Scheduling choices under rank dependent utility maximization

Qian Wang – KTH
Marcus Sundberg – KTH
Anders Karlström – KTH

CTS Working Paper 2013:16

Abstract

Rank dependent utility maximization is applied in maximizing a linear and a quadratic scheduling model considering a subjective weighting over uncertain outcomes. The optimal departure time and maximal utility are different from that under expect utility maximization in the transformed travel time density function. Probability weighting is found when estimating the linear model and the estimated weighting function suggests optimism behaviour of respondents. The results also reveal the evidence of heterogeneity in scheduling preferences. Moreover, evidence for the variable of excessive travel time beyond the traditional scheduling model specification is found even with controlling for probability weighting. Our results also show no empirical equivalence between the scheduling model and its derived forms.

Keywords: Scheduling models, Rank dependent utility, Risk attitudes, Heterogeneity

Centre for Transport Studies
SE-100 44 Stockholm
Sweden
www.cts.kth.se

Scheduling choices under rank dependent utility maximization *

Qian Wang
Marcus Sundberg
Anders Karlström†

KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden

*The abstract of this paper has been presented at the 1st European Symposium on Quantitative Methods in Transportation Systems (LATSIS) 2012 Sep. The full paper has been presented on Euro Working Group on Transportation (EWGT) 2013 Sep., and a compact version of this paper has been submitted to EWGT2013 proceeding: Elsevier Procedia - Social and Behavioral Sciences

†Corresponding author: amail@kth.se; qianw@kth.se

1 Introduction

The dispersion in travel time, usually notated as travel time variability (TTV), has been acknowledged as an indispensable part of travel cost. Transport economics research has been interested in the assessment of value of travel time variability (VTTV) and its integration into traffic assignment and Cost Benefit Analysis (Eliasson (2006); Eliasson (2009); Fosgerau et al. (2008)). One way of evaluating VTTV is to assume that individuals value characteristics of the travel time distribution per se. For example, individuals trade-off between mean travel time and travel time variance in mean-variance model. Peer et al. (2012) used the stadard deviation as the alternative measures of the variability and Lam and Small (2001) used the difference between the 90th percentile and the median. Another approach is to use scheduling models, assuming that individuals adjust their departure times with respect to random travel times (Gaver (1968)). Scheduling utility specifications have various forms and there are two frequently applied scheduling models: one has a linear function of travel time (Vickrey (1969); Small (1982)) and the other has a quadratic form (Vickrey (1973)). Noland and Small (1995) extended the linear scheduling model in the context of stochastic travel time, and Bates et al. (2001) deduced a more general conclusion that VTTV can be inferred from the scheduling preferences estimated in a scheduling model. Recently a simplified form of the linear scheduling model is studied by Fosgerau and Karlström (2010); and the quadratic scheduling model is studied by Fosgerau and Engelson (2011). Both provided a theoretical basis that the derived forms of scheduling models can be reconciled with the traditional mean-variance models [i]. The scheduling approach is considered to be more behaviorally consistent in evaluating VTTV.

The scheduling models were derived, however, under the assumption that individuals are maximizing Expected Utility (EU). Such an EU hypothesis (Bernoulli David (1954); von Neumann and Morgenstern (1947)) is convenient but not always "true". Allais (1953) already displayed many cases where the axioms of EU are violated by actual human behavior. In the face of TTV, Individuals could have different risk attitudes (risk aversion, risk taking, risk neutrality) and subjective probability weighting. In EU, risk attitudes depend completely on the curvature of value function which can be transformed into constant absolute risk aversion (CARA) (Pratt (1964) and Arrow (1965)) or constant relative risk aversion (CRRA). Risk attitudes were found in many TTV studies. Senna (1994) proposed a non-linear value function in which the risk attitude parameter is fixed. Polak et al. (2008) applied CARA to the linear scheduling model (Small (1982)) and found mildly risk aversion among their sample. Significant risk taking was found by Li et al. (2012) when applying CRRA to the same scheduling model.

Meanwhile, there are also other alternatives which relax the independence axiom of EU, such as Prospect Theory (PT) (Kahneman and Tversky (1979)) and Rank Dependent Utility

[i]The theoretical equivalence is proved in Fosgerau and Karlström (2010) and Fosgerau and Engelson (2011).The derived form of scheduling model is equivalent to the mean-variance approach

(RDU) (Quiggin (1982); Wakker (2010); Diecidue and Wakker (2001)). RDU has some favorable properties over PT for transforming the cumulative probability of ranked outcomes. Tversky and Kahneman (1992) advanced their PT by incorporating the cumulative probability weighting, and develop Cumulative Prospect Theory (CPT). These promising alternatives accommodate anomalies of travel behavior under risks (Avineri and Prashker (2004)).

For the purpose of evaluating VTTV, these extended theories are often applied to the mean-variance approach. Hjorth (2011) applied the CPT in estimating the mean-variance model and found that the economic significance of probability weighting was however ascertain since the shape of weighting function varied considerably when applying different function forms (Tversky and Kahneman (1992); Prelec (1998)). Similarly, different shapes of weighting functions were also identified by Hensher et al. (2011), where risk attitude, probability weighting and heterogeneity were considered in estimating the mean-variance model.

Little research has been made in applying sophisticated frameworks, such as PT, RDU, and CPT, in the context of scheduling model, particularly the quadratic one. It is considered that scheduling approach is better alternative than the mean-variance model in VTTV evaluation. It is of interest to investigate how the departure time choice and its maximal scheduling utility change under RDU maximization. Koster and Verhoef (2010) provided an analytical framework for deriving the linear scheduling model under RDU maximization. They performed a sensitivity analysis on the optimal departure time considering the probability weighting function of Prelec (1998). There is no analytical framework developed for the quadratic model, and a more flexible probability weighting function is on demand Rieger and Wang (2006). In addition, the two scheduling models were extended by allowing for discrete penalties for excessive travel time (Wang et al. (2012)). Such an idea of evaluating the shape of travel time distribution for VTTV is consistent with the argument in van Lint et al. (2008), for the reason that the travel time distribution is often skewed. Therefore, this paper is motivated in deriving the extended forms of linear and quadratic scheduling models accounting the flexible weighting function of Rieger and Wang (2006) in a RDU framework.

In the following sections, we firstly show how to derive the linear and the quadratic models using RDU maximization; and how the optimal departure times and optimal scheduling utilities are affected by the probability weighting function of Rieger and Wang (2006). The derived scheduling models are estimated on a Stated Preference data set, which was designed for measuring VTTV. Scheduling parameter estimates from multinomial logit (MNL) estimation and mixed multinomial logit (MXL) are reported. In particular, the parameters indicating the probability weighting function and the random parameters for heterogeneity are presented. An estimation technique, applied in estimating the derived liner scheduling model, is also discussed.

2 Theoretical Model

2.1 Rank dependent utility maximizing

RDU assumes that the value of an outcome depends on both the probability of such an outcome and the ranking of this outcome in comparison to other outcomes in the same lottery set. Following a new version of RDU derived by Diecidue and Wakker (2001), all possible outcomes are ordered from best to worst, i.e., $x_1 > ...x_n$ with corresponding probability $p_1 > ...p_n$. The ranked dependent utility of such a ranked lottery $(p_1, x_1; ...; p_n, x_n)$ is given by

$$RDU(p_1, x_1; ...; p_n, x_n) = \sum_i \pi_i * U(x_i), \tag{1}$$

where, for each i

$$\begin{aligned} \pi_i &= w(p_1 + ... + p_i) - w(p_1 + ... + p_{i-1}) \\ \pi_1 &= w(p_1) \end{aligned} \tag{2}$$

Graphically, Diecidue and Wakker (2001) showed the probabilities are transformed by the decision weight π in Fig.1.

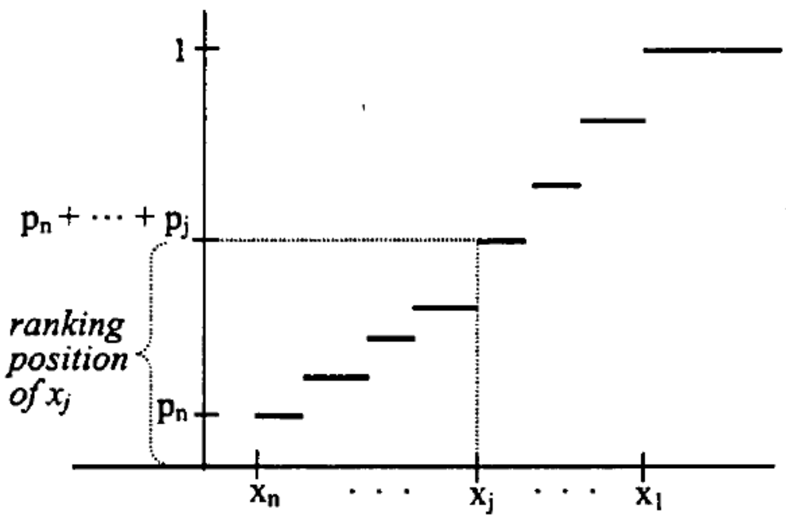

Figure 1: Graph of distribution function of lottery($p_1, x_1; ...; p_n, x_n$) source: Diecidue and Wakker (2001)

RDU is determined by the utility and the corresponding decision weight . A probability weighting function w transforms the probability into non-linear. As discussed in the introduction, there are some alternative parametric functions of w, such as Tversky and Kahneman (1992) and Prelec (1998). Each of those classic weighting functions has its own limitation, and Rieger and Wang (2006) suggested a new type of weighting function which avoids infinite values for the subjective utility.

$$w(F) = \frac{3-3b}{a^2-a+1}[F^3-(a+1)F^2+aF]+F, \tag{3}$$

w in Eq. 3 transforms probability F with two parameters $a \in 0,1$ and $b \in 0,1$. "a is the point on which w changes from over-weighting to under-weighting; and b represents the curvature of w", as stated in Rieger and Wang (2006)). It possesses some favorable properties, and some could be violated by the classic weighting functions. At the same time, w has a concave-convex structure so that it is limited to inverse S-shaped weighting function. To give a intuition of such a new weighting function, plots of w for different values of a and b are displayed in Fig.2

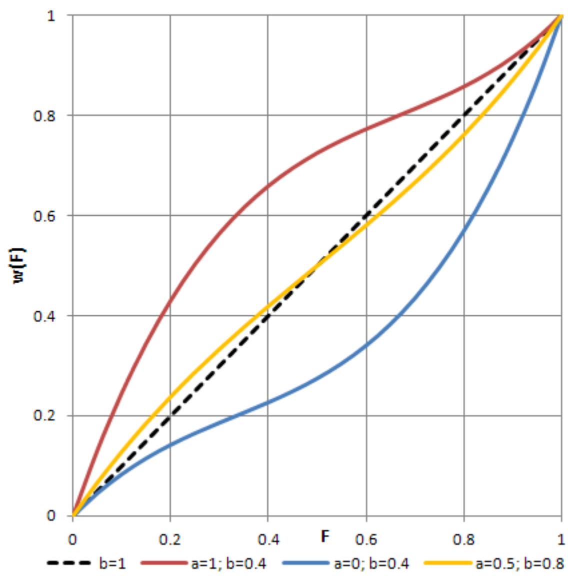

Figure 2: Example of weighting function w by Rieger and Wang (2006)

w is linear when $b = 1$ since $w(F) = F$, as the black dashed line displays. The yellow curve depicts a concave-convex w with $0 < b < 1$ and it crosses at a point $w(F) = F = 0.5$. The smaller b is, the more twisted w becomes. In this case, extreme outcomes are over-weighted. In the area of concave, the transformed cumulative probability function indicates optimism, since the probability of good outcomes are over-weighted. Reversely, pessimism corresponds to the area of convex, that the probability of bad outcomes are over-weighted instead.

Parameter a indicates the point n which the curve changes from concave to convex. If we allow $a = 1$, w always over-weight probabilities as the red line shows. The estimated weighting function on empirical data,as displayed later in Fig.3, has the parameter a statistically indifferent from 1. It implies optimism when bad outcomes are under-weighted, even though the curve changes from concave to convex at $F = 2/3$. Likewise, the bad outcomes are overweighted if parameter a is fixed at 0, as the blue curve shows. Its curvature changes from concave to convex at $F = 1/3$.

2.2 Scheduling models under RDU

Scheduling model can take different specification (see Vickrey (1973); Small (1982); Polak (1987)), In this paper, two frequently applied scheduling models, a linear one and a quadratic one are analyzed accounting probability weighting from RDU. Considering a trip departures at time D, travel time of length T, and arrives at time $D+T$, the scheduling models are given by

$$U(T,D) = \begin{cases} \alpha T + \beta \max(0, -(D+T)) + \gamma \max(0, (D+T)) + \kappa J & \text{Linear} \\ \eta T - \nu/2D^2 + \omega/2(D+T)^2 + \kappa J & \text{Quadratic} \end{cases} \tag{4}$$

The linear model provides the scheduling preference parameter α to travel time T, β to arrival earlier than preferable arrival time (PAT) and γ to arrival later than PAT. PAT is normalized at 0 without loss of generality. As we known that only the difference in utility matters. When the difference in departure time and travel time are used as the input in utility, $PAT = 0$ also means that the current arrival time is considered as preferable. The quadratic scheduling model provides scheduling preferences parameters η, ν and ω. The penalty caused by excessive travel time κJ is included according to Wang et al. (2012). The dummy variable J is defined by

$$J = \begin{cases} 1 & \text{if} \quad T > \tau \\ 0 & \text{otherwise} \end{cases} . \tag{5}$$

The threshold τ determines the point at which a discrete penalty occurs. If T is continuous distributed, τ could be estimated.

Assume that the random travel time T is distributed with a probability density function of $\phi(T)$ and cumulative probability function of $\Phi(T)$, which is independent of time-of-day [ii]. The expected scheduling utility is simply $EU(D) = \int_0^\infty U(T,D)\phi(T)dT$, and the expected utility is

[ii]The assumption that the travel time distribution is independent of the departure time could be unrealistic and more discussion can be found in Fosgerau and Karlström (2010) and Wang et al. (2012)

maximized by choosing the optimal departure time. The derivation of the linear and quadratic scheduling models in Eq. 4 under EU maximization are shown in Appendix A.1 and Appendix B.1 respectively. The maximal expected scheduling utilities LEU^* and QEU^* are yielded in Eq.6 and Eq.7. The scheduling models was once presented in Wang et al. (2012), in which the standardized travel time X with cumulative probability $\Phi(X)$ was used to represent the travel time distribution[iii].

$$LEU^* = (\alpha - \beta)\mu + (\beta + \gamma) \int_{\frac{\gamma}{\beta+\gamma}}^{1} \Phi^{-1}(F)dF + \kappa(1 - \Phi(\tau))dF \tag{6}$$

$$QEU^* = \eta\mu + \frac{\omega^2\mu^2}{2(\nu - \omega)} + \frac{\omega}{2}(\mu^2 + \sigma^2) + \kappa(1 - \Phi(\tau))dF \tag{7}$$

Rank depended scheduling utility is different from the expected one in the transformed probability function. If we denote the weighted cumulative probability function as $w[\Phi(T)]$, the probability density becomes

$$dw(\Phi(T)) = \frac{\partial w(\Phi(T))}{\partial T} = \frac{\partial w[\Phi(T)]}{\partial \Phi(T)}\phi(T)dT. \tag{8}$$

The travel time is ranked from best to worst, i.e., from 0 to positive infinity. Then rank dependent utility is given by

$$RDU(D) = \int_0^\infty U(T, D)\frac{\partial w(\Phi(T))}{\partial \Phi(T)}\phi(T)dT. \tag{9}$$

By inserting the scheduling utilities of Eq.4 in Eq.9, the rank dependent linear scheduling utility $LRDU(D)$ and the rank dependent quadratic scheduling utility $QRDU(D)$ are given by Eq.10 and Eq.11.

$$\begin{aligned} LRDU(D) &= \alpha \int_0^\infty T\frac{\partial w[\Phi(T)]}{\partial \Phi(T)}\phi(T)dT + \beta \int_0^{-D} -(D+T)\frac{\partial w[\Phi(T)]}{\partial \Phi(T)}\phi(T)dT \\ &+ \gamma \int_{-D}^\infty (D+T)\frac{\partial w[\Phi(T)]}{\partial \Phi(T)}\phi(T)dT + \kappa \int_\tau^\infty \frac{\partial w[\Phi(T)]}{\partial \Phi(T)}\phi(T)dT, \end{aligned} \tag{10}$$

$$\begin{aligned} QRDU(D) &= \eta \int_0^\infty T\frac{\partial w[\Phi(T)]}{\partial \Phi(T)}\phi(T)dT - \nu/2D^2 + \frac{\omega}{2}\int_0^\infty (D+T)^2\frac{\partial w[\Phi(T)]}{\partial \Phi(T)}\phi(T)dT \\ &+ \kappa \int_\tau^\infty \frac{\partial w[\Phi(T)]}{\partial \Phi(T)}\phi(T)dT \end{aligned} \tag{11}$$

[iii]Fosgerau and Karlström (2010) defined $T = \mu + \sigma X$, where X is a standardize random variable with mean 0, variance 1, density $\phi(X)$, and cumulative distribution $\Phi(x)$. Their results suggested that the derived form of a linear scheduling model is a linear function of μ and σ. Wang et al. (2012) used $\Phi(X)$ in their analysis to examine whether the favorable linear property is kept in the extended model specification. Transformation is provided in Appendix A.3.

In line with random utility maximization (RUM), individuals are assumed to have preferences over scheduling choices and they would adjust departure times D to maximize their rank dependent utilities.

$$RDU^* = \max_D RDU(D) \tag{12}$$

By taking the first order condition of $LRDU$ over D, the optimal departure time LRD^* is obtained in Eq.13. Likewise, QRD^* is given in Eq.14. Derivations are found in Appendix A.2 and Appendix B.2

$$LRD^* = -\Phi^{-1}\left[w^{-1}(\frac{\gamma}{\beta+\gamma})\right] \tag{13}$$

$$QRD^* \quad = \quad \frac{\omega}{\nu-\omega}\int_0^1 \Phi^{-1}(F)\frac{\partial w(F)}{\partial F}dF \tag{14}$$

where the cumulative probability is notated as $F = \Phi(T)$. By inserting the optimal departure time into the rank dependent utility, the optimal rank dependent utility $LRDU^*$ and $QRDU^*$ are obtained in Eq.15 and Eq.16. The derivations are found in Appendix A.2 and Appendix B.2 respectively.

$$\begin{aligned} LRDU^* \quad &= \quad (\alpha-\beta)\int_0^1 \Phi^{-1}(F)\frac{\partial w(F)}{\partial F}dF \\ &+ \quad (\beta+\gamma)\int_{w^{-1}(\frac{\gamma}{\beta+\gamma})}^1 \Phi^{-1}(F)\frac{\partial w(F)}{\partial F}dF + \kappa\int_{\Phi(\tau)}^1 \frac{\partial w(F)}{\partial F}dF \end{aligned} \tag{15}$$

$$\begin{aligned} QRDU^* \quad &= \quad \eta\int_0^1 \Phi^{-1}(F)\frac{\partial w(F)}{\partial F}dF + \frac{\omega^2}{2(\nu-\omega)}[\int_0^1 \Phi^{-1}(F)\frac{\partial w(F)}{\partial F}dF]^2 \\ &+ \quad \frac{\omega}{2}\int_0^1 [\Phi^{-1}(F)]^2\frac{\partial w(F)}{\partial F}dF + \kappa\int_{\Phi(\tau)}^1 \frac{\partial w(F)}{\partial F}dF \end{aligned} \tag{16}$$

3 Data

A stated preference study was conducted to commuters, traveling towards and from the center of Stockholm city by the subway and commuter trains during peak hours. The questionnaire started with questions relating to their current trips, including travel time duration, ticket price, constraints at origin or destination, safety margin, etc. Then respondents were faced with two stated choice experiments, and they were asked to make a series of binary choices in each experiment.

Experiment 1 served for the purpose of estimating a scheduling model which assumed respondents would trade-off departure times for travel times. The experiment contained binary choices differing in the departure time d, travel time t and the travel cost c. Travel times T was deterministic in *Experiment 1*. *Experiment 2* was designed for estimating the derived-form model which assumed the trade-off was made among travel time distribution characteristics. Respondents were asked to make binary choices differing in the travel time t, travel cost c, the probability of a delay p and its duration l. The travel times T was stochastic in experiment 2, and it followed a binary travel time distribution defined as

$$\Phi(T) = \begin{cases} 0 & \text{for} \quad T < t^* + t \\ 1-p & \text{for} \quad t^* + t <= T < t^* + t + l \\ 1 & \text{for} \quad T >= t^* + t + l \end{cases}, \tag{17}$$

where t^* is the observed travel time [iv].

An example of the binary choice *Experiment 1* and *Experiment 2* are displayed respectively in table 1 and 2. Note that to cancel the trip is also offered as an option. More information about the experimental design and sample statistics is provided by Börjesson et al. (2012)).

	Departure1	**Departure2**
Start Time	**25 min**later than today	**5 min** later than today
Travel time	**15 min**longer than today	**45 min** longer than today
Ticket price	**€0.70** higher than today	**€0.40** lower than today
Arrival Time	(45 min later than today)	(50 min later than today)
I choose	1	2
	Cancel the trip 3	

Table 1: Stated preference choice experiment 1

4 Estimation

Scheduling preference parameters for *experiment 2* has been reported and discussed in Wang et al. (2012), where estimation for *experiment 1* is not included. Wang et al. (2012) focused on the prof, whether the derived forms of extended scheduling models retain the linear property in characteristics of travel time distributions. The hypothesis on the equivalence between scheduling models and corresponding derived forms was once examined by Börjesson et al. (2012). Their findings opposed the hypothesized equivalence. Since the conventional specifications of scheduling models have been extended, the comparison are carried out in this paper.

[iv] Does the observed travel time t^* affect scheduling choices of respondents? In other words, do the respondents mentally consider that the travel time distribution depends on the observed travel time? This can make a difference in the estimation for the quadratic model which is nonlinear in T

	Departure1	**Departure2**
Delay(if you made this trip every day)	Once every other month, the train is **45 min** delayed All other trip are on-time	Once every other week, the train is **10 min** delayed All other trip are on-time
Travel time according to the timetable:	**3 min** shorter than today	**10 min** shorter than today
Ticket price	**€0.20** higher than today	**€1.00** higher than today
I choose	1	2
	Cancel the trip 3	

Table 2: Stated preference choice experiment 2

The ratios between the scheduling preference parameter and the coefficient for travel cost are presented and discussed in 4.4.

To make comparison for the effect RDU maximization on scheduling parameter estimates, estimations on derived forms under EU maximization are presented in the first place. Both MNL estimations and MXL estimations are displayed. Last but most importantly, we estimate the derived scheduling models considering the probability weighting function w under RDU maximization. The parameter estimates with robust t-statistics value are reported. Values of Null Log-Likelihood (Null LL) and Final Log-Likelihood (Final LL), Adjusted Rho Square (Adj. RS) and Akaike information Criterion with Correction (AICC) are computed as the criteria for evaluating model performance [v]. To be clear, the abbreviations of estimated models are explained in Tab.3.

4.1 Logit estimation

In *Experiment 1*, individuals are assumed to trade off their departure time D for travel time T, as modeled by Eq.4. We estimate the models on data from *Experiment 1*, using Eq.18

$$U = \begin{cases} \alpha T + \beta \max(0, -(D+T)) + \gamma \max(0, (D+T)) + \theta D_L + \kappa J + \lambda c & \text{linear} \\ (\eta + \frac{\nu\omega}{\nu-\omega} t^*)T - \nu/2D^2 + \omega/2(D+T)^2 + \kappa J + \lambda c & \text{quadratic} \end{cases}, \tag{18}$$

where λ measures the marginal utility for travel cost c, and θ to the variable D_L is the lateness dummy for actual arrival time D+T surpassing the PAT at time 0 (Small (1982)). The formulation for the quadratic model is derived (see derivation in Börjesson et al. (2012)) on the assumption that individual chooses an optimal departure time given the current travel time duration t^*.

The derived forms of scheduling models are estimated on data from *Experiment 2*. Given

[v] $RS = 1 - \frac{L^*}{L^0}$; $Adj.RS = 1 - \frac{L^*-k}{L^0}$; $AICC = 2(k - L^*) + \frac{2k(k+1)}{n-k-1}$;
L^* is Final Log-Likelihood; L^0 is Null Log-Likelihood; k is number of parameters; n is sample size.

Abbreviation	Explanation
L or Q	Linear or quadratic specification
U	Scheduling models
EU^*	The derived forms of the conventional scheduling models
$EU^*(\theta)$	The derived form of the conventional linear scheduling model with the coefficent θ for the lateness variable
$EU^*(\kappa)$	The derived forms of the extended scheduling models with the coefficent κ for the excessive travel time variable
RDU^*	The derived forms of the extended scheduling models under rank dependent utility maximization
MXL	Mixed logit estimation

Table 3: Explanation on abbreviations

the binary distribution, the derived models in Eq.(6) and Eq.(7) can be formulated as

$$EU^* = \begin{cases} \alpha\mu + \gamma pl + \theta p * 100 + \kappa p * 100 + \lambda c & \text{linear, for } \frac{\gamma}{\beta+\gamma} < 1 - p \\ \alpha\mu + \beta(1-p)l + \kappa p * 100 + \lambda c & \text{linear, for } \frac{\gamma}{\beta+\gamma} \geq 1 - p \\ \eta\mu + \frac{\omega^2}{2(\nu-\omega)}\mu^2 + \frac{\omega}{2}(\mu^2 + \sigma^2) + \kappa p * 100 + \lambda c & \text{quadratic} \end{cases}, \tag{19}$$

where $\mu = t^* + t + pl$ and $\sigma = l\sqrt{p(1-p)}$. Because of the binary travel time distribution, the linear model has a piece-wise function. In addition,the variable of excessive travel time frequency is equal to the probability of being delay p, provided that we have assumed the lateness threshold $\tau \in (t, t+l)$. The maximized expected value of lateness dummy D_L equals also to p. By scaling p by 100 in Eq.19, the parameter κ represents the marginal value to percentage of excessive delays. The scale in p is kept for estimations afterwards.

We estimate scheduling models on data from *experiment 1* and the corresponding derived forms under EU maximization on data from *experiment 2*. Results of linear and quadratic models are presented in Tab.4 and Tab.5.

All parameters reported are statistically significant different from 0. The discrete variable J turns out to be zero in *experiment 1* . I.e., we can not find a value for variable τ so that the dummy variable J is active. Parameter κ estimated on *experiment 2* appears as an important explanatory. The contribution by allowing for the penalty of excessive travel time is sufficiently discussed in Wang et al. (2012). If compared to parameter θ estimated in the derived form of the linear specification, parameter κ has stronger power in fitting the scheduling preferences when travel time is uncertain. Another difference between model $LEU^*(\theta)$ and $LEU^*(\kappa)$ is that the marginal utility of γ. Compared to γ in the original specification LEU^*, $LEU^*(\theta)$ has lower absolute value of γ while $LEU^*(\kappa)$ has much higher value. Adding the penalty for excessive travel time frequency in the linear scheduling model result in the high preference spending time at the destination after the PAT rather than at the origin. It implies that

	Experiment 1		Experiment 2					
Models	*LU*		LEU^*		$LEU^*(\theta)$		$LEU^*(\kappa)$	
Results	Value	Robust-t	Value	Robust-t	Value	Robust-t	Value	Robust-t
α	-0.092	-16.952	-0.128	-7.702	-0.139	-8.011	-0.166	-14.496
β	-0.062	-15.861	-0.096	-5.182	-0.089	-4.778	-0.035	-6.849
γ	-0.058	12.822	-0.476	-4.829	-0.400	-3.844	-1.030	-6.978
λ	-0.082	-15.017	-0.123	-9.51	-0.134	-9.584	-0.151	-15.063
κ or θ	-	-	-	-	-0.018	-2.153	-0.100	-8.887
No. of observations	3226		2996		2996		2996	
Null Log-Likelihood	-2236.09		-2076.70		-2076.70		-2076.70	
Final Log-Likelihood	-1666.71		-1783.42		-1781.10		-1757.17	
Rho square	0.255		0.141		0.142		0.154	
Adj. Rho square	0.252		0.139		0.140		0.152	
AICC	3347.46		3574.86		3572.33		3524.13	

Table 4: Results of the MNL estimation of linear scheduling models

individuals have the stronger tendency to departure earlier to avoid uncertain delay.

Similar results are found in quadratic scheduling model. Parameter ω estimated on data from *experiment 1*, representing the willingness of staying at work other than in the vehicle, is insignificant different from 0. In other words, the slope of marginal utility of being at work is constant when travel time T varies with certainty. So ω is fixed at 0 when estimating the quadratic model in *experiment 1.* Both the derived forms of the original specification and the extended specification with κ are estimated. Penalty of κJ for excessive travel time frequency makes the preference of being early stronger and the slope of decreasing marginal utility of ω is bigger. Note that results of quadratic scheduling model on *experiment 1* are different from the values in Börjesson et al. (2012), since the latter estimated an extra coefficient to replace $\frac{\nu\omega}{\nu-\omega}$ in Eq.18. Comparing model performances of the linear and the quadratic model, the quadratic formulation fits the data of *experiment 1* better than the linear formulation with one parameter estimate less; whilst they achieve similar goodness of fit to the empirical data of *experiment 2.*

4.2 Mixed Logit estimation

The standard MNL estimation does not considered the random taste variation among individuals. It is natural to allow for heterogeneity for individual scheduling preferences of spending time at different activities. The panel effect is taken into consideration for the repeated choices made by each individual. We assume the scheduling parameters are normal distributed, and we use statistical tests to find significant standard deviation parameters of scheduling parameters.

	Experiment 1		Experiment 2			
Models	QU		QEU^*		$QEU^*(\kappa)$	
Results	Value	Robust-t	Value	Robust-t	Value	Robust-t
η	-0.125	-21.3082	-0.095	-4.4122	-0.102	-4.7422
ν	0.003	17.780	0.002	4.356	0.001	2.577
ω	-	-	-0.009	-2.970	-0.026	-6.349
λ	-0.104	-18.16	-0.154	-13.802	-0.137	-12.213
κ	-	-	-	-	-0.066	-7.265
No. of observations	3226		2996		2996	
Null Log-Likelihood	-2236.09		-2076.70		-2076.70	
Final Log-Likelihood	-1621.91		-1783.45		-1755.44	
Rho square	0.275		0.141		0.155	
Adj. Rho square	0.272		0.139		0.152	
AICC	3257.86		3574.82		3520.90	

Table 5: Results of the MNL estimation of quadratic scheduling models

Börjesson et al. (2012) tested a triangular distributions in which scheduling parameters were constrained to be negative. They conclude the normal distribution yields better results.

An exception is made for the MXL estimation of derived linear model. The linear model has a piecewise function, as formulated in Eq.19. The piecewise function generates difficulty in estimation because the simulated likelihood function is not differentiable everywhere. In Appendix C, we discuss the problem and apply a sampling technique, Gibbs sampling, to draw random values for scheduling parameters. We thus assume that α, β and γ in the derived linear model are log normal distributed, even through Börjesson et al. (2012) argued that the data do not support a log normal distribution with long and fat tail.

Tab.6 presents results MXL estimations of the linear models. In *Experiment 1*, σ_α and σ_β are found significant different from 0. The final log-likelihood is -1628.35, significantly better than -1666.71 for MNL estimation in Tab.4 to, with two more parameters. No heterogeneity is detected for the marginal utility of being at the destination γ, when travel time is delayed with certainty. MXL estimations on *Experiment 2* find statistical significant σ_α and insignificant σ_β. Though the random parameter σ_γ is statistically insignificant different from 0 in $LEU^*\_MXL_1$, excluding σ_γ in $LEU^*\_MXL_2$ deteriorates the model performance, according to decreased final log-likelihood. There is taste variation in the marginal utility of staying at the destination γ, but it is difficult to capture such a variation with a distribution parameter. So the random parameter σ_γ is insignificant and its value is unexpectedly high. The final Log-likelihood is increased by nearly 100, comparing the MNL estimation of -1757.17 in 4 and the MXL estimation of -1647.67. Such results show the significance in considering the heterogeneity of

	Experiment 1		Experiment 2			
Models	LU_MXL		$LEU^{*}\_MXL_1$		$LEU^{*}\_MXL_2$	
Results	Value	Robust-t	Value	Robust-t	Value	Robust-t
α	-0.123	-14.134	-0.375	-9.650	-0.416	-11.474
β	-0.075	-14.302	-0.097	-6.503	-0.049	-7.087
γ	-0.071	-12.903	-2.677	-3.293	-1.325	-6.568
λ	-0.102	-14.074	-0.222	-13.226	-0.244	-16.043
κ	-	-	-0.130	-7.749	-0.117	-7.473
σ_α	-0.067	-8.521	0.556	4.967	0.536	5.854
σ_β	0.021	2.369	-	-	-	-
σ_γ	-	-	4.569	1.593	-	-
No. of observation	3226		2996		2996	
Null Log-Likelihood	-2236.09		-2076.70		-2076.70	
Final Log-Likelihood	-1628.35		-1647.67		-1656.95	
Rho square	0.272		0.207		0.202	
Adj. Rho square	0.269		0.203		0.199	
AICC	3268.73		3309.37		3325.93	

Table 6: Results of the MXL estimation for linear scheduling models.

scheduling parameters.

MXL estimations of the quadratic models are reported in Tab.7. σ_η is statiscally significant estimated on *Experiment 1*, improving the final log-likelihood of -1621.09 in Tab.5 to -1590.79. The hypothesis on that parameters ν and ω follow normal distribution is rejected, indicating that there is neither individual taste preferences to the slope of marginal utility of staying at home other than in vehicle, nor to the slope of marginal utility of staying at work other than in vehicle. The MXL estimation for *Experiment 2* shows that parameter ν becomes insignificant, if allowing for normal distribution for parameters η and ω. When travel time varies with certainty, individuals are indifferent to the scheduling preference parameter ω, the slope of marginal utility of beging at work other than in vehicle. On the other hand, individuals are indifferent to the scheduling preference parameter ν, the slope of marginal utility of being at home other than in vehicle when travel time varies with uncertainty. Again, including random parameters improves the performance of the quadratic model in fitting the data. The *final Log-likelihood* of the quadratic model is increased by 100 in the MXL estimation than in the MNL estimation.

If the performance of linear and quadratic scheduling models are compared. MXL estimation on *Experiment 1* shows that quadratic one with 4 parameters outperforms the linear one

	Experiment 1		Experiment 2	
Models	QU_MXL		QEU^{*}_MXL	
Results	Value	Robust-t	Value	Robust-t
η	-0.161	-16.821	-0.285	-10.737
ν	0.003	15.910	-	-
ω	-	-	-0.063	-7.499
λ	-0.128	-15.942	-0.211	-11.857
κ	-	-	-0.087	-6.907
σ_η	-0.070	-8.713	-0.257	-11.080
σ_ν	-	-	-	-
σ_ω	-	-	0.062	6.101
No. of observation	3226		2996	
Null Log-Likelihood	-2236.09		-2076.70	
Final Log-Likelihood	-1590.79		-1656.53	
Rho square	0.289		0.202	
Adj. Rho square	0.287		0.199	
AICC	3189.59		3325.09	

Table 7: Results of the MXL estimation for linear scheduling models.

with 6 parameters. Whilst MXL estimation on *Experiment 2* draws different conclusion that the linear one is slight better than the quadratic one, considering the degree of freedom. By taking the heterogeneity, the quadratic model lose its advantage of nonlinearity to the linear scheduling model. The MXL estimations provide us the knowledge on the heterogeneity that allowed in the preference parameters. The next step is to investigate the proof of the probability weighting in scheduling choices.

4.3 Derived models under RDU maximization

To remind us, the cumulative probability of outcomes, ranked from good to bad, are transformed by the weighting function w. Provided the binary travel time distribution in *Experiment 2*, the cumulative probability is the probability of being on time $1 - p$. There are only four points for the variable $1 - p$ in the data. This could cause difficulty in identifying parameters a and b in the parametric weighting function w. Especially, the four discrete points are gathered around small probabilities of being delayed p, and we might not be able to estimate the parameter a, which determines the point where the probability weighting function crosses between over-weighting and under-weighting.

The general formulation of derived model taking into consideration of probability weighting in Eq.15 and Eq.16 are interpreted with respect to the binary travel time distribution in *Experiment 2*, we have

$$\begin{cases} \alpha\tilde{\mu} + \gamma(1-w(1-p))l + \kappa(1-w(1-p)) * 100 + \lambda c & \text{Linear, for } \frac{\gamma}{\beta+\gamma} < w(1-p) \\ \alpha\tilde{\mu} + \beta w(1-p)l + \kappa(1-w(1-p)) * 100 + \lambda c & \text{Linear, for } \frac{\gamma}{\beta+\gamma} \geq w(1-p) \\ \eta\tilde{\mu} + \frac{\omega^2}{2(\nu-\omega)}\tilde{\mu}^2 + \frac{\omega}{2}(\tilde{\mu}^2+\tilde{\sigma}^2) + \lambda c + \kappa(1-w(1-p)) * 100 & \text{Quadratic} \end{cases} \tag{20}$$

where $\tilde{\mu} = t^* + t + (1-w(1-p))l$ and $\tilde{\sigma} = l\sqrt{(1-w(1-p))w(1-p)}$

Models	$LRDU^*\_MXL_1$		$LRDU^*\_MXL_2$		$LRDU^*\_MXL_3$	
ML estimates	Value	Robust-t	Value	Robust-t	Value	Robust-t
α	-0.386	-9.957	-0.386	-10.081	-0.421	-11.219
β	-0.090	-5.699	-0.090	-5.725	-0.047	-5.868
γ	-5.137	-1.267	-5.189	-1.397	-1.477	-2.073
λ	-0.232	-12.653	-0.232	-13.167	-0.242	-14.310
κ	-0.205	-1.882	-0.209	-4.918	-0.208	-2.168
σ_α	0.531	4.879	0.529	5.279	0.553	5.271
σ_γ	14.681	0.730	14.672	0.741	-	-
a	0.991	3.611	1	-	1	-
b	0.410	2.224	0.409	2.265	0.154	0.959
No. of observation	2996		2996		2996	
Null Log-Likelihood	-2076.7		-2076.7		-2076.7	
Final Log-Likelihood	-1644.9		-1644.9		-1649.3	
Rho square	0.208		0.208		0.205	
Adj. Rho square	0.204		0.204		0.202	
AICC	3305.93		3305.37		3314.57	

Table 8: Results of the MXL estimation for linear scheduling models with probability weighting

We estimate the derived models under RDU maximization. The scheduling preference parameters α and γ are assumed log normal distributed, according to the previous MXL estimations. As aforementioned, a Gibbs sampling technique may be required in MXL estimation of the linear model. Tab.8 shows the results of the linear model. $LRDU^*\_MXL_1$ estimates the parameters a and b of the probability weighting function in Eq.3, parameter a is found insignificant different from 1 and b is about 0.4. Thus we assume $a = 1$ in the model $LRDU^*\_MXL_2$, and the model performance to the data is not deteriorated by fixing a at 1. Again, the random parameter σ_γ is statistically indifferent from 0, but removing it in the model $LRDU^*_MXL_3$degrades the model performance. The taste variation in the marginal utility of staying at the destination other than at the origin should be kept. Thus, the model

$LRDU^{*}\_MXL_2$ is chosen. The estimated probability function w in the derived liner scheduling model is parameterized by $a = 1$ and $b = 0.4$.

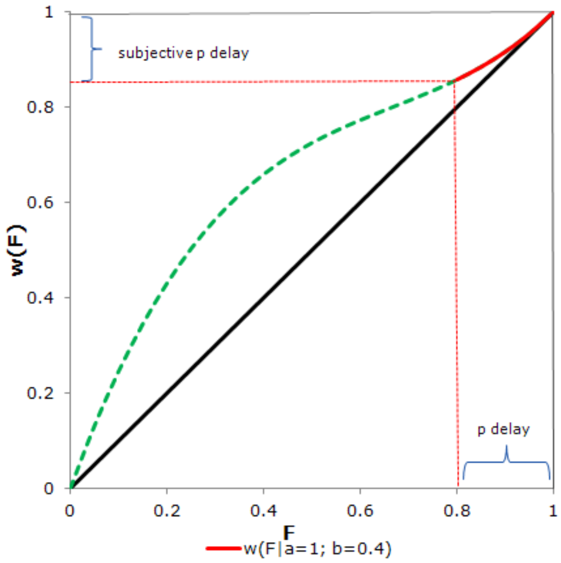

Figure 3: Estimated weighting function w

The green curve in Fig.3 depicts the shape of the estimated weighting function. The probability of being delay in the data is $p = \{0.025; 0.05; 0.1; 0.2\}$, that is, the probability of being on time $1 - p = \{0.975; 0.95; 0.9; 0.8\}$. Thus partial curve in the domain $[0.8, 1]$ is highlighted with solid red line. A point of $F = 1 - p = 0.8$ is analyzed. The objective probability of being delayed is 0.2 and the subjected probability of being delayed is smaller than 0.2, because of the transformation of weighting function. In other words, probabilities of being on time are over-weighted. Such behavior corresponds to optimism, provided that the uncertain travel time are ranked from best to worst in the domain of the analysis. This finding is consistent with other empirical results, for instance, Hensher et al. (2011) found risk-taking attitude in his data set where respondents were faced with risky route choices.

When estimating the same probability weighting function w with the quadratic specification, the parameter b is statistically indifferent from 1. The nonlinearity in probability function is statistically insignificant with the the quadratic specification. I.e., we cannot reject the expected utility maximization in favor of rank dependent utility maximization. One reason could be that the quadratic scheduling utility is concave per se and risk attitude is taken into account. On the other hand, the risk attitude is captured by the probability weighting function when the utility specification is linear. It is also difficult to estimate parameters in the weighting function

w, provided the lack of variation in probability.

Evidently, values of scheduling preference parameters are affected by the transformed probability in rank dependent scheduling model. Thus the value of travel time variability by scheduling preference could be different between EU assumption and RDU assumption. Another observation with RDU estimation is parameter κ retains its significance after introducing probability weighting. In other words, the right tail of the travel time distribution is still evaluated as part of value of travel time variability

4.4 *Experiment 1* vs *Experiment 2*

One of the purposes in this paper is to analyze the equivalence between the scheduling model and its derived form. I.e., the scheduling preference parameters estimated on *Experiment 1* and *Experiment 2* are compared. Theoretically, they could be reconciled. But such an argument is denied by empirical results from Börjesson et al. (2012). The scheduling models are extended by allowing for penalty of excessive travel time, considering random taste in scheduling preferences, and including probability weighting in utility maximization. It is of interest to make the comparison again for empirical evidence of equivalence. The ratio between the scheduling parameter and the coefficient for travel cost is used for indication. Ratios by estimating the scheduling models on *Experiment 1* are computed and considered as the reference values. Ratios by estimating the derived forms of scheduling models on *Experiment 2* are computed and compared with the corresponding reference values. For the MXL estimation, the mean of the parameter is used for computing the ratio and the variance is not considered here.

Ratio	LU	LU_MXL	LEU^*	$LEU^*(\theta)$	$LEU^*(\kappa)$	$LEU^*\_MXL_2$	$LRDU^*\_MXL_2$
α/λ	1.122	1.206	1.041	1.037	1.099	1.689	1.664
β/λ	0.756	0.735	0.781	0.664	0.232	0.437	0.388
γ/λ	0.707	0.696	3.870	2.985	6.821	12.059	22.366
κ/λ or θ/λ	-	-	-	0.134	0.662	0.586	0.901

Table 9: Ratios from the linear scheduling model

Ratio	QU	QU_MXL	QEU^*	$QEU^*(\kappa)$	QEU^*_MXL
η/λ	1.202	1.258	0.617	0.745	1.351
ν/λ	-0.029	-0.023	-0.013	-0.007	-
ω/λ	-	-	0.058	0.190	0.299
κ/λ	-	-	-	0.482	0.412

Table 10: Ratios from the quadratic scheduling model

Ratios from the linear scheduling model are presented in Tab.9. Ratios in the first two columns are the reference values,which are compared with the corresponding ratios in the rest columns. The ratio γ/λ changes drastically across columns. γ/λ from model LEU^* is 5 times higher than the reference value from model LU or LU_MXL. If the lateness penalty Small (1982) is considered in $LEU^*(\theta)$, the difference is diminished by a small amount. The model $LEU^*(\kappa)$ is considered more consistent with the behavior in *Experiment 2*, but the ratio γ/λ has even larger disparity, compared with the reference value. Such difference is enlarged in the model $LEU^*(\kappa)$ and $LEU^*\_MXL_2$. The difference among the ratios α/λ is rather small, because they represents the value of mean travel time. The ratios β/λ are also similar across columns. Another interesting finding is that the ratio κ/λ is larger than the ratio θ/λ.

Similar results are found with the quadratic models, displayed in Tab.10. The reference values from the model QU and QU_MXL show that the ratio ω/λ is 0 in *Experiment 1*. While in *Experiment 2*, ω/λ becomes significant. In contrast, the ratio ν/λ decreases and become insignificant, estimated in the model QEU^*_MXL.

In conclusion, no reconcile is found from the scheduling models, after considering the shape of travel time distribution, random taste in scheduling preferences parameters, and probability weighting. Individuals behave differently between *Experiment 1* and *Experiment 2*, thus the theoretical equivalence is not satisfied.

5 Conclusion

Rank dependent utility maximization is applied in deriving a linear and a quadratic scheduling model in this paper. The optimal departure time and maximal utility are different from that under expect utility maximization in the transformed travel time density function. Probability weighting is found when estimating the linear model and the estimated weighting function suggests optimism behavior of respondents. One should interpret this finding with caution. The nonlinearity in subjective probability weighting over ranked outcomes is only confirmed in the derived linear scheduling specification. The data cover only a restricted domain where the probabilities of delay are small. Respondents could behave as risk-taking when probabilities of bad outcomes are small. The results also reveal the evidence of heterogeneity in scheduling preferences. Moreover, evidence for the variable of excessive travel time beyond the traditional scheduling model specification is found in all estimation, even with controlling for probability weighting. Ratio between scheduling parameters and cost parameter are computed, and significant differences are found between the scheduling models and its derived forms.

In conclusion, by allowing for penalties for excessive travel time and heterogeneity in scheduling preferences parameters, the model performances of both the linear and quadratic model are improved. The linear scheduling specification outperforms the quadratic one in the expected utility framework. Since the linear model is further improved by subjective probability weighting, the linear scheduling specification still outperforms the quadratic one in the rank dependent utility framework. The linear scheduling specification, including random parameters and an appropriate probability weighting, is thus recommended in studying departure time choices and value of travel time variability. Individuals behave differently when facing uncertain travel time, so the theoretical equivalence is not satisfied. For future work, it is of interest to explore the linear scheduling model on different travel time distributions while considering different weighting functions.

References

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'ecole americaine. *Econometrica*, 21(4):pp. 503–546.

Arrow, K. (1965). The theory of risk aversion. In *Essays in the Theory of Risk-Bearing*, pages 90–109. North-Holland.

Avineri, E. and Prashker, J. N. (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.

Bates, J., Polak, J., Jones, P., and Cook, A. (2001). The valuation of reliability for personal travel. *Transportation Research Part E: Logistics and Transportation Review*, 37(23):191 – 229.

Bernoulli David, . (1954). Exposition of a new theory on the measurement of risk. *Econometrica*, 22. translated from Latin into English by Dr. Louise Sommer.

Börjesson, M., Eliasson, J., and Franklin, J. P. (2012). Valuations of travel time variability in scheduling versus meanvariance models. *Transportation Research Part B: Methodological*, 46(7):855 – 873.

Diecidue, E. and Wakker, P. P. (2001). On the intuition of rank dependent utility. *Journal of Risk and Uncertainty*, 23(3):281–298.

Eliasson, J. (2006). Forecasting travel time variability. *Proceedings of the European Transport Conference, Association of European Transport.*

Eliasson, J. (2009). A cost-benefit analysis of the stockholm congestion charging system. *Transportation Research Part A: Policy and Practice*, 43(4):468–480.

Fosgerau, M. and Engelson, L. (2011). The value of travel time variance. *Transportation Research Part B: Methodological*, 45(1):1 – 8.

Fosgerau, M., Hjorth, K., Brems, C., and Fukuda, D. (2008). Travel time variability: Definition and valuation. Technical report, DTU Transport, Denmark.

Fosgerau, M. and Karlström, A. (2010). The value of reliability. *Transportation Research Part B: Methodological*, 44(1):38 – 49.

Gaver, D.P., J. (1968). Headstart strategies for combating congestion. *Transportation Science*, 2(3):172–181.

Hensher, D. A., Greene, W. H., and Li, Z. (2011). Embedding risk attitude and decision weights in non-linear logit to accommodate time variability in the value of expected travel time savings. *Transportation Research Part B: Methodological*, 45(7):954 – 972.

Hjorth, K. (2011). Cumulative prospect theory applied to stated preference data with travel time variability. Working paper,Department of Transport, Technical University of Denmark.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–91.

Koster, P. R. and Verhoef, E. T. (2010). A rank dependent scheduling model. Tinbergen institute discussion papers, Tinbergen Institute.

Lam, T. and Small, K. (2001). The value of time and reliability: measurement from a value pricing experiment. *Transportation Research Part E: Logistics and Transportation Review*, 37(2-3):231251.

Li, Z., Tirachini, A., and Hensher, D. A. (2012). Embedding risk attitudes in a scheduling model: application to the study of commuting departure time. *Transportation Science*, 46(2):170 – 188.

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

Peer, S., Koopmans, C. C., and Verhoef, E. T. (2012). Prediction of travel time variability for cost-benefit analysis. *Transportation Research Part A: Policy and Practice*, 46(1):79–90.

Polak, J. (1987). A more general model of individual departure time choice transportation planning methods. *PTRC Education and Research Services Limited*, P290(1):249–258.

Polak, J. W., Hess, S., and Liu, X. (2008). Characterising heterogeneity in attitudes to risk in expected utility models of mode and departure time choice. Presented in the 87th Transportation Research Board meeting.

Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica*, 32(1/2):pp. 122–136.

Prelec, D. (1998). The proability weighting function. *Econometrica*, 66(3):497–527.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization*, 3(4):323–343.

Rieger, M. O. and Wang, M. (2006). Cumulative prospect theory and the st. petersburg paradox. *Economic Theory*, 28(3):665–679.

Senna, L. A. D. S. (1994). The influence of travel time variability on the value of time. *Transportation*, 21(2):203–228.

Small, K. A. (1982). The scheduling of consumer activities: Work trips. *The American Economic Review*, 72(3):pp. 467–479.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323.

van Lint, J., van Zuylen Henk J., and Tu, H. (2008). Travel time unreliability on freeways: Why measures based on variance tell only half the story. *Transportation Research Part A: Policy and Practice*, 42(1):258–277.

Vickrey, W. S. (1969). Congestion theory and transport investment. *American Economic Review*, 59(2):251–60.

Vickrey, W. S. (1973). Pricing, metering and efficiently using urban transportation facilities. *Highway Research Record*, 476:36–48.

von Neumann, J. and Morgenstern, O. (1947). *Theory Of Games And Economic Behavior*. Princeton University Press.

Wakker, P. P. (2010). *Prospect theory: For risk and ambiguity.* Cambridge University Press.

Wang, Q., Sundberg, M., and Karlström, A. (2012). On the other half of the story: allowing for discrete penalties for excessive travel times in scheduling. Presented on Kuhmo Nectar 2012 and Latsis 2012.

A Linear model

A.1 Expected utility theory

$$LEU = \alpha \int_0^{\infty} T\phi(T)dT + \beta \int_0^{-D} -(D+T)\phi(T)dT + \gamma \int_{-D}^{\infty} (D+T)\phi(T)dT + \kappa \int_{\tau}^{\infty} \phi(T)dT \tag{21}$$

Take derivative of EU over departure time D

$$\begin{aligned}
\frac{\partial LEU}{\partial D} &= \beta \left[-1 * (-(D-D)) * \phi(D) + \int_0^{-D} -\phi(T)dT \right] + \gamma \left[-(-1) * (D-D) * \phi(-D) + \int_{-D}^{\infty} \phi(T)dT \right] \\
&= -\beta \int_0^{-D} -\phi(T)dT + \gamma \int_{-D}^{\infty} \phi(T)dT \\
&= -\beta \left[\Phi(-D) - \Phi(0) \right] + \gamma \left[\Phi(\infty) - \Phi(-D) \right] \\
&= -(\beta+\gamma)\Phi(-D) + \gamma
\end{aligned} \tag{22}$$

F.O.C to maximize the EU

$$LD^* = -\Phi^{-1}(\frac{\gamma}{\beta+\gamma}) \tag{23}$$

The maximal utility is

$$\begin{aligned}
LEU^* &= \alpha\mu + \beta \int_0^{\Phi^{-1}(\frac{\gamma}{\beta+\gamma})} [\Phi^{-1}(\frac{\gamma}{\beta+\gamma}) - T]\phi(T)dT + \gamma \int_{\Phi^{-1}(\frac{\gamma}{\beta+\gamma})}^{\infty} [-\Phi^{-1}(\frac{\gamma}{\beta+\gamma}) + T]\phi(T)dT + \kappa \int_{\tau}^{\infty} \phi(T)dT \\
&= \alpha\mu + \beta\Phi^{-1}(\frac{\gamma}{\beta+\gamma}) \int_{\Phi(0)}^{\frac{\gamma}{\beta+\gamma}} dF - \beta \int_{\Phi(0)}^{\frac{\gamma}{\beta+\gamma}} \Phi^{-1}(F)dF \\
&- \gamma\Phi^{-1}(\frac{\gamma}{\beta+\gamma}) \int_{\frac{\gamma}{\beta+\gamma}}^{\Phi(\infty)} dF + \gamma \int_{\frac{\gamma}{\beta+\gamma}}^{\Phi(\infty)} \Phi^{-1}(F)dF + \kappa \int_{\Phi(\tau)}^{\Phi(\infty)} dF \\
&= \alpha\mu - \beta \int_0^{\frac{\gamma}{\beta+\gamma}} \Phi^{-1}(F)dF + \gamma \int_{\frac{\gamma}{\beta+\gamma}}^{1} \Phi^{-1}(F)dF + \kappa(1-\Phi(\tau)) \\
&= (\alpha-\beta)\mu + (\beta+\gamma) \int_{\frac{\gamma}{\beta+\gamma}}^{1} \Phi^{-1}(F)dF + \kappa(1-\Phi(\tau))
\end{aligned} \tag{24}$$

A.2 Rank dependent utility theory

a rank dependent utility is evaluated

$$
\begin{aligned}
LRDU &= \alpha \int_0^{\infty} T \frac{\partial w[\Phi(T)]}{\partial \Phi(T)} \phi(T) dT + \beta \int_0^{-D} -(D+T) \frac{\partial w[\Phi(T)]}{\partial \Phi(T)} \phi(T) dT \\
&+ \gamma \int_{-D}^{\infty} (D+T) \frac{\partial w[\Phi(T)]}{\partial \Phi(T)} \phi(T) dT + \kappa \int_{\tau}^{\infty} \frac{\partial w[\Phi(T)]}{\partial \Phi(T)} \phi(T) dT
\end{aligned} \tag{25}
$$

To derive $LRDU$ regarding D

$$
\begin{aligned}
\frac{\partial LRDU}{\partial D} &= \beta \left[\int_0^{-D} - \frac{\partial w[\Phi(T)]}{\partial \Phi(T)} \phi(T) dT \right] + \gamma \left[\int_{-D}^{\infty} \frac{\partial w[\Phi(T)]}{\partial \Phi(T)} \phi(T) dT \right] \\
&= -\beta \int_{\Phi(0)}^{\Phi(-D)} \frac{\partial w[\Phi(T)]}{\partial \Phi(T)} d\Phi(T) + \gamma \int_{\Phi(-D)}^{\Phi(\infty)} \frac{\partial w[\Phi(T)]}{\partial \Phi(T)} d\Phi(T) \\
&= -\beta \int_{w[\Phi(0)]}^{w[\Phi(-D)]} dw[\Phi(T)] + \gamma \int_{w[\Phi(-D)]}^{w[\Phi(\infty)]} dw[\Phi(T)] \\
&= -\beta \left\{ w[\Phi(-D)] - 0 \right] + \gamma \left[1 - w[\Phi(-D)] \right\} \\
&= -(\beta + \gamma) w[\Phi(-D)] + \gamma
\end{aligned} \tag{26}
$$

F.O.C to maximize the utility

$$
LRD^{*} = -\Phi^{-1} \left[w^{-1} \left(\frac{\gamma}{\beta + \gamma} \right) \right] \tag{27}
$$

The maximal rank dependent utility is

$$
\begin{aligned}
LRDU^* &= \alpha \int_0^\infty T \frac{\partial w[\Phi(T)]}{\partial \Phi(T)} \phi(T) dT + \beta \int_0^{\Phi^{-1}\left[w^{-1}(\frac{\gamma}{\beta+\gamma})\right]} \left\{ \Phi^{-1} \left[w^{-1}(\frac{\gamma}{\beta+\gamma}) \right] - T \right\} \frac{\partial w[\Phi(T)]}{\partial \Phi(T)} \phi(T) dT \\
&+ \gamma \int_{\Phi^{-1}\left[w^{-1}(\frac{\gamma}{\beta+\gamma})\right]}^\infty \left\{ -\Phi^{-1} \left[w^{-1}(\frac{\gamma}{\beta+\gamma}) \right] + T \right\} \frac{\partial w[\Phi(T)]}{\partial \Phi(T)} \phi(T) dT + \kappa \int_\tau^\infty \frac{\partial w[\Phi(T)]}{\partial \Phi(T)} \phi(T) dT \\
&= \alpha \int_0^1 \Phi^{-1}(F) \frac{\partial w(F)}{\partial F} dF - \beta \int_0^{w^{-1}(\frac{\gamma}{\beta+\gamma})} \Phi^{-1}(F) \frac{\partial w(F)}{\partial F} dF \\
&+ \gamma \int_{w^{-1}(\frac{\gamma}{\beta+\gamma})}^1 \Phi^{-1}(F) \frac{\partial w(F)}{\partial F} dF + \kappa \int_{\Phi(\tau)}^1 \frac{\partial w(F)}{\partial F} dF \\
&= (\alpha - \beta) \int_0^1 \Phi^{-1}(F) \frac{\partial w(F)}{\partial F} dF + (\beta + \gamma) \int_{w^{-1}(\frac{\gamma}{\beta+\gamma})}^1 \Phi^{-1}(F) \frac{\partial w(F)}{\partial F} dF + \kappa \int_{\Phi(\tau)}^1 \frac{\partial w(F)}{\partial F} dF \qquad (28)
\end{aligned}
$$

A.3 Transformation of the linear scheduling model

If we standardized random variable X to present random travel time as $T = \mu + \sigma * X$, with CDF $s = \Phi_X(x)$

$$LD^* = -\mu - \sigma \Phi_X^{-1}(\frac{\gamma}{\beta + \gamma}) \tag{29}$$

$$\begin{aligned} LEU^* &= (\alpha - \beta)\mu + (\beta + \gamma) \int_{\frac{\gamma}{\beta+\gamma}}^{1} (\mu + \sigma \Phi_X^{-1}(s)) ds + \kappa(1 - \Phi_X(\frac{\tau - \mu}{\sigma})) \\ &= \alpha\mu + (\beta + \gamma)\sigma \int_{\frac{\gamma}{\beta+\gamma}}^{1} \Phi_X^{-1}(s) ds + \kappa(1 - \Phi_X(\frac{\tau - \mu}{\sigma})) \end{aligned} \tag{30}$$

$$LRD^* = -\mu - \sigma \Phi_X^{-1}\left[w^{-1}(\frac{\gamma}{\beta + \gamma})\right] \tag{31}$$

$$LRDU^* = \alpha \int_0^1 \Phi_X^{-1}(s) \frac{\partial w(s)}{\partial s} ds + (\beta + \gamma)\sigma \int_{w^{-1}(\frac{\gamma}{\beta+\gamma})}^{1} \Phi_X^{-1}(s) \frac{\partial w(s)}{\partial s} ds + \kappa \int_{\Phi_X(\frac{\tau-\mu}{\sigma})}^{1} \frac{\partial w(s)}{\partial s} ds \tag{32}$$

B Quadratic model

B.1 Expected utility theory

$$QEU = \eta \int_0^\infty T\phi(T)dT - \nu/2D^2 + \omega/2 \int_0^\infty (D+T)^2 \phi(T)dT + \kappa \int_\tau^\infty \phi(T)dT \tag{33}$$

Take derivative of QEU over the departure time D

$$\begin{aligned}
\frac{\partial QEU}{\partial D} &= -\nu D + \omega/2 \int_0^\infty (2D + 2T)\phi(T)dT \\
&= -\nu D + \omega D \int_0^\infty \phi(T)dT + \omega \int_0^\infty T\phi(T)dT \\
&= -\nu D + \omega D + \omega\mu \\
&= (\omega - \nu)D + \omega\mu
\end{aligned} \tag{34}$$

F.O.C to maximize QEU

$$QD^* = \frac{\omega}{\nu - \omega}\mu \tag{35}$$

The maximal utility regarding QD^*

$$\begin{aligned}
QEU^* &= \eta\mu - \nu/2(\frac{\omega}{\nu-\omega}\mu)^2 + \omega/2 \int_0^\infty \left\{ (\frac{\omega}{\nu-\omega}\mu)^2 + 2\frac{\omega}{\nu-\omega}\mu T + T^2 \right\} \phi(T)dT + \kappa[1 - \Phi(\tau)] \\
&= \eta\mu - \frac{\nu\omega^2\mu^2}{2(\nu-\omega)^2} + \frac{\omega\omega^2\mu^2}{2(\nu-\omega)^2} + \frac{\omega^2\mu^2}{\nu-\omega} + \omega/2 \int_0^\infty T^2\phi(T)dT + \kappa[1-\Phi(\tau)] \\
&= \eta\mu + \frac{\omega^2\mu^2}{2(\nu-\omega)} + \omega/2(\mu^2 + \sigma^2) + \kappa[1-\Phi(\tau)] \\
&= \eta\mu + \frac{\nu\omega\mu^2}{2(\nu-\omega)} + \omega/2\sigma^2 + \kappa[1-\Phi(\tau)]
\end{aligned} \tag{36}$$

B.2 Rank dependent utility theory

$$QRDU = \eta \int_0^\infty T \frac{\partial w[\Phi(T)]}{\partial \Phi(T)} \phi(T) dT - \nu/2 D^2 + \omega/2 \int_0^\infty (D+T)^2 \frac{\partial w[\Phi(T)]}{\partial \Phi(T)} \phi(T) dT + \kappa \int_\tau^\infty \frac{\partial w[\Phi(T)]}{\partial \Phi(T)} \phi(T) dT \quad (37)$$

Take derivative of $QRDU$ over departure time D

$$\begin{aligned}
\frac{\partial QRDU}{\partial D} &= -\nu D + \omega \int_0^\infty (D+T) \frac{\partial w[\Phi(T)]}{\partial \Phi(T)} \phi(T) dT \\
&= -\nu D + \omega D \int_0^\infty \frac{\partial w[\Phi(T)]}{\partial \Phi(T)} \phi(T) dT + \omega \int_0^\infty T \frac{\partial w[\Phi(T)]}{\partial \Phi(T)} \phi(T) dT \\
&= -\nu D + \omega D \int_{\Phi(0)}^{\Phi(\infty)} \frac{\partial w(F)}{\partial F} dF + \omega \int_{\Phi(0)}^{\Phi(\infty)} \Phi^{-1}(F) \frac{\partial w(F)}{\partial F} dF \\
&= -\nu D + \omega D \int_{w(0)}^{w(1)} dw(F) + \omega \int_0^1 \Phi^{-1}(F) \frac{\partial w(F)}{\partial F} dF \\
&= (\omega - \nu) D + \omega \int_0^1 \Phi^{-1}(F) \frac{\partial w(F)}{\partial F} dF \qquad (38)
\end{aligned}$$

F.O.C to maximize $QRDU$

$$Q\tilde{R}D^* = \frac{\omega}{\nu - \omega} \int_0^1 \Phi^{-1}(F) \frac{\partial w(F)}{\partial F} dF \qquad (39)$$

The maximal rank dependent utility regarding QRD^*

$$
\begin{aligned}
Q\tilde{RDU}^* &= \eta\int_0^1 \Phi^{-1}(F)\frac{\partial w(F)}{\partial F}dF - \frac{\nu}{2}\frac{\omega^2}{(\nu-\omega)^2}[\int_0^1 \Phi^{-1}(F)\frac{\partial w(F)}{\partial F}dF]^2 \\
&+ \omega/2\int_0^1 \left\{\frac{\omega^2}{(\nu-\omega)^2}[\int_0^1 \Phi^{-1}(F)\frac{\partial w(F)}{\partial F}dF]^2\right\}\frac{\partial w(F)}{\partial F}dF \\
&+ \omega/2\int_0^1 \left\{[\frac{2\omega}{\nu-\omega}\int_0^1 \Phi^{-1}(F)\frac{\partial w(F)}{\partial F}dF]\Phi^{-1}(F)\right\}\frac{\partial w(F)}{\partial F}dF \\
&+ \omega/2\int_0^1 [\Phi^{-1}(F)]^2\frac{\partial w(F)}{\partial F}dF + \kappa\int_{\Phi(\tau)}^1 \frac{\partial w(F)}{\partial F}dF \\
&= \eta\int_0^1 \Phi^{-1}(F)\frac{\partial w(F)}{\partial F}dF - \left\{\frac{\nu}{2}\frac{\omega^2}{(\nu-\omega)^2} - \frac{\omega}{2}\frac{\omega^2}{(\nu-\omega)^2} - \frac{\omega^2}{\nu-\omega}\right\}[\int_0^1 \Phi^{-1}(F)\frac{\partial w(F)}{\partial F}dF]^2 \\
&+ \omega/2\int_0^1 [\Phi^{-1}(F)]^2\frac{\partial w(F)}{\partial F}dF + \kappa\int_{\Phi(\tau)}^1 \frac{\partial w(F)}{\partial F}dF \\
&= \eta\int_0^1 \Phi^{-1}(F)\frac{\partial w(F)}{\partial F}dF + \frac{\omega^2}{2(\nu-\omega)}[\int_0^1 \Phi^{-1}(F)\frac{\partial w(F)}{\partial F}dF]^2 \\
&+ \omega/2\int_0^1 [\Phi^{-1}(F)]^2\frac{\partial w(F)}{\partial F}dF + \kappa\int_{\Phi(\tau)}^1 \frac{\partial w(F)}{\partial F}dF \qquad (40)
\end{aligned}
$$

C Gibbs Sampling

In the linear scheduling model, an individual would prefer to depart early (or late) if

$$\frac{\gamma}{\beta+\gamma}(>) < 1-p, \tag{41}$$

where β and γ are the individual's preference parameters and p is the probability of a delay, given the travel time distribution. The probability p is different between different alternatives, depending on the question posed in *Experiment 2*. An individual with own preference β and γ can switch between his or her behavior. I.e., the individual departure early in some cases (*Case I*), whilst departure late in some other cases (*Case II*). When allowing distribution for scheduling preference parameters, to maximize a standard simulated likelihood function becomes infeasible. Because the standard simulated likelihood is continuous but non-smooth, thus a maximum may be found but it is likely to be located at a kink of the simulated likelihood function. Consider individual i has a likelihood contribution, depending on the distribution of the two parameters and the probability P_i reflecting the individuals behavior conditional on β and γ.The likelihood function of individual i is given by

$$L_i(\theta) = \int\int P_i(\beta,\gamma)f(\beta,\gamma;\theta)d\beta d\gamma$$

where θ represents the distribution parameters, such as mean and variance. It could be approximated by simulating D draws of (β_n, γ_n) from the distribution $f(\beta,\gamma;\theta)$, as formulated by

$$L_i(\theta) \approx \frac{1}{D}\sum_{d=1}^{D} P_i(\beta_n, \gamma_n)$$

Now we illustrate how does the draws change as the mean of γ changes in Fig. 4

We postulate an artificial threshold line $\beta = \gamma$, splitting domain into *Case I* and *Case II.* The initial draws are all shifted upwards in the γ-dimension, when holding $\beta = 1.6$ and increasing the mean of γ. As shown in the figure, one draw, out of five draws, cross the threshold line we supposed. For this particular draw, the individual change behavior from *Case I* to *Case II.* Thus the utility function is continuous moving across the threshold, but its derivative is not, generating a discontinuity in the derivative of the individual's likelihood contribution. Such a kink at the optimum causes the problem in applying simulated maximum likelihood estimation.

As a remedy to this problem, we propose an estimator based on breaking up the integration of the likelihood function into domains, where none of the individual's draws cross any thresholds. Such conditional draws generate conditional likelihood contributions, constructing smooth simulated likelihood function in the parameters of interest. First of all, domain a is

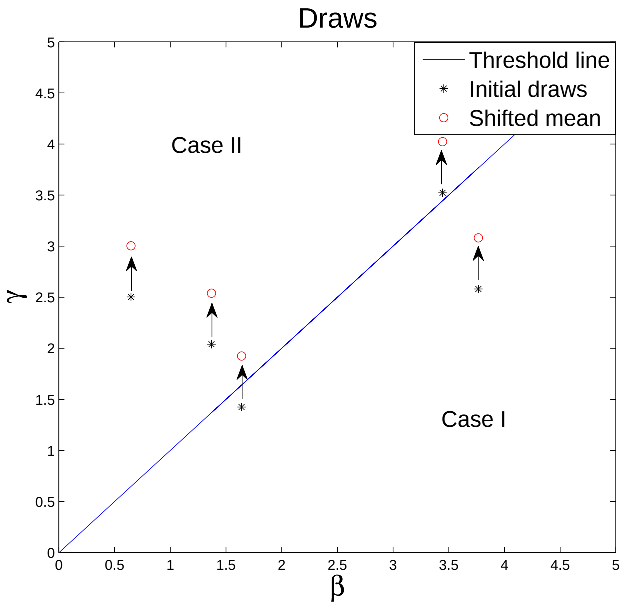

Figure 4: One critical draw, where $\beta \approx 1.6$, which crosses the threshold line when the mean of γ is increased. * denotes draws before increase in mean, and o denotes draws after the increase.

separated by the thresholds, which depends on the probability of being delay p and scheduling preference parameter $(\beta,\ \gamma)$. In the questionnaires, there are four distinct probabilities and they are ordered from the largest to the smallest, i.e. $p = (p_1 > p_2 > p_3 > p_4)$. Thus the space is partitioned into five disjunct domains as

$$a = \begin{cases} 1: & \{(\beta,\gamma)|(\beta,\gamma) \geq 0, \text{ and } \gamma < \beta(1-p_1)/p_1\} \\ 2: & \{(\beta,\gamma)|(\beta,\gamma) \geq 0, \text{ and } \beta(1-p_1)/p_1 \leq \gamma < \beta(1-p_2)/p_2\} \\ 3: & \{(\beta,\gamma)|(\beta,\gamma) \geq 0, \text{ and } \beta(1-p_2)/p_2 \leq \gamma < \beta(1-p_3)/p_3\} \\ 4: & \{(\beta,\gamma)|(\beta,\gamma) \geq 0, \text{ and } \beta(1-p_3)/p_3 \leq \gamma < \beta(1-p_4)/p_4\} \\ 5: & \{(\beta,\gamma)|(\beta,\gamma) \geq 0, \text{ and } \gamma \geq \beta(1-p_4)/p_4\} \end{cases}$$

Each question with a specific probability p makes a threshold line. Draws within each domain do not cross the threshold line, and the probability of each domain is given by

$$P_a(\theta) = \int\int_a f(\beta,\gamma;\theta)d\beta d\gamma. \tag{42}$$

Then the individual likelihood contribution is integrated over the domains

$$L_i(\theta) = \sum_a P_a(\theta) L_{i|a}(\theta), \tag{43}$$

Given the draws in domain a, The conditional likelihood function in this domain is approximated by the simulation

$$L_{i|a}(\theta) \approx \frac{1}{D} \sum_{d=1}^{D} \left(P_i(\beta_{d|a}, \gamma_{d|a}) \right) \tag{44}$$

As a result, individual will not cross between *Case I* and *Case II*. Such conditional draws of $(\beta_{d|a}, \gamma_{d|a})$ requires a technique in sampling, and *Gibbs Sampling* is used. We assume that β and γ are independently distributed with marginal distributions $logN(\mu_\beta, \sigma_\beta)$, and $logN(\mu_\gamma, \sigma_\gamma)$. For each domain, we alternate between drawing either a β or a γ, with truncation limits conditional on the proceeding draw and the thresholds. E.g, to make draws in domain $a = 1$, we start with a draw β_n from $logN(\mu_\beta, \sigma_\beta)$, and alternate to draw γ_n from $logN(\mu_\gamma, \sigma_\gamma)$, which is truncated by $\gamma < \beta(1 - p_1)/p_1$. Typical draws are displayed in Fig 5.

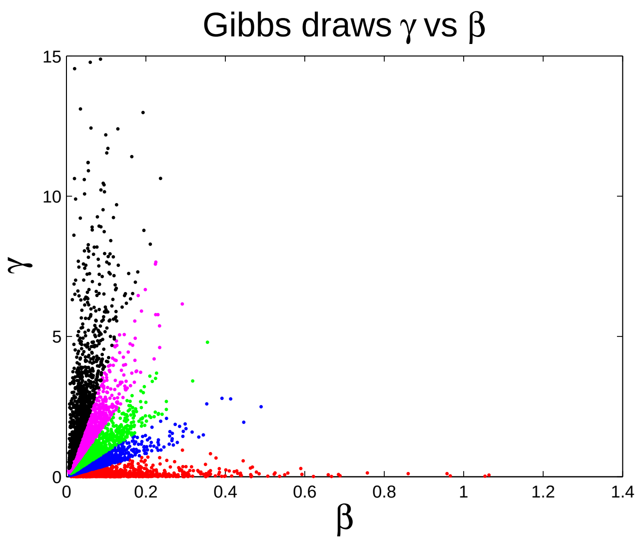

Figure 5: Draws with Gibbs Sampling within the five domains of (β, γ)