Article:
Batley, R.P. (2007) Marginal valuations of travel time and scheduling, and the reliability premium. Transportation Research E, 43 (4). pp. 387-408. ISSN 1366-5545

https://doi.org/10.1016/j.tre.2006.06.004

Published paper

Batley, R.P. (2007) *Marginal valuations of travel time and scheduling, and the reliability premium* - Transportation Research Part E: Logistics and Transportation Review 43(4) pp387-408

Marginal valuations of travel time and scheduling, and the reliability premium

Richard Batley

Institute for Transport Studies, University of Leeds, Leeds LS2 9JT, UK

rbatley@its.leeds.ac.uk

Telephone +44 (0) 113 343 6607

Fax +44 (0) 113 343 5334

Abstract

Previous works have established synonymity between the notions of uncertainty and unreliability, exploiting this in deriving marginal valuations of travel time and scheduling under uncertainty. Whilst valid for forecasting demand, such valuations fail to illuminate the costs of bearing unreliability - herein referred to as the 'reliability premium'. The paper derives marginal valuations of travel time and scheduling at the certainty equivalent, showing these to diverge from those under uncertainty. That divergence, which represents the marginal valuation of reliability, raises the possibility of bias should the costs of unreliability not be included in appraisal.

Keywords

Reliability, Valuation, Expected Utility, Risk Aversion, Risk Premium

1. Introduction

Although a precise understanding has often seemed elusive, it is widely accepted that the reliability of transport systems may impact upon the choices of travellers. Previous research has illuminated several facets of this proposition, but often without the authority of comprehensive evidence on the *value* of reliability to travellers. That such evidence is lacking can perhaps, in turn, be attributed to the difficulty of formulating a research apparatus that carries theoretical validity, is insightful, but remains practicable. These aspirations are the concern of the present paper.

The task of reviewing relevant literature is well-served by the recent contributions of Noland and Polak (2002) and De Jong *et al.* (2004), and it would seem unnecessary to offer further commentary in this regard. Suffice to say, De Jong *et al.* distinguish between three approaches to the valuation of reliability, specifically: I) mean *vs.* variance of the travel time distribution, II) percentiles of the travel time distribution, and III) scheduling models. The present paper exploits the third approach in particular, which is founded on the hypothesis that travellers may accommodate expectations of unreliability through their trip scheduling.

In the analysis of trip scheduling, Small's (1982) approach has received considerable support. Small extends the microeconomic theory of time allocation (Becker, 1965; De Serpa, 1971), supplementing the usual objective problem of utility maximisation subject to money and time constraints with a trip scheduling constraint deriving from Vickrey (1969). Implicit in Small's approach, however, is the assumption that scheduling choices are made under certainty, and this would seem to impose considerable restriction on its applicability. The usual accommodation of uncertainty - at least in terms of microeconomic theory - is to reformulate the objective problem from the maximisation of utility to one of maximising expected utility, as first proposed by von Neumann & Morgenstern (1947). The latter is indeed exploited by Noland & Small (1995), who establish merger between the works of von Neumann & Morgenstern and Small.

Two particular, but related, properties of Noland & Small's analysis might be noted. First, both the choice (i.e. departure time) and pay-off (i.e. arrival time) dimensions are represented continuously; this carries the attraction of permitting easy calculation of the optimal departure time. Second, interest is restricted to the morning commute of car travellers. Continuity in departure time would appear more reasonable for car travellers than for users of public transport services, since the latter are typically constrained by fixed service intervals. Bates *et al.* (2001) develop Noland & Small further, first considering its amenability to public transport users, and then applying the analysis to an interest in marginal valuations of travel time and scheduling under uncertainty. These marginal valuations derive from choices between two public transport services, wherein each service offers a range of departure times, and the consequent arrival times are characterised by uncertainty.

The present paper follows the basic thesis of Noland & Small and Bates *et al*., but pursues a number of extensions. Time is represented as a discrete variable in both departure and arrival dimensions. Not only is this more faithful to von Neumann & Morgenstern (1947), but it permits ready accommodation of public transport users. The discrete representation, furthermore, is amenable to one of the principal analytical tools of travel demand analysis, namely the Random Utility Model (RUM). Aside from this presentational distinction, the substantive contribution of the paper is to apply the workings of Noland & Small and Bates *et al.* to an interest in travellers' attitudes to risk. Specifically, the paper reconciles the transport planner's notion of unreliability with the microeconomist's notion of risk (e.g. Pratt, 1964; Arrow, 1970), and in so doing reveals the treatment of risk within marginal valuations of travel time and scheduling under uncertainty. This provokes the proposition of a further metric for valuing reliability, referred to as the 'reliability premium'. The reliability premium, which is drawn from analogy with Pratt's 'risk premium', isolates those costs arising specifically from unreliability and, equivalently, the benefits that would transpire should unreliability be eliminated. The paper exploits the reliability premium in deriving a further set of marginal valuations of travel time and scheduling, this time at the 'certainty equivalent'. The latter valuations are then compared with those derived under uncertainty, an exercise that serves to yield marginal valuations of reliability *per se*. The paper considers the relevance of marginal valuations of reliability for both forecasting and economic appraisal.

2. Theory of individual choice under uncertainty

Microeconomic theory of individual choice under uncertainty is founded on the proposition that there exists some relation between an individual's choices under risk or uncertainty and a distribution of outcomes. This has been exploited in the reliability literature; in the present paper we shall examine the particular proposition that travellers choose a time of departure on the basis of a distribution of the consequent arrival times. The interpretation of the probability distribution has been the source of some contention in the microeconomic literature, since it is embroiled with the dichotomy between risk and uncertainty. Keynes (1921, 1936) and Knight (1921) are helpful in this regard, characterising risk as situations where probabilities of outcomes are known (or knowable), and uncertainty as situations where such probabilities may be neither knowable nor definable. Though it offers an appealing clarity, it would not seem crucial to the subsequent analysis that one commits to this or any other typology. Rather the terms risk and uncertainty will, in what follows, be used interchangeably without implication.

More central to our interest is the precise nature of the relation between an individual's choices under uncertainty and the distribution of outcomes; this may be formalised in the following terms:

Let E be a finite and exhaustive set of mutually exclusive 'events':

$$E = \{e_1, \ldots, e_K\}$$

Let E be associated with a vector $\mathbf{w}$, which is referred to as a 'prospect', and gives the probability p_k that each event $e_k \in E$ will occur, together with the pay-off w_k to the individual should that event indeed occur, thus:

$$\mathbf{w} = (w_1, \ldots, w_K; p_1, \ldots, p_K)$$

Following the usual rules of probability, it is necessarily the case that $\sum_{k=1}^{K} p_k = 1$, implying that $0 \leq p_k \leq 1$ for $k = 1, \ldots, K$.

Let S be a finite and exhaustive set of such prospects, from which the individual is invited to choose his or her preferred prospect:

$$S = \{\mathbf{w}_1, \ldots, \mathbf{w}_N\}$$

The seminal exposition of von Neumann & Morgenstern (1947) established a set of necessary and sufficient axioms on the above definitions such that an individual could be represented as if choosing the prospect $\mathbf{w}_n \in S$ that yields maximum expected utility. These axioms have subsequently been adapted in various ways; of particular note in this regard are the contributions of Marschak (1950), Herstein and Milnor (1953) and Fishburn (1970). In what follows, however, we remain faithful to the original exposition, which provides a basis for the following proposition:

For any pair of prospects $\mathbf{w}_q, \mathbf{w}_r \in S$:

$$\mathbf{w}_q \succeq \mathbf{w}_r \text{ iff } Y(\mathbf{w}_q) \geq Y(\mathbf{w}_r)$$

where $Y(\mathbf{w}_n)$ is the expected utility of prospect $\mathbf{w}_n$, and is itself given by:

$$Y(\mathbf{w}_n) = \sum_{k=1}^{K} p_{kn} U(w_{kn}) \text{ for all } \mathbf{w}_n \in S \tag{1}$$

where $U(w_{kn})$ is the utility deriving from pay-off w_{kn}.

Having summarised the theory, let us conclude this section with discussion of several issues that follow. A first point to make is that von Neumann & Morgenstern's (1947) axioms are not sufficient to support a continuous representation of the pay-off dimension, as adopted by Noland & Small (1995) and Bates *et al.* (2001), for which the expectations operation in (1) would require integration. A continuous representation is in fact feasible, but must be supported by a set of axioms that diverges from von Neumann & Morgenstern's; see Fishburn (1970) for further instruction.

A second point is that a discrete representation of the choice dimension promotes ready application to RUM, as demonstrated by Marschak *et al.* (1963). This contrasts with the inherently ambiguous process of 'discretising' a continuous dimension (e.g. Jotisankasa *et al.*, 2004). In exploiting Marschak *et al.*'s demonstration, it is useful to articulate their basis for adopting a probabilistic representation, which is as follows. Consider an individual faced with a repeated choice task under uncertainty. On any given repetition of the choice task, he or she is able to order a set of prospects in terms of expected utility, but on successive repetitions this ordering may show variability. More formally, the probability of choosing any prospect $\mathbf{w}_q \in S$ can be expressed as RUM, such that:

$$P(\mathbf{w}_q | S) = \Pr\{Y(\mathbf{w}_q) \geq Y(\mathbf{w}_r)\} \text{ for all } \mathbf{w}_r \in S, q \neq r \qquad (2)$$

A third point concerns the comparability of expected utility $Y(\mathbf{w}_n)$ across the prospects $\mathbf{w}_n \in S$. Baumol (1958) dismisses the conventional wisdom that choice under uncertainty permits the mutation of utility from an ordinal metric to a cardinal one; that is to say (1) relies indeed on the proposition that $U(w_{kn})$ is cardinal, but this does not distract from the requirement that $Y(\mathbf{w}_n)$ is an entirely ordinal construct. Moreover, there is no basis for interpreting differences in expected utility (or any ratio thereof) across prospects. Whilst the definition of RUM (2) adheres to this requirement, it is important to ensure that applications of RUM to forecasting or appraisal do not inadvertently stray into cardinality. The latter possibility is considered by Batley (2006), who demonstrates that 'log sum' measurements of consumer surplus derived from RUM carry an inference of cardinality.

Finally, it might be acknowledged that von Neumann & Morgenstern's (1947) analysis has been the subject of sustained assault almost ever since it was conceived. In the contemporary literature, experimental economists have been active contributors, presenting instances of individual choices that apparently violate expected utility maximisation (Kahneman & Tversky's (2000) compendium includes several such works). This has prompted generalisations of von Neumann & Morgenstern's analysis; for example, several of the identified violations may be accommodated through non-linear forms on the event probabilities p_{kn} in (1); see Edwards (1955) and Kahneman & Tversky (1979). Whilst important to acknowledge, this critique has not to date succeeded in deposing the paradigm of expected utility maximisation, and it would therefore seem entirely reasonable to adhere to von Neumann & Morgenstern in what follows.

3. The theory applied to trip scheduling

The contributions of Noland & Small (1995) and Bates *et al.* (2001) establish precedence in applying the theory of individual choice under uncertainty to the context of trip scheduling. The purpose of the present section is to explicate this application; the subsequent analysis follows these precedent works reasonably faithfully, save for the representation of time as a discrete variable.

3.1. Preliminary definitions

In applying the theory outlined in section 2 to trip scheduling, it would seem reasonable to represent choice in terms of departure time and events in terms of arrival time, and to postulate that uncertainty derives from the distribution of arrival times for any given departure time. With reference to Noland & Small's typology, this distribution could feasibly derive from recurrent delay, from incident-related delay, or from both of the aforementioned. More formally:

Let A be a finite and exhaustive set of arrival times:

$$A = \{a_1, ..., a_K\}$$

Let D be a finite and exhaustive set of departure times:

$$D = \{d_1, ..., d_N\}$$

The latter corresponds to the choice set S, as defined in section 2:

$S = \{\mathbf{w}_1, ..., \mathbf{w}_N\}$

Let the expected utility Y_n of any prospect $\mathbf{w}_n \in S$ be given by:

$$Y_n = \sum_{k=1}^{K} p_{kn} U_{kn} \tag{3}$$

where U_{kn} is the utility deriving from the arrival time a_k, having departed at d_n, and p_{kn} is the associated event probability.

3.2 Utility function

It remains to equip utility U_{kn} with more precise form. Whilst the theory of section 2 would seem to impose few restrictions on such form, the trip scheduling literature demonstrates considerable support for Small's (1982) utility function, and this function is straightforwardly adopted by Noland & Small (1995) and Bates *et al.* (2001). In what follows, we shall ourselves follow this convention, but in so doing demonstrate that Small's utility function implies particular properties with respect to travellers' attitudes to unreliability in arrival time.

With reference to Small (1982), define utility:

$$U_{kn} = \alpha T_{kn} + \beta SDE_k + \gamma SDL_k + \delta L_k \quad (4)$$

where:

T is travel time

SDE is schedule delay early

SDL is schedule delay late

L is a dummy variable that is unitary in the case of late arrival, and zero otherwise

Hence Small's utility function (4) can be seen to be a linear function of four attributes - travel time, schedule delay early, schedule delay late and a lateness dummy - where the latter three are conditioned by the 'preferred arrival time' (PAT) of the traveller, which we take as given. On this basis, let us re-express the attributes of (4) in terms of our dimensions of interest - arrival time and departure time - for given PAT :

$T_{kn} = a_k - d_n$

$SDE_k = \max[(PAT - a_k),0]$

$SDL_k = \max[(a_k - PAT),0]$

$L_k = 1$ if $(a_k - PAT) > 0$, $= 0$ otherwise

Now completing our application of the theory of individual choice under uncertainty to the context of trip scheduling, let us explicate the definition of the prospect vector, thus:

$$\mathbf{w}_n = \left[(T_{1n}, SDE_1, SDL_1, L_1), \ldots, (T_{Kn}, SDE_K, SDL_K, L_K); p_{1n}, \ldots, p_{Kn}\right]$$

Moreover the adoption of Small's utility function implies the proposition that pay-offs are given by an amalgam of the four attributes of that function, three of which are measured in time units, and the fourth as a dummy variable. Thus while previous applications of von Neumann & Morgenstern (1947) have focussed heavily on monetary pay-offs, it would in the present context seem natural, and not outside the remit of the theory, to represent pay-offs in time units. In any case, travel costs do not routinely vary by arrival time, such that the arrival time dimension may not be readily amenable to monetisation. That does not preclude the possibility that arrival time (in particular, lateness) may incur incidental costs. Rather any such costs will arise as a function of arrival time, as would seem intuitive in the case of lateness, and the attitudes of travellers to such costs will manifest in their attitudes to the uncertainty of arrival time.

Let us derive a schematic representation of the utility function (4) for any departure time $d_n \in D$, but restrict attention to a binary subset of arrival times $\tilde{A} \subseteq A$. In particular, let $\tilde{A} = \{a_i, a_j\}$, wherein the following relations hold: $a_{\min} \leq a_i < PAT < a_j$. That is to say, let $a_{\min}$ be the earliest feasible arrival time (i.e. in free flow conditions), and let a_i and a_j be two further arrival times that are defined arbitrarily save for the requirement that a_i falls before the PAT and a_j after the PAT.

Figure 1: Small's utility function, for given departure time, with $\alpha < \beta$ (ABOUT HERE)

This is illustrated by Figure 1, which represents arrival time on the horizontal axis, and utility on the vertical. All attributes of the utility function are 'bad', i.e. $\alpha, \beta, \gamma, \delta < 0$ in (4), and U is therefore drawn in the lower right quadrant of the figure; note also that U originates at the earliest feasible arrival time $a_{\min}$. Now consider the utility derived at arrival times a_i and a_j, respectively:

At $a = a_i$: $U = \alpha(a_i - d) + \beta(PAT - a_i)$

At $a = a_j$: $U = \alpha(a_j - d) + \gamma(a_j - PAT) + \delta$

Consider also the utility derived as the arrival time approaches the PAT from a_i and a_j, respectively:

As $(PAT - a_i) \to 0$: $U \to \alpha(PAT - d) \to \alpha(a_i - d) + \alpha(PAT - a_i)$

As $(a_j - PAT) \to 0$: $U \to \alpha(PAT - d) + \delta \to \alpha(a_j - d) - \alpha(a_j - PAT) + \delta$

We can now establish the slope of the utility function U, which for arrival times before the PAT must be $(\alpha - \beta)$, and for arrival times after the PAT must be $(\alpha + \gamma)$. For purposes of illustration, Figure 1 exploits Small's (1982) empirical finding that $\alpha < \beta$, i.e. that travel time imposes greater disutility than schedule delay early. This engenders the property that U is strictly decreasing in utility, and steeper after the PAT than before. Contrariwise, if it instead held that $\beta < \alpha$, then the two portions of U would show opposing slopes, with arrival times before the PAT characterised by positive slope, and arrival times after the PAT by negative slope. In this latter case, the absolute slope of the first portion of U may be greater than or less than the second portion, depending on the relative disutility of schedule delay early. Figure 2 embodies the relations $\beta < \alpha$ and $|(\alpha - \beta)| < |(\alpha + \gamma)|$, such that the two portions of U show opposing slopes, with the portion before the PAT carrying lower absolute slope than the portion after the PAT.

Figure 2: Small's utility function, for given departure time, with $\beta < \alpha$ (ABOUT HERE)

Finally, it might be observed that Figures 1 and 2 diverge from the more usual presentation of Small's utility function, as perhaps illustrated by Figure 1 of Bates *et al.* (2001). This is because Figures 1 and 2 of the present paper consider both travel time and scheduling parameters, whereas Figure 1 of Bates *et al.* considers only scheduling parameters. The presentation adopted in the present paper is motivated by a desire to clearly articulate the relation of the *complete* utility function to the expected utility function, an interest that we shall pursue in the following sections. Moreover, it is perhaps helpful - for purposes of distinction - to refer to Figure 1 of Bates *et al.* as Small's 'scheduling function', and to Figures 1 and 2 of the present paper as Small's 'utility function'.

3.3 Expected utility function

For reasons of brevity, this and subsequent sections will focus attention on Small's utility function with the property $\alpha < \beta$, i.e. as in Figure 1, particularly as this property carries the empirical support of Small (1982) himself. Whilst it should be reassured that the general principles of the analysis extend also to functions with the property $\beta < \alpha$, i.e. as in Figure 2, some differences in analysis do arise, and these are remarked upon where relevant.

In order to promote expositional clarity, let us once again restrict consideration to a binary subset of arrival times $\tilde{A} \subseteq A$, where $\tilde{A} = \{a_i, a_j\}$, this time imposing (at least for the moment) the restriction only that $a_i < a_j$. For each departure time $d_n \in D$, define the prospect:

$$\mathbf{w}_n = \left[(T_{in}, SDE_i, SDL_i, L_i), (T_{jn}, SDE_j, SDL_j, L_j); p_{in}, (1 - p_{in})\right]$$

Now propagating expected utility (3) with the utility function (4), let us write the expected utility deriving from the above prospect:

$$\begin{aligned} Y_n = &\left\{p_{in}\left[\alpha T_{in} + \beta SDE_i + \gamma SDL_i + \delta L_i\right]\right\} \\ &+ \left\{(1 - p_{in})\left[\alpha T_{jn} + \beta SDE_j + \gamma SDL_j + \delta L_j\right]\right\} \end{aligned} \quad (5)$$

Or more generally:

$$Y_n = \alpha E(T_n) + \beta E(SDE_n) + \gamma E(SDL_n) + \delta E(L_n) \quad (6)$$

Figure 3: Utility and expected utility functions, for given departure time, with $\alpha < \beta$

(ABOUT HERE)

The binary subset of arrival times in (5) permits easy translation to the diagrammatic analysis of Figure 3, wherein three cases might usefully be identified depending on the relation of the arrival times to the PAT, thus:

Case 1: $a_i < a_j \leq PAT$

Case 2: $PAT \leq a_i < a_j$

Case 3: $a_i < PAT < a_j$

Let us draw some observations relating to each of the three cases, as follows:

Case 1: $a_i < a_j \leq PAT$

Expected utility - labelled $Y^{\text{Case 1}}$ in the figure - will in this case fall somewhere on the section of the utility function U preceding the PAT; precisely where will depend on the probability p_i that arrival time a_i actually occurs.

Case 2: $PAT \leq a_i < a_j$

By contrast to the previous case, expected utility $Y^{\text{Case 2}}$ will now fall somewhere on the section of the utility function U that comes after the PAT. Since utility is strictly decreasing in travel time, it must hold that $Y^{\text{Case 2}} < Y^{\text{Case 1}}$; hence Case 1 dominates Case 2.

Case 3: $a_i < PAT < a_j$

Again with reference to Figure 3, consider the pair of arrival times $a_i^{\text{Case 3}}$ and $a_j^{\text{Case 3}}$, which satisfy the requirements of Case 3, but are otherwise defined arbitrarily. If we identify the point at which $a_i^{\text{Case 3}}$ and $a_j^{\text{Case 3}}$ intersect the utility function U, then expected utility $Y^{\text{Case 3}}$ will fall somewhere on the straight line joining these two points. The discontinuity of U introduces some complexities, however, as we shall see in due course.

3.4 Choice between prospects

Having derived expected utility for any departure time $d_n \in D$, let us consider the choice between a pair of prospects $\mathbf{w}_q, \mathbf{w}_r \in S$, which correspond to the departure times $d_q, d_r \in D$. Furthermore, let $d_q < d_r$, such that d_q is the earlier of the two departures. Since von Neumann & Morgenstern's (1947) axioms refer entirely to the properties of pairs, a binary choice would not seem unreasonably restrictive. It might also be remarked that binary choice would seem particularly amenable to implementation within Stated Preference (SP).

Following section 2, it must hold that:

$\mathbf{w}_q \succeq \mathbf{w}_r$ iff $Y_q - Y_r \geq 0$

$\mathbf{w}_r \succeq \mathbf{w}_q$ iff $Y_r - Y_q \geq 0$

$\mathbf{w}_q \sim \mathbf{w}_r$ iff both $Y_q - Y_r \geq 0$ and $Y_r - Y_q \geq 0$

Since these three preference relations are dictated by the difference in expected utility, let us derive:

$$Y_q - Y_r = \alpha(d_q - d_r) + \left[(p_{ir} - p_{iq})(\alpha\Delta a + \beta\Delta SDE + \gamma\Delta SDL + \delta\Delta L)\right] \qquad (7)$$

where:

$\Delta a = (a_j - a_i) > 0$

$\Delta SDE = (SDE_j - SDE_i) \leq 0$

$\Delta SDL = (SDL_j - SDL_i) \geq 0$

$\Delta L = 1$ if $a_i < PAT < a_j$, $= 0$ otherwise

It may therefore be seen, with reference to (7), that the difference in expected utility is a function of the difference in departure time, as well as differences in expected arrival time, expected schedule delay early, expected schedule delay late and the expected lateness dummy (where the latter may otherwise be referred to as the probability of lateness).

Now directing attention to Figure 4, let us translate the preceding algebraic analysis to a diagrammatic one. The reader's attention is first drawn to the utility functions U_q and U_r, which pertain to the departure times d_q and d_r respectively. With reference to the horizontal axis, note that the two departures will, assuming constant free flow travel time, have different times for their earliest feasible arrivals (i.e. $a_{\min}(d_q) < a_{\min}(d_r)$). The two utility functions originate, therefore, from different points on the arrival time axis. With reference to the vertical axis, the two functions are separated by the constant $\alpha(d_q - d_r)$, which represents the utility difference arising from the difference in their respective departure times. Otherwise the two functions carry the same properties, with common scheduling parameters and PAT .

Figure 4: Choice between departure times, with $\alpha < \beta$ (ABOUT HERE)

Having defined the relevant events, we can now derive the expected utility functions relating to the prospects $\mathbf{w}_q$ and $\mathbf{w}_r$; these functions are labelled Y_q and Y_r respectively in the figure. Precisely where expected utility falls on these functions will be determined by the expected arrival times $E(a_q)$ and $E(a_r)$. If however it holds that $a_{\min} \leq a_i$, i.e. the earlier arrival time within the prospects falls at or after the earliest feasible arrival time, it would seem uncontroversial to make the assertion that $p_{iq} > p_{ir}$, i.e. the earlier departure

d_q is more likely to arrive at a_i than the later departure d_r. Though it is left in practice to empirical investigation to identify the expected arrival times $E(a_q)$ and $E(a_r)$, the figure illustrates a situation where $E(a_q) < PAT < E(a_r)$, with an outcome that $Y_q > Y_r$. In this situation, therefore, prospect $\mathbf{w}_q$ would be chosen over prospect $\mathbf{w}_r$.

4. Marginal valuations of travel time and scheduling, and the reliability premium

The previous section applied the theory of individual choice under uncertainty to trip scheduling. This equips us with the necessary theory to now develop the principal interest of the paper, which is in the value of reliability to an individual traveller. Whilst acknowledging that the value of reliability has been variously defined (see de Jong *et al*., 2004), the following analysis will exploit and extend Bates *et al*.'s (2001) definition, which is grounded in the theory of section 3. Let us first summarise Bates *et al*.'s definition, before considering its extension.

4.1 Bates et al.'s marginal valuations of travel time and scheduling under uncertainty

Following Bates *et al.*, the reliability of arrival time may be represented in terms of the event probabilities of von Neumann & Morgenstern's (1947) analysis; with reference to (5), for example, any change in reliability (i.e. p_i and p_j) will impact upon expected arrival time and, by implication, expected travel time, expected schedule delay early, expected schedule delay late, and the expected lateness dummy. Applying this representation of reliability to empirical investigation, Bates *et al.* devise and implement a binary choice SP experiment, and from this data infer marginal valuations of expected travel time, expected schedule delay early, expected schedule delay late and the expected lateness dummy.

Let us illustrate the nub of Bates *et al.*'s empirical investigation, but within the context of our own working. We begin by supplementing the expected utility function (5) with a variable representing travel cost, noting importantly that cost is indexed by departure time but not by arrival time. More formally, for any prospect $\mathbf{w}_n \in S$ let:

$$\hat{Y}_n = Y_n + \phi c_n \tag{8}$$

where c_n is the travel cost of prospect $\mathbf{w}_n$. Adopting (8), let us return to the choice between the pair $\mathbf{w}_q, \mathbf{w}_r \in S$. Critical to this choice is the point at which the individual is indifferent between the prospects; re-working (7) we can establish that:

If $\hat{Y}_q - \hat{Y}_r = 0$ then:

$$\phi(c_r - c_q) = \alpha(d_q - d_r) + [(p_{ir} - p_{iq})(\alpha\Delta a + \beta\Delta SDE + \gamma\Delta SDL + \delta\Delta L)] \qquad (9)$$

The proposition now emerges that an individual traveller might be willing to exchange an adjustment in cost difference, i.e. the left-hand side of (9), for adjustments in one or more of the travel time and scheduling differences, i.e. the right-hand side of (9). Accepting this proposition of exchange, it would seem a relatively small extension to isolate the rate of exchange between travel cost and each constituent of expected utility, and thereby derive marginal valuations of expected travel time, expected schedule delay early, expected schedule delay late, and the expected lateness dummy. For example, let us derive marginal valuation of expected travel time $Vo(E(T))$, thus:

If $\hat{Y}_q - \hat{Y}_r \big|_{\beta,\gamma,\delta=0} = 0$ then:

$$\phi(c_r - c_q) = \alpha\{[E(a_q) - E(a_r)] + (d_r - d_q)\}$$

$$Vo(E(T)) = \frac{\alpha}{\phi} = \frac{(c_r - c_q)}{[E(a_q) - E(a_r)] + (d_r - d_q)} \qquad (10)$$

Marginal valuations of expected schedule delay early $Vo(E(SDE))$, expected schedule delay late $Vo(E(SDL))$, and the expected lateness dummy $Vo(E(L))$ can be derived analogously, as shown in Appendix A.

4.2 Attitudes to unreliability

The exposition of section 3 would, in applying the theory of individual choice under uncertainty to the reliability of arrival time, seem to establish synonymity between the microeconomist's notion of uncertainty (or risk) and the transport planner's notion of unreliability. This in turn provokes an interest in reconciling Bates *et al.*'s marginal valuations of travel time and scheduling under uncertainty with established microeconomic methods for valuing the risk inherent in prospects. In pursuing this interest one might, rekindling the earlier discussion in section 3.2, distinguish between attitudes to risky monetary outcomes and attitudes to risky time outcomes. Whilst the former have attracted considerable research attention, the latter have not, hence the opportunity for further investigation.

To this end, let us consider the particular concept of the 'risk premium' (Pratt, 1964), which carries an appealing intuition and commands considerable support across the microeconomic community. The risk premium arises from the relation between the expected utility function and its underlying utility function, and is crucially dictated by the individual's attitude towards risk. Before proceeding to the risk premium, let us then consider attitudes to risk as they apply to our interest in the reliability of arrival time. We begin by deriving the utility of the expected arrival time, thus:

$$U(E(a_n)) = \alpha[E(a_n) - d_n] + \beta \max[(PAT - E(a_n)),0] + \gamma \max[(E(a_n) - PAT),0] + \delta L(E(a_n))$$

where:

$$E(a_n) = p_{in} a_i + (1 - p_{in}) a_j \qquad (11)$$

$L(E(a_n)) = 1$ if $(E(a_n) - PAT) > 0$, $= 0$ otherwise

Then comparing the utility of the expected arrival time to expected utility, Jensen's inequality (see for example Johansson, 1991) provides a basis for the following inferences:

- If $Y_n = U(E(a_n))$ then the traveller is 'risk neutral in arrival time'
- If $Y_n > U(E(a_n))$ then the traveller is 'risk preferred in arrival time'
- If $Y_n < U(E(a_n))$ then the traveller is 'risk averse in arrival time'

It remains to establish the prevalence of these three relations, which again invokes the same three cases as section 3.3:

Case 1: $a_i < a_j \leq PAT$

This always yields the equality $Y_n = U(E(a_n))$, such that the traveller exhibits risk neutrality.

Case 2: $PAT \leq a_i < a_j$

The equality $Y_n = U(E(a_n))$ again holds, and an inference of risk neutrality may therefore be drawn.

Case 3: $a_i < PAT < a_j$

This case is more ambiguous, and it is instructive to introduce the further dichotomy:

Case 3.1: $E(a_n) < PAT$

Where Case 3 holds and the expected arrival time is earlier than the PAT, the difference between expected utility and utility of the expected arrival time is given by the quantity:

$$Y_n - U(E(a_n)) = (1 - p_{in})\left[\beta(a_j - PAT) + \gamma(a_j - PAT) + \delta L_j\right]$$

Since $\beta, \gamma, \delta < 0$ and $0 \leq p_{in} \leq 1$ by definition, and $(a_j - PAT) > 0$ by assumption, it must be the case that $Y_n < U(E(a_n))$. The traveller therefore exhibits risk aversion. This can be confirmed through reference to Figure 3, wherein the utility function U lies above the expected utility function Y throughout the interval $[a_i, PAT]$.

Case 3.2 $PAT < E(a_n)$

Where Case 3 continues to hold but the expected arrival time is now later than the PAT, the difference between expected utility and utility of the expected arrival time becomes:

$$Y_n - U(E(a_n)) = p_{in}[\beta(PAT - a_i) + \gamma(PAT - a_i) - \delta L]$$

where $L = L_j = L(E(a_n))$

In a similar manner to before, $\beta, \gamma, \delta < 0$ and $0 \leq p_{in} \leq 1$ by definition, and $(PAT - a_i) > 0$ by assumption. Unlike Case 3.1, however, it cannot be determined *a priori* which of Y_n and $U(E(a_n))$ will be the greater. Rather one must defer to the empirical outcome, which will be dictated by the proximity of the arrival times a_i and a_j to the PAT. With reference to Figure 3, for example, the utility function U lies above the expected utility function Y for any arrival time in the interval $\lfloor PAT, a_j \rfloor$; i.e. risk aversion. Contrast this with Figure 5, wherein the expected utility function intersects the vertical segment of the utility function, and the utility function therefore lies below the expected utility function throughout the interval $\lfloor PAT, a_j \rfloor$; i.e. risk preference.

Figure 5: Utility and expected utility functions, for given departure time, with $\alpha < \beta$ *(ABOUT HERE)*

Before moving on, it should be acknowledged that Polak (1987) pursues similar interests to the above, albeit with different focus. Polak devotes careful attention to the form of his utility function, postulating that the function should be monotonically decreasing in travel cost and demonstrate a constant degree of risk aversion to travel cost (see Pratt (1964) and Arrow (1970) for discussion of measures of risk aversion); contrast this with the present paper, which is founded on attitudes to time risk rather than cost risk. Adopting an exponential form as a working representation of the postulated properties, Polak seeks to identify, in a similar manner to section 3.4 above, the preferred departure time of an individual traveller. Then developing ideas further, he introduces the notion of the 'safety margin', as proposed by Gaver (1968) and Knight (1974), and establishes a relationship between the magnitude of the safety margin and the degree of risk aversion.

4.3 The reliability premium

Returning to our own analysis, let us now develop Pratt's (1964) concept of the risk premium in the context of trip scheduling. As a precursor, let us first introduce the concept of a 'certainty equivalent', which in the present context may be defined as follows. The certainty equivalent relating to any prospect $\mathbf{w}_n \in S$ is the arrival time $\tilde{a}_n$ that yields the same utility with certainty as the expected utility of the prospect. In Cases 1 and 2 of the previous section, both of which imply risk neutrality, utility and expected utility are one and the same, and the notion of a certainty equivalent is therefore somewhat tautological.

Of rather more interest is Case 3, since it carries a potential for risk aversion. Subsequent discussion will reveal an intuition that, in developing the certainty equivalent for Case 3, the certainty equivalent should be constrained to fall within the interval of arrival times defined by the prospect, i.e. $a_i \leq \tilde{a}_n \leq a_j$. Let us for the moment simply accept this requirement and, by analogy to Cases 3.1 and 3.2, distinguish between cases where the certainty equivalent falls before and after the PAT, as follows:

Case 3.3: $a_i \leq \tilde{a}_n < PAT$

In developing this case, let us establish - by way of assertion - equivalence between expected utility and the utility deriving from the certainty equivalent, i.e.

Let $Y_n = \alpha(\tilde{a}_n - d_n) + \beta(PAT - \tilde{a}_n)$, where $a_i \leq \tilde{a}_n < PAT$

Rearranging, we can then identify the certainty equivalent:

$$\tilde{a}_n = \frac{Y_n + \alpha d_n - \beta PAT}{(\alpha - \beta)} \qquad (12)$$

Case 3.4: $PAT < \tilde{a}_n \leq a_j$

Now repeating the same exercise, but for a certainty equivalent later than the PAT:

Let $Y_n = \alpha(\tilde{a}_n - d_n) + \gamma(\tilde{a}_n - PAT) + \delta L(\tilde{a}_n)$

where $L(\tilde{a}_n) = 1$ if $(\tilde{a}_n - PAT) > 0$, $= 0$ otherwise, and $PAT < \tilde{a}_n \leq a_j$

Rearranging:

$$\tilde{a}_n = \frac{Y_n + \alpha d_n + \gamma PAT - \delta L(\tilde{a}_n)}{(\alpha + \gamma)}$$

Note that, with reference to earlier discussion in section 3, the denominator of the certainty equivalent for Cases 3.3 and 3.4 is given by the slope of the relevant section of the utility function. Whilst the above working carries an apparent clarity, it might however be cautioned that the properties of the function - more particularly, its discontinuity - carry the implication that an exact certainty equivalent is not empirically guaranteed. Nevertheless, let us proceed to the definition of the risk premium.

In general terms, the risk premium measures the individual's willingness-to-pay, in units of the pay-off, to avoid the risk of choosing an uncertain prospect. Or more succinctly, the risk premium measures the 'cost of risk bearing'. Applying these definitions to our interest in the reliability of arrival time, the risk premium - or in present parlance, the 'reliability premium' - measures, for a given departure time, the delay in arrival time (with its consequent impacts on travel time, schedule delay early, schedule delay late and the lateness dummy) that the individual would be willing-to-pay in exchange for eliminating the unreliability. The reliability premium thus measures the costs borne by the traveller that arise specifically from unreliability in arrival time.

Now adopting greater formality, define the reliability premium:

$$K_n = \max[(\tilde{a}_n - E(a_n)),0] \qquad (13)$$

Since the pay-off is in this instance defined on a bad (i.e. arrival time), the expected pay-off must be subtracted from the certainty equivalent in order to yield the reliability premium K_n. Note the requirement on the sign of K_n, which implies that a non-zero reliability premium is representative of risk averse behaviour. The reliability premium is illustrated diagrammatically in Figures 6 and 7, again with $\alpha < \beta$. Figure 6 considers an expected late arrival, and demonstrates that a traveller would yield equal utility from the prospect Y_n and the certain arrival time $\tilde{a}_n$, thereby identifying the reliability premium to be the distance $\tilde{a}_n - E(a_n)$. Figure 7 applies analogously to the case of an expected early arrival, with similar result; the traveller would be indifferent between the prospect and a certain arrival time just early of the PAT. Whilst Figures 6 and 7 would seem reasonably clear, the discontinuity of U introduces some complexities; if in particular Y intersects the vertical segment of U in the manner of Figure 5, then the certainty equivalent may fall before the expected arrival time, such that the risk premium becomes zero in accordance with (13).

Figure 6: The reliability premium of an expected late arrival, for given departure time, with $\alpha < \beta$ (ABOUT HERE)

Figure 7: The reliability premium of an expected early arrival, for given departure time, with $\alpha < \beta$ (ABOUT HERE)

If we now extend discussion of the reliability premium to consider the relation $\beta < \alpha$, then it is apparent that further complexities arise. With reference to Figure 8, which omits some labelling in the hope of promoting clarity, the opposing slopes of the two segments of U yield the pair of certainty equivalents $\tilde{a}_n$ and $\tilde{a}_n'$. It might be noted that, whereas $\tilde{a}_n$ falls within the interval $\lfloor a_i, a_j \rfloor$, $\tilde{a}_n'$ does not. Furthermore, whilst $\tilde{a}_n$ falls after the expected arrival time and therefore yields the possibility of a non-zero reliability premium (acknowledging again the possibility of a zero reliability premium should Y intersect the vertical segment of U), $\tilde{a}_n'$ falls before the expected arrival time and will never therefore yield a non-zero reliability premium. Hence, we arrive at an intuition for the constraint $a_i \leq \tilde{a}_n \leq a_j$ introduced at the outset of this section.

Figure 8: The reliability premium of an expected early arrival, for given departure time, with $\beta < \alpha$

Before proceeding, it is useful to distinguish the notion of the reliability premium from the aforementioned notion of the safety margin, in that each derives from a different reference point. Whereas the reliability premium derives from the reference point of the certainty equivalent, the safety margin derives from the reference point of the PAT. This provokes the important distinction that the safety margin pertains only to early arrivals, whereas the reliability premium may pertain to either early (i.e. Figure 7) or late (i.e. Figure 6) arrivals. Indeed, the necessarily theoretical discussion above might be embellished with the remark that the reliability premium offers rationale for a practice that is widely applied in the public transport industry, thus. An operator, if faced with the situation presented in Figure 6, could introduce an increased journey time (i.e. $\tilde{a}_n - d_n$, such that the individual incurs a late arrival) and still maintain market share (i.e. maintain the individual's level of utility at Y_n), provided it could ensure full reliability of service (i.e. move from the expected utility function Y to the utility function U).

4.4 Reconciling the reliability premium with Bates et al.

Having introduced two alternative notions of the value of reliability - the reliability premium (section 4.3) and Bates et *al.*'s (2001) marginal valuations of travel time and scheduling under uncertainty (section 4.1) - let us now seek to reconcile them. To this end, and exploiting the analysis of the previous section, we can re-express expected utility as the utility at the certainty equivalent:

$$Y_n = U(\tilde{a}_n) = \alpha'(\tilde{a}_n - d_n) + \beta' \max[(PAT - \tilde{a}_n),0] + \gamma' \max[(\tilde{a}_n - PAT),0] + \delta' L(\tilde{a}_n)$$

If the individual traveller is risk averse then we can substitute for $\tilde{a}_n$ using the risk premium K_n, thus:

$$\begin{aligned} Y_n = \alpha'[E(a_n) + K_n - d_n] + \beta' \max[(PAT - E(a_n) - K_n),0] \\ + \gamma' \max[(E(a_n) + K_n - PAT),0] + \delta' L(E(a_n) + K_n) \end{aligned} \tag{14}$$

Should we now introduce travel cost to (14) in the manner of (8), then we can derive a further set of marginal valuations, but this time at the certainty equivalent. For example, let us derive marginal valuation of travel time at the certainty equivalent:

If $\hat{Y}_q - \hat{Y}_r \Big|_{\beta,\gamma,\delta=0} = 0$ then:

$$\phi(c_r - c_q) = \alpha'\{(K_q - K_r) + \lfloor E(a_q) - E(a_r) \rfloor + (d_r - d_q)\}$$

$$Vo(T) = \frac{\alpha'}{\phi} = \frac{(c_r - c_q)}{(K_q - K_r) + \lfloor E(a_q) - E(a_r) \rfloor + (d_r - d_q)} \quad (15)$$

Marginal valuations - at the certainty equivalent - of schedule delay early, schedule delay late and the lateness dummy are more complicated. Whilst it is reasonable to restrict attention to Case 3 (i.e. $a_i < PAT < a_j$), remembering that the reliability premia K_q and K_r are pertinent only to conditions of risk aversion, this does not constrain the certainty equivalents $\tilde{a}_q$ and $\tilde{a}_r$ to fall either side of the PAT. Rather it is necessary to define three further possibilities, as extensions of Cases 3.3 and 3.4, thus:

Case 3.5: $\tilde{a}_q, \tilde{a}_r \leq PAT$

Case 3.6: $PAT \leq \tilde{a}_q, \tilde{a}_r$

Case 3.7: $\tilde{a}_q < PAT < \tilde{a}_r$

Table I displays the complete set of marginal valuations at the certainty equivalent, with the first three columns discriminating by the above three cases. Of particular interest is whether these marginal valuations at the certainty equivalent accord with those previously derived under uncertainty. For purposes of this comparison, it is useful to consider how (10) and (A1) to (A3) apply to Case 3, and these are given in the final column of Table I.

Table I: Marginal valuations of travel time and scheduling under uncertainty and at the certainty equivalent (ABOUT HERE)

With reference to Table I, comparison of the marginal valuations of travel time under uncertainty and at the certainty equivalent, which are given by α/ϕ (i.e. from equation (15)) and α'/ϕ (i.e. from equation (10)) respectively, reveals that the denominators of the two valuations differ by the quantity $(K_q - K_r)$. These valuations will therefore be equal only if $K_q = K_r$. Whilst the latter equality may feasibly hold, there is no *a priori* basis for necessarily expecting this result, and it would seem quite possible that marginal valuation of travel time under uncertainty will differ from that at the certainty equivalent.

Now consider marginal valuation of the lateness dummy. In Cases 3.5 and 3.6, the marginal valuations under uncertainty and at the certainty equivalent (i.e. δ/ϕ and δ'/ϕ respectively) will be equal if $c_q = c_r$. In Case 3.7, similar equality will arise if both $p_{iq} = 1$ and $p_{ir} = 0$. Note in passing that, as section 3.4 has already considered, one would in practice expect the inequality $p_{iq} > p_{ir}$ to hold, which would be consistent with equality between marginal valuations of the lateness dummy under uncertainty and at the certainty equivalent.

Finally, the marginal valuations of schedule delay early $(\beta/\phi, \beta'/\phi)$ and schedule delay late $(\gamma/\phi, \gamma'/\phi)$ show similar differences in their formulae when comparing uncertainty with the certainty equivalent. It would however seem more difficult to predict the outcome *a priori*, and one must instead defer to empirical investigation.

That notwithstanding, the preceding theoretical argument has illuminated the possibility that marginal valuations of travel time and scheduling at the certainty equivalent will show discrepancy from the same valuations under uncertainty. Whether theoretical discrepancy results in empirical discrepancy remains to be seen, and this provokes a call for future work addressing such matters. Should the discrepancy be confirmed empirically, then it carries important interpretation as the individual traveller's willingness-to-pay (now in monetary units) to eliminate unreliability in arrival time. For example, the quantity $(\alpha'/\phi - \alpha/\phi)$ represents the individual's willingness-to-pay to eliminate unreliability in arrival time, specifically as it impacts upon travel time; it might therefore be referred to as the 'marginal valuation of reliability in travel time'. Similar interpretations apply to the other constituents of

Small's (1982) utility function; the quantity $(\beta'/\phi - \beta/\phi)$ might be referred to as the 'marginal valuation of reliability in schedule delay early', and so on.

5. The prevalence and distribution of benefits from the reliability premium

Whilst marginal valuations of travel time and scheduling under uncertainty are in themselves perfectly adequate for purposes of forecasting individual choice under uncertainty, it is only through their comparison to marginal valuations of travel time and scheduling at the certainty equivalent that the effect of unreliability in arrival time on choice can be revealed. Such insight might usefully inform public transport operators as to the effects of changes in reliability on the individual traveller's choice of departure time. More generally, the minimisation of risk may be an important policy aspiration, and marginal valuations of reliability could, to this end, offer a useful means of discriminating between alternative investment options.

Though intrinsic to marginal valuations of reliability, arguably the more significant contribution of the reliability premium is to economic appraisal. As Pearce and Nash (1981) observe, the projected benefits of any transport scheme - which may include benefits from improved reliability - should mitigate for the cost of inherent risk, and failure to do so introduces the possibility of bias. Hence it would seem crucial to make explicit statement of the cost of unreliability in arrival time to the individual traveller (or equivalently, the benefit of eliminating that unreliability), which for any prospect $\mathbf{w}_n \in S$ is given by:

$$\begin{aligned} Vo(K_n) = &\alpha'/\phi * [\tilde{a}_n - E(a_n)] + \\ &\beta'/\phi * \{\max[(PAT - \tilde{a}_n),0] - \max[(PAT - E(a_n)),0]\} + \\ &\gamma'/\phi * \{\max[(\tilde{a}_n - PAT),0] - \max[(E(a_n) - PAT),0]\} + \\ &\delta'/\phi * [L(\tilde{a}_n) - L(E(a_n))] \end{aligned} \qquad (16)$$

Noting importantly that the potential benefit of eliminating unreliability manifests only under circumstances of risk aversion, let us consider the prevalence of such benefit in relation to the same three cases considered earlier.

Case 1: $a_i < a_j \le PAT$

In this case, there would be nil benefit (i.e. $Vo(K_n) = 0$ in (16)) to the traveller from the elimination of unreliability; this is because the traveller exhibits risk neutrality.

Case 2: $PAT \le a_i < a_j$

This would yield the same outcome as Case 1.

Case 3: $a_i < PAT < a_j$

The traveller might or might not realise benefit from the elimination of unreliability, as follows:

Case 3.1: $E(a_n) < PAT$

There would be unambiguous benefit (i.e. $Vo(K_n) > 0$), since the traveller exhibits risk aversion.

Case 3.2 $PAT < E(a_n)$

The prevalence of benefit would depend on the proximity of the arrival times a_i and a_j to the PAT, and their consequent effect on attitudes to unreliability.

Since the above outcomes are dictated by the relation of the arrival times a_i and a_j to the PAT, one could therefore expect the prevalence of benefit to vary by the departure times $d_n \in D$. That is to say, one might reasonably expect relatively early departures to pertain to Case 1, relatively late departures to Case 2, and intermediate departures to Case 3.

It should be emphasised, however, that the above conclusions apply specifically to the case of an individual traveller. If, as is more common in practice, one is interested in a sample of travellers, say the users of a particular public transport departure, then the sample would typically demonstrate some heterogeneity in respect of the PAT. The implication follows that, for any improvement in the reliability of arrival time for that service, some users might realise a benefit from that improvement, whilst others might not. The total benefit to the sample would therefore depend on the number of beneficiaries, and their individual valuations of the reliability premium. With reference to (2), interest in a sample of individuals is often developed through a re-interpretation of RUM, with 'repetitions' becoming 'individuals', and probabilities of choice deriving from differences in the preferences of individuals across the sample. Irrespective of whether it derives from intra-individual or inter-individual variability in preferences, extension of the reliability premium to accommodate RUM would give rise to distributed marginal valuations of travel time and scheduling, as well as to distributed marginal valuations of reliability in travel time and scheduling.

Aside from the matter of aggregation, a further restriction of the above analysis is to a binary subset of arrival times, and it is appropriate to acknowledge that the clarity of Case 3 becomes compromised once the set of arrival times is extended beyond the binary. In such circumstances, the relation between the utility of the expected arrival time and expected utility remains pertinent but must be revealed empirically. That said, and with reference to the aspirations outlined in the introduction to this paper, it might be argued that a binary subset of arrival times would bring appealing convenience to SP analysis of the reliability premium. In short, the binary case is far from abstract. In contrast to Case 3, the unambiguous results of Cases 1 and 2 readily extend to trinomial or larger arrival time sets.

6. Worked example of the reliability premium

Let us now illustrate the theoretical exposition of the reliability premium by means of a worked example, noting that we will restrict attention to a single individual and therefore avoid the complications of aggregation just mentioned. With reference to Table II, which quantifies all times in minutes after midnight, consider a one-way commute with a departure time profile of 420 (i.e. 7:00am) to 495 (i.e. 8:15am), in increments of 5 minutes. Arrival times are similarly defined in increments of 5 minutes, and reveal a minimum journey time of 30 minutes (i.e. $a_{\min} = 450$). This could be representative of a high-frequency scheduled public transport service; alternatively, it could be a discrete approximation to a car-based journey. The body of the table displays the event probabilities by departure and arrival times. It might be observed that the subset of arrival times varies by departure time, and contains between two and five possible arrivals. Since the problem is more general than the binary subset of arrival times considered above, analysis of Case 3 must therefore defer to the empirical results that follow.

Table II: Pay-off matrix for worked example (ABOUT HERE)

In focussing on the reliability premium, it is unnecessary to explicitly consider travel cost, and we therefore proceed with the formulations of utility and expected utility given by (4) and (6) respectively. Let us populate these with the estimates of α, β, γ and δ from Model (1) of Small (1982), specifically $\alpha = -0.106$, $\beta = -0.065$, $\gamma = -0.254$ and $\delta = -0.58$, noting that $\alpha < \beta$. Let us assume also that $PAT = 525$ (i.e. 8:45am). Figure 9 plots the various attributes of expected utility - expected travel time, expected schedule delay early, expected schedule delay late, and the expected lateness dummy - against expected arrival time. The properties of this figure accord with those of previous presentations in the literature, for example Figure 2 of Bates *et al.* (2001), and confirm the position of the PAT at 525.

Figure 9: Expected travel time, SDE, SDL and lateness dummy vs. expected arrival time (ABOUT HERE)

Figure 10 plots, for each departure time, expected utility Y and utility of the expected arrival time $U(E(a))$. It might be remarked that many of the departure times are inferior, in that their maximum utility falls short of the minimum utility of other departures. On this basis, we can restrict attention to departure times in the range 450 to 475. Indeed expected utility is maximised within this range, specifically at 465. Comparing the plots of Y and $U(E(a))$, it can be observed that expected utility and utility of the expected arrival time coincide for the vast majority of departure times; these departures pertain to Cases 1 and 2 (i.e. their respective arrival times fall either always before or always after the PAT). By contrast, the two plots diverge for the 465, 470 and 475 departures, each of which pertains to Case 3 (i.e. their respective arrival times straddle the PAT). More specifically, expected utility is less than the utility of the expected arrival time for the 465 and 470 departures, whereas the reverse applies for the 475 departure. Hence with reference to the discussion of section 4.2, it may be seen that the preferred departure time of 465 is relatively risky in comparison to other available departures.

Figure 10: Expected utility and utility of expected arrival time, by departure time (ABOUT HERE)

Finally, let us consider an example of the reliability premium, taking the particular case of the 465 departure (since this departure is characterised by risk aversion). The empirical utility function for this departure is shown in Figure 11; this follows the characteristic shape of the theoretical utility functions in Figures 1 and 3 to 7, with distinct sections before and after the PAT, and $\alpha < \beta$. The empirical expected utility function, by contrast, cannot be shown in the manner of the theoretical examples, since we have expanded the set of arrival times beyond the binary. Suffice to say, the arrival time window extends from 510 to 530, hence the points labelled Y.

Figure 11: Utility and expected utility functions for $d = 465$ *(ABOUT HERE)*

For the 465 departure, we can calculate the expected arrival time (11), giving $E(a_{465}) = 515.25$, and the certainty equivalent (12), giving $\tilde{a}_{465} = 517.90$. Then applying these to the reliability premium (13), we can calculate $K_{465} = \tilde{a}_{465} - E(a_{465}) = 2.65$, such that a certain arrival time 2.65 minutes later than the expected arrival time would yield the same utility as the expected utility of the prospect. Finally, the value of this reliability premium is, with reference to (16), given by:

$$Vo(K_{465}) = \left[\frac{\alpha' - \beta'}{\phi}\right] \times 2.65$$

Contrast this with the 460 departure, where $Y_{460} = U(E(a_{460}))$ and $Vo(K_{460}) = 0$. Hence elimination of unreliability in arrival time would for the 465 departure (i.e. Case 3) yield an additional benefit to the individual, but there is no possibility of similar benefit for the 460 departure (i.e. Case 1).

7. Summary and conclusion

Unreliability is endemic in many transport systems, and this stimulates interest in whether and how unreliability impacts upon the choices of travellers. The paper pursued specific interest in the effect of unreliability in arrival time on scheduling choice. Following the precedent of Noland & Small (1995) and Bates *et al.* (2001), this was developed through the marriage of Small's (1982) utility function with von Neumann & Morgenstern's (1947) theory of individual choice under uncertainty. Arising from the latter union is the proposition that unreliability imposes disutility on the traveller. Hence the potential for benefit should unreliability be reduced or indeed eliminated.

In contrast to the precedent works on reliability, the present paper adopted a discrete representation of time, motivated in particular by a desire to promote implementation within the apparatus of RUM and SP. The substantive contribution of the paper, however, was the scope of the theoretical exposition, which offered significant extensions beyond the extant reliability literature. With reference to the theoretical literature on attitudes to risk (e.g. Pratt, 1964; Arrow, 1970), the paper considered the implications of Small's utility function for travellers' attitudes to unreliability in arrival time, and in particular identified circumstances under which travellers would be risk averse. In response to the latter observation, and drawing analogy with Pratt's (1964) concept of the risk premium, the paper introduced the notion of the 'reliability premium'. This measures, for a given departure time, the delay in arrival time that the individual would be willing-to-pay in exchange for eliminating unreliability in arrival time.

The paper then sought to reconcile this consideration of attitudes to risk with Bates *et al.*'s (2001) marginal valuations of travel time and scheduling under uncertainty. The latter arise from the proposition that an individual traveller would, in choosing between prospects, be willing to exchange travel cost for expected travel time, expected schedule delay early, expected schedule delay late, and the probability of late arrival. Exploiting the reliability premium, the paper established a basis for comparison between marginal valuations of travel time and scheduling under uncertainty, and analogous valuations derived at the certainty equivalent. This comparison revealed the theoretical possibility that the two sets of valuations might show discrepancy. Should this theoretical discrepancy manifest in empirical discrepancy - which remains to be seen - then it carries important interpretation as the marginal valuation of reliability in arrival time.

Whilst marginal valuations of travel time and scheduling under uncertainty are adequate for demand forecasting, economic appraisal should mitigate the projected benefits of a scheme against the costs of risk bearing, and this is where the reliability premium becomes pertinent. It is crucial to acknowledge that reliability benefits arise only under the particular circumstances of risk aversion, and that risk aversion is, in the terms of the utility function, dictated by the relation of the possible arrival times to the preferred arrival time. Moreover, the prevalence of benefit will likely vary by departure time for given preferred arrival time. This assumes a single individual however; once the analysis is extended to a sample of individuals, the outcome will be complicated by heterogeneity in the preferred arrival time and, it follows, heterogeneity in the prevalence and magnitude of reliability benefits. Thus for any particular departure time, an improvement in the reliability of arrival time might yield benefit for some travellers but no benefit for others.

Acknowledgements

The author was supported by a Departmental Research Fellowship at the Institute for Transport Studies (ITS), University of Leeds, and by the Platform Grant GR/S46291/01 'Towards a unified theoretical framework for transport network and choice modelling', which is funded by the UK Engineering and Physical Sciences Research Council and also held at ITS. The author gratefully acknowledges the comments of Professor David Watling, Professor Gerard de Jong and Mr James Laird on earlier drafts, as well as the comments of two anonymous referees. Any remaining errors are the sole responsibility of the author.

Appendix A: Derivation of $Vo(E(SDE))$, $Vo(E(SDL))$ and $Vo(E(L))$

If $\hat{Y}_q - \hat{Y}_r \Big|_{\alpha,\gamma,\delta=0} = 0$ then: (A1)

$$Vo(E(SDE)) = \frac{\beta}{\phi} = \frac{(c_r - c_q)}{(p_{ir} - p_{iq})\Delta SDE}$$

If $\hat{Y}_q - \hat{Y}_r \Big|_{\alpha,\beta,\delta=0} = 0$ then: (A2)

$$Vo(E(SDL)) = \frac{\gamma}{\phi} = \frac{(c_r - c_q)}{(p_{ir} - p_{iq})\Delta SDL}$$

If $\hat{Y}_q - \hat{Y}_r \Big|_{\alpha,\beta,\gamma=0} = 0$ then: (A3)

$$Vo(E(L)) = \frac{\delta}{\phi} = \frac{(c_r - c_q)}{(p_{ir} - p_{iq})\Delta L}$$

References

Arrow, K.J. (1970) Essays in the Theory of Risk-Bearing. North Holland, Amsterdam.

Bates, J., Polak, J., Jones, P. and Cook, A. (2001) The valuation of reliability for personal travel. Transportation Research Part E, 37 (2-3), 191-229.

Batley, R. (2006) On ordinal utility, cardinal utility, and random utility. Working paper, Institute for Transport Studies, University of Leeds.

Baumol, W. (1958) The cardinal utility which is ordinal. The Economic Journal, 68 (272), 665-672.

Becker, G.S. (1965) A theory of the allocation of time. The Economic Journal, 75 (299), 493-517.

De Jong, G., Kroes, E., Plasmeijer, R., Sanders, P. and Warffemius, P. (2004) The value of reliability. Paper presented at the European Transport Conference, Strasbourg, 4-6 October 2004.

De Serpa, A.C. (1971) A theory of the economics of time. The Economic Journal, 81 (324), 828-846.

Edwards, W. (1955) The prediction of decisions among bets. Journal of Experimental Psychology, 50 (3), 201-214.

Fishburn, P.C. (1970) Utility Theory for Decision-Making. John Wiley & Sons, New York.

Gaver, D.P. Jr. (1968) Headstart strategies for combating congestion. Transportation Science, 2 (2), 172-181.

Herstein, I.N. and Milnor, J. (1953) An axiomatic approach to measurable utility. Econometrica, 21 (2), 291-297.

Johansson, P.O. (1991) An Introduction to Modern Welfare Economics. Cambridge University Press, Cambridge.

Jotisankasa, A., Hess, S. and Polak, J.W. (2004) On the effects of alternative discretisation strategies in departure time choice modelling. Paper presented at the 36th Annual Conference of the Universities' Transport Study Group, University of Newcastle-upon-Tyne, 5-7 January 2004.

Kahneman, D. and Tversky, A. (1979) Prospect theory: an analysis of decision under risk. Econometrica, 47 (2), 263-292.

Kahneman, D. and Tversky, A. (2000) Choices, Values and Frames. Cambridge University Press, Cambridge.

Keynes, J.M. (1921) A Treatise on Probability. Macmillan, London.

Keynes, J.M. (1936) The General Theory of Employment, Interest and Money. Macmillan, London.

Knight, F.H. (1921) Risk, Uncertainty and Profit. Houghton Mifflin, Boston.

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

Marschak, J. (1950) Rational behaviour, uncertain prospects and measurable utility. Econometrica, 18 (2), pp111-141.

Marschak, J., Becker, G.M. and DeGroot, M.H. (1963) Stochastic models of choice behavior. In Marschak, J. (1974) Economic Information, Decision and Prediction: Selected Essays (Volume 1). D. Reidel, Dordrecht.

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

Noland, R. and Small, K. (1995) Travel time uncertainty, departure time choice and the cost of the morning commute. Working paper UCI-ITS-WP-95-1, Institute of Transportation Studies, University of California, Irvine.

Pearce, D.W. and Nash, C.A. (1981) The Social Appraisal of Projects: A Text in Cost-Benefit Analysis. Macmillan, London.

Polak, J. (1987) Travel time variability and departure time choice. Transport Studies Group Discussion Paper No. 15, Polytechnic of Central London.

Pratt, J.W. (1964) Risk aversion in the small and in the large. Econometrica, 32 (1-2), 122-136.

Small, K. (1982) The scheduling of consumer activities: work trips. American Economic Review, 72 (3), 467-479.

Vickrey, W.S. (1969) Congestion theory and transport investment. American Economic Review, 59 (2), 251-260.

von Neumann, J. and Morgenstern, O. (1947) Theory of Games and Economic Behavior (Second Edition). Princeton University Press, Princeton.

Figure 1: Small's utility function, for given departure time, with $\alpha < \beta$

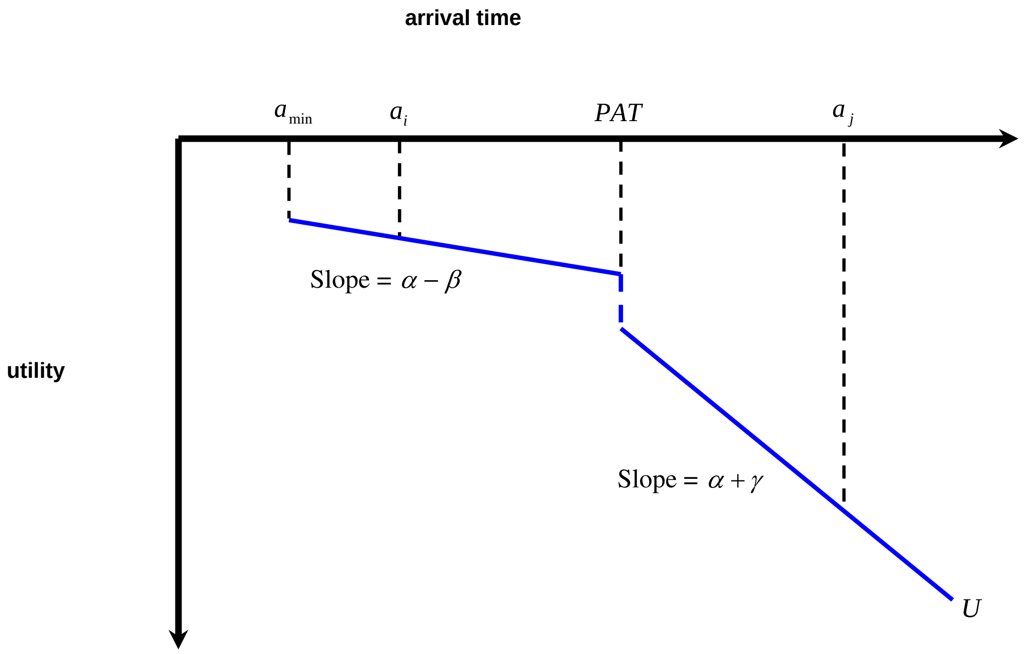

Figure 2: Small's utility function, for given departure time, with $\beta < \alpha$

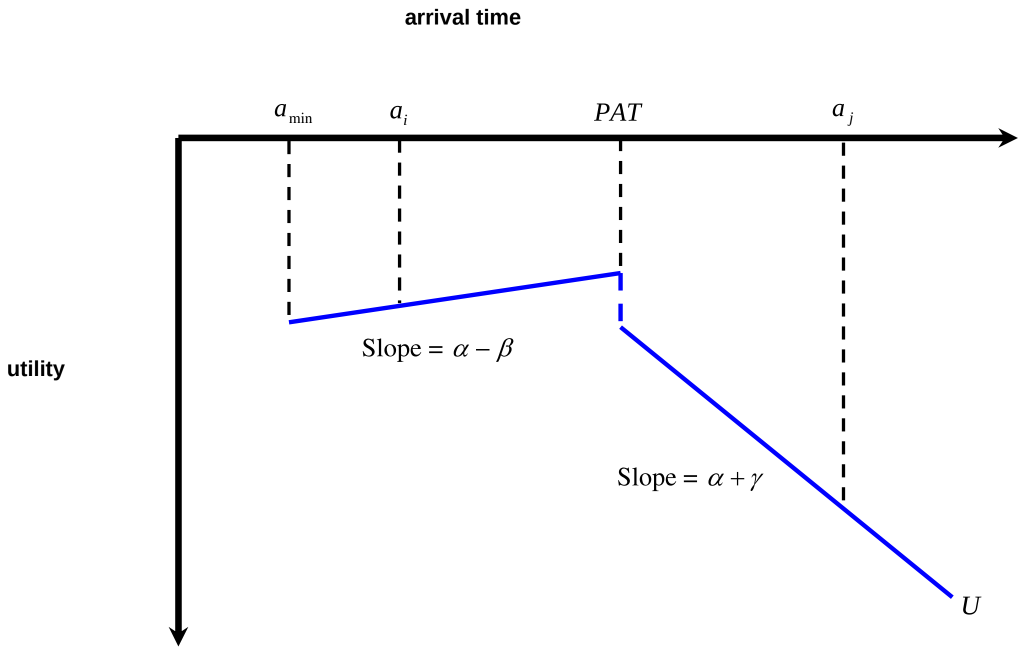

Figure 3: Utility and expected utility functions, for given departure time, with $\alpha < \beta$

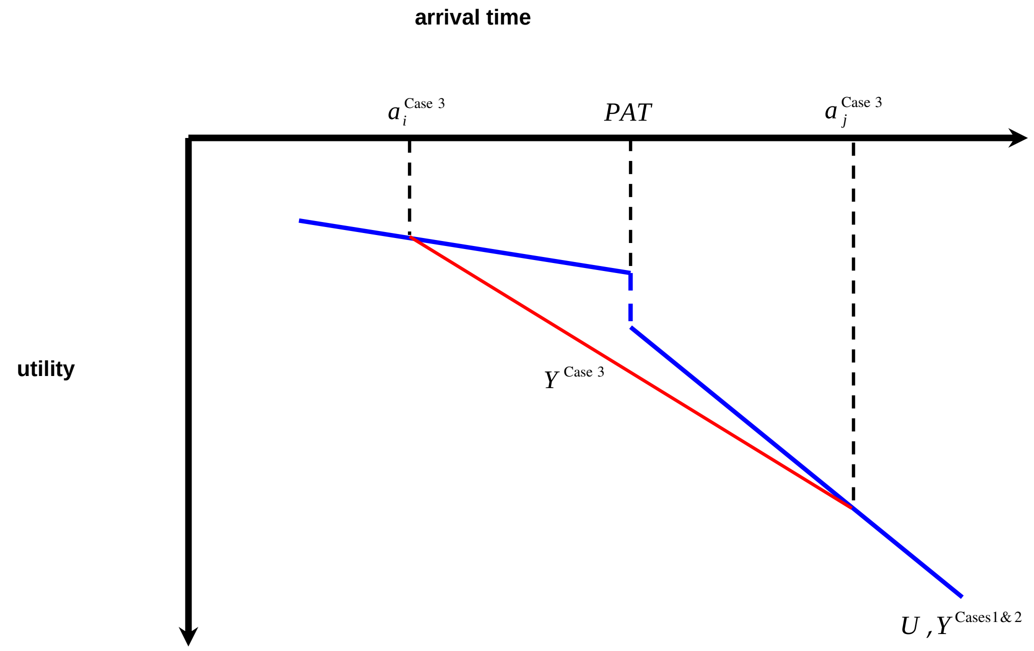

Figure 4: Choice between departure times, with $\alpha < \beta$

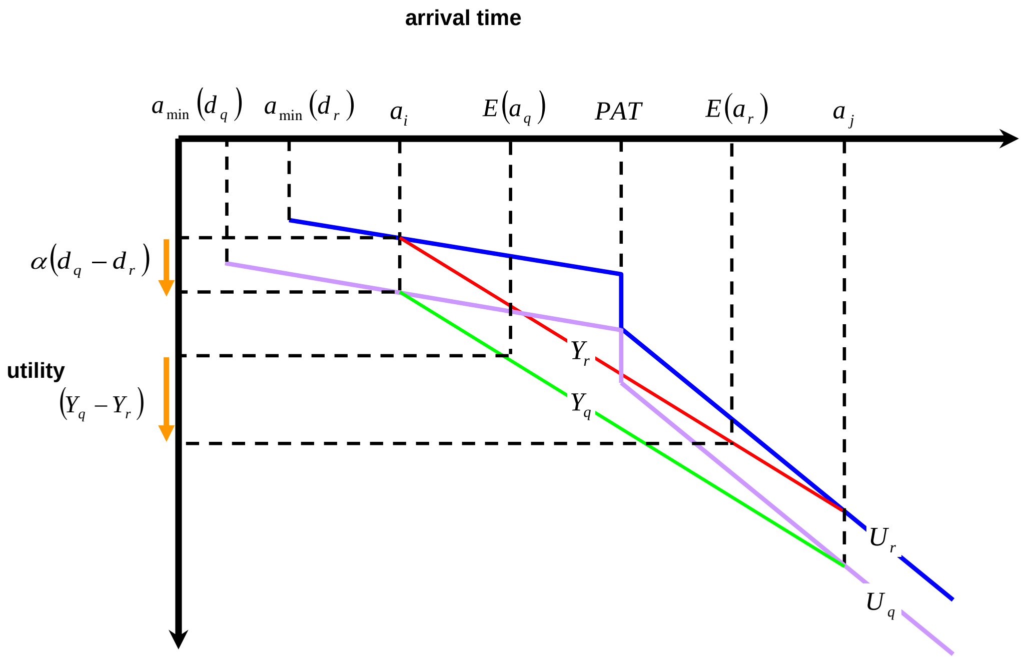

Figure 5: Utility and expected utility functions, for given departure time, with $\alpha < \beta$

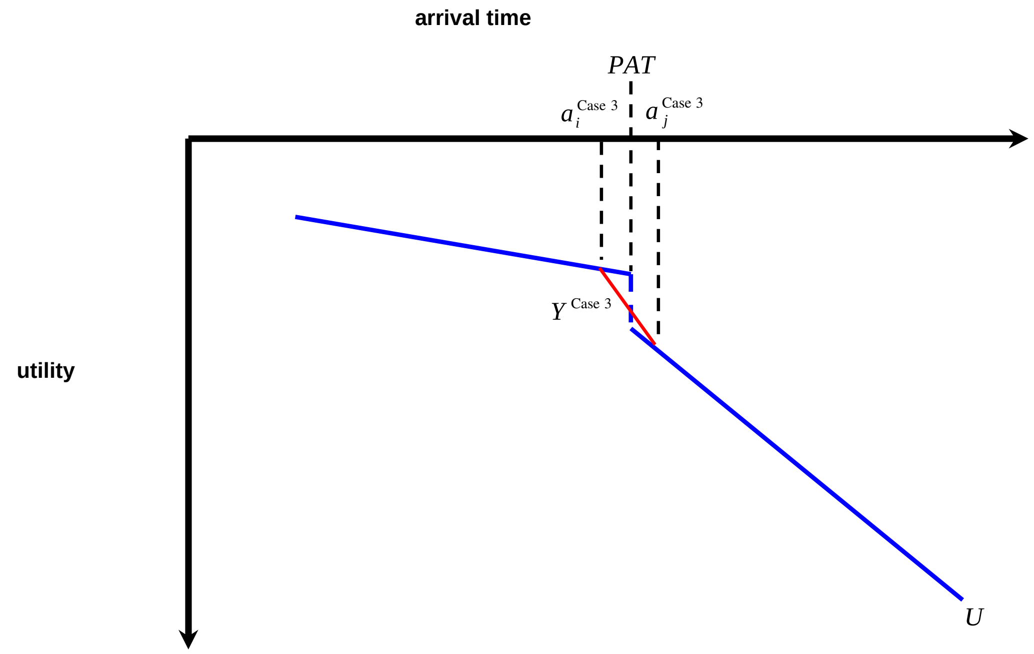

Figure 6: The reliability premium of an expected late arrival, for given departure time, with $\alpha < \beta$

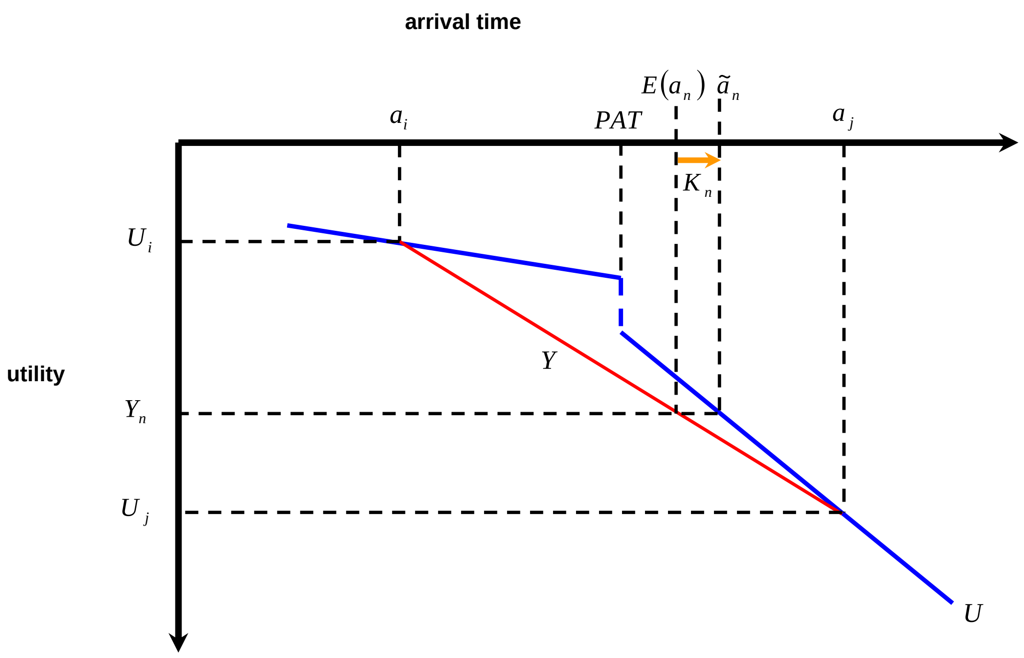

Figure 7: The reliability premium of an expected early arrival, for given departure time, with $\alpha < \beta$

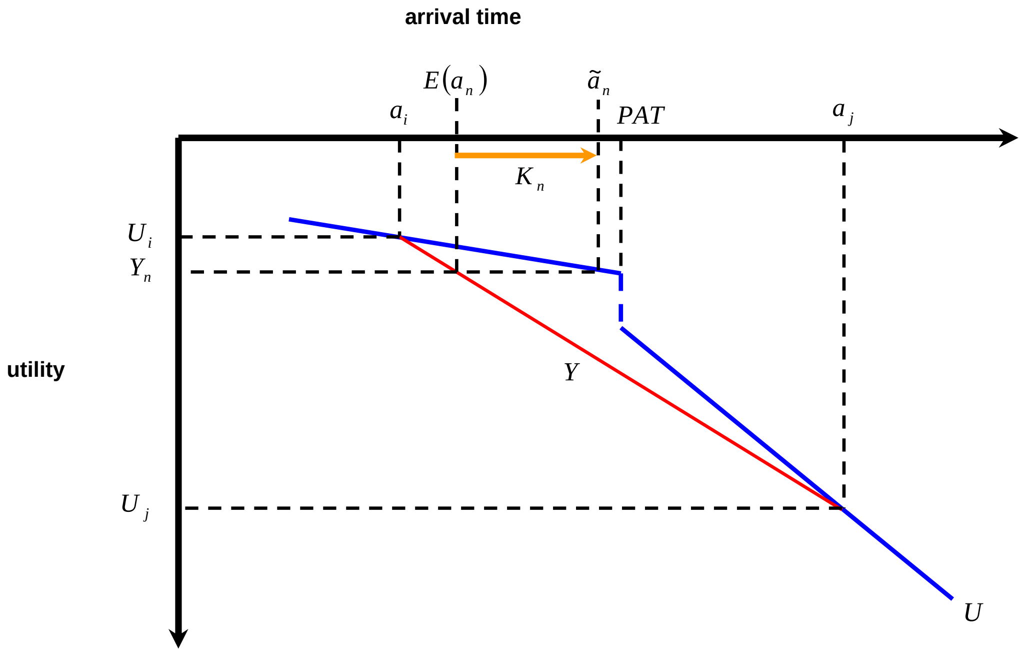

Figure 8: The reliability premium of an expected early arrival, for given departure time, with $\beta < \alpha$

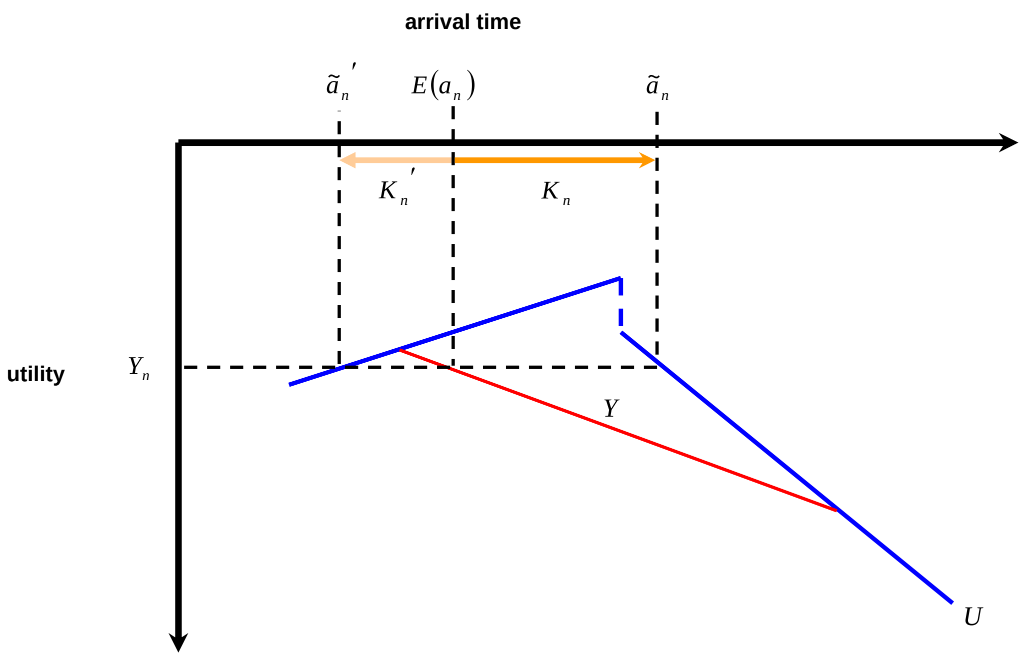

Figure 9: Expected travel time, SDE, SDL and late penalty vs. expected arrival time

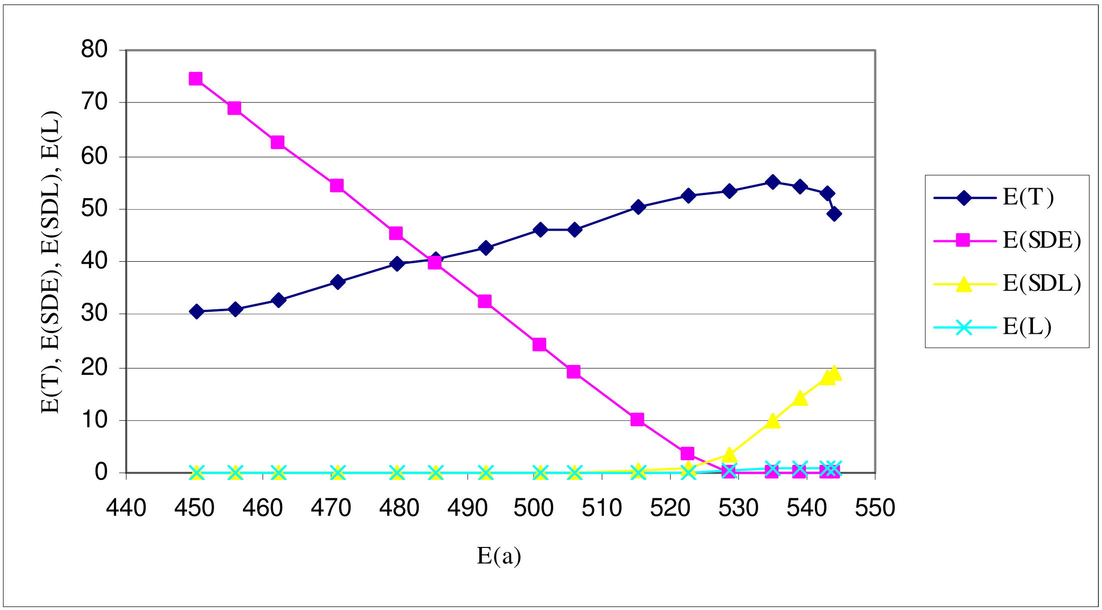

Figure 10: Expected utility and utility of expected arrival time, by departure time

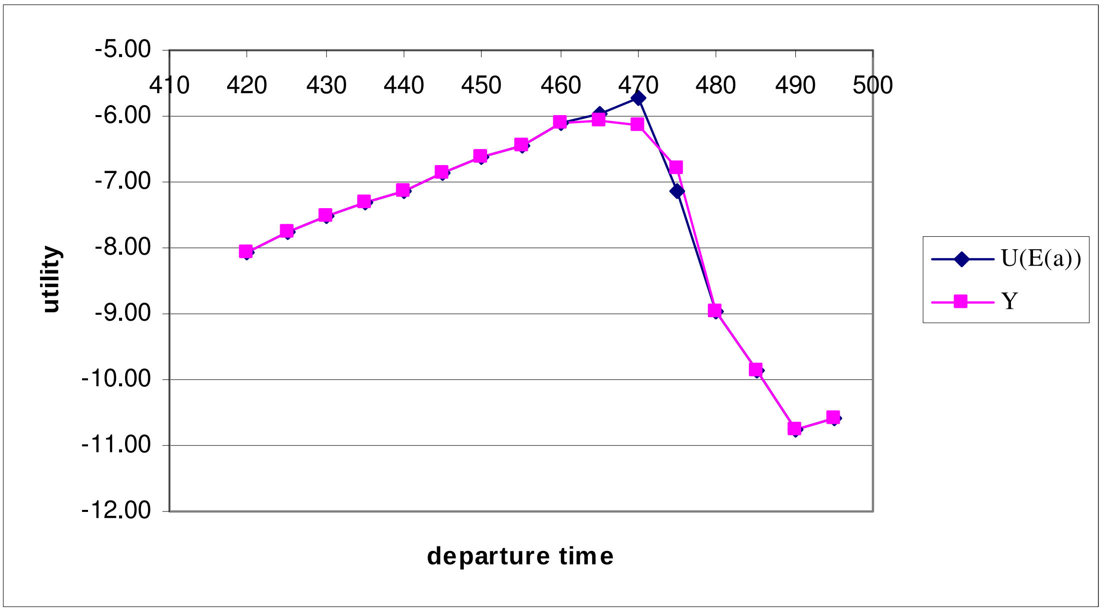

Figure 11: Utility and expected utility functions for $d = 465$

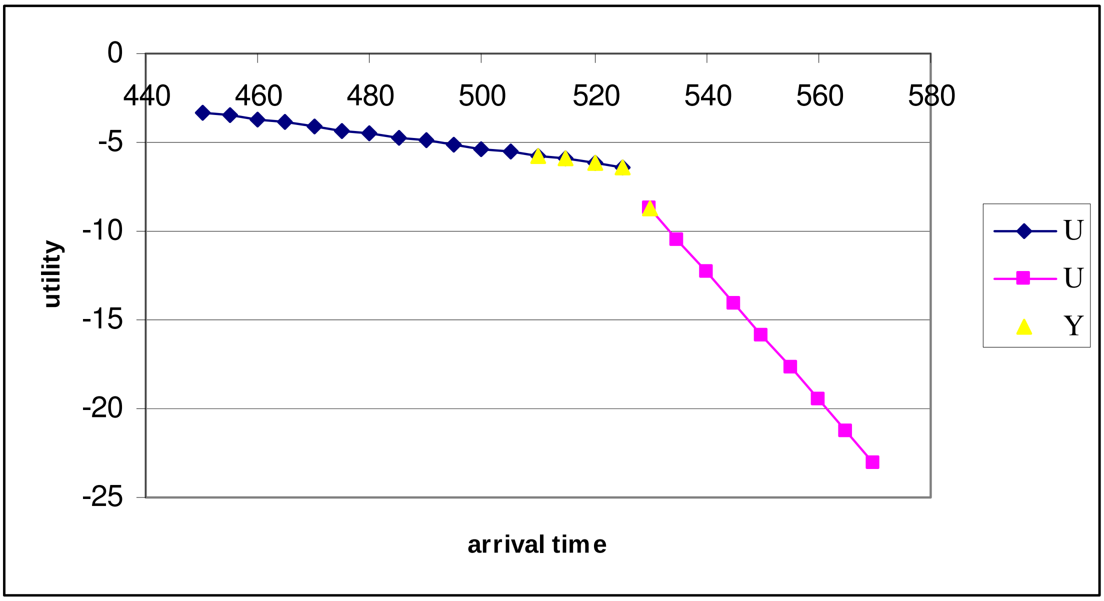

Table I: Marginal valuations of travel time and scheduling under uncertainty and at the certainty equivalent

CERTAINTY EQUIVALENT				UNCERTAINTY	
	Case 3.5: $\tilde{a}_q, \tilde{a}_r < PAT$	Case 3.6: $PAT < \tilde{a}_q, \tilde{a}_r$	Case 3.7: $\tilde{a}_q < PAT < \tilde{a}_r$	Case 3: $a_i < PAT < a_j$	
$\frac{\alpha'}{\phi}$	$\frac{(c_r - c_q)}{(K_q - K_r) + [E(a_q) - E(a_r)] + (d_r - d_q)}$			$\frac{(c_r - c_q)}{[E(a_q) - E(a_r)] + (d_r - d_q)}$	$\frac{\alpha}{\phi}$
$\frac{\beta'}{\phi}$	$\frac{(c_r - c_q)}{(K_r - K_q) + [E(a_r) - E(a_q)]}$	0	$\frac{(c_r - c_q)}{PAT - K_q - E(a_q)}$	$\frac{(c_r - c_q)}{[(p_{iq} - p_{ir})(PAT - a_i)]}$	$\frac{\beta}{\phi}$
$\frac{\gamma'}{\phi}$	0	$\frac{(c_r - c_q)}{(K_q - K_r) + [E(a_q) - E(a_r)]}$	$\frac{(c_q - c_r)}{K_r + E(a_r) - PAT}$	$\frac{(c_r - c_q)}{[(p_{ir} - p_{iq})(a_j - PAT)]}$	$\frac{\gamma}{\phi}$
$\frac{\delta'}{\phi}$	0	0	$(c_q - c_r)$	$\frac{(c_r - c_q)}{(p_{ir} - p_{iq})}$	$\frac{\delta}{\phi}$

Table II: Pay-off matrix for worked example

	a 450.00	a 455.00	a 460.00	a 465.00	a 470.00	a 475.00	a 480.00	a 485.00	a 490.00	a 495.00	a 500.00	a 505.00	a 510.00	a 515.00	a 520.00	a 525.00	a 530.00	a 535.00	a 540.00	a 545.00	a 550.00	a 555.00
d 420.00	0.90	0.10																				
d 425.00		0.85	0.10	0.05																		
d 430.00			0.60	0.30	0.10																	
d 435.00				0.20	0.50	0.20	0.10															
d 440.00						0.30	0.50	0.15	0.05													
d 445.00							0.20	0.60	0.10	0.10												
d 450.00								0.05	0.50	0.30	0.15											
d 455.00										0.20	0.50	0.20	0.10									
d 460.00											0.10	0.70	0.10	0.10								
d 465.00													0.30	0.50	0.10	0.05	0.05					
d 470.00														0.05	0.60	0.20	0.10	0.05				
d 475.00																0.60	0.20	0.10	0.10			
d 480.00																	0.30	0.50	0.10	0.10		
d 485.00																		0.50	0.30	0.10	0.10	
d 490.00																			0.60	0.20	0.20	
d 495.00																			0.50	0.30	0.10	0.10