Transportation Research Part B 104 (2017) 583–601

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

# Testing the proportionality condition with taxi trajectory data

Jun Xie[a], Yu (Marco) Nie[b,*], Xiaobo Liu[a]

[a] School of Transportation and Logistics, Southwest Jiaotong University, Chengdu, Sichuan 610031, China
[b] Department of Civil and Environmental Engineering, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, USA

ARTICLE INFO

*Article history:*
Received 16 December 2016
Revised 2 May 2017
Accepted 3 May 2017
Available online 13 May 2017

*Keywords:*
User equilibrium
The proportionality condition
Paired alternative segments
Taxi trajectory data

ABSTRACT

The proportionality condition has been widely used to produce a unique path flow solution in the user equilibrium traffic assignment problem. However, it remains an open question whether and to what extent this condition accords to real travel behavior. This paper attempts to validate the behavioural realism of the proportionality condition using more than 27 million route choice observations obtained by mining a large taxi trajectory data set. A method is first developed to uncover more than three hundred valid paired alternative segments (PAS), on which the proportionality condition is tested by performing linear regression analysis and chi-square tests. The results show that the majority of the PASs tested (up to 85%) satisfy the proportionality condition at a reasonable level of statistical significance.



## 1. Introduction

Traffic assignment is one of the most widely used tools in transportation planning. While the standard user-equilibrium (UE) traffic assignment model (Beckmann et al., 1956) uniquely determines the total flow on each network link, it is generally unable to produce unique route flows or multiple-class link flows (see e.g. Chapter 3 Sheffi, 1985). However, many useful planning applications, such as *select link analysis*, do use route flows. The lack of uniqueness thus can result in inconsistency and instability in the analyses that depend on these solutions (Lu and Nie, 2010; Bar-Gera et al., 2012).

Additional assumptions are needed to assure that route flows and multi-class link flows are uniquely determined. A generally accepted assumption is to find the *most likely route flows*, which is the solution to an entropy maximization problem (Rossi et al., 1989). Bar-Gera and Boyce (1999) derive the proportionality condition from entropy maximization, which states that *flows should be distributed to two paired alternative segments (PAS) with equal cost according to the same proportion regardless of user class, origin or destination*. By enforcing this condition, Bar-Gera (2010) develops a new assignment algorithm - known as Traffic Assignment by Paired Alternative Segments (TAPAS) - that is capable of determining *nearly unique* route flows. By "nearly unique" we highlight the fact that the proportionality condition is not sufficient to ensure entropy maximization. As pointed out by Bar-Gera (2006) and Borchers et al. (2015), *higher-order* proportionality conditions may be needed to secure a solution to the entropy maximization problem. Nevertheless, empirical evidence suggests that higher-order proportionality conditions rarely arise in real-life networks, and that a solution satisfying the proportionality condition offers approximation of high quality to the maximum entropy solution for most practical purposes (Bar-Gera, 2006).

* Corresponding author.
*E-mail address:* y-nie@northwestern.edu (Y. (Marco) Nie).

http://dx.doi.org/10.1016/j.trb.2017.05.003
0191-2615/

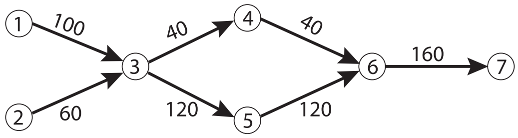


**Fig. 1.** Illustration of proportionality with one set of paired alternative segments (link labels are link flows at UE).

**Table 1**
Multiple UE path flow solutions for the UE link flow solution reported in Fig. 1 ($h^*$ is the unique path flow solution that also satisfies proportionality).

| ID | Path | $h^*$ | $h_1$ | $h_2$ |
|---|---|---|---|---|
| 1 | 1-3-4-6-7 | 25 | 40 | 0 |
| 2 | 1-3-5-6-7 | 75 | 60 | 100 |
| 3 | 2-3-4-6-7 | 15 | 0 | 40 |
| 4 | 2-3-5-6-7 | 45 | 60 | 20 |

Not surprisingly, the concept of proportionality has been gradually introduced into mainstream transportation planning software[1] since 2010 (Boyce et al., 2010; Bar-Gera et al., 2012).

One may argue that the proportionality condition is merely an artificial instrument for producing a unique solution that can be consistently used as a benchmark for scenario comparison. In theory, Lu and Nie (2010) showed that a unique UE route flow may be determined by maximizing any objective function that possesses certain properties. Among these competing "unique" route flow solutions, there seems no sound reason (other than for the sake of convenience) to prefer one to the other. It remains an open question whether and to what extent the proportionality condition, which is the choice of the current practice, accords to real travel behavior.

In light of the above, we document in this paper the first attempt to validate the behavioral realism of the proportionality assumption. The validation study lags behind in the literature mainly because it requires a large amount of route choice data, which is not readily available from traditional data sources. This critical gap is filled in this study thanks to the availability of GPS log data collected by the GPS receiver installed in taxis, referred to as *taxi trajectory data* hereafter. Taxi trajectory data have several appealing characteristics that make it a desirable choice for studying human mobility patterns in general (see e.g. Li et al., 2011; Yuan et al., 2010; Zhang et al., 2016; Wang et al., 2014; Yue et al., 2009; Liu and Qu, 2016; Zheng et al., 2013; Liu et al., 2010), and route choice behavior in particular. First, it does not have the privacy issue, because no personal information can be recovered from such data (unlike mobile phone or social media data). Second, taxi drivers can be seen as "experts" in path planning (Lin et al., 2015) because they usually know the road network and traffic conditions well. Thus, the path taken by a taxi driver is usually the "optimal" choice, which is a prerequisite for applying the proportionality condition. Third, thanks to passenger boarding information, it is relatively easy to break the continuous trajectory data into taxi trips with a specific OD pair. Last but not least, the real travel time of a taxi through any path segment can be easily estimated based on the position and time information, which is essential in guaranteeing equal travel time on selected PASs.

The reminder of this paper is organized as follows. Section 2 briefly reviews the proportionality condition and Section 3 introduces the proposed procedure for its validation. Section 4 presents the details of the taxi trajectory data, as well as the necessary data processing procedures. Section 5 reports and discusses the validation results, and Section 6 concludes the paper with a summary of main findings.

## 2. Concept of proportionality

If we assume that paths with identical costs are equally attractive to all travelers, then their probability of choosing either segment of any two paired alternative segments (PAS) with equal cost would be identical, regardless of their origin, destination and user class (provided the class identification does not affect a traveler's evaluation of segment costs). Accordingly, the *most likely* UE path solution, or the UE path flow solution that has the highest probability to realize (or be observed), must satisfy the proportionality condition.

Fig. 1 illustrates the concept of proportionality. The link flows shown in Fig. 1 satisfy the user equilibrium (UE) condition, that is, the paired alternative segments (3-4-6 and 3-5-6) both have positive flows and exactly the same cost. Given the UE link flows, multiple path flow solutions exist that all produce the same link flow solution. Table 1 lists three feasible UE path flow solutions, but only the path flow solution $h^*$ is determined according to the proportionality condition. The reader can verify that the ratio between the flow on path 1 and 2 (as well as path 3 and 4) is 1:3, equal to the ratio of the flow

[1] see e.g. https://www.inrosoftware.com/en/news-and-events/posts/traffic-assignment-choices-at-emme-4-1/; http://www.trafikanalysforum.se/sites/default/files/bibliotek/latest_developments_in_visum_-_klaus_nokel.pdf.