Journal of Urban Economics 76 (2013) 1–14

Contents lists available at SciVerse ScienceDirect

Journal of Urban Economics

www.elsevier.com/locate/jue

City size, network structure and traffic congestion

Theodore Tsekeris [a,*], Nikolas Geroliminis [b]

[a] Centre of Planning and Economic Research (KEPE), 11 Amerikis, 10672 Athens, Greece

[b] Urban Transport Systems Laboratory (LUTS), École Polytechnique Fédérale de Lausanne (EPFL), Room GC C2 389, Station 18, CH-1015 Lausanne, Switzerland

ARTICLE INFO

Article history:
Received 9 June 2011
Revised 20 December 2012
Available online 1 February 2013

JEL classification:
L9
R1
R3
R4

Keywords:
City size
Land use
Transport network
Traffic congestion dynamics
Macroscopic fundamental diagram

ABSTRACT

This paper presents an alternative approach for analyzing the relationship between land use and traffic congestion by employing the Macroscopic Fundamental Diagram (MFD). The MFD is an empirically observed relationship between traffic flow and traffic density at the level of an urban region, including hypercongestion, where flow decreases as density increases. This approach is consistent with the physics of traffic and allows the parsimonious modeling of intra-day traffic dynamics and their connection with city size, land use and network characteristics. The MFD can accurately measure the inefficiency of land and network resource allocation due to hypercongestion, in contrast with existing models of congestion. The findings reinforce the 'compact city' hypothesis, by favoring a larger mixed-use core area with greater zone width, block density and number of lanes, compared to the peripheral area. They also suggest a new set of policies, including the optimization of perimeter controls and the fraction of land for transport, which constitute robust second-best optimal strategies that can further reduce congestion externalities.

1. Introduction

The increasing economic and environmental concerns raised by the growth of private vehicle use in urban areas have resulted in the design and implementation of a number of planning and management strategies on the supply side (control of traffic signals, ramp metering, capacity enhancement, etc.) or the demand side (congestion pricing, parking restriction, etc.) to diminish efficiency losses. From the planning perspective, policies have favored more compact development patterns by revitalizing the city center and restricting urban sprawl, through density and boundary growth controls (Anas et al., 1998; McConnell et al., 2006). In this context, the appropriate selection of network design parameters is crucial for the efficient allocation of road investment in the early stages of planning, or when updating the urban master plan. Such design parameters may encompass the number of road links, average link length, block area and average number of lanes. Particularly, the question of the allocation of resources to large urban clusters or more spatially dispersed metropolitan areas is critical for the development of countries with a rapidly growing urban population, such as China and India (Henderson, 2010).

The proper modeling, interpretation and treatment of the relationship between urban land use and congestion are necessary to address the above question. However, existing traffic models in urban economics pose severe theoretical and empirical limitations in realistic applications. This is because they employ link travel cost functions which cannot accurately specify the intra-day traffic dynamics and relate them to land use and urban-scale network characteristics in a way that is computationally tractable and consistent with the physics of traffic. This failure hinders the ability of economic models to support accurate and robust design proposals for the allocation of urban land and network resources to diminish congestion externalities.

Specifically, the traditional models of congestion simplistically assume that the travel time on each link is separable and monotonically increasing with link flow. This assumption is adopted by the Bureau of Public Roads (Branston, 1976) and the Vickrey congestion function. The form of such travel time functions implies the existence of a stationary traffic equilibrium regime and steady-state volume-delay relationships. Several studies have challenged the use of these functions because of the need to account for the non-monotonicity of travel time with traffic flow (McDonald et al., 1999) and showed the intrinsic inconsistency, infeasibility

* Corresponding author.
E-mail addresses: tsek@kepe.gr (T. Tsekeris), nikolas.geroliminis@epfl.ch (N. Geroliminis).

0094-1190/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jue.2013.01.002

and instability of static models of congestion (Verhoef, 1999, 2001, 2005). As more recent studies (Geroliminis and Daganzo, 2008) have pointed out, plots between pertinent traffic variables (flow, speed, delay) on a spatially disaggregated (link) level are not in steady state, but are actually in highly-scattered conditions and do not follow a well-defined curve.

The general equilibrium models of urban land use and transport typically assume an uncongested regime or point-queuing. The latter assumption does not account for the temporal and spatial dimensions of congestion and is not consistent with the laws of physics, because travel speed is entirely determined by traffic density at a specific location and time (Ross and Yinger, 2000; Arnott, 2007). Even recent, conceptually sophisticated models of this type, such as the Regional Economy, Land Use and Transportation (RELU-TRAN) model (Anas and Liu, 2007; Anas, 2011), do not consider the intra-day traffic dynamics. Moreover, such models involve increased calibration costs for network modeling (including partitioning into a considerable number of zones) and require the availability of origin–destination trip matrices and traffic assignment procedures to produce a user equilibrium pattern of link travel times. The traffic assignment procedures are associated with strong assumptions about the route choice behavior of consumers, intense computational burden and mathematical properties that are difficult to analyze in realistically large networks (Peeta and Ziliaskopoulos, 2001; Chiu et al., 2011).

The bottleneck model (Vickrey, 1969; Arnott et al., 1993; Arnott, 1998), in which peak-period traffic congestion is represented as a queue behind a bottleneck with fixed flow capacity at the edge of CBD, provides an alternative congestion technology. However, an important problem pertaining to all the above models is that they ignore the downward-sloping part of the curve between flow and density, known as hypercongestion, on a single link that is typically homogenous with uniformly distributed capacity (Lo and Szeto, 2005). In hypercongestion, the flow decreases from the point where the density reaches a particular critical value that maximizes flow, until the flow falls to zero when density reaches its maximum value (jam density).

Furthermore, only a few studies have attempted to abstractly represent the network-wide relationships between traffic variables, by extending the constraints and dynamics involved in the bottleneck model. In some of them, like that of Small and Chu (2003), which allowed for hypercongestion, the dynamic analysis cannot ensure a stable macroscopic relationship that is consistent with the physics of traffic. In others, such a relationship cannot be generalized (Arnott and Inci, 2010), or it becomes computationally intractable for large-scale urban networks. There are also studies (e.g. Lago and Daganzo, 2007; Arnott and de Palma, 2011) which consider the suburbs and CBD as points in space, i.e., without having a physical dimension. The latter assumption hinders the analysis of intersection-based control strategies (e.g. metering of network access) and land use policies for reducing congestion costs.

The approach proposed here treats the above theoretical and computational shortcomings by abstracting the complexity of network traffic dynamics into a single graphical expression. Fig. 1 illustrates a stable and low-scattered graphical relationship, referred to as the Macroscopic Fundamental Diagram (MFD),[1] between space–mean flow (or produced amount of travel), the system-wide vehicle-kilometers traveled per hour (veh-km/h), and vehicle density, the system accumulation in vehicles per kilometer per lane (veh/km/lane), for two regions (neighborhoods) of a hypothetical concentric city (presented in Section 2). The MFD recognizes the inherent dynamics of macroscopic phenomena of congestion, the dependence of travel delays on the initial density conditions and the non-linearity of the average throughput-accumulation relationship. The system throughput is defined as the network outflow in vehicles per hour (veh/h), which is the rate that vehicles exit the network by driving to other neighborhoods or reaching their destinations.

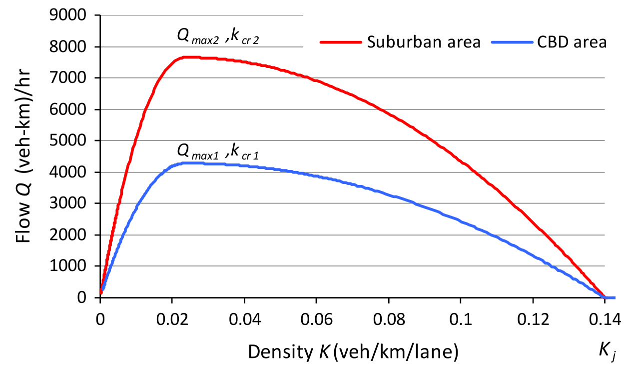

Fig. 1. Typical macroscopic fundamental diagram (network flow Q *vs.* network traffic density K) for the concentric city.

In methodological terms, the MFD represents macroscopically (at each region) the flow of vehicles from their origin to destination at each instant (very short time interval) of the period of analysis. The modeling encompasses the rates at which vehicle trips enter and exit the network, and the queuing dynamics, according to the flow-density relationship given by the MFD curve of each region. This process ensures the conservation of flows between entry and exit rates and accounts for both states of congestion and hypercongestion at the rising and falling portion of the curve, respectively (see Fig. 1). The traffic capacity in some regions may change over time, when during hypercongestion the vehicle density (or accumulation) degrades throughput.

The MFD can also integrate the physical and functional characteristics of the urban land and transport network. This allows the expression of the spatial–temporal patterns of congestion as a function of city size, land use, network topology and traffic control. Hence, it is shown how different land use and network structures affect the formation and dynamics of congestion. Therefore, the MFD offers a tractable and parsimonious approach for modeling congestion that is consistent with the physics of traffic and relevant to economic analysis of policies to reduce it. The proposed alternative paradigm addresses the evaluation of land use policies and the design of transport systems, and suggests suitable combinations of strategies to alleviate congestion.

Existing studies in economics (see above) have mostly focused on the provision of new road capacity and imposition of tolls to reduce congestion externalities. Nonetheless, the implementation of these strategies is usually restricted by limited funding and enforcement capacities, particularly in developing countries. The MFD approach permits the analysis of a set of second-best policies, which could not be investigated in models that fail to include hypercongestion or are intractable for deployment in large-scale urban areas. Such policies include the use of advanced technology for metering of access (CBD perimeter control) and the reallocation of the existing network capacity and land use among urban zones. The resulting plans recognize that (i) commuting costs are not only a function of inflow, but also (non-linearly) depend on the previous level of congestion and (ii) an optimal solution may be reached even in hypercongestion.

The paper proceeds as follows. Section 2 describes the structure of a simple monocentric city model with concentric neighborhoods, which is used to apply the MFD approach. Section 3 pre-

[1] The existence of MFD has been empirically verified through aggregating highly scattered plots of flow *vs.* density from individual links as measured by fixed detectors and other readily available monitoring technologies in uniformly congested urban areas (Geroliminis and Daganzo, 2008).