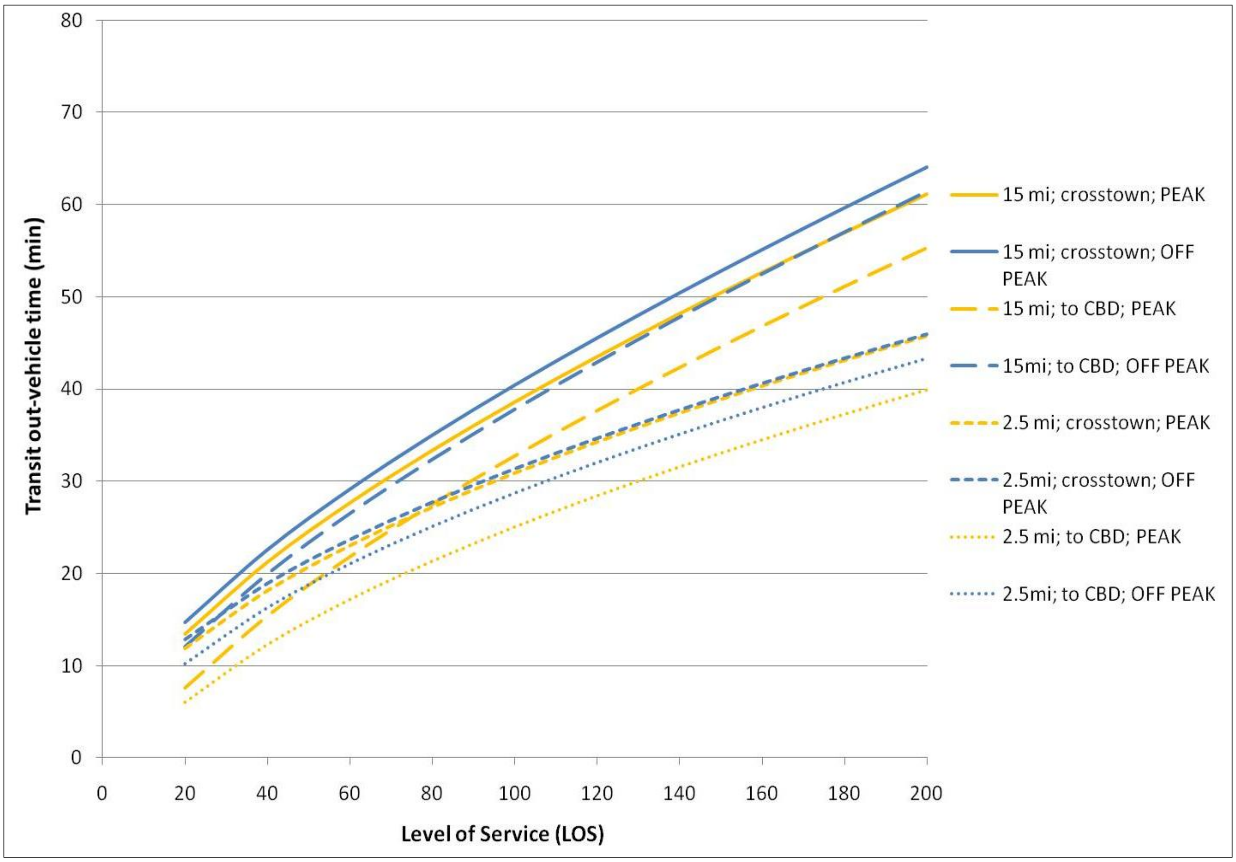

FIGURE 4 Transit Out-of-vehicle Time Curves for four examples of trip types, by LOS (horizontal axis) and for Peak and Off-Peak Time

The combination of the IVT and OVT functions allows evaluating the effects of transit investments. In particular, the functional form and the estimated parameters suggest a diminishing return to the investments in service: doubling the service in a region allows improving transit travel times, although more limited reductions in travel times are obtained in those areas in which the LOS is already quite good.

6. Application in the CSTDM Framework

The local transit functions were integrated in the CSTDM framework using four CUBE scripts, one for each time of the day. This allowed the provision of a multi-modal transit system, including the possibility to take local bus to and/or from a rail station, or to take the local bus all of the way. The model provides a reasonable representation of transit, and is sensitive to land use patterns and congestion in addition to the policy decisions to alter the level of local transit service provided, or to construct rail infrastructure. Moreover, the model sensitivity to population and employment density can be used as policy variables in future scenarios.

The whole approach is optimized for an efficient allocation of resources in the development of the model, and for the generation of future scenarios. The CUBE scripts for the local transit functions are the same in each scenario. However, a separate input file is generated for each scenario to provide the required information on catchment areas, LOS and fares. Policy testing can be carried out varying the input information on the catchment areas, LOS and fares for each service area, and population and employment densities.

7. Conclusions

This paper discusses the use of a simplified methodology for the representation of the local transit system in a statewide travel demand model. The proposed approach provides a robust representation of bus services in the Short Distance Personal Travel Model of the California Statewide Travel Demand Modeling (CSTDM) Framework.

The proposed approach involves the use of econometric models to express the local transit attributes in dependence of other transportation and land use variables. Two separate functions are estimated for the In-Vehicle Time (IVT) and the Out-of-Vehicle Time (OVT). Different models are also estimated for Peak and Off-Peak time, to account for variability in services during the day. The models are sensitive to different levels of service and local conditions through the definition of the *catchment areas* for local transit operators, and the use of a *Level of Service* index.

The model is estimated using observed data collected from the internet platform Google, with a total sample of more than 90,000 records. All models have a very good goodness of fit, with measures of R-square that respectively exceed 0.9 for the *In-Vehicle Time* functions, and 0.8 for the *Out-of-vehicle Time*.

The methodology has several advantages compared to other explicit methodologies for representing transit in a statewide model. It eliminates the need for explicit coding of all bus links and connections in the state, and is appropriate also in consideration of the limited mode share of bus services in California. Thus, it optimizes the use of resources, providing a robust model for local transit with efforts that are consistent with the output benefit in a large scale Statewide Travel Demand Model.

The simplified methodology permits easier data input, and it allows several possibilities of policy testing through the simulation of their impact on the LOS of the operators in each service area. Finally, the proposed approach allows easier maintenance of the transit network, with reduces the efforts required for the definition of future scenarios to study in the modeling framework.

References

1. Albalate, D. and G. Bel (2010) What Shapes Local Public Transportation in Europe? Economics, Mobility, Institutions, and Geography. *Transportation Research Part E*, **46(5)**, 775-790.
2. Badoe, D. A. and E. J. Miller (2000) Transportation-Land-Use Interaction: Empirical Finding in North America, and Their Implications for Modeling. *Transportation Research Part D*, **5**, 235-263.
3. Buehler, R. (2009) Promoting Public Transportation: Comparison of Passengers and Policies in Germany and the United States. *Transportation Research Records*, **2110**, 60-68.
4. Cervero, R. (1994) Transit-Based Housing in California: evidence on ridership impacts. *Transport Policy*, **1(3)**, 174-183.
5. Cervero, R. (2007) Transit-Oriented Development’s Ridership Bonus: a Product of Self-Selection and Public Policies. *Environment and Planning A*, **39**, 2068-2085.
6. Cervero, R. and K. Kockelman (1997) Travel Demand and the 3Ds: Density, Diversity, and Design. *Transportation Research Part D*, **2(3)**, 199-219.
7. Crawford J. H. (2000) *Carfree Cities*. Utrecht: International Books.
8. Dittmar, H. and G. Ohland (2004) *The New Transit Town. Best Practices in Transit-Oriented Development*. Washington (DC): Island Press.
9. Dunphy, R. T. and K. Fisher (1996) Transportation, Congestion and Density: New Insights. *Transportation Research Records*, **1552**, 89-96.
10. Giuliano, G. and J. Dargay (2005) Car Ownership, Travel and Land Use: A Comparison of the US and Great Britain. *Transportation Research Part A*, **40(2)**, 106-124.
11. Hendrickson, C. (1985) A Note on Trends in Transit Commuting in the United States Relating to Employment in the Central Business District. *Transportation Research Part A*, **20(1)**, 33-37.

12. Kain, J. F. and Z. Liu (1999) Secret of Success: Assessing the Large Increase in Transit Ridership Achieved by Houston and San Diego Transit Providers. *Transportation Research Part A*, **33(7/8)**, 601-624.
13. Karlaftis, M. G. and P. McCarthy (1998) Operating Subsidies and Performance in Public Transit: An Empirical Study. *Transportation Research Part A*, **32(5)**, 359-375.
14. Kenworthy, J. R. and F. B. Laube (1996) Automobile Dependence in Cities: An International Comparison of Urban Transport and Land Use Patterns with Implications for Sustainability. *Environmental Impact Assessment Review*, **16(4/6)**, 279-308.
15. Lane, B. (2009) The Relationship between Recent Gasoline Price Fluctuations and Transit Ridership in Major US Cities. Journal of Transport Geography, 18(2), 214-225.
16. Litman, T. (2007) Transit Price Elasticities and Cross-elasticities. *Journal of Public Transportation*, **7(2)**, 37-58.
17. Maley, D. W. and R. Weinberger (2009) Rising Gas Price and Transit Ridership: Case Study of Philadelphia, Pennsylvania. *Transportation Research Records*, **2139**, 183-188.
18. Newman, P. and J. R. Kenworthy (1989) Gasoline Consumption and Cities: A Comparison of U.S. Cities with a Global Survey. *Journal of the American Planning Association*, **55**, 24-37.
19. Paulley, N., R. Balcombe, R. Mackett, H. Titheridge, J. Preston, M. Wardman, J. Shires and P. White (2006) The Demand for Public Transport: The Effects of Fares, Quality of Service, Income and Car Ownership. *Transport Policy*, **13**, 295-306.
20. Schimek, P. (1996) Automobile and Public Transit Use in the United States and Canada: Comparison of Postwar Trends. *Transportation Research Records*, **1521**, 3-11.
21. Taylor, B. D., D. Miller, H. Iseki and C. Fink (2009) Nature and/or Nurture? Analyzing the Determinants of Transit Ridership across US Urbanized Areas. *Transportation Research Part A*, **43(1)**, 60-77.
22. Wang, G. H. K. and D. Skinner (1984) The Impact of Fare and Gasoline Price Changes on Monthly Transit Ridership: Empirical Evidence from Seven U.S. Transit Authorities. *Transportation Research Part B*, **18(1)**, 29-41.