A portfolio theory of route choice

David Levinson [a], Shanjiang Zhu [b,*]

[a] University of Minnesota, Department of Civil Engineering, 500 Pillsbury Drive SE, Minneapolis, MN 55455, USA
[b] George Mason University, Sid and Reva Dewberry Department of Civil, Environmental, and Infrastructure Engineering, 4400 University Drive, MS6C1 Fairfax, VA 22030, USA

ARTICLE INFO

ABSTRACT

Although many individual route choice models have been proposed to incorporate travel time variability as a decision factor, they are typically still deterministic in the sense that the optimal strategy requires choosing one particular route that maximizes utility. In contrast, this study introduces an individual route choice model where choosing a portfolio of routes instead of a single route is the best strategy for a rational traveler who cares about both journey time and lateness when facing stochastic network conditions. The proposed model is compared with UE and SUE models and the difference in both behavioral foundation and model characteristics is highlighted. A numerical example is introduced to demonstrate how such model can be used in traffic assignment problem. The model is then tested with GPS data collected in metropolitan Minneapolis–St. Paul, Minnesota. Our data suggest there is no single dominant route (defined here as a route with the shortest travel time for a 15 day period) in 18% of cases when links travel times are correlated. This paper demonstrates that choosing a portfolio of routes could be the rational choice of a traveler who wants to optimize route decisions under variability.

1. Introduction

Route choice is a daily decision travelers make under variable traffic conditions. Traffic patterns emerge from individual decisions, and each day's collective decisions update the travel experience of all travelers. In the long run, we expect that each traveler will develop an explicit or implicit strategy to guide individual route decisions. Conventional User Equilibrium (UE) models assume that travelers seek to minimize individual travel time with perfect knowledge of network conditions. In equilibrium, "the journey times in all routes actually used are equal and less than those which would be experienced by a single vehicle on any unused route" (Wardrop, 1952). Although this shortest-path (usually measured as shortest travel time path) assumption and the resulting aggregate UE approach is simple, intuitive, and easy to implement (efficient solutions are widely available), it has been criticized for ignoring the heterogeneity in individual preferences among travelers and limitations in their spatial knowledge. Given the stochasticity in network conditions and potential penalties for being late or early, travel time reliability has been widely identified as an important factor in route decisions (e.g. Bekhor et al., 2006; Brownstone and Small, 2005; Noland and Polak, 2002; Small et al., 2005 among others).

Wardrop's UE principle requires travelers chose the shortest time path. Although several paths may have equal journey time in equilibrium, only one can be chosen by an individual for a given trip. However empirical evidence finds individual

* Corresponding author. Address: George Mason University Sid and Reva Dewberry Department of Civil, Environmental, and Infrastructure Engineering Volgenau School of Engineering, Nguyen Engineering Building, Room 1408 4400 University Drive, MS6C1 Fairfax, VA 22030, United States. Tel.: +1 703-993-1797; fax: +1 703 993 9790.

E-mail addresses: dlevinson@umn.edu (D. Levinson), szhu3@gmu.edu (S. Zhu).
URL: http://civil.gmu.edu/people/shanjiang-zhu/ (S. Zhu).

travelers chose multiple routes between a given origin–destination pair through repeated choices (Jan et al., 2000). We discuss this in Section 2.1. While some travelers only had minor deviations, most travelers followed routes that deviated significantly from the shortest time path.

Although many individual route choice models have been proposed to incorporate travel time variability as a decision factor, they are typically still deterministic in the sense that the optimal strategy requires choosing one particular route that maximizes utility.

The Stochastic User Equilibrium (SUE) model adds a random component to the expected travel time. Under SUE, "no user believes he can improve his travel time by unilaterally changing routes" (Daganzo and Sheffi, 1977). The stochasticity is due to either some traveler characteristics not observable by the modeler or other randomness on the network.

The SUE theory however does not explain why some travelers prefer multiple routes over a time period instead of choosing one optimal route. Route choice models based on prospect theory further argue that travelers usually perceive uncertainty in travel cost asymmetrically and human choices usually deviate from what is predicted by expected utility based in empirical studies (Tversky et al., 2005; Kahneman and Tversky, 1979). People are found to underweight high probability events when certainty is not guaranteed (Allais Paradox, Allais, 1979) and inflate the larger gain when facing alternatives with small probabilities. Parthasarathi (2011) found that traffic network structure variables (such as intersection density, street density, proportion of limited access roads, and route complexity) can also affect travel time perception. Prospect theory has been applied to route choice by investigating the value function and appropriate reference point (Avineri and Prashker, 2004; de Palma and Picard, 2006; Katsikopoulos et al., 2002). However, given an estimated value function, we would expect a pure strategy of choosing the route that minimizes the relative utility when compared to the reference point.

Some researchers approached this multiplicity problem by arguing that travelers are boundedly rational (Lou et al., 2010; Mahmassani and Chang, 1987) and may use one of multiple acceptable routes. Differences in travel cost between these routes and the shortest route are tolerable or not noticeable by travelers. Under this theory, the route in the acceptable set to be chosen will depend on some random events or personal experience. However, no theory is provided to determine the probability of choosing each route. In the context of transit route choice, Spiess and Florian (1989) proposed that the chance of taking a particular transit line among several attractive ones is proportionate to their service frequency. Its applicability to vehicular route choice problem has yet to be explored.

Models such as SUE still treat link travel costs as a deterministic value. In contrast, Watling (2002) assumes network conditions are stochastic and proposes more complicated equilibrium models. Facing such stochasticity, travelers could also change route through a day-to-day learning process, or simply react to previous bad experience. One recent example of that day-to-day dynamics could be the significant link flow oscillation observed after the 2007 I-35 Bridge collapse in Minneapolis, Minnesota (Zhu et al., 2010). However, as travelers accumulate more network knowledge through day-to-day experience, especially for commute trips, deterministic route choice models predict a single optimal route based on the perceived travel time distribution. For example, Mirchandani and Soroush (1987) considered both stochastic link travel time and individual travel time perception error. Although travelers with different risk-taking preferences, thus different utility functions, would take different routes, the final choice for each individual is still deterministic.

To provide such an explanation to the phenomenon that travelers chose multiple routes between a given origin–destination pair through repeated choices, this study introduces an individual route choice model where choosing a portfolio of routes instead of a single route may be the best strategy for a rational traveler trying to satisfy multiple criteria (trading-off journey time and lateness) facing stochastic network conditions. The next section provides empirical evidence of people choosing multiple routes between the same origin and destination, employing GPS data collected in metropolitan Minneapolis–St. Paul, Minnesota. A portfolio theory of route choice (RPT) is then proposed and tested with the field data. Findings from this paper may inform future travel demand models.

2. Empirical evidence of route portfolios

2.1. GPS data

This study investigates commuters' day-to-day route choices by analyzing a large set of GPS data collected during a 13-week long study targeting behavioral reactions to the I-35W Bridge reopening on September 18th, 2008. Details about this behavioral study and data collection process are provided by Zhu et al. (2010). Participants were randomly selected commuters in the Minneapolis, St. Paul, Minnesota metropolitan area (Twin Cities). Either a logging Global Positioning System (GPS) devices (QSTARZ BT-Q1000p GPS Travel Recorder powered by DC output from in-vehicle cigarette lighter) or a real-time communicating GPS device (adapted from the system deployed in the Commute Atlanta study (Rates, 2007)) were installed in the vehicle of study participants. The GPS device is non-intrusive and unlikely to affect the behavior of participants. No instructions were given and participants were free to make travel choices. In all, 190 subjects participated in this study. However, only 143 GPS records were recovered due to the failure of devices (the data from GPS loggers can only be checked at the end of the study. Some of them failed because of power supply problems, such as being disconnected by subjects.).

The logging GPS devices accurately monitored the travel trajectories of each probe vehicle at a frequency of one point per 25 m up to 13 weeks, about 3 weeks before the reopening of the bridge and between 8 and 10 weeks after it. The real-time communicating GPS device recorded the position of instrumented vehicles every second. The geographic location and time

stamps of each point were documented and projected onto a GIS map for post-processing. The GPS data were then matched to the 2009 Twin Cities Regional Planning network, which has been conflated to real road geometry.

An algorithm was developed and applied to ensure all points have been snapped to the nearest link which:

- is directly connected to the upstream link previously identified;
- is consistent with the travel direction of nearby GPS points; and
- is connected to the downstream link which is also consistent with travel direction of downstream GPS points.

This algorithm rules out the possibility of incorrectly snapping the GPS point to the link in the opposite direction and changing directions mid-link. The high resolution of one point every 25 m (the real-time communicating GPS provided an even higher resolution) reduces the possibility of holes and keeps discontinuity in identified routes to a minimum. In rare cases of data losses due to the communication difficulties with satellites, the shortest time path was used to connect the different segments of the same trip. This algorithm, combined with accurate GIS files, ensures that the right links will be identified for each trip. It also helps to ensure that the speed estimated from vehicle trajectories will later be assigned to the link through which travelers passed. A visual check was conducted for all trips of two random subjects during the entire study period, and confirms the accuracy of the algorithm.

2.2. Diversity of commute trips

This study focuses on commute trips because (1) a large number of trips could be observed between the same origin and destination; (2) travelers are likely to gain enough experience through daily commuting to develop a reasonable estimation of the relevant part of the network; (3) people are more concerned about lateness and travel time reliability, for commute trips than for discretionary trips. To keep the problem simple, we only consider home-to-work trips here, although the same analysis could also be applied for work-to-home trips. Home-to-work trips are defined as any trips starting within a 600 m radius from home and ending in a 600 m radius from the workplace during a weekday, without any stop longer than 5 min. The threshold of 600 m represents approximately four city blocks, which is chosen by observing parking and workplaces for a subset of subjects. To make all trips comparable in the following analysis, minor changes have been made to ensure trips made by the same subjects always start from the same origin node and end at the same destination node. Very few changes resulted, since parking locations at both home and work places were stable for most subjects.

The reopening of I-35W Mississippi River Bridge during the study period represents a major change of network condition, which may affect people's route choice behavior. To avoid this confounding factor, we only use data collected during the three weeks before the bridge reopening. Since we only focus on the Twin Cities (7 County) area, subjects who live outside of the region are excluded from the study. In total, 657 home-to-work trips made by 95 subjects have been identified. These trips are then compared segment-by-segment using GIS and different home-to-work routes are identified for each subject. Although the problems of route overlapping and trivial alternatives have been discussed by many researchers under various contexts (e.g. (Bovy, 2009; Frejinger and Bierlaire, 2007)), no consensus has been reached for the threshold to define distinct routes. Therefore, a series of threshold values have been tested. Fig. 1 summarizes the percentage of subjects with different number of distinct home-to-work routes observed during three weeks.
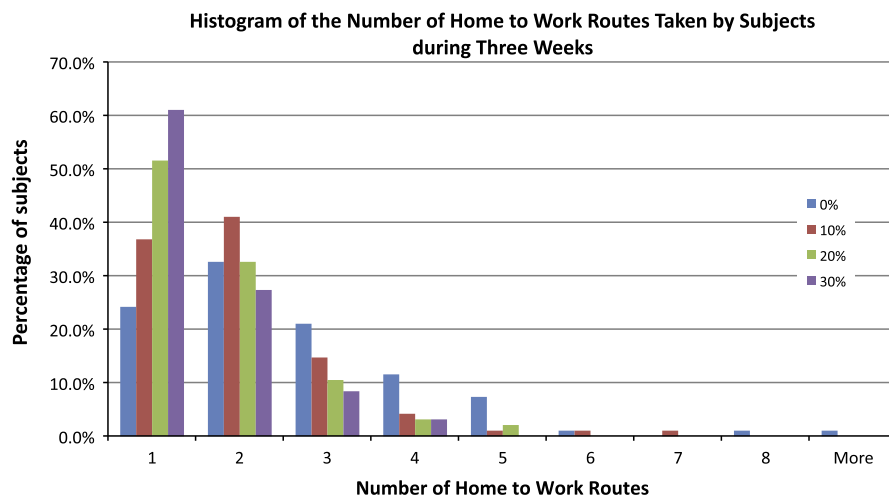

Fig. 1. The morning commute route diversity among 95 subject during 3 weeks. Percentage indicates share of distance which may differ without routes being considered "different". Data were collected by in-vehicle GPS devices during September, 2008 for a study focusing on route choice behavior before and after the reopening of I-35W Mississippi River Bridge in Minneapolis. *Source*: Authors.

If routes with any different segments are treated as different routes, then more than three quarters of all subjects used more than one route during 3 weeks. Some subjects traveled on more than eight different routes. As the threshold of minimum difference in length to define distinct routes increases, home-to-work route choices exhibit less diversity. However, even if more than 30% difference in distance is required to define a different route, about 40% of all subjects followed more than one route during the study period. Therefore, a significant fraction of subjects chose a portfolio of routes for their morning commute trips. Many reasons could help to explain the behavior of choosing multiple commute routes during a period of time. The next section addresses this problem by investigating route decisions of a rational traveler under uncertain network conditions.

3. Portfolio theory of route choice

In his seminal work *Risk, Uncertainty, and Profit*, Frank Knight (1921) established the distinction between risk and uncertainty.

"... *Uncertainty* must be taken in a sense radically distinct from the familiar notion of *Risk*, from which it has never been properly separated. The term *risk*, as loosely used in everyday speech and in economic discussion, really covers two things which, functionally at least, in their causal relations to the phenomena of economic organization, are categorically different. ... The essential fact is that *risk* means in some cases a quantity susceptible of measurement, while at other times it is something distinctly not of this character; and there are far-reaching and crucial differences in the bearings of the phenomenon depending on which of the two is really present and operating. ... It will appear that a measurable uncertainty, or *risk* proper, as we shall use the term, is so far different from an unmeasurable one that it is not in effect an uncertainty at all. We ... accordingly restrict the term *uncertainty* to cases of the non-quantitive type." (Knight, 1921).

This study investigates the route choice behavior of a rational user who seeks to maximize utility under network variability. Many studies have dealt with risk and uncertainty in route choice, either due to perception errors or stochasticity in network conditions (in most cases travel time). Some researchers followed the expected utility approach which was originally proposed by Bernoulli (1954) (originally in Latin and translated by Dr. Louis Sommer) and later popularized by Von Neumann et al. (1947). For example, Pells (1987) assumes that travelers' utility is a linear combination of the generalized travel cost and a slack time that travelers allocate to avoid arriving late (dubbed a "safety margin" by Knight (1974)). Polak (1987) further defined a safety margin as the difference between the mean arrival time and the work start time. The problem with risk and uncertainty was implicitly addressed since travelers have to reserve a larger safety margin with lower travel time reliability.

In contrast, other researchers insist that travel time reliability has intrinsic value (you still prefer reliability even when you are flexible with arrival time) and should be modeled explicitly. Research work in this direction follows the two-parameter approach (mean–variance in most cases) which was originated by Markowitz (1952a,b) and Hester and Tobin (1967) in portfolio studies and then introduced to transportation by Jackson and Jucker (1982). Given an *a priori* estimate of network conditions (mean travel time and the variance), two-parameter models usually define the objective for a rational traveler as:

$$Min \ U = \alpha E(t) + \tau V(t) + \delta C \tag{1}$$

where $E(t)$ is the expected travel time, the $V(t)$ is the variance of travel time, and C summarizes other generalized costs associated with each route. The relative importance of travel time reliability is captured by the parameter τ. Given an individual whose value of time (measured as α/δ), value of reliability (measured as τ/δ), and perception (or prediction for a specific day) of network conditions are fixed, a deterministic choice would be generated in previous studies.

Although these studies reveal that travelers have strong preference for travel time reliability, they provide limited information about why travelers would choose multiple routes over time under the same condition. In this section, a model applying RPT will be developed and we will show that under certain conditions, choosing a set of routes over one unique route (which UE or SUE model would predict) becomes the rational choice. We will start with a simple model in which travelers have a strong preference on that their travel time variance should be smaller than a threshold. This model would later be extended to incorporate a more general structure similar to the two-parameter approach in Jackson and Jucker (1982). Sources of travel time uncertainty will be discussed and a numerical example will be provided on a stylized network to demonstrate how RPT can be applied to solve the route choice problem. Since the Stochastic User Equilibrium (SUE) models also consider uncertainty in route choice, a comparison will be provided between the proposed portfolio model and the SUE model. As will be shown in this section, these two models differ in both behavioral foundation and predicted route choice patterns.

3.1. Portfolio route choice model

Previous studies show that travelers trade-off between travel time and travel time reliability because there is a disutility associated with either arriving too early or arriving too late and because of variability of network conditions. Therefore, we first assume that the objective of travelers is to minimize travel time while keeping travel time variability under a certain threshold. Consider a traveler m who faces N alternative routes whose travel time $t' = \{t_1, t_2, \ldots, t_N\}$ are believed to have ex-

pected values $E(t)' = \{E(t_1), E(t_2), \ldots, E(t_N)\}$ and a covariance matrix $\Sigma = (\sigma_{i,j})$. It has to be indicated that route travel time t_n is stochastic, which differs from many previous models such as SUE which assumes deterministic network condition. Here the subscript m is omitted to keep the expression succinct. A rational traveler is to select daily routes according to $p' = \{p_1, p_2, \ldots, p_N\}$ in order to

$$\text{Min} \quad U = E(p't) \tag{2}$$

$$\text{s.t.} : \quad Var(p't) \leqslant v_c \tag{3}$$

$$\sum_i p_i = 1 \tag{4}$$

$$p_i \in [0, 1], \quad \forall i \in N \tag{5}$$

where p_i is the probability of choosing route i on a given day and v_c is the maximum travel time variance the traveler can tolerate. Given a network condition t and a personal preference v_c, the optimal strategy $\hat{p}$ can be derived by solving the problem. If $\hat{p}$ has more than one non-zero member, then the optimal strategy is to choose a route portfolio according to $\hat{p}$ instead of sticking to a single route.

To illustrate the idea, consider the simplest case where the rational traveler faces only two alternative routes: 1 and 2. Fig. 2 illustrates a possible distribution of travel time on routes 1 and 2 where the traveler has to trade-off between travel time and travel time reliability.

Depending on the travel time distributions and the tolerance for travel time reliability (or variability), a rational traveler could have a different strategy.

For convenience, assume t_1 and t_2 are independent. Then

$$Var(p't) = p'Var(t)p = p_1^2 Var(t_1) + p_2^2 Var(t_2) = p_1^2 Var(t_1) + (1 - p_1)^2 Var(t_2) \tag{6}$$

Here $p_2 = 1 - p_1$ because of Eq. (4). Therefore, the travel time variance by following strategy p is a quadratic function of p_1. Without losing generality, assume $Var(t_1) \geqslant Var(t_2)$. Then by evaluating Eq. (6) on the range $[0,1]$,

$$Var(p't) \in \left[\frac{Var(t_1)Var(t_2)}{Var(t_1) + Var(t_2)}, Var(t_1) \right] \tag{7}$$

and as shown in Fig. 3, the minimum is achieved when

$$p_1 = \frac{Var(t_2)}{Var(t_1) + Var(t_2)} \tag{8}$$

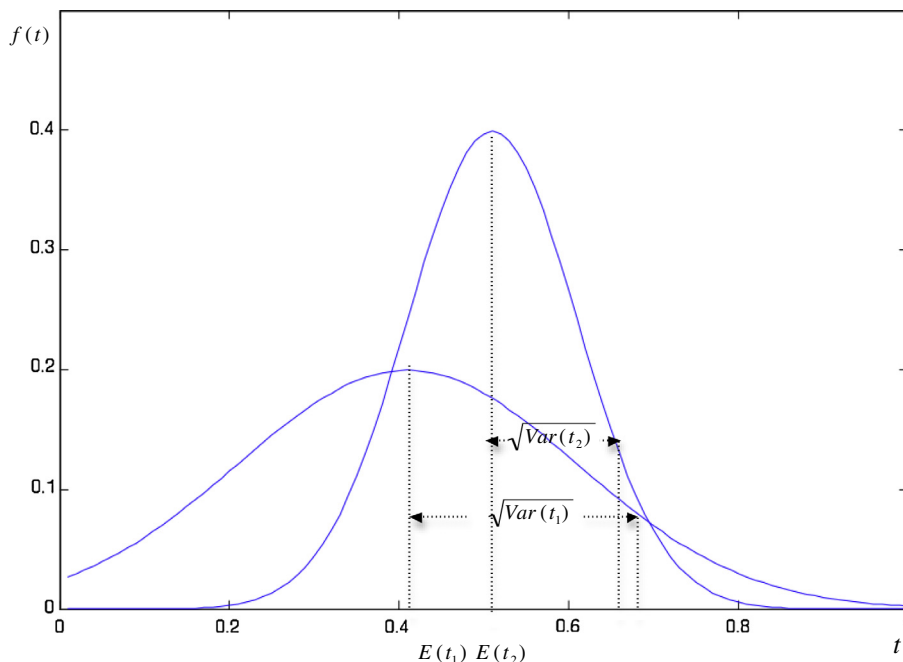Depending on the value of v_c, there are 4 situations.

Fig. 2. A possible distribution of travel times on route 1 and route 2. Route 1 has a smaller expected travel time, but larger travel time variability.
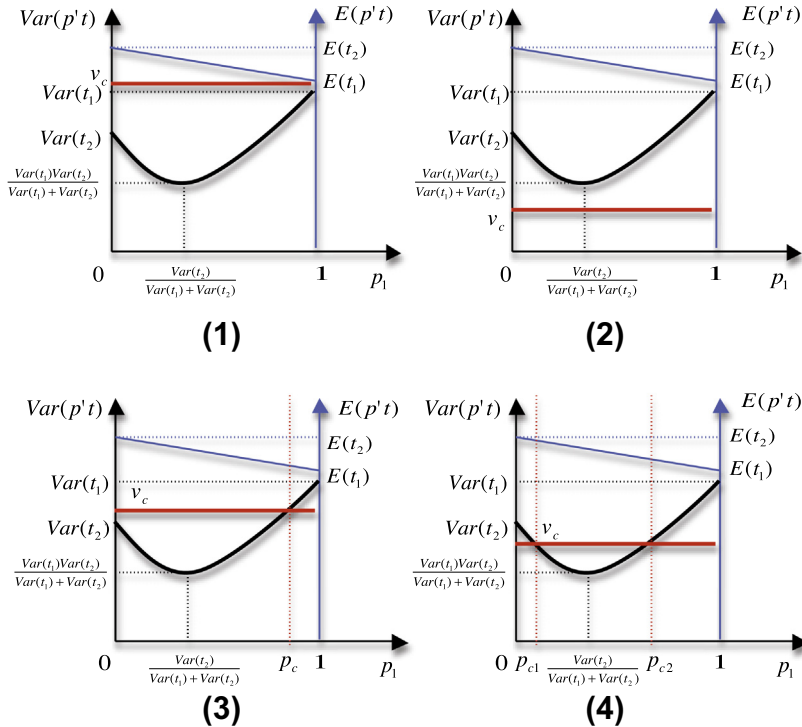
Fig. 3. Optimal strategy under different travel time reliability criteria. (1) Any p_1 is feasible and optimal is achieved at $p_1 = 1$. (2) There is no feasible p_1 that satisfy $Var(p't) < v_c$. (3) p_1 is feasible in $[0, p_c]$ and optimal is achieved at $p_1 = p_c$. (4) p_1 is feasible in $[p_{c1}, p_{c2}]$ and optimal is achieved at $p_1 = p_{c2}$.

1. If $v_c \geqslant Var(t_1)$, all possible strategies p are feasible and the best strategy is to always select the route with smaller expected travel time because constraint (3) is always satisfied.
2. If $v_c < \frac{Var(t_1)Var(t_2)}{Var(t_1)+Var(t_2)}$, there is no solution to the problem because no strategy p can satisfy constraint (3), in which case the traveler needs to adapt v_c or not travel.
3. If $v_c \in [Var(t_2), Var(t_1)]$, a feasible strategy p must satisfy $p_1 \in [0, p_c]$, while p_c is the strategy when $Var(p't)$ equals v_c. Therefore, if $E(t_1) < E(t_2)$, the best strategy should be to always select route 1. However, when $E(t_1) > E(t_2)$, the optimal is achieved by selecting route 1 p_c of the time and route 2 $1 - p_c$ of the time. A route portfolio serves better the objective than a strategy of always choosing a single route. When $E(t_1) = E(t_2)$, the traveler is indifferent.
4. If $v_c \in [\frac{Var(t_1)Var(t_2)}{Var(t_1)+Var(t_2)}, Var(t_2))$, the feasible strategy is depicted by $[p_{c1}, p_{c2}]$ and the best strategy is to always select a route portfolio. The minimum expected travel time is achieved on either p_{c1} or p_{c2}, depending on which route has smaller mean travel time.

Therefore, under some circumstances (such as that of cases 3 and 4), the proposed model predicts that choosing a route portfolio over time represents a better strategy compared to that of always choosing a single route. The independence of travel time on alternatives is not a required condition, but only helps to simplify the presentation. Actually, when the decision maker holds a belief of travel time correlation (captured by Σ), the only difference is that Eq. (6) becomes

$$Var(p't) = p'\Sigma p = p_1^2 Var(t_1) + p_2^2 Var(t_2) + 2p_1 p_2 Cov(t_1, t_2) \tag{9}$$

The conclusion may differ depending on the new quadratic curve depicted by (9). Under certain conditions, a route portfolio could become a dominant strategy.

The results could be further extended to the case of route choice when facing *N* alternatives. Because the covariance matrix Σ is positive semi-definite, the feasible set defined by Eq. (3) is convex. The objective function is a linear combination of expected travel time of all alternatives, so the optimal solution will always fall on the boundary of the feasible set. As long as the optimal solution is achieved at a point other than such corner points that one of the $p_i = 1$ and other members of $\hat{p}$ equal zero, a route portfolio becomes a dominant strategy. Moreover, since the feasible set defined by Eq. (3) is convex, the objective function Eq. (2) could be a non-linear function as long as it is also a convex function. By following the same reasoning as presented in this section, situations under which a route portfolio dominates a single-route strategy can be derived. The math is likely to be more complex.

As a further extension, other criteria regarding travel time reliability are also applicable. For example, travelers might prefer that the trip duration stay less than 5 min longer than the average 95% of the time. Given the travel time co-variance

matrix of alternative routes, this constraint can be easily translated into forms similar to Eq. (3) ($Pr([t - E(t)] \leqslant 5) \geqslant 0.95$) $\iff Var(t) \leqslant \left(\frac{5}{1.645}\right)^2$) if we assume t is normal. This too is likely to result in a route portfolio being preferred in circumstances similar to cases 3 and 4.

3.2. Portfolio theory under two-parameter models

The behavior of choosing a portfolio of routes is still rational for a traveler who prefers travel time certainty when a more general two-parameter model is used. Following Eq. (1), we assume that the objective of a rational traveler is:

$$Min \ U = \frac{\alpha}{\delta}E(t) + \frac{\tau}{\delta}V(t) + C \tag{10}$$

where $\frac{\alpha}{\delta}$ represents the value of time and $\frac{\tau}{\delta}$ presents the value of reliability. To simplify the presentation and focus on the portfolio issue, we ignore the toll C and assume $\delta = 1$. The preference for travel time certainty is now captured by the value of reliability instead of a threshold on overall travel time variance. Following the reasoning presented in the previous subsection, the problem becomes:

$$Min \quad U = \alpha E(p't) + \tau Var(p't) \tag{11}$$
$$s.t. : \quad \sum_i p_i = 1 \tag{12}$$
$$p_i \in [0, 1], \quad \forall i \in N \tag{13}$$

On the same stylized network with only one OD and two independent routes, the utility becomes:

$$U = \alpha((p_1 t_1 + (1 - p_1)t_2) + \tau\left(p_1^2 Var(t_1) + (1 - p_1)^2 Var(t_2)\right) \tag{14}$$

The first derivative of U over p_1 becomes:

$$\frac{\partial U}{\partial p_1} = \alpha(t_1 - t_2) + \tau(2p_1(Var(t_1) + Var(t_2)) - 2Var(t_2)) \tag{15}$$

The first-order condition $\frac{\partial U}{\partial p_1} = 0$ gives that:

$$p_1 = \frac{Var(t_2) - \frac{\alpha}{2\tau}(t_1 - t_2)}{Var(t_1) + Var(t_2)} \tag{16}$$

If p_1 falls in $(0, 1)$, then it implies that a traveler with a specific structure of preference (described by the value of time and value of reliability) facing the choice between two routes $(t_1, Var(t_1))$ and $(t_2, Var(t_2))$ would include both routes in a route portfolio and choose the first route p_1 percent of time.

In contrast, if such p_1 cannot be found given a preference structure and a choice set, it means $\frac{\partial U}{\partial p_1}$ is always positive or negative. If it is always positive, then the disutility function is always increasing within the range of possible p and a rational traveler should always choose Route 1 ($p_1 = 1$) to minimize the disutility. Similarly, if $\frac{\partial U}{\partial p_1}$ is always negative, then a rational traveler should always choose Route 2 to minimize the disutility. Therefore, under this more general two-parameter model, a rational traveler may prefer a portfolio of routes over a unique route.

The value of time (α) and value of reliability (τ) affect the route choice portfolio. When a traveler has a very strong preference for travel time reliability (large τ), $\frac{\alpha}{2\tau}$ becomes very small and the second term in the numerator of Eq. (16) is negligible. The traveler would always prefer a portfolio of routes and choose the first route with a probability of $\frac{Var(t_2)}{Var(t_1)+Var(t_2)}$. In contrast, if α is much larger than τ, the traveler values travel time much more than travel time reliability. In this case, Eq. (16) predicts that the travel would always choose Route 1 if $t_1 < t_2$ and Route 2 if $t_1 > t_2$, which is consistent with UE assumption. Following Jackson and Jucker (1982), many researchers have investigated individual preference for travel time reliability both empirically and theoretically. In many empirical studies, the value of reliability is defined as $\frac{\partial U/\partial \sigma_T}{\partial U/\partial C}$ (as a comparison, Jackson and Jucker (1982) used σ_T^2 instead of σ_T) and a new term, Reliability Ratio, defined as the value of reliability over the value of time, measures the relative preference of travel time reliability over travel time. For example, Small et al. (2005) reported a median value of time of \$21.46/h and a median value of reliability of \$19.56/h using revealed preference data collected in Los Angeles area. Smaller values (\$11.92/h for VOT and \$5.40) were found based on stated preference data collected in the same study. Carrion and Levinson (2012) reviewed 12 such studies and found the reliability ratio varies from 0.1 in Hollander (2006) to 1.47 in Yan (2002), with an average of 0.83 from all 12 studies. Therefore, empirical evidence shows that $\frac{\alpha}{2\tau}$ is nether going to be too large nor too small, and choosing a portfolio of routes becomes a rational choice under RPT when Eq. (16) can generate a probability between 0 and 1.

3.3. Comparison with UE and SUE Models

Although the proposed RPT may generate similar aggregate travel demand, it differs from Conventional User Equilibrium (UE) or Stochastic User Equilibrium (SUE) models through several fundamental behavior assumptions.

UE models would always have a traveler choose a route that minimizes disutility. In equilibrium, several routes may provide the same level of utility and a traveler becomes indifferent between these routes. A traveler may randomly pick any route among them. However, this randomness is not caused by a rational decision as under RPT and the probability of choosing one route for a specific traveler cannot be predicted. In contrast, the behavior of choosing route randomly according to a probability predicted by the route portfolio theory is a rational decision in reaction to uncertainty in network conditions. In reality, given the same travel demand, travel time could still fluctuate significantly for reasons such as signal control, freeway bottleneck activation when demand is close to capacity, etc. When the travel time fluctuation becomes small, the reliability constraint imposed by (3) is no longer binding. Similarly, if a traveler does not value travel time reliability (negligible $\frac{\tau}{\alpha}$ in Eq. (10)), the proposed route portfolio theory collapses to UE models.

The SUE model assumes travelers cannot perceive route travel time accurately, either because of stochasticity in network conditions, or limited human capability in perceiving route travel time. Such imperfections are summarized by an error term ξ in route travel time. For our one OD pair, two routes case, the perceived travel time on two alternative routes becomes $T_1 = t_1 + \xi_1$ and $T_2 = t_2 + \xi_2$. SUE models assume the individual error term $\xi = (\xi_1, \xi_2)$ follows uncertain distribution among the population. For example, Multi Variate Normal distribution. Following the standard Probit SUE model, the probability of choosing alternative 1 becomes:

$$p_1 = \Phi\left(\frac{t_2 - t_1}{\sqrt{Var(\xi_1) + Var(\xi_2)}}\right) \tag{17}$$

Although both portfolio theory and SUE models may generate similar route choice probabilities given a set of parameters, they are fundamentally different in their behavioral foundation. For SUE models, the randomness in route choice comes from the imperfection in travel time perception (either because of network uncertainty or because perception capability) instead of an intrinsic desire for travel time certainty. Consequently, when travel time variance for one route goes to infinity, travelers become indifferent between routes, which contradicts empirical observations. In contrast, a large travel time variance for one route would generate a strong preference for routes with better travel time certainty under RPT. In the two-route example, $\lim Var(t_2) \to \infty$ would generate $p_1 = 1$ no matter how small t_2 is, which is more reasonable behavioral reaction. More advanced SUE models may also consider travel time variance as disutility. However, this would create correlation between explanatory variables and the random error term, which is troublesome for model calibration.

Another key difference in behavioral assumptions between the models is that in SUE models, only relative value matters. Therefore, the route travel time variance is usually normalized. However, in RPT, the absolute value is also important in the route decision. One constraint in implementing RPT as a model would limit the overall travel time variance, which may relate to physical constraints of work schedules.

Moreover, the interpretation of route choice probabilities from the two models are different. RPT allows individual travelers different routes on different days (all else equal), while SUE gives each traveler the same route probabilistically. This difference is important for the following reasons:

1. modeling traveler learning behavior: in RPT, travelers explicitly learn about some alternatives as they are actually experienced; in SUE, travelers cannot know about the routes that are not traveled on, and some unreasonable assumptions are required about travelers possessing perfect information about alternatives never chosen,
2. modeling the behavior of individuals, which is important for air quality, pricing, and many other applications,
3. understanding the underlying logic of traveler behavior.

Future research is required to estimate individual preferences. The objective is not necessarily to have a better aggregate traffic assignment, it is to have a better disaggregate route choice. This will begin to matter as HOT lanes and other differentiated pricing schemes on roads are deployed, where ignoring preference for travel time reliability would result in misestimation of benefits.

3.4. Numerical example

This subsection provides more insights on characteristics of RPT through a numerical example. Travelers face travel time uncertainty when making day-to-day route choice decisions. SUE models assume that such uncertainty is due to imperfection in travel time estimation. Such imperfection is usually summarized by an error term (ξ_i) in route travel time. This term usually has a mean of 0 and a variance that is independent to route travel time. The scale of such error term is not going to be affected by individual route decision either in most models. In reality, travel time uncertainty comes from many sources, both internal and external of the route choice process.

For example, if $x_1^i = 1$ represents the fact that the ith traveler chooses Route 1 and $Pr(x_1^i = 1) = p_1^i$, then x_1^i is a Bernoulli variable. If everyone has the same probability of choosing Route 1, the aggregate flow $X_1 = \sum_{i=1}^{n} x_1^i$ follows a binomial dis-
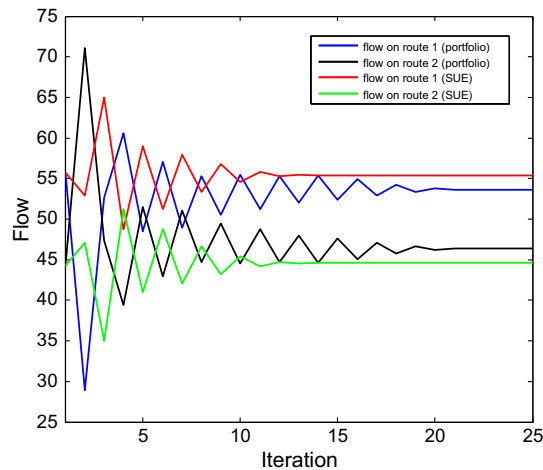
Fig. 4. Route traffic flow from RPT and SUE traffic assignment models.

tribution and $Var(X_1) = np_1(1 - p_1)$. If Vickrey's model (Vickrey, 1963) (in which travel time is a linear function of travel demand) is used as the link performance function, the stochasticity in route choice decision would generate a variance of $T_1^0 \beta Var(X_1)$, where T_1^0 is the free flow travel time on Route 1 and β is a parameter. If a link performance function of higher order (like BPR function) is used, the variance will be different but will not disappear. Heterogeneity among travelers (which help to generate different route choice probability p_1) will make the variance of route travel time more complicated. This variance in route travel time is internal to the travel choice process, which is ignored by the SUE model. Sources of travel time uncertainty could also be external to the route choice processing. Examples include signal control, freeway bottleneck activation when demand is close to capacity, stochasticity in background travel demand, among others. Most existing models focus on travel time prediction, and future research is required to develop an accurate travel time variance prediction model. To avoid unnecessary complexity in demonstrating RPT, this research adopts a Monte Carlo simulation approach to generate travel time variance in both the simple numerical example and the field study.

To demonstrate how to apply the RPT in the traffic assignment problem, we develop an iterative process on the stylized network of one OD pair and two links previously used in this section. We assume Route 1 has a free flow travel time of 10 min and a capacity of 40 veh/h; Route 2 has free flow travel time of 15 min and a capacity of 60 veh/h. Totally 100 travelers are going to make route decisions. For demonstration purpose, we assume $\alpha = 1$ and $\tau = 0.83$ for all travelers, although careful calibration is required before any application in real world. The standard BPR function is used to calculate route travel time. To simulate travel time uncertainty caused by various sources, we assume that the demand on any route could vary in the range $[X - 5, X + 5]$ under a uniform distribution, where X is the travel demand generated from the last iteration. Travel time associated with each possible travel demand value could be calculated and the overall travel time variance can be evaluated. With BPR function, the variance in travel time is small when demand is also small and increases rapidly when travel demand approaches or exceeds capacity. This process is created for demonstration purposes and a more accurate travel time variance model based on field data will be developed for future study. The assignment process is summarized below:

1. Step 1: Travel time and travel time variance are estimated for each route given route travel demand X_i.
2. Step 2: Each traveler would select an individual route choice strategy according to the two-parameter model of RPT (although reliability threshold can also be used, the two-parameter model is simpler in numerical evaluation). Individual alternatives include choosing any route between the two, or picking a route portfolio and choosing routes according to a probability.
3. Step 3: The total number of users of each route is aggregated from individual decision.
4. Step 4: Check if any traveler still has motivation to change strategy compared to last iteration. If yes, goes back to Step 1; if not, stop.

In order to compare the results with SUE, we conduct the SUE assignment on the same network using a Probit SUE model (route choice probability is determined by Eq. (17)). The results are summarized in Fig. 4. Because we assume travelers are homogeneous in both travel time (and its variance) perception and preference, they will follow the same strategies in this numerical example. Fig. 4 shows that travelers actually followed mixed strategy over a portfolio of two routes (because all 100 travelers would use one route if this route dominates the other under RPT with homogenous travelers). Neither route dominates the other in both travel time and travel time variance through all iterations. The probability of randomly choosing one route for each iteration can be calculated by dividing the number of route users in that iteration by the total OD demand. After about 20 iterations, the traffic assignment based on RPT converges to a flow pattern that differs from what is predicted by the Probit SUE model. The Probit Model also converges faster under current settings, although a different set of param-
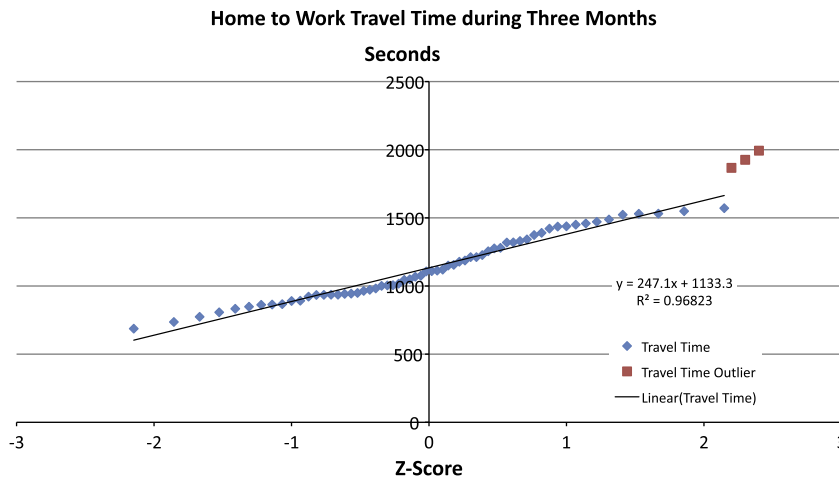
Home to Work Travel Time during Three Months

Fig. 5. Normal probability plot of home-to-work travel time observations for one subject during 13 weeks.

eters might generate different results. More importantly, RPT allows heterogenous travel time perception and preference for travel time reliability. Future empirical studies will extend this research.

4. Field study

The RPT we propose shows when the strategy of randomly choosing a route among alternatives is superior to the strategy of always using one route, given travelers' belief of network conditions. The posterior outcome of perceived average travel time could differ from the perceived expected travel time based on prior information. In the long term, however, especially for commute trips where enough experience has been gained to develop consistent perceptions of network conditions, we anticipate convergence between prior and posterior estimates. To check the theoretical reasoning, this section tests the proposed model against field data.

4.1. Route travel time and variance

This study uses the same GPS data presented in Section 2.1. In order to evaluate whether a rational decision maker guided by our theoretical model is likely to choose a portfolio of routes due to concerns of travel time reliability, travel time distributions of different routes are required. Ideally, we can obtain this information by observing day-to-day route choices during a period sufficiently long that we can collect enough samples for each route to empirically establish its travel time distribution. However, this is infeasible due to limited resources for most studies, especially for those subjects with very diverse route choices. Moreover, relying exclusively on direct observations also limits our ability to extend our analysis to the general population and to inform travel demand modeling efforts. Instead, route travel time distributions are generated based on average link traffic speed using a Monte Carlo simulation strategy.

Given the mean and variance of link speed observed from a large set of GPS data, we can simulate route travel time by generating random link speed and then summing up link travel time for all links along a specific route. Although route travel time has been widely assumed to follow a normal distribution in previous research (e.g. (Liu et al., 2004; Ryuichi and Mohamed, 1997)), this assumption is empirically tested here using GPS observations. Fig. 5 provides the normal probability plot of home-to-work travel time observations for one subject over 13 weeks. The *y*-axis represents route travel time in seconds, while the *x*-axis represents the *Z*-score of corresponding points ordered from small to large. The normality of data is established if a straight line can be fit to the points. Three points in red are clearly outliers according to the plot. According to the original data, the unusually long travel time in these three cases are due to stops near the destination. These stops could be due to activities such as searching for parking, visiting a coffee shop, or making a phone call along the route. However, GPS data alone cannot detect the causes and future research is required to define those trips.

Although the normal probability plot provides an intuitive illustration of how well a normal distribution fits the data, a more robust statistical test is required. We apply the Shapiro–Wilk test to all subjects and choose 0.05 as the critical value to reject normal assumption. In total, the assumption that route travel time follows a normal distribution has been rejected for 22 out of 95 subjects (23%). Future research efforts for better detecting side trips can help to exclude outliers in current commute time data set and the normal assumption could be more convincingly supported. Given that the majority of evidence does not reject the normal assumption, we assume route travel time follows a normal distribution in the following sections.

In order to simulate route travel time from random draws of link speed, assumptions about speed interdependency among different links on the network have to be made. Some studies conveniently assumed link travel time are independent and identically distributed (IID), which implies no correlation across links. However, there is clearly speed correlation across

links, presumably due in part to exogenous factors like weather, holidays, etc. affecting overall demand and in part to congestion (recurring and non-recurring) causing link interactions. Although many previous studies investigate the short-term spatiotemporal pattern of traffic flow (e.g. Kalman filtering approach by Whittaker et al. (1997)), there is no consensus on how to extend these models to a real network where the number of variables to be estimated become prohibitive (Kamarianakis and Prastacos, 2005). The reality is somewhere between the IID and perfection correlation. This study would use both assumptions and leave the development of more accurate route travel time prediction models to future research.

4.2. Mixed strategy in route choice

Many reasons explain the choice of multiple commute routes during a period of time. This study adds one more explanation by assuming that some travelers seek to minimize travel time while maintaining an appropriate level of travel time reliability. In this study, 60 subjects used more than 1 morning commute route during the study period. We then compare the predicted travel time of the two most frequently used alternatives suggested by GPS data.

Under the IID assumption, random numbers are drawn separately for each link from a standard–normal distribution and a normal distributed link travel time is calculated using the mean link travel time and the variance previously derived from GPS data. A minimal speed of 12.8 km h^{-1} is assumed to truncate the extremely long travel time. Path travel time can then be obtained by summing link travel time along the path. In all, 15 random draws are conducted to simulate random commute times for 15 days and the mean and variance of the travel time for each path can be calculated and compared. Similarly, the same process is followed to derive path travel time and variance in the case of perfect correlation, except the same random number is used for all links in each day so that link travel time are perfectly correlated across the network.

Under the IID assumption, there is no single dominant route (travel time always shorter during 15 days) in 38% of cases. It drops to 18% under perfect correlation conditions. If we define a dominant route as the those which possess both shorter travel time and smaller travel time variance, 16 out of 60 subjects do not have a dominant route under IID conditions (12 under the perfect correlation condition). Thus, it is possible for these people to choose a portfolio of routes in order to trade-off between travel time and travel time reliability, as illustrated by our theoretical model. However, better designed experiments controlling for more confounding factors are required to establish a causal relationship between the process of seeking travel time reliability and the choice of a route portfolio.

5. Conclusions

Many travelers use multiple routes to connect the same origin and destination on different days. This paper demonstrates that choosing a portfolio of routes could be the rational choice of a traveler with multiple criteria (e.g. minimizing journey time subject to avoiding frequent lateness) who wants to optimize route decisions under variability. This result can be extended to the choice among N alternatives and by following a more sophisticated objective function.

An implementation of RPT with a two-parameter model would collapse to the UE model when the value of reliability is small. In contrast, travelers would always choose a portfolio of routes and adopt a mixed strategy within the route portfolio if the value of reliability is significantly larger than the value of time. Empirical studies from the literature suggest the value of reliability is comparable to the value of time and choosing a portfolio of routes is more preferable when conditions specified in this paper are satisfied.

Although both RPT and SUE can generate stochasticity in route decisions, they differ in nature. Stochasticity in SUE models is due to randomness in choices when travelers cannot tell which route is better. However, adopting a mixed strategy over a portfolio of routes is a rational choice under RPT. When travel time variance of one route goes to infinity, travelers cannot distinguish it from other routes according to SUE models and will choose them indifferently. However, RPT predicts travelers would never choose such as a route, which is more plausible in behavioral theory. Both models also differ in the way stochasticity in route choice is interpreted.

RPT can be applied to the traffic assignment problem. A simple numerical example is presented to predict traffic flow patterns based on individual preferences for travel time and travel time reliability. The modeling structure can also consider impact of individual travel experience and heterogeneity in preference. More data and calibration efforts are needed before any field application.

This paper further tested the proposed model against field data, and concludes the RPT is consistent with observations that many travelers chose a set of routes where no route dominates other routes in both travel time and travel time variance. This implies a trade-off between them is possible. However, many other factors may influence route choice decisions and make travelers choose different routes over time between the same origin–destination pair. For example, they may have to reroute to drop-off a child or spouse on certain days but not the other days. Some routes may be a better choice only certain days of the week. Travelers may also prefer different routes due to different weather conditions. This study provides an additional explanation for the stochasticity in individual route choice decisions. However, more data and research efforts are needed to evaluate these possibilities.

Acknowledgments

References

Allais, M., 1979. Expected Utility Hypotheses and the Allais Paradox: Contemporary Discussions of Decisions Under Uncertainty with Allais' Rejoinder. Chapter The So-called Allais Paradox and Rational Decisions under Uncertainty. Reidel Publishing Company, Dordrecht, pp. 434–698.

Avineri, E., Prashker, J., 2004. Violations of expected utility theory in route-choice stated preferences: certainty effect and inflation of small probabilities. Transportation Research Record: Journal of the Transportation Research Board 1894 (-1), 222–229.

Bekhor, S., Ben-Akiva, M., Ramming, M., 2006. Evaluation of choice set generation algorithms for route choice models. Annals of Operations Research 144 (1), 235–247.

Bernoulli, D., 1954. Exposition of a new theory on the measurement of risk. Econometrica: Journal of the Econometric Society, 23–36.

Bovy, P., 2009. On modelling route choice sets in transportation networks: a synthesis. Transport Reviews 29 (1), 43–68.

Brownstone, D., Small, K., 2005. Valuing time and reliability: assessing the evidence from road pricing demonstrations. Transportation Research Part A 39 (4), 279–293.

Carrion, C., Levinson, D., 2012. Value of travel time reliability: a review of current evidence. Transportation Research Part A: Policy and Practice 46 (4), 720–741.

Daganzo, C.F., Sheffi, Y., 1977. On stochastic models of traffic assignment. Transportation Science 11 (3), 253–274.

de Palma, A., Picard, N., 2006. Route choice behavior with risk averse users. Spatial Evolution and Modelling, 139–178.

Frejinger, E., Bierlaire, M., 2007. Capturing correlation with subnetworks in route choice models. Transportation Research Part B: Methodological 41 (3), 363–378.

Hester, D., Tobin, J. (Eds.), 1967. Risk Aversion and Portfolio Choice. John Wiley & Sons Inc., New York, London, Sidney.

Hollander, Y., 2006. Direct versus indirect models for the effects of unreliability. Transportation Research Part A: Policy and Practice 40 (9), 699–711.

Jackson, W., Jucker, J., 1982. An empirical study of travel time variability and travel choice behavior. Transportation Science 16 (4), 460–475.

Jan, O., Horowitz, A., Peng, Z., 2000. Using global positioning system data to understand variations in path choice. Transportation Research Record: Journal of the Transportation Research Board 1725 (-1), 37–44.

Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. Econometrica 47 (2), 263–291.

Kamarianakis, Y., Prastacos, P., 2005. Space–time modeling of traffic flow. Computers & Geosciences 31 (2), 119–133.

Katsikopoulos, K., Duse-Anthony, Y., Fisher, D., Duffy, S., 2002. Risk attitude reversals in drivers' route choice when range of travel time information is provided. Human Factors 44 (3), 466–473.

Knight, F., 1921. Risk, Uncertainty, and Profit. Houghton Mifflin Company, Boston, MA.

Knight, T., 1974. An approach to the evaluation of changes in travel unreliability: a "safety margin" hypothesis. Transportation 3 (4), 393–408.

Liu, H., Recker, W., Chen, A., 2004. Uncovering the contribution of travel time reliability to dynamic route choice using real-time loop data. Transportation Research Part A: Policy and Practice 38 (6), 435–453.

Lou, Y., Yin, Y., Lawphongpanich, S., 2010. Robust congestion pricing under boundedly rational user equilibrium. Transportation Research Part B 44 (1), 15–28.

Mahmassani, H., Chang, G., 1987. On boundedly rational user equilibrium in transportation systems. Transportation Science 21 (2), 89–99.

Markowitz, H., 1952a. Portfolio selection. The Journal of Finance 7 (1), 77–91.

Markowitz, H., 1952b. The utility of wealth. The Journal of Political Economy 60 (2), 151–158.

Mirchandani, P., Soroush, H., 1987. Generalized traffic equilibrium with probabilistic travel times and perceptions. Transportation Science 21 (3), 133–152.

Noland, R., Polak, J., 2002. Travel time variability: a review of theoretical and empirical issues. Transport Reviews 22 (1), 39–54.

Parthasarathi, P., 2011. Network Structure and Travel. Ph.D. Thesis, University of Minnesota, Twin Cities.

Pells, S., 1987. Evaluation of reductions in travel time variability. Unpublished PhD Thesis, Institute for Transport Studies, School of Economic Studies, University of Leeds.

Polak, J., 1987. A more general model of individual departure time choice. In: Transportation Planning Methods, Proceedings of seminar C held at the PTRC Summer Annual Meeting, University of Bath, Great Britain, vol. 290, pp. 247–258.

Rates, E., 2007. Atlanta Commute Vehicle Soak and Start Distributions and Engine Starts per Day Impact on Mobile Source, Atlanta.

Ryuichi, A., Mohamed, A., 1997. Using stated preference data for studying the effect of advanced traffic information on drivers' route choice. Transportation Research Part C: Emerging Technologies 5 (1), 39–50.

Small, K., Winston, C., Yan, J., 2005. Uncovering the distribution of motorists' preferences for travel time and reliability. Econometrica 73 (4), 1367–1382.

Spiess, H., Florian, M., 1989. Optimal strategies: a new assignment model for transit networks. Transportation Research Part B: Methodological 23 (2), 83–102.

Tversky, A., Kahneman, D., Kahneman, D., Slovic, P., Tversky, A., 2005. Judgment under uncertainty: heuristics and biases. In: Hamilton, D.L. (Ed.), Social Cognition: Key Readings in Social Psychology. Psychology Press, New York, pp. 167–177.

Vickrey, W.S., 1963. Pricing in urban and suburban transport. American Economic Review 53 (2), 452.

Von Neumann, J., Morgenstern, O., Kuhn, H., Rubinstein, A., 1947. Theory of Games and Economic Behavior. Princeton University Press, Princeton, NJ.

Wardrop, J., 1952. Some theoretical aspects of road traffic research. In: Proceedings of the Institute of Civil Engineers, Pt. II, vol. 1, pp. 325–378.

Watling, D., 2002. A second order stochastic network equilibrium model, I: Theoretical foundation. Transportation Science 36 (2), 149–166.

Whittaker, J., Garside, S., Lindveld, K., 1997. Tracking and predicting a network traffic process. International Journal of Forecasting 13 (1), 51–61.

Yan, J., 2002. Heterogeneity in Motorists' Preferences for Travel Time and Time Reliability: Empirical Finding from Multiple Survey Data Sets and its Policy Implications. Ph.D. Thesis, University of California.

Zhu, S., Levinson, D., Liu, H., 2010. Measuring winners and losers from the new I-35W Mississippi River Bridge. In: The 89th Annual Conference of Transportation Research Board, Washington, DC.

Zhu, S., Levinson, D., Liu, H., Harder, K., 2010. The traffic and behavioral effects of the I-35W Mississippi River bridge collapse. Transportation Research Part A: Policy and Practice 44, 771–784.