Made available by Hasselt University Library in https://documentserver.uhasselt.be

Exploratory analysis of Zipf's universal power law in activity schedules

Peer-reviewed author version

ECTORS, Wim; KOCHAN, Bruno; JANSSENS, Davy; BELLEMANS, Tom & WETS, Geert (2018) Exploratory analysis of Zipf's universal power law in activity schedules. In: TRANSPORTATION, 46 (5), p. 1689-1712.

DOI: 10.1007/s11116-018-9864-9
Handle: http://hdl.handle.net/1942/26102

Exploratory analysis of Zipf's universal power law in activity schedules

Wim Ectors [(a)] (*corresponding author)
Email: wim.ectors@uhasselt.be
Office: +32-11-269114
ORCID: orcid.org/0000-0001-5340-084X

Bruno Kochan [(a)]
Email: bruno.kochan@uhasselt.be
ORCID: orcid.org/0000-0001-6456-2799

Davy Janssens [(a)]
Email: davy.janssens@uhasselt.be
ORCID: orcid.org/0000-0003-4809-5363

Tom Bellemans [(a)]
Email: tom.bellemans@uhasselt.be
ORCID: orcid.org/0000-0001-6839-3072

Geert Wets [(a)]
Email: geert.wets@uhasselt.be
ORCID: orcid.org/0000-0002-5485-9705

[a] Hasselt University, Transportation Research Institute (IMOB)
Agoralaan, 3590 Diepenbeek, Belgium

Abstract:
People's behavior is governed by extremely complex, multidimensional processes. This fact is well-established in the transportation research community, which has been working on travel behavior (travel demand) models for many years. The number of degrees of freedom in a person's activity schedule is enormous. However, the frequency of occurrence of day-long activity schedules obeys a remarkably simple, scale-free distribution. This particular distribution has been observed in many natural and social processes and is commonly referred to as Zipf's law, a power law distribution.

This research provides evidence that activity schedules from various study areas exhibit a universal power law distribution. To this end, an elaborate analysis using 13 household travel surveys from diverse study areas discusses the effect of proportional outlier removal on the power law's exponent value. Statistical evidence is provided for the hypothesis that activity schedules in all these datasets exhibit a power law distribution with a common exponent value.

The study proposes that a Zipf power law could be used as an additional dimension within a travel demand model's validation process. Contrary to other validation methods, no new data is required. The observation of a Zipf power law distribution in the generated schedules appears to be a necessary condition. Additionally, the universal activity schedule distribution might enable the full integration of activity schedules in models based on universal mobility patterns.

Keywords: *Zipf's law, power law, activity schedules, universal distributions*.

1. Introduction

People's behavior is governed by extremely complex, multidimensional processes, characterized by many degrees of freedom. Behavioral research makes it possible to model the participation in activities in e.g. so-called activity-based models, which can be used to model personal transportation as a derived demand. Such research has revealed that modeling the four basic dimensions of activities (activity type, duration, location and sequence (order)) are subject to many types of constraints and optimization problems. Additionally, it has become clear that any reliable behavioral model actually requires the understanding of activity decisions considering the *whole household*. This makes understanding and modeling people's activities (i.e. the schedules of people during the day) a daunting task. (Arentze et al. 2000; Doherty and Axhausen 1999)

Yet despite the apparent complexity of activity schedules, this paper will demonstrate that the frequency of occurrence of day-long activity schedules obeys a remarkably simple, scale-free distribution. This is a noteworthy finding given the apparent countless number of constraints and optimization requirements for people's activities. By analysing several datasets from different study areas, evidence is provided that it is a *universal* distribution, i.e. it is valid for activity schedule frequencies in many, if not all, study areas. The discovery of this new universal law provides further insight in mobility behavioral patterns. Other universal laws within this domain (in the form of power law-like distributions) have been evidenced in i.a. displacement distance, activity duration, gyration radius and location visiting frequency (González et al. 2008; Noulas et al. 2012; Riccardo et al. 2012; Zheng et al. 2016), as well as in location visiting duration (Brockmann et al. 2006). Power laws were found in transportation networks as well Yang et al. (Jiang 2009; Paleari et al. 2010; Saberi et al. 2017; Wang et al. 2015; Yang et al. 2014). Another well-known universal law is the BreVer law stating a universal daily travel time budget on an aggregated level (Hupkes 1982).

Observed universal laws are also being used in modeling applications. For example, Simini et al. (2012) proposed the universal radiation model which predicts mobility and migration patterns. Jiang et al. (2016) presented a framework to simulate spatiotemporal patterns of human mobility (TimeGeo) based on universal mobility pattern distributions. Pappalardo and Simini (2017) developed a conceptually very similar framework (DITRAS), yet in a more data-driven fashion being completely parameter-free.

The validation of microscopic simulation models, i.e. the verification that a model sufficiently replicates reality, is not straightforward (Toledo and Koutsopoulos 2004). Almost all current validation strategies seem to involve the collection of new (unseen) data to compare with the model's output, whether it be a visual or statistical validation (Boyce and Bar-Gera 2003; Park, B.; Schneeberger 2003; Toledo and Koutsopoulos 2004). However, the universal distribution revealed in the current paper could be used as an additional dimension, a *necessary* condition, within a travel demand model's validation process. Contrary to other validation methods, such additional validation instrument would not require any new data.

In conclusion, the motivation for this work is threefold: (i) investigating universal laws advances the understanding of mobility behavior. The demonstrated universal distribution for daily activity schedules (a very complex subject) is highly remarkable. Additionally, (ii) universal mobility patterns are being used in modeling travel demand or other mobility-related models as they facilitate nearly parameter-free, lightweight models. The revealed universal law may facilitate the inclusion of detailed activity patterns in these models. Finally, (iii) it is suggested that a Zipf power law could be used as an additional

dimension within a travel demand model's validation process as a *necessary* condition, not requiring any additional validation data.

The remainder of the paper is structured as follows: after the Introduction section the Literature review section discusses Zipf's law and revises related work in more detail. Next, the Methodology section describes the data used in this study, as well as the methods used to estimate the power laws and to conduct the statistical analyses. The Results section illustrates the outcomes of applying this methodology to the 13 datasets in this research. This section is followed by a Discussion section on potential applications, a hypothetical generation mechanism and outlier data. The Conclusion section finalizes the paper.

2. Literature review

2.1. Zipf's law

This particular distribution, now also observed in day-long activity schedules, has intrigued researchers from various domains and is commonly referred to as Zipf's law. It obeys a power law distribution and has been observed in many natural and social processes. It was first observed by Felix Auerbach in (1913) (Auerbach, cited in Zipf (1949) and Newman (2005)). He discovered that city size is governed by such a power law. Willis (1922), cited in Chen (1980), noted in 1922 that the size distribution of biological genera follows a power law distribution. Zipf, an American linguist, described a power law distribution in word frequency in 1949 (although it had first been noticed by Estroup in 1916 (Estroup, cited in KiBaek et al. (2011)). Zipf famously investigated this distribution in more detail, revealing that the same power law distribution holds for a large number of events in different domains, ranging from sizes of earthquakes, annual income of companies, solar flares, to the number of citations received on papers, social media check-in data (Fujiwara 2004; Furusawa and Kaneko 2003; Ma et al. 2017; Maillart et al. 2008; Newman 2005; Okuyama et al. 1999). More examples are listed in e.g. Clauset (2009) and Newman (2005).

The rank-size interpretation of Zipf's law is most common. For example: within the context of city sizes, the size of a city at rank r_i scales with a factor $1/r_i$ relative to the size of the largest city. The second largest city is for example half the first city's size, the third largest one-third its size etc.:

$$f(r_i) = \frac{f(r_1)}{r_i} \tag{1}$$

where f represents *frequency* and r the *rank*. In other words, the size of a city is inversely proportional to its rank. The data obeys a power law distribution with exponent close to 1.0.

In its more formal form of a Pareto distribution, Zipf's law states that the probability for a city to have a size greater than S decreases as $1/S$:

$$P(s > S) = aS^{-\varsigma} \tag{2}$$

where s is the size of a particular city, ς denotes the exponent (~1 for Zipf's law) and a a scaling factor. This relation should be valid over a large range of sizes S. (Gabaix 1999)

A power law distribution has characteristics of a heavy-tailed distribution, meaning that the tail of rare events has a higher frequency than in an exponential distribution (which is characterized by a stronger frequency decay). Another property is its scale invariance. The Pareto principle, or 80-20 rule, states that for many events in nature or society, approximately 80% of the effects come from 20% of the causes. A similar behavior is present for a power law distributed quantity.

Thus far no conclusive proof exists (to the authors' best knowledge) rejecting the existence of a natural power law mechanism, nor does a final consensus exist on why the

empirical Zipf's law seems to surface in a multitude of domains. Without going into detail, mechanisms leading to an observed power law distribution have been described by (Chen 1980; Corominas-Murtra and Solé 2010; Gabaix 1999; Ioannides and Overman 2003; Jiang and Jia 2010; Ki Baek et al. 2011; Li 1992; Marsili and Zhang 1998; Newman 2005; Reed 2001).

The fact that the law is so ubiquitous and many of the power law observations share the same exponent desires a universal mechanism explaining the peculiar findings. Most researchers however agree that several mechanisms lead to the power law distributions in different domains. Newman (2005) presents an overview of several mechanisms leading to a power law distribution which were previously proposed in literature.

Some work objects against the general applicability of such power laws. Ioannides and Overman (2003) built upon Gabaix (1999)'s work and investigated the validity of Zipf's law in case of non-constant city growth rates conditional on city size (i.e. as a time-varying geometric Brownian motion, also depending on city size). In a large-scale study by Soo (2005), city sizes in 73 countries from across the world were analyzed for conformity with Zipf's law. Two methods were utilized: OLS and the Hill estimator. Both methods reject Zipf's law in more cities than expected. Nitsch (2005) arrives at the same conclusion after a meta study including 515 estimates from 29 studies. He concludes that the power law exponent is statistically different from Zipf's value of 1.0, actually being closer to 1.1. Lü, Zhang and Zhou (2013) evidenced that not all language display a Zipf’s law rank-frequency relation. Non-Indo-European languages such as Chinese, Japanese and Korean have an exponential decay instead of a power law.

2.2. Power laws within the domain of transportation

References to Zipf's law within the domain of transportation sciences are thin, though there are many references to universal laws (in many cases with a power law distribution).

For example, Guidotti et al. (2015) proposed a two-step clustering algorithm for personal locations detection (*TOSCA*). They created synthetic data to compare their algorithm to that of other researchers. In this dataset, the number of visits in every location was governed by a Zipf power law distribution. This property of human trajectories was described by González et al. (2008). Based on two large long-term trajectory datasets, they found that the probability of a user visiting a location of a certain rank L follows $P(L) \sim 1/L$, i.e. Zipf's law. They also revealed that the distribution of displacements (not trips) obeys a truncated power law of the form:

$$P(\Delta r) = (\Delta r + \Delta r_0)^{-\beta} e^{-\Delta r/\kappa} \quad (3)$$

where Δr is the displacement, β a power law exponent, and Δr_0 and κ cutoff parameters. This behavior can be well-approximated by:

$$P(\Delta r) \sim |\Delta r|^{-1-\alpha} \quad (4)$$

with $0 < \alpha \leq 2$. Random trajectories governed by this equation are also called Lévie flights (González et al. 2008; Klafter et al. 1996). Noulas et al. (2012) found universal laws in rank-distance data from social network check-ins (based on the number of places between origin and destination). Their finding was inspired by the theory of intervening opportunities. Riccardo et al. (2012) discovered universal distributions while studying human mobility characteristics such as the trip length distribution, activity duration and the rank distribution of the visitation frequency. Preferential attachment is clearly observed. Also on the topic of trip lengths, Zheng et al. (2016) found that short (in the sense of duration) trips follow an exponential power law, and medium to long trips obey a truncated Pareto model. Trips were separated by travel purpose and the fitted distributions, as well as the breakpoints between the two regimes, were discussed. Their research, in contrast to the current research, only

considers trip duration distributions at the highest level of aggregation and it does not consider personal schedules.

A truncated power law was observed in the distribution of the radius of gyration. Brockmann et al. (2006) additionally found that the time spent by an individual at a location, Δt, follows a fat-tailed, power law-like distribution

$$\mathrm{P}(\Delta t) \sim |\Delta t|^{-1-\beta} \qquad (5)$$

with $0 < \beta \leq 1$. Song et al. (2010) combined (4) and (5) with concepts of preferential return and exploration to obtain a microscopic model for individual human mobility.

There is evidence for power law distributions in betweenness centrality and traffic flow in the road network (Jiang 2009). This relates to the concept of space syntax, which encompasses a set of theories to analyze space and its configuration while considering the implications on i.e. social interactions and pedestrian movement. The *integration*, i.e. a centrality measure expressing whether a node in the spatial network is integrated or segregated from the rest of the urban texture, follows a rank-size power law distribution similar to Zipf's law. Other parameters in the space syntax domain also seem to obey a power law distribution (Makse et al. 1998; Shen and Karimi 2016; Volchenkov and Blanchard 2008). In his research, M. Batty investigated fractal geometry and its relation to i.a. urban dynamics and scaling laws in cities (e.g. Batty 2008; Batty 2005).

Also considering networks, Yang (2014) found power law distributions in their study of bus transport networks (BTNs) in China. They described an ESL network in which they observed a power law distributed edge length. The edge lengths in other network types (ESP and ESW) were found to have exponential and linear distributions. A power law distribution was also observed in airport networks: the distribution of degree (number of edges starting from a node or airport) has a Pareto double power law distribution (Paleari et al. 2010). Wang et al. (2015) applied a place rank algorithm to the Singapore public transport network. They found that place rank (inspired by the PageRank approach and representing the accessibility of an area) and travel time have a power-law relationship. Saberi et al. (2017) investigated origin-destination matrices of multiple cities as large-scale directed graphs. They found that these travel demand networks exhibit remarkably similar properties (displaying universal, scale-free power law distributions) despite differences in underlying topography and urban structure.

Universal laws are also adopted in the travel demand modeling community. Simini et al. (2012) proposed the universal radiation model which predicts mobility and migration patterns. The model's universality comes from the fact that it is parameter-free (it only depends on population density). Jiang et al. (2016) presented a framework to simulate spatiotemporal patterns of human mobility (TimeGeo). Call detailed records (CDR) are used to extract and calibrate universal mobility pattern distributions, which are then used to generate the duration of activities, the number of visited locations per day and the daily mobility networks for synthetic individuals. Few tunable parameters are required to generate individual mobility patterns. The framework does not account for different activity types other than `home', `work' and `other'. Similarly, Pappalardo and Simini (2017) developed a conceptually related framework to simulate spatiotemporal patterns of human mobility (DITRAS) by first generating mobility schedules and subsequently converting those into mobility trajectories. These two components operate in a very data-driven fashion and are parameter-free. The synthetic mobility trajectories are compared with universally observed aggregated mobility characteristics. Yet, also this model only considers time and duration of visits at various locations, without accounting for the activity type.

In addition to advancing the understanding of mobility behavior, and as discussed in the previous paragraph, several researchers investigate universal mobility patterns and attempt to use them in modeling travel demand or other mobility-related components. The current research illustrates an apparent universal property of activity schedules which was not observed before, which could also benefit models such as discussed above since these do not take into account detailed daily activity schedules.

Comparable to the current research, some other studies also directly worked on generating activity schedules. Schneider et al. (2013) studied *daily* mobility patterns as they observed how many of the universal laws only capture long-term mobility behavior. They focused on temporal and spatial properties of day-long schedules, discarding i.a. activity type, travel time and activity duration characteristics. By contrast, the current research explicitly takes into account the activity type. Schneider et al. found how up to 90% of all schedules can be described in 17 distinct mobility patterns. This is consistent with findings from Ectors et. al. (2016b), who proposed a clustering method to determine activity skeleton schedules, though taking into account the activity type rather than trip destination.

Another study by Li and Lee (2017) used context-free grammar techniques to generate sequences of activities (similar to the activity schedules in the current research). A grammar in this context is a set of rules for combining elements (words/activities) into sequences (sentences/schedules). They found considerable analogies between these schedules and natural languages. The proposed grammar was able to generate activity sequences with a very similar distribution as the source household travel survey (HTS), though admittedly only three distinct schedules already represented 80% of the HTS observations. This work did not reveal the universality in the distribution of these schedules.

3. Methodology

This paper provides statistical evidence that activity schedules exhibit a power law distribution with a common exponent value, i.e. that out-of-home activity schedules obey a universal power law distribution. In the methodology section, first the data from multiple study areas will be described. Secondly, the methodology for establishing the fact that a power law distribution governs a dataset is discussed. Thirdly, statistical tests combined with a proportional cutoff are discussed as an approach to thoroughly analyze the power law exponent estimate across different study areas.

3.1. Data description

In this research, an out-of-home activity schedule is constructed out of trip purpose information from HTSs. Trip purposes are concatenated into a sequence which represents a schedule of the main out-of-home activities. Location information is used as well in cases where trip purpose used a tour-definition rather than a trip-definition. An example of a tour-definition in the trip purpose variable is the series of trip motives ‘commuting-commuting’ for an individual going to work and afterwards returning from work. In this research, such a series of trip motives is interpreted as ‘work-home’ by taking into account location information as well.

Throughout this work, a trip is denoted as the movement from one location where an activity was performed to the next. A trip may consist out of several trip components, for instance because of mode changes. Trip components are not considered in this research. A tour is defined as a chain of trips, usually starting and ending at the home location.

Because usually only the main trip purpose is recorded in a HTS, a known error occurs whenever multiple activities were performed at one location without an intermediate trip. The activity schedules used within this research represent chains of out-of-home

activities, depending on trip purposes. Additionally, the first activity of the day is assumed to be 'home' as this information is not consistently present in the HTS data. This is a known (small) error.

Frequency tables can be generated from the activity schedules. They represent the observed distribution of activity schedules within a certain study area. Table 1 lists the top-10 of most frequent schedules within the US HTS (USA NHTS 2009) dataset as an example. In total this dataset contains more than 83,000 distinct schedules based on 257,586 travel diaries. In case individual weights are available, these are taken into account. Individual weights correct any biasness which might be present in the travel survey's sample (i.e. from surveying too many persons of a particular profile). Some of the weights not only correct the biasness, but also augment the records to year-long observations for all such individuals in the population.

A considerable amount of different datasets was collected for the purpose of this research. This set includes 13 different HTSs from across the world. The dataset names, country of origin and the number of (valid) extracted single-day schedules are listed in Table 2. Some datasets contain multiple years of observations. Careful testing (not described in this paper) revealed that activity schedule distributions for a particular study area do not significantly change in a period of a few years, and hence there was no need to separate them. However, activity schedule distributions from different study areas do exhibit (slightly) different activity schedules at each rank, although this research will demonstrate that their rank-frequency distributions are very similar. The selection of datasets to be used in this research was based on ease of access and cost. Apart from data freely available online and data available to the department, several institutions were contacted to obtain HTS microdata, also guided by the useful report by Ahern et al. (2013).

Table 1: Top-10 most frequent activity schedules in USA NHTS 2009. A 'home' activity is abbreviated as 'H'.

Schedule	Interpretation	Frequency. [%]	Rank
1-11-1	H – Work - H	8.82	1
1-21-1	H – Education - H	4.39	2
1-41-1	H - Buying groceries/clothing/… - H	3.17	3
1-53-1	H - Visiting friends/relatives - H	1.67	4
1-51-1	H - Going to the gym/exercising/… - H	1.21	5
1-22-1	H - Religious activity - H	1.16	6
1-30-1	H - Medical/dental service - H	0.85	7
1-82-1	H - Going out to eat - H	0.83	8
1-41-41-1	H - Buying groceries/clothing/… - Buying groceries/clothing/… - H	0.78	9
1-54-1	H - Going out (entertainment, going to a bar…) - H	0.63	10

Table 2: HTS datasets used within this research.

Dataset	Country of origin	Number of extracted schedules
USA NHTS 2009	United States of America	257,586
NLD OViN 2013	The Netherlands	34,710
BEL OVG 3.0-4.5	Belgium (Flanders)	13,522
SVN Ljubljana 2013	Slovenia	3,426
GBR NTS 2009-2014	United Kingdom	551,234
KOR Seoul HTS 2010	South Korea	219,269

DEU Mobidrive 1999	Germany	13,244
CHE Thurgau 2003	Switzerland	8,522
BEL Beldam 2010	Belgium	11,279
IRL NTS 2009	Ireland	5,023
FIN HLT 2010-2011	Finland	10,137
SWE RVU 2011-2014	Sweden	31,457
AUS VISTA 2007 & 2009	Australia	67,060

3.2. Power law estimation methods

The PDF of a power law distribution takes the form of Equation (6):

$$p(x) = Cx^{-\alpha} \tag{6}$$

where C a constant and α the exponent of the power law. When fitting a power law on non-numeric data such as activity schedules in this research (similar to words in a text; see also Equation (1)), one fits a power law on rank-ordered (frequency) distributions of the data. Consequently, one will estimate the parameters in $f(n) = C'n^{-\tau}$ where n the (relative) frequency in the rank-ordered distribution and τ the so-called Zipf exponent. These two exponents, α and τ, are related by Equation (7) (Adamic and Huberman 2002; Hanel et al. 2017):

$$\alpha = 1 + \frac{1}{\tau} \tag{7}$$

The estimates in this paper are those based on Equation (6), that is the estimates tabulated are $\hat{\alpha}$. Zipf's exponent $\tau = 1$ yields an expected $\alpha = 2.0$ according to Equation (7) in order to confirm Zipf's law in activity schedules.

In order to claim with relative certainty that the empirical Zipf's law can be observed in the domain of transportation behavior, a power law distribution should be fitted to the data. Often this fitting exercise is performed as a linear regression (least-squares) with logarithmic transformed variables, however this method is flawed as explained by (Clauset et al. 2009; Newman 2005; Urzúa 2011). The slope estimate may suffer from systematic, large errors. Clauset et al. (2009) proposed a method based on a maximum likelihood estimation (MLE) combined with the Kolmogorov-Smirnov (KS) goodness-of-fit statistic. The R package called "poweRlaw" (Gillespie 2015) was developed to automate this process.

As the power law probability distribution $p(x) = Cx^{-\alpha}$ with $\alpha \geq 1$ diverges for $x \to 0$, resulting in an infinite area under the distribution, an additional parameter x_{min} needs to be specified. This parameter depicts the fact that few observed distributions follow a power law distribution across their entire range; most only have a power law in the (upper) tail. Observations with a value smaller than the cutoff parameter x_{min} will be excluded. The x_{min} parameter may be optimized using the KS goodness-of-fit statistic. Under this optimization, x_{min} is selected such that the maximum distance between the observed distribution and the hypothesized power law reference distribution is minimized.

A bootstrapping procedure allows to investigate uncertainty in the exponent estimate. It randomly selects data with replacement, and then executes a MLE step with KS cutoff on that sample. This is repeated a number of times to evaluate uncertainty (Gillespie 2015). In this paper, 2000 iterations were used on all datasets (in some cases even 5000 iterations).

The PoweRlaw R package also conducts a (special) hypothesis test and reports the resulting p-value. The null hypothesis states that a power law distribution is a good fit, the alternative hypothesis states that another distribution might be better. Clauset (2009) recommends a 10% significance level in this test.

3.3. Statistical analyses

To investigate the general applicability of Zipf's law, one approach could be to use the bootstrap samples of two (or more) study areas to infer information about the true (i.e. population) exponent mean. A Welch t-test, which does not assume normality and does not assume equal, known variances, can be used to test whether the mean power law exponent values of different study areas are equal. Alternatively, the Wilcoxon rank sum test (unpaired test), a non-parametric, distribution-free rank-based test, may be used. It assumes that two distributions have the same shape and spread, but they don't necessarily have to be normally distributed. It too can be used to test for equal means in the bootstrapping results for the power law exponent. Under a second approach, the hypothesis that the true power law exponent value of a study area is equal to the value expected under Zipf's law (i.e. 2.0; see section 3.2) could be tested.

The above tests are very sensitive to the sample size. In the bootstrapping procedure, this sample size can be chosen at will by changing the desired number of iterations, being only limited by computing power and available time. During experimentation with the above hypothesis tests it was found that sample sizes of less than ± 200 will not reject the null hypotheses and sample sizes of over ± 200 will lead to a rejection of the null hypothesis. Each bootstrap sample is entirely based on a single observed distribution, hence the bootstrapping procedure does not add information, as is the case in other examples where an increase sample size will lead to hypothesis test with greater confidence level. As a consequence, the use of the bootstrapping samples in the above tests might intrinsically be flawed and should be avoided, as one can manipulate the outcome by controlling the sample size.

Additionally, as the KS cutoff approach is used in the bootstrapping uncertainty procedure, one implicitly also tests the performance of this cutoff criterion in the above tests. Clearly a different approach should be followed.

Summarizing, another method which avoids the use of the KS criterion is needed. Still, some sort of x_{min} cutoff approach is necessary. For this, a proportional cutoff methodology was used. The value of x_{min} (in a particular dataset) is determined so that a certain percentage of outlier-like schedules is excluded. In other words, a certain cumulative percentage of the most rare schedules is excluded. If for example an exclusion fraction of 0.2 (20%) is specified for the USA NHTS 2009 dataset (see Table 2), 51,517 of the most rare schedules will be excluded. Note that these are *not 51,517 distinct* schedules, but rather the bottom cumulative 20% in the frequency table. The proportional cutoff criterion was used as an alternative to the KS cutoff in the PoweRlaw R package.

There is no rationale on what exact value the exclusion fraction should take. Instead, this value was varied and for each exclusion fraction the corresponding power law exponent value was estimated through MLE. Each estimate of a particular dataset can be considered as a single independent data point. Estimating the exponents of the 13 datasets yields a (small) sample for each given exclusion fraction, upon which statistical tests were performed. Student's t-tests and weighted t-tests (using 'wtd.t.test' from the 'weights' R package) were conducted. The number of extracted schedules listed in Table 2 (a measure for the dataset's size) serves as the weight variable.

4. Results

This section presents the results from the methodology discussed before. This research attempts to provide evidence that activity schedules exhibit a power law distribution with a common exponent value (i.e. a universal power law distribution). To this end, activity

schedules are constructed out of trip purpose information in HTS data. Fig. 1 shows the frequency distribution of the schedules from each data source on a log-log plot. One observes the remarkable consistency across the various study areas. A curiously uniform linear trend exhibits itself starting at the most frequent schedule (at rank 1) and continuing up to schedules with a frequency of several orders of magnitude lower. At this lower end of the frequency spectrum, the linear trend breaks down. The plot exhibits the characteristics of a power law distribution. Such distribution is easily recognized by its linear manifestation on a log-log plot (although this fact alone is inconclusive). As discussed in the methodology section, there often exists an x_{min} below which a power law distribution cannot be maintained. Such x_{min}, or final schedule rank in this case, may visually be interpreted as the point where the linear trend in Fig. 1 breaks down.

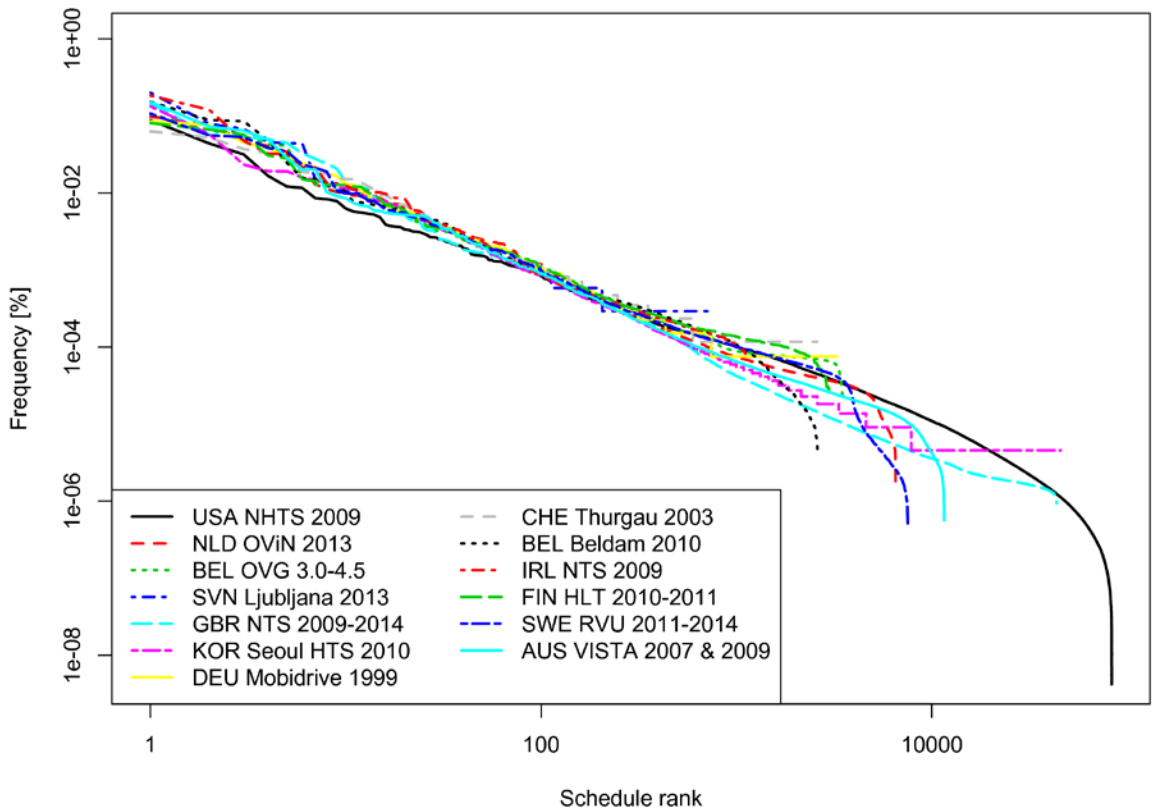


Fig. 1 Activity schedule frequency distribution for several HTS datasets as a function of schedule rank. For each survey area, the original activity type encoding was used. Multiple-day survey data was converted into single-day schedules. The horizontal axis shows schedule rank (ID within a descending frequency table) and the vertical axis the relative frequency as a percentage of the total number of schedules.

4.1. MLE combined with KS estimation results

Apart from the visual analysis in Fig. 1, the distributions may also be analyzed quantitatively. As explained in the methodology section, a MLE step combined with the KS cutoff criterion was used to most optimally fit a power law distribution to the observed activity schedule distribution of each dataset. The R package poweRlaw (Gillespie 2015) was used to this end. Table 3 summarizes the most relevant results for each of the 13 datasets. The first part of the table lists the MLE + KS cutoff results, while the second part lists the results from the bootstrapping uncertainty evaluation.

To confirm a Zipf power law distribution, an exponent value of 2.0 should be observed in this table. All but two of the 13 datasets in the first part of Table 3 yield an estimated exponent value of approximately 2.0, which is an extraordinary observation. The mean value of the exponent estimates is 2.153.

Table 3: Estimation results for activity schedule distributions in several HTS datasets. α represents the estimated power law exponent value, x_{min} the cutoff value, 'Cum. Pct rejected' the cumulative percentage of data rejected because of the cutoff value, KS the Kolmogorov-Smirnov distant measure, AM the arithmetic mean and SD the standard deviation.

	poweRlaw estimations (MLE + KS)				Bootstrapping uncertainty evaluation		
Dataset name	α	x_{min} [c]	**Cum. Pct rejected**	**KS**	**AM(α)**	**SD(α)**	**P-value**
USA NHTS 2009	2.003	36,809,977	55%	0	2.006	0.070	0.255
NLD OViN 2013 [a]	1.885	1325	20%	0.034	1.877	0.058	0.005
BEL OVG 3.0-4.5 [a]	1.947	2	23%	0.031	1.953	0.051	0.149
SVN Ljubljana 2013 [a]	1.995	9	31%	0.079	2.028	0.208	0.449
GBR NTS 2009-2014	1.862	4.071	9%	0.023	1.869	0.117	0 [d]
KOR Seoul HTS 2010	1.859	26	32%	0.031	1.869	0.048	0.156
BEL Beldam 2010	2.067	869	7%	0.025	2.067	0.055	0.097
IRL NTS 2009 [b]	1.842	1,520	20%	0.042	1.983	0.118	0.415
DEU Mobidrive 1999	2.053	23	52%	0.054	2.002	0.133	0.714
CHE Thurgau 2003	1.929	16	49%	0.074	2.009	0.113	0.317
FIN HLT 2010-2011 [b]	1.965	1,175	30%	0.045	1.980	0.096	0.016
SWE RVU 2011-2014 [b]	3.309	1,049,401	75%	0	2.096	0.172	0.401
AUS VISTA 2007 & 2009	3.270	360,498	66%	0	2.477	0.838	0.783

[a] *Based on 5000 bootstrapping iterations instead of 2000*
[b] *Trip purpose encoding made compatible (tour- to trip-definitions; see methodology)*
[c] *The large differences in the magnitude of x_{min} are due to different kind of weights used*
[d] *Could not be calculated reliably*

The first outlying observation is that belonging to the Swedish HTS (SWE RVU 2011-2014) dataset. The estimated exponent value is much higher than initially expected. However, upon closer inspection, this most likely is caused by the KS cutoff criterion. Because of this criterion, 75% of the data was excluded in the MLE step and consequently the exponent estimate was based only on the 4 most frequent schedules (combined accounting for about 25% of the schedules). The KS distant measure is 0, indicating a (nearly) perfect fit to the observed 4 schedules. In this particular case the KS criterion is inappropriate, a fact confirmed by the bootstrapping uncertainty evaluation where a mean exponent value of 2.096 was calculated for this dataset based on 2000 iterations. A cutoff based on the KS criterion is sensitive to small irregularities in the distribution and will aggressively exclude more data in favor of minimizing the maximal difference between fitted and observed distribution. For example, irregularities such as small 'bumps' might cause the exclusion of all data after (and including) the 'bump'.

The same explanation holds for the second outlying observation, the Australian HTS (AUS VISTA 2007 & 2009). Also in this case the KS cutoff criterion was too stringent and a high exponent value of 3.27 was estimated. The bootstrapping algorithm yields a much lower mean exponent value of 2.477. This value is still somewhat higher than expected and is

accompanied by a relatively high standard deviation. This is due to a slight sinusoidal-shaped pattern in the high-frequency region of the distribution.

One also observes how the KS cutoff criterion is responsible for removing 7% up to 55% of the data (not including the two outlying observations), indicating that the KS criterion may be quite stringent in some cases where one would visually not expect it (see Fig. 1). The KS criterion is consequently quite good (well-optimized) and low values are tabulated.

The bootstrapping procedure yields results which favor the claim a universal exponent value of approximately 2.0, as listed in Table 3. The mean of the AM(α) values is 2.016, again extremely close to the expected value of 2.0. Fig. 2 shows the bootstrapping estimation results in a graphical manner. These results seem to cluster around the expected value of 2.0.

Table 3 additionally lists the p-values resulting from a bootstrapping procedure. As detailed by Clauset (2009), a p-value greater than 0.10 would not reject the null hypothesis of a power law distribution being a good distribution to fit the data. Only 3 datasets would reject the null hypothesis. The Dutch HTS (NLD OViN 2013) has a p-value of 0.005 and the Finnish HTS (FIN HLT 2010-2011) has a p-value of 0.016. The UK HTS (GBR NTS 2009-2014) yielded a p-value of 0, which is interpreted skeptically as this p-value (for an unknown reason) showed no evolution at all during the 2000 bootstrapping iterations. A possible cause for the result for the other two datasets could be that a power law distribution is only appropriate for the upper tail, and another distribution (lognormal distribution?) is more fitting for the rest of the data. A lognormal and power law distribution have very similar characteristics in the upper tail of such distributions. Yet, the large majority of the datasets does not reject the hypothesis that the power law distribution is an appropriate distribution to describe the data.

On a side note, Table 3 lists the results of in total several *weeks* of computing time on two servers running 20 threads at 80% CPU usage. The first one is equipped with 2x Intel Xeon (X5670) CPUs, the second one with 2x Intel Xeon E5-2643 v2 CPUs. The bootstrapping procedure, and in particular the use of the KS cutoff optimization, is the main responsible for the significant computational effort. The required computing time increases strongly with increasing number of distinct schedules in a dataset.

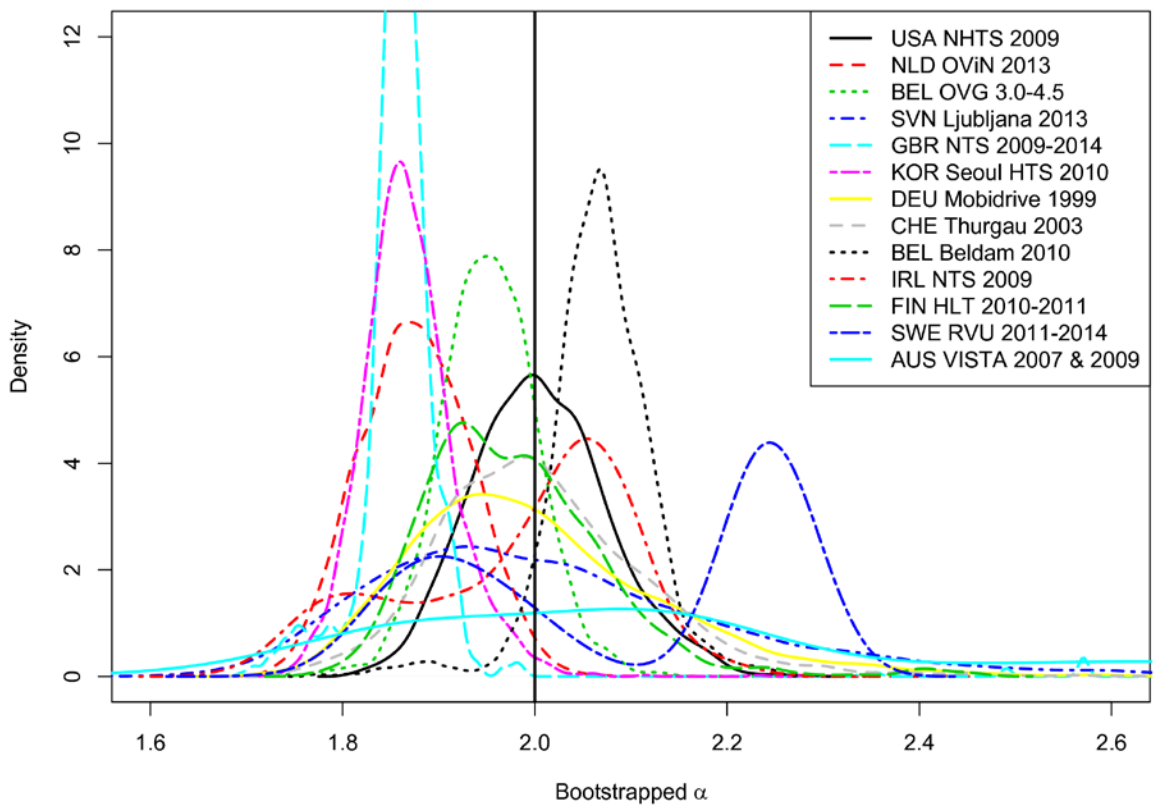


Fig. 2 Kernel density plot of the bootstrapped power law exponent estimates. Visually, one observes a large cluster of exponents around the value of 2.0.

4.2. MLE combined with a proportional cutoff

As explained in the methodology section, the x_{min} cutoff value may be varied in function of a certain cumulative percentage of data to be excluded (instead of using the KS criterion). The power law exponent value can subsequently be estimated through MLE. For each value of x_{min} this results in a small sample of exponent values from the different datasets, which can be used to perform an hypothesis test for a true mean value of 2.0. In this approach, one considers each power law exponent estimate of a particular dataset as a single independent data point.

Fig. 3 illustrates the estimated exponent values in function of the cumulative fraction of data that was excluded. It shows that for all datasets the exponent estimation stabilizes after removing approximately 20% of the outliers. It is a satisfying finding that there is a good correspondence regarding this cutoff point between the datasets of different study areas. Excluding outliers is necessary in power law fitting, as also widely recognized in literature and discussed before. In the poweRlaw R package, this cutoff point is determined using a KS criterion (the x_{min} parameter).

Excluding to much of the data, i.e. more than 60%, makes the estimates unstable again, especially in case of small-sample size datasets. This is to be expected, as only few distinct schedules remain in this case. In the 'stable area' where between 20% and approximately 60% of the outliers is excluded, the exponent estimates visually seem to center around Zipf's value of 2.0.

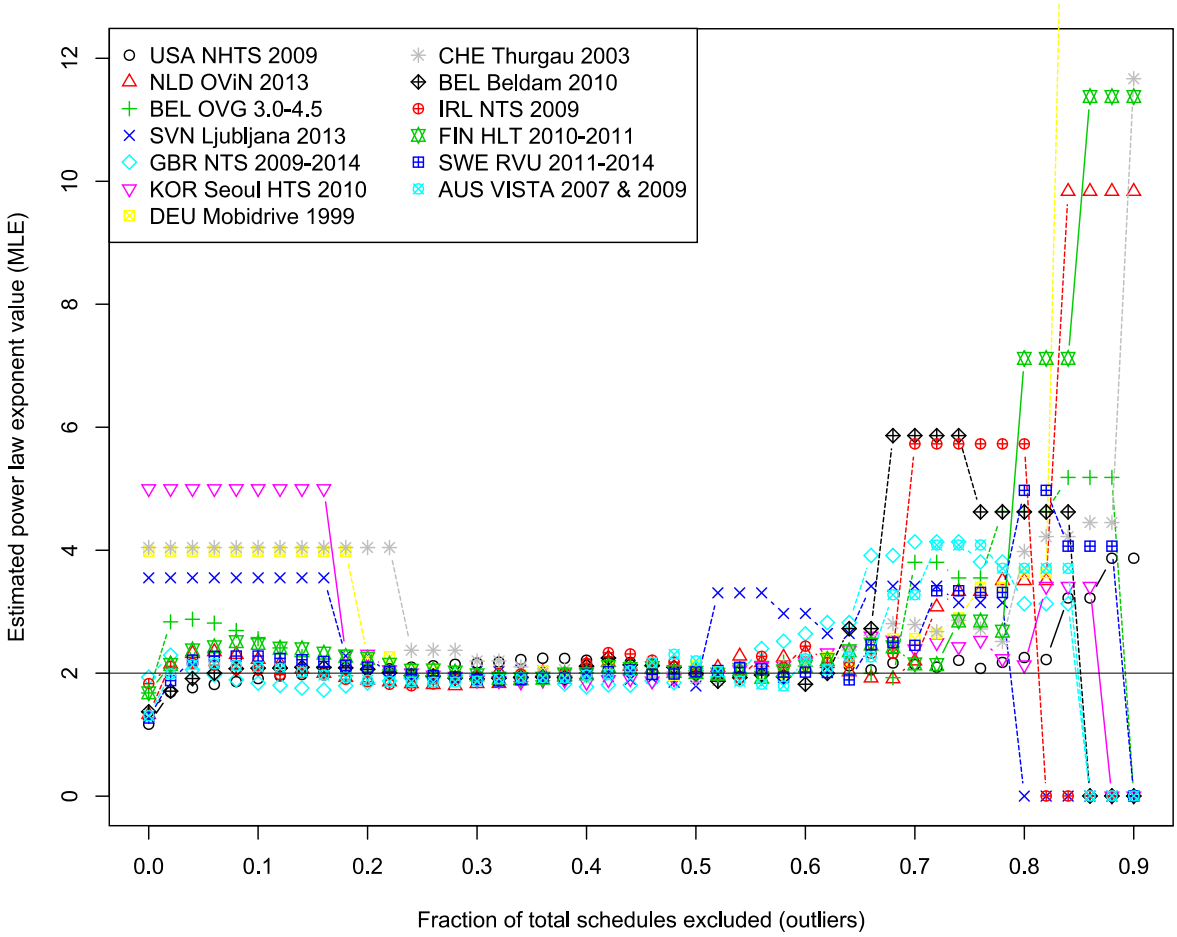


Fig. 3 Power law exponent values for several HTS datasets in function of the exclusion fraction (of outliers). This overview plot shows a range in exclusion fraction between 0.0 (no data excluded) and 0.9 (90% of the schedule data excluded).

As can be observed from the first part of Fig. 3, large datasets might already have a quite stable exponent estimate at lower exclusion levels than the 20% level, whilst smaller datasets might require a slightly larger exclusion level. One also clearly observes how some datasets display plateaus. This is due to the fact that no individual weight variable was available in these datasets, and that the schedule frequency distribution is also characterized by these plateaus of discrete frequencies. Datasets where individual weights were available are recognized by their smooth distributions. Possibly, even more consistent results would have been obtained if all datasets contained individual weights. Fig. 4 illustrates the schedule frequency distributions after removing 20% of the data. The fraction of data where the linear trend breaks down in Fig. 1 appears to have been excluded.

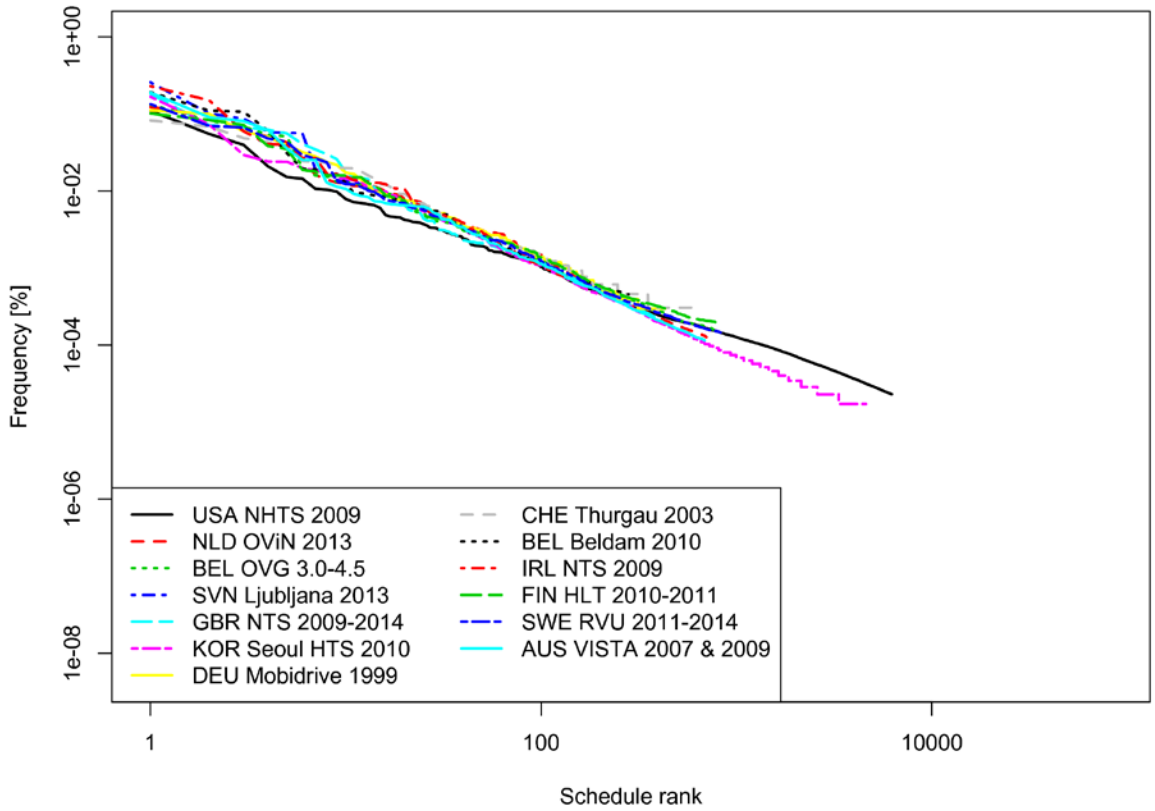
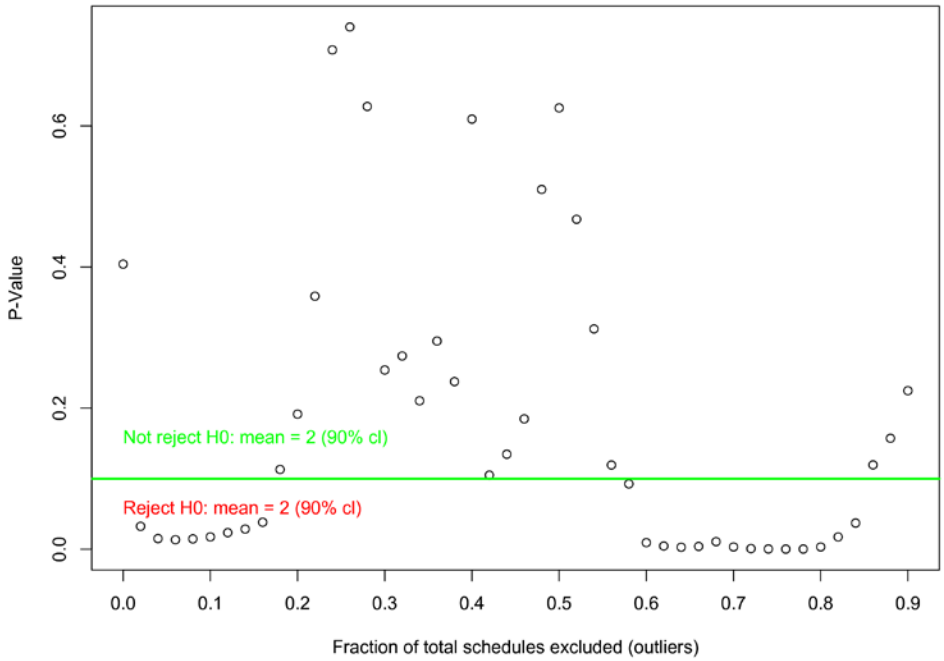


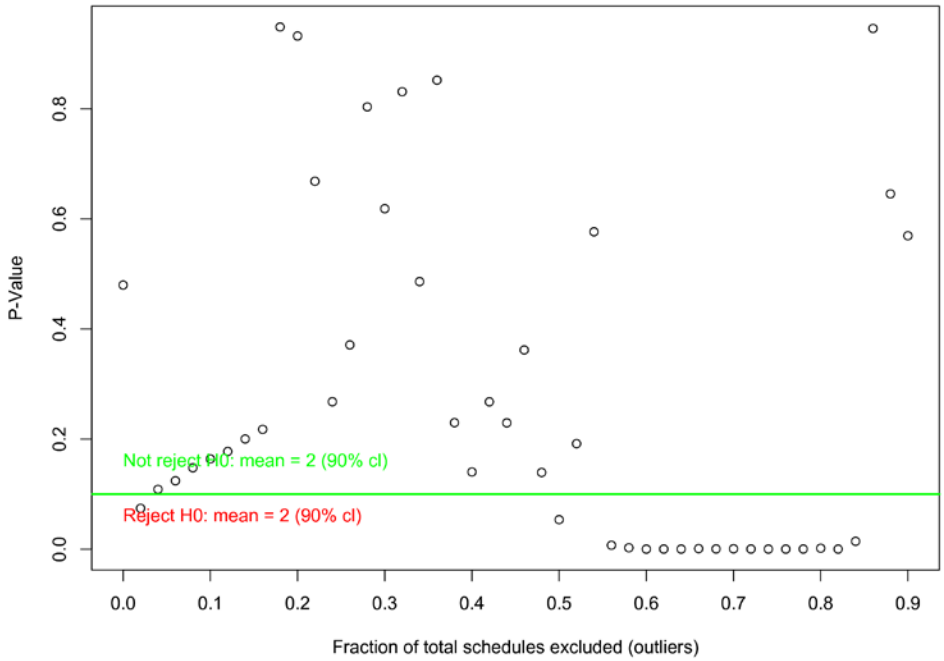
Fig. 4 Similar as in Fig. 1, this figure shows the activity schedule frequency distribution for several HTS datasets as a function of schedule rank. For each dataset, the 20% least occurring schedules have been removed as outliers.

The goal of this paper is to demonstrate that Zipf's law governs the schedule frequency distribution, the latter thus exhibiting a power law distribution with exponent value 2.0. For this reason, statistical tests were conducted. The methodology section describes the use of each power law exponent estimate of a particular dataset as a single independent data point in a (small) sample upon which statistical tests may be performed.

For each exclusion level, a Student's t-test was performed based on 13 data points (the exponent estimates from all datasets). Fig. 5 (a) shows the p-values. A significance level of 10% is indicated for convenience. Within this research, a 10% significance level is a more stringent test than the more common significance level of 5%, as it is easier to reject the null hypothesis of a universal exponent value of 2.0. It appears that, for most of the outlier exclusion levels in the ‘stable region’ (between 20% and approximately 55% of the data excluded), one cannot reject the null hypothesis that the mean is equal to 2.0, i.e. Zipf's value.



(a)



(b)

Fig. 5 P-values in function of the exclusion fraction (of outliers). The corresponding hypothesis test is: H0: The true mean power law exponent value of all study areas is 2.0. In (a) no weights were used, while in (b) a weighted t-test was used with as weights the number of schedules as in Table 2. A range in exclusion fraction between 0.0 (no data excluded) and 0.9 (90% of schedules excluded) is shown.

Considering the hypothesis test above, one might recognize that some datasets are much larger than others. This fact was not accounted for. Therefore, in Fig. 5 (b) the p-values from *weighted* Student's t-tests are shown. Each exponent estimate is weighted according to the dataset's size, i.e. according to the number of extracted schedules listed in Table 2. Using weighted t-tests, one no longer rejects (at 10% significance level) the null hypothesis for exclusion levels lower than 20%. One in general only rejects the null hypothesis for exclusion levels greater than approximately 55%. Comparing Fig. 5 (a) and (b) provides the insight that large datasets generally obey a power law distribution for a larger part of their data range than the smaller datasets. When taking into account dataset size, (almost) no data needs to be excluded in order to not reject the hypothesis of a universal exponent value of 2.0.

Fig. 6 shows the (weighted) mean and (weighted) standard deviation values for the different exclusion levels upon which the Student's t-test and its weighted variant were conducted.

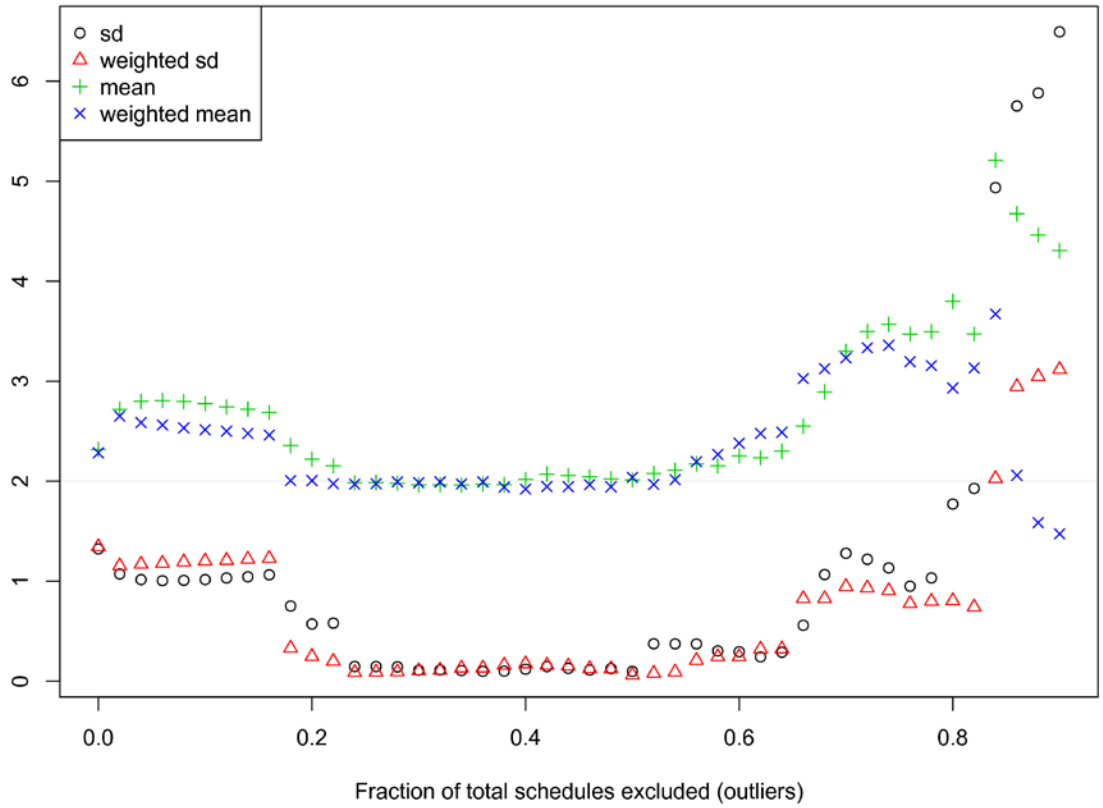


Fig. 6 Descriptive statistics for the power law exponent value of the sample of activity schedules of 13 HTSs, in function of the exclusion fraction (of outliers). For the weighted mean and standard deviation, the number of schedules as in Table 2 was used as weight variable.

5. Discussion

5.1. Application

This paper establishes that out-of-home activity schedules obey a power law distribution. This knowledge can be employed in applications such as for example model validation. The validation of microscopic simulation models, i.e. the verification that a model sufficiently

replicates reality, is not straightforward (Toledo and Koutsopoulos 2004). Almost all current validation strategies seem to involve the collection of new (unseen) data to compare with the model's output, whether it be a visual or statistical validation (Boyce and Bar-Gera 2003; Park, B.; Schneeberger 2003; Toledo and Koutsopoulos 2004). In some cases a historical validation is possible, in which the model is trained and executed for historical cases and compared to census and HTS data (from that time) which are known in the current time period (Miller et al. 2012).

However, the authors envision that a Zipf power law could be used as an additional dimension within a travel demand model's validation process. Contrary to other validation methods, such additional validation instrument would not require any new data, thus not bringing any additional cost to the project or research. The observation of a Zipf power law distribution in the generated schedules appears to be a *necessary* condition: if it cannot be observed, one can confidently state something is wrong in the model. However, one has to remark that such an observation is not a *sufficient* condition to conclude a model's validity, as a set of unrealistic schedules could still exhibit a Zipf power law distribution.

As described in the Literature review section, observed universal laws are also being used in modeling applications. Modeling frameworks such as the radiation model by Simini et al. (2012), TimeGeo by Jiang et al. (2016) and DITRAS by Pappalardo and Simini (2017) typically don't (fully) encompass activity schedules. The distribution from the current paper might enable such integrations in models based on universal mobility patterns.

5.2. Generation mechanism

At this point there is no conclusive explanation for the power law in the distribution of activity type schedules. As stated in the literature review section, many mechanisms leading to a power law distribution have been described in literature. The work by Newman (2005) presents an overview of the most often published mechanisms for generating power law distributions. One promising generation mechanism is based on the interaction of two exponential distributions. Activity schedules are i.a. characterized by (i) the number of activities performed during the day, and (ii) the probability of each activity type to occur. Looking at the available schedule data it becomes clear that schedules consisting of few out-of-home activities (in other words 'short' schedules) occur at a higher frequency than more complex schedule consisting of several activities. Additionally, some activity types take place much more often than others (think for example of 'work' vs 'medical visit').

The number of activities (or trips) is approximately exponentially distributed. In fact, one could think of it as a geometric distribution, where each additional performed activity reduces the probability on that schedule by a constant factor. The exponential distribution is a good approximation of the discrete geometric distribution. In tests on the USA NHTS 2009 data, it was found that the exponential or geometric distributions offer a good fit for a large proportion of the data. A function of the form $Ae^{(Bl+Cl^2)}$, where l represents the number of activities and A, B, C constants, offers an excellent fit across the entire range.

The probability of activity types is approximately exponentially distributed (on a rank-frequency plot) for the most frequent activity types. The entire range of activity types in USA NHTS 2009 is described by a truncated power law (a power law with an exponential cutoff) of the form $Ae^{Br}r^{C}$ where r the rank of a particular activity type and A, B, C constants.

Schedules may be generated by simply sampling from the distribution for the number of activities in a schedule, and by sampling the distribution for the activity types in order to populate that schedule (using either the two simple exponential functions, or the two more complex functions). Though perhaps different from the mechanism in Newman (2005),

generating schedules by simply sampling from these distributions yields a power law for the activity schedule frequency. It resembles constructing a probability tree where each node represents a schedule and each link adds an activity to the previous schedule with a certain probability.

A problem with this mechanism is the fact that it fails to reproduce the power law regime in the high frequency regime of the distribution, which is characterized by schedules with only one out-of-home activity (it is obvious that simple schedules occur at the highest frequency). The sampling mechanism will result in the approximate exponential distribution of the activity types for the upper range of the activity schedule frequency distribution (since the schedules here consists of only one out-of-home activity), after which it transforms into a power law distribution for the rest of the schedule data.

This issue may for example be addressed by investigating the interdependence of both distributions in future research. This research focusses however on the remarkable observation, and the demonstration of the evidence for a universal law. A conclusive explanation may be formulated in future research.

5.3. Outlier data

One might wonder why outliers might be present in the activity schedule distributions. Under the above hypothesis, the data collection might be the cause. The ‘outlier’ regime can to a large degree be explained by HTS sample size in combination with the number of distinct activity types in the study. For example, the USA NHTS 2009 HTS has 37 distinct activity types. The sample size of 257,586 schedules (see Table 2) seems however insufficient to allow ample occurrences of all combinations with these activity types. Therefore, ultimately the tail will decay faster than in the power law regime, which is seen as outlier behavior. The GBR NTS 2009-2010 HTS has 23 distinct activity types and a sample size of 551,234 schedules. As can be observed in Fig. 1, the power law regime visually extends almost across the entire range, without displaying outlier-like behavior. When investigating Fig. 3 into the detail, this is confirmed by the fact that the estimated power law exponent hardly changes across the range of excluding no data at all, to excluding approximately 55%. The personal weights may have a large influence on this part of the distribution as well.

6. Conclusion

People’s behavior is governed by extremely complex, multidimensional processes, characterized by many degrees of freedom. This paper demonstrated that the frequency of occurrence of day-long activity schedules obeys a remarkably simple, scale-free distribution which is commonly referred to as Zipf's law, a power law distribution which has intrigued researchers from various domains. By analyzing 13 datasets from different study areas from across the world, evidence is provided that it is a universal distribution, i.e. a common exponent value of approximately 2.0 defines the power law distribution in activity schedule frequencies of many, if not all, study areas.

First, activity schedules were composed from trip purpose (and activity location) data from 13 household travel surveys. Subsequently, a power law distribution was fitted to this activity schedule data using a maximum likelihood estimation approach combined with a Kolmogorov-Smirnov cutoff (lower bound on the considered data) criterion using the poweRlaw R package (Gillespie 2015). For the thirteen datasets, a mean exponent value of 2.153 was found, which is very close to the expected value of 2.0 under Zipf’s law. A bootstrapping procedure analyzed the uncertainty in the power law exponent estimates (a mean exponent value of 2.016 was found) and tested whether a power law distribution is an

appropriate distribution for the data (at least 10 out of 13 datasets did not reject this hypothesis).

Next, recognizing the limitations of using the Kolmogorov-Smirnov criterion, a cumulative percentage of ‘outlier’ data to be excluded (the *exclusion level*) was used as an alternative cutoff criterion. The exclusion level was varied between 0.0 (no data excluded) and 0.9 (90% of schedule data excluded). It shows that for all datasets the exponent estimation stabilizes near Zipf’s value after excluding approximately 20% of the outliers.

At each exclusion level the corresponding power law exponent values were estimated. A (weighted) hypothesis test for a true mean exponent value of 2.0 was performed at every exclusion level, based on the sample of 13 power law exponent estimates. It appears that, for most of the outlier exclusion levels between 20% and approximately 55%, one cannot reject the null hypothesis that the mean is equal to 2.0, i.e. Zipf's value. In a weighted test, this is the case even onwards from an exclusion level of 0%. In conclusion: at least 80% of the activity schedule distribution obeys a power law with a universal exponent value of 2.0.

Almost all current validation strategies seem to involve the collection of new (unseen) data to compare with the model’s output. The authors envisioned that a Zipf power law could be used as an additional dimension within a travel demand model’s validation process, not requiring any additional validation data and thus not bringing any additional cost to the project or research. Additionally, the distribution from the current paper might enable the integration of activity schedules in models based on universal mobility patterns.

Future research will focus on a detailed modeling of the mechanism that leads to a Zipf power law distribution for activity schedules. Additionally, the law’s extend will be put to the test in further analyses and applications will be investigated in depth.

Acknowledgements

The authors are deeply grateful to the original data creators, depositors and/or copyright holders for making the microdata available for this research. The copyright and all other intellectual property rights in the data and associated documentation are vested in the original data creators or depositors. The original data creators and analyzers bear no responsibility for the further analysis or interpretation of the data in this research.

The authors would like to acknowledge the researches, works, individuals and institutions supporting the following data collections for making this research possible: US NHTS 2009 (U.S. Department of Transportation and Federal Highway Administration 2009), NLD OViN 2013 (Centraal Bureau voor de Statistiek (CBS) and Rijkswaterstaat (RWS) 2014), BEL OVG 3.0-4.5 (Janssens et al. 2014), SVN Ljubljana 2013 (Klemenčič et al. 2014) [Acknowledgements go out to the City Municipality of Ljubljana], GBR NTS 2009-2014 (Department for Transport 2015), KOR Seoul HTS 2010 (Korea Transportation Institute 2011; Metropolitan Transport Authority 2012), DEU Mobidrive 1999 (Chalasani and Axhausen 2004), CHE Thurgau 2003 (Loechl 2005), FRA ENTD 2008 (Armoogum et al. 2011), BEL Beldam 2010 (Cornelis et al. 2012) [financed by BELSPO, FOD Mobiliteit & Vervoer and others. Coordinated by GRT (Université de Namur) in cooperation with IMOB (Universiteit Hasselt) and CES (FUSL)], IRL NTS 2009 (Central Statistics Office 2011) [Accessed via the Irish Social Science Data Archive - www.ucd.ie/issda], FIN HLT 2010-2011 (Liikennevirasto - Finnish Transport Agency) [Finnish National Travel Survey 2010-2011 / Finnish Transport Agency and Wim Ectors], SWE RVU 2011-2014 (Trafik Analys 2015), AUS VISTA 2007 & 2009 (Department of Economic Development; Jobs; Transport and Resources (DEDJTR) 2009; Department of Economic Development; Jobs; Transport and Resources (DEDJTR) 2007). Part of this work was presented at the hEART 2016 conference

in Delft, The Netherlands, offering valuable reflections during the preparation of this work. (Ectors et al. 2016a)

Author contributions

W. Ectors: Literature Search and Review, Manuscript writing and Editing
B. Kochan: Meta-Analysis, Content planning and Detailed Manuscript Review
D. Janssens: Meta-Analysis and Content planning
T. Bellemans: Meta-Analysis and Content planning
G. Wets: Meta-Analysis and Content planning

Conflict of interest statement

On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Adamic, L.A., Huberman, B.A.: Zipf's law and the Internet. Glottometrics 3. 143–150 (2002)

Ahern, A., Weyman, G., Redelbach, M., Schulz, A., Akkermans, L., Vannacci, L., Anoyrkati, E., Von grinsven, A.: Analysis of National Travel Statistics in Europe. (2013)

Arentze, T., Hofman, F., van Mourik, H., Timmermans, H.J.P.: ALBATROSS: Multiagent, Rule-Based Model of Activity Pattern Decisions. Transportation Research Record. 1706, 136–144 (2000). doi:10.3141/1706-16

Armoogum, J., Hubert, J.-P., Francois, D., Roumier, B., Robin, M., Roux, S.: Enquête nationale transports et déplacements 2007-2008 (ENTD 2007-2008) (Rapport technique). (2011)

Auerbach, F.: Das Gesetz der Bevölkerungskonzentration. Petermanns Geographische Mitteilungen. 59, 74–76 (1913)

Batty, M.: Agents, cells, and cities: New representational models for simulating multiscale urban dynamics. Environment and Planning A. 37, 1373–1394 (2005). doi:10.1068/a3784

Batty, M.: The size, scale, and shape of cities. Science (New York, N.Y.). 319, 769–771 (2008). doi:10.1126/science.1151419

Boyce, D., Bar-Gera, H.: Validation of multiclass urban travel forecasting models combining origin-destination, mode, and route choices. Journal of Regional Science. 43, 517–540 (2003). doi:10.1111/1467-9787.00309

Brockmann, D., Hufnagel, L., Geisel, T.: The scaling laws of human travel. Nature. 439, 462–5 (2006). doi:10.1038/nature04292

Centraal Bureau voor de Statistiek (CBS), Rijkswaterstaat (RWS): Onderzoek Verplaatsingen in Nederland 2013 - OViN 2013, http://dx.doi.org/10.17026/dans-x9h-dsdg, (2014)

Central Statistics Office: National Travel Survey 2009. , Dublin, Ireland (2011)

Chalasani, V.S., Axhausen, K.W.: Mobidrive: A six week travel diary, https://www.ethz.ch/content/dam/ethz/special-interest/baug/ivt/ivt-dam/vpl/tsms/tsms2.pdf, (2004)

Chen, W.-C.: On the Weak Form of Zipf's Law. Journal of Applied Probability. 17, 611–622 (1980). doi:10.2307/3212955

Clauset, A., Shalizi, C.R., Newman, M.E.J.: Power-Law Distributions in Empirical Data. SIAM Review. 51, 661 (2009). doi:10.1137/070710111

Cornelis, E., Hubert, M., Hyunen, P., Lebrun, K., Patriarche, G., De Witte, A., Creemers, L., Declercq, K., Janssens, D., Castaigne, M., Hollaert, L., Walle, F.: La mobilité en Belgique en 2010 : résultats de l'enquête BELDAM, (2012)

Corominas-Murtra, B., Solé, R. V.: Universality of Zipf's law. Physical Review E. 82, 9 (2010). doi:011102 10.1103/PhysRevE.82.011102

Department for Transport: National Travel Survey, 2002-2014 [computer file]. 9th Edition., (2015)

Department of Economic Development; Jobs; Transport and Resources (DEDJTR): Victorian Integrated survey of Travel and Activity 2007, www.economicdevelopment.vic.gov.au/vista, (2007)

Department of Economic Development; Jobs; Transport and Resources (DEDJTR): Victorian Integrated survey of Travel and Activity 2009, www.economicdevelopment.vic.gov.au/vista, (2009)

Doherty, S.T., Axhausen, K.W.: The Development of a Unified Modeling Framework for the Household Activity-Travel Scheduling Process. Traffic and Mobility: Simulation-Economics-Environment. 35–36 (1999)

Ectors, W., Kochan, B., Janssens, D., Bellemans, T., Wets, G.: Zipf's law in activity schedules. In: hEART 2016: The 5th symposium arranged by European Association for Research in Transportation, Delft, The Netherlands, 14-16 September 2016 (2016)(a)

Ectors, W., Kochan, B., Knapen, L., Janssens, D., Bellemans, T.: A Generic Data-driven Sequential Clustering Algorithm Determining Activity Skeletons. In: Procedia Computer Science. pp. 34–41. Elsevier Masson SAS (2016)(b)

Fujiwara, Y.: Zipf law in firms bankruptcy. Physica A: Statistical Mechanics and its Applications. 337, 219–230 (2004). doi:10.1016/j.physa.2004.01.037

Furusawa, C., Kaneko, K.: Zipf's Law in Gene Expression. Physical Review Letters. 90, 1–11 (2003). doi:10.1103/PhysRevLett.90.088102

Gabaix, X.: Zipf ' S Law for Cities : an Explanation. Quarterly Journal of Economics. 114, 739–767 (1999)

Gillespie, C.S.: Fitting Heavy Tailed Distributions: The poweRlaw Package. Journal of Statistical Software. 64, 1–16 (2015)

González, M.C., Hidalgo, C.A., Barabási, A.-L.: Understanding individual human mobility patterns. Nature. 453, 779–782 (2008). doi:10.1038/nature07850

Guidotti, R., Trasarti, R., Nanni, M.: TOSCA : TwO-Steps Clustering Algorithm for Personal Locations Detection. Proceedings of the 23nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. (2015). doi:10.1145/2820783.2820818

Hanel, R., Corominas-Murtra, B., Liu, B., Thurner, S.: Fitting power-laws in empirical data with estimators that work for all exponents. PLoS ONE. 12, 1–15 (2017). doi:10.1371/journal.pone.0170920

Hupkes, G.: The law of constant travel time and trip-rates. Futures. 14, 38–46 (1982). doi:https://doi.org/10.1016/0016-3287(82)90070-2

Ioannides, Y.M., Overman, H.G.: Zipf's law for cities: An empirical examintion. Regional Science and Urban Economics. 33, 127–137 (2003). doi:10.1016/S0166-0462(02)00006-6

Janssens, D., Declercq, K., Wets, G.: Onderzoek Verplaatsingsgedrag Vlaanderen 4.5 (2012-2013). (2014)

Jiang, B.: Street Hierarchies: A Minority of Streets Account for a Majority of Traffic Flow. International Journal of Geographical Information Science. 23, 1033–1048 (2009). doi:10.1080/13658810802004648

Jiang, B., Jia, T.: Zipf's Law for All the Natural Cities in the United States: A Geospatial Perspective. International Journal of Geographical Information Science. 10 (2010). doi:10.1080/13658816.2010.510801

Jiang, S., Yang, Y., Gupta, S., Veneziano, D., Athavale, S., González, M.C.: The TimeGeo modeling framework for urban motility without travel surveys. Proceedings of the National Academy of Sciences. 113, E5370–E5378 (2016). doi:10.1073/pnas.1524261113

Ki Baek, S., Bernhardsson, S., Minnhagen, P.: Zipf's law unzipped. New Journal of Physics. 13, 043004 (2011). doi:10.1088/1367-2630/13/4/043004

Klafter, J., Shlesinger, M.F., Zumofen, G.: Beyond Brownian Motion. Physics Today. 49, 33 (1996). doi:10.1063/1.881487

Klemenčič, M., Lep, M., Mesarec, B., Žnuderl, B.: Potovalne navade prebivalcev v Mestni občini Ljubljana in Ljubljanski urbani regiji, (2014)

Korea Transportation Institute: National Transportation Demand Survey and Database Establishment in 2010: Passenger O/D Survey on the National Area. (2011)

Li, S., Lee, D.H.: Learning daily activity patterns with probabilistic grammars. Transportation. 44, 49–68 (2017). doi:10.1007/s11116-015-9622-1

Li, W.: Random Texts Exhibit Zipf's-Law-Like Word Frequency Distribution. IEEE Transactions on Information Theory. 38, 1842–1845 (1992). doi:10.1109/18.165464

Liikennevirasto - Finnish Transport Agency: National Travel Survey 2010–2011.

Loechl, M.: Stability of Travel Behaviour: Thurgau 2003, http://archiv.ivt.ethz.ch/vpl/publications/tsms/tsms16.pdf, (2005)

Lü, L., Zhang, Z.-K., Zhou, T.: Deviation of Zipf's and Heaps' Laws in human languages with limited dictionary sizes. Scientific reports. 3, 1082 (2013). doi:10.1038/srep01082

Ma, D., Sandberg, M., Jiang, B.: A Socio-geographic Perspective on Human Activities in Social Media. Geographical Analysis. 49, 328–342 (2017). doi:10.1111/gean.12122

Maillart, T., Sornette, D., Spaeth, S., von Krogh, G.: Empirical Tests of Zipf's Law Mechanism in Open Source Linux Distribution. Physical Review Letters. 101, 218701 (2008). doi:10.1103/PhysRevLett.101.218701

Makse, H.A., Andrade Jr., J.S., Batty, M., Havlin, S., Eugene Stanley, H., Jr, S.A., Stanley, H.E., Batty, M., Havlin, S., Stanley, H.E., Makse, H.A., Andrade Jr., J.S., Batty, M., Havlin, S., Eugene Stanley, H., Jr, S.A., Stanley, H.E.: Modeling urban growth patterns with correlated percolation. Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics. 58, 7054–7062 (1998). doi:10.1103/PhysRevE.58.7054

Marsili, M., Zhang, Y.: Interacting Individuals Leading to Zipf ' s Law. Physical Review Letters. 80, 2741–2744 (1998). doi:10.1103/PhysRevLett.80.2741

Metropolitan Transport Authority: The Report of Household travel survey in Seoul Metropolitan Area [In Korean]. , Seoul (2012)

Miller, E.J., Farooq, B., Chingcuanco, F., Wang, D.: Historical Validation of Integrated Transport-Land Use Model System. Transportation Research Record: Journal of the Transportation Research Board. 2255, 91–99 (2012). doi:10.3141/2255-10

Newman, M.: Power Laws, Pareto Distributions and Zipf's Law. Contemporary physics. 46, 323–351 (2005). doi:10.1080/00107510500052444

Nitsch, V.: Zipf zipped. Journal of Urban Economics. 57, 86–100 (2005). doi:10.1016/j.jue.2004.09.002

Noulas, A., Scellato, S., Lambiotte, R., Pontil, M., Mascolo, C.: A tale of many cities: Universal patterns in human urban mobility. PLoS ONE. 7, (2012). doi:10.1371/journal.pone.0037027

Okuyama, K., Takayasu, M., Takayasu, H.: Zipf's law in income distribution of companies. Physica A: Statistical Mechanics and its Applications. 269, 125–131 (1999). doi:10.1016/S0378-4371(99)00086-2

Paleari, S., Redondi, R., Malighetti, P.: A comparative study of airport connectivity in China, Europe and US: Which network provides the best service to passengers? Transportation Research Part E: Logistics and Transportation Review. 46, 198–210 (2010). doi:10.1016/j.tre.2009.08.003

Pappalardo, L., Simini, F.: Data-driven generation of spatio-temporal routines in human mobility. Springer US (2017)

Park, B.; Schneeberger, J.D.: Microscopic Simulation Model Calibration and Validation: Case Study of VISSIM Simulation Model for a Coordinated Actuated Signal System. Transportation Research Record. 1856, 185–192 (2003). doi:10.3141/1856-20

Reed, W.J.: The Pareto, Zipf and other power laws. Economics Letters. 74, 15–19 (2001). doi:10.1016/S0165-1765(01)00524-9

Riccardo, G., Armando, B., Sandro, R.: Towards a Statistical Physics of Human Mobility. International Journal of Modern Physics C. 23, 1250061 (2012). doi:10.1142/S0129183112500611

Saberi, M., Mahmassani, H.S., Brockmann, D., Hosseini, A.: A complex network perspective for characterizing urban travel demand patterns: graph theoretical analysis of large-scale origin–destination demand networks. Transportation. 44, 1383–1402 (2017). doi:10.1007/s11116-016-9706-6

Schneider, C.M., Belik, V., Couronné, T., Smoreda, Z., González, M.C.: Unravelling daily human mobility motifs. Journal of the Royal Society, Interface / the Royal Society. 10, 20130246 (2013). doi:10.1098/rsif.2013.0246

Shen, Y., Karimi, K.: Urban function connectivity: Characterisation of functional urban streets with social media check-in data. Cities. 55, 9–21 (2016). doi:10.1016/j.cities.2016.03.013

Simini, F., González, M.C., Maritan, A., Barabási, A.-L.: A universal model for mobility and migration patterns. Nature. 484, 96–100 (2012). doi:10.1038/nature10856

Song, C., Koren, T., Wang, P., Barabasi, A.-L.: Modelling the scaling properties of human mobility. Nature Physics. 6, 1–6 (2010). doi:10.1038/NPHYS1760

Soo, K.T.: Zipf's Law for cities: A cross-country investigation. Regional Science and Urban Economics. 35, 239–263 (2005). doi:10.1016/j.regsciurbeco.2004.04.004

Toledo, T., Koutsopoulos, H.: Statistical Validation of Traffic Simulation Models. Transportation Research Record. 1876, 142–150 (2004). doi:10.3141/1876-15

Trafik Analys: RVU Sverige 2011–2014 - Den nationella resvaneundersökningen (RVU Sweden 2011-2014 - national travel survey). , Stockholm (2015)

U.S. Department of Transportation, Federal Highway Administration: 2009 National Household Travel Survey, http://nhts.ornl.gov, (2009)

Urzúa, C.M.: Testing for Zipf's law: A common pitfall. Economics Letters. 112, 254–255 (2011). doi:10.1016/j.econlet.2011.05.049

Volchenkov, D., Blanchard, P.: Scaling and universality in city space syntax: Between Zipf and Matthew. Physica A: Statistical Mechanics and its Applications. 387, 2353–2364 (2008). doi:10.1016/j.physa.2007.11.049

Wang, G., Zhong, Y., Teo, C.-P., Liu, Q.: Flow-based accessibility measurement: The Place Rank approach. Transportation Research Part C: Emerging Technologies. 56, 335–345 (2015). doi:10.1016/j.trc.2015.04.017

Willis, J.C.: Age and Area. Cambridge University Press. (1922)

Yang, X.H., Chen, G., Chen, S.Y., Wang, W.L., Wang, L.: Study on some bus transport networks in China with considering spatial characteristics. Transportation Research Part A: Policy and Practice. 69, 1–10 (2014). doi:10.1016/j.tra.2014.08.004

Zheng, Z., Rasouli, S., Timmermans, H.: Two-regime Pattern in Human Mobility: Evidence from GPS Taxi Trajectory Data. Geographical Analysis. 48, 157–175 (2016). doi:10.1111/gean.12087

Zipf, G.K.: Human Behaviour and the Principle of Least Effort. Addison-Wesley, Reading (1949)

List of Figures:

Fig. 1 Activity schedule frequency distribution for several HTS datasets as a function of schedule rank. For each survey area, the original activity type encoding was used. Multiple-day survey data was converted into single-day schedules. The horizontal axis shows schedule rank (ID within a descending frequency table) and the vertical axis the relative frequency as a percentage of the total number of schedules. 11

Fig. 2 Kernel density plot of the bootstrapped power law exponent estimates. Visually, one observes a large cluster of exponents around the value of 2.0. 14

Fig. 3 Power law exponent values for several HTS datasets in function of the exclusion fraction (of outliers). This overview plot shows a range in exclusion fraction between 0.0 (no data excluded) and 0.9 (90% of the schedule data excluded). 15

Fig. 4 Similar as in Fig. 1, this figure shows the activity schedule frequency distribution for several HTS datasets as a function of schedule rank. For each dataset, the 20% least occurring schedules have been removed as outliers. 16

Fig. 5 P-values in function of the exclusion fraction (of outliers). The corresponding hypothesis test is: H0: The true mean power law exponent value of all study areas is 2.0. In

(a) no weights were used, while in (b) a weighted t-test was used with as weights the number of schedules as in Table 2. A range in exclusion fraction between 0.0 (no data excluded) and 0.9 (90% of schedules excluded) is shown. 17

Fig. 6 Descriptive statistics for the power law exponent value of the sample of activity schedules of 13 HTSs, in function of the exclusion fraction (of outliers). For the weighted mean and standard deviation, the number of schedules as in Table 2 was used as weight variable. 18

List of Tables:

Table 1: Top-10 most frequent activity schedules in USA NHTS 2009. A 'home' activity is abbreviated as 'H'. 8

Table 2: HTS datasets used within this research. 8

Table 3: Estimation results from the R package poweRlaw for activity schedule distributions in several HTS datasets. α represents the estimated power law exponent value, xmin the cutoff value, 'Cum. Pct rejected' the cumulative percentage of data rejected because of the cutoff value, KS the Kolmogorov-Smirnov distant measure, AM the arithmetic mean and SD the standard deviation. 12