TI 2011-181/3
Tinbergen Institute Discussion Paper

Long-Run vs. Short-Run Perspectives on Consumer Scheduling: Evidence from a Revealed-Preference Experiment among Peak-Hour Road Commuters

*Stefanie Peer**
*Erik Verhoef**
Jasper Knockaert
*Paul Koster**
*Yin-Yen Tseng**

Faculty of Economics and Business Administration, VU University Amsterdam.

** Tinbergen Institute.*

LONG-RUN VS. SHORT-RUN PERSPECTIVES ON CONSUMER SCHEDULING: EVIDENCE FROM A REVEALED-PREFERENCE EXPERIMENT AMONG PEAK-HOUR ROAD COMMUTERS*

Stefanie Peer†
Erik T. Verhoef
Jasper Knockaert
Paul Koster
Yin-Yen Tseng

May 16, 2012

Abstract

Theoretical and empirical studies of consumer scheduling behavior usually ignore that consumers have more flexibility to adjust their schedule in the long run than in the short run. We are able to distinguish between long-run choices of travel routines and short-run choices of departure times due to an extensive panel dataset of commuters who participate in a real-life peak avoidance experiment. We find that the participants, who obtain a monetary reward for not traveling along a camera-observed highway link during the morning peak, value travel time higher in the long-run context compared to the short run, as changes in travel time are more permanent and can be exploited better through the adjustment of routines. Schedule delays are, in contrast, valued higher in the short-run model, reflecting that scheduling restrictions are typically more binding in the short run. Since the short-run and the long-run shadow prices differ by factors ranging from 2 to 5 in our basic model, our results may have substantial impacts on optimal choices for transport policies such as pricing and investment.

JEL codes: C25, D03, D80, R48

---

*This study is financially supported by the Dutch Ministry of Infrastructure and the Environment, as well as by the TRANSUMO foundation. Also the financial support from the ERC (Advanced Grant OPTION #246969) for the research of Paul Koster and Erik Verhoef is gratefully acknowledged. The authors would like to thank Carl Koopmans, Hans Koster and Piet Rietveld for very helpful comments.

†Corresponding author (*Adress:* Department of Spatial Economics, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands; *E-mail:* s.peer@vu.nl; *Tel.:*+31 20 59 85 232)

1 Introduction

Peak-hour traffic congestion is a major problem in most urban areas, leading to substantial economic costs. Policy measures that aim at relieving congestion have therefore received much attention. In particular, road pricing schemes have been proposed, leading to the implementation of pricing both along corridors (e.g. high occupancy toll (HOT) lanes at various locations in the US) as well as within specific areas (e.g. London, Singapore, Stockholm). The expansion of infrastructure capacity is another, more widely applied policy.

Peak-hour congestion usually arises when a large number of people share the preference to be at the same place at the same time. In such cases, individuals face a trade-off in their departure time decisions. If they depart at the begin or at the end of the peak, they will encounter modest or no congestion; however, they will arrive at their destination at a time quite different from their preferred arrival time. The longest travel times will be faced by those who choose a departure time that results in an arrival time close to their preferred arrival time. Vickrey (1969) was the first to account for such within-the-day dynamics, using a dynamic equilibrium model of queuing at a bottleneck. Small (1982), building on the works of Becker (1965) and DeSerpa (1971) on the allocation of time in non-work activities, allowed for disutility of arriving early or late at work in the utility function of commuters. He also provided empirical estimates of the values attached to these so-called schedule delays. His work gave rise to numerous studies that aim at deriving monetary valuations of travel time savings and schedule delays, and more recently, also of travel time reliability (e.g. Noland and Small, 1995). Most of them use disaggregate data on departure time choices collected from stated preference (SP) or revealed preference (RP) experiments, or combine data from both sources (e.g. Small et al., 2005). Random-utility discrete choice models are then usually used to estimate the model parameters (McFadden, 1974).

Even though the estimated values of travel time savings (henceforth, the 'value of time') and schedule delays vary widely across studies, they are usually found to be substantial. The value of time usually ranges between 20 and 90% of the gross wage rate, while the costs of arriving one minute early tend to be lower, and the costs of arriving one minute late tend to be higher than the costs of an additional minute of traveling (see Small and Verhoef (2007) and Li et al. (2010) for overviews). These time-related costs thus compose a large share of generalized travel costs, and are therefore key ingredients in the evaluation of policy options. To identify the appropriate values of time and schedule delays, and to use them for policy evaluation, a good understanding of the underlying scheduling behavior of travelers is crucial.

In this paper, we introduce a distinction between short-run and long-run scheduling decisions. The distinction reflects that commuters typically have more flexibility in adapting their daily schedules in the longer run than in the shorter run. We expect this to be reflected by different valuations of times and schedule delays in the long run compared to the short run. We propose a framework that decomposes scheduling decisions into long-run choices of what we will call 'routines', and short-run choices of departure times. Routines are considered fixed in the short run, but are endogenous in the long run. We also distinguish between long-run and short-run travel time expectations, taking into consideration that more information on possible travel time realizations becomes available in the short run. This distinction is an important one to make, not in the least place for identification purposes. Most current models implicitly assume that successive days are exact replicas, which implies that the long-run choice of routines, and the short-run choices of departure time are identical, and could not be meaningfully disentangled. The same is in fact true for stochastic models when these assume that travel time expectations do not vary between days (hence, when ignoring learning or information-updating (e.g. Arnott et al., 1999)). In this case, departure time choices will not be adjusted between days, and again, long-run and short-run behavior coincide. But when travel conditions vary between days, we can distinguish between the two concepts, assuming that long-run choices are based on expectations of travel times over the day, whereas short-run choices also consider deviations from these expectations.

We assume that in the long run, drivers choose a preferred arrival time by trading off the disutility of expected travel time losses against deviations of this short-run preferred arrival time from an exogenously determined long-run preferred arrival time. In the short run, i.e. on the day of departure, drivers have a more accurate forecast of travel times, and in the light of that may choose an actual arrival time that differs from the short-run preferred arrival time.[1]

The data used in this study are drawn from a large-scale, revealed-preference peak avoidance experiment among commuters in the Netherlands. Over a period of several months, participants were able to obtain a daily reward of 4 Euro if they avoided traveling on a specific, frequently congested highway link during morning peak hours. Only few comparable RP experiments that also account for departure time choice behavior have been conducted (e.g Small, 1982; Lam and Small, 2001; Börjesson, 2008), mainly because they require a setup with a monetary attribute, a toll or a reward, that varies over the time of the day. Unlike the datasets used in most of the previous studies, the current dataset has a panel structure, allowing for the possibility to observe commuters over time and to identify both the long-run and short-run dimensions of their scheduling behavior. Moreover, detailed data on observed departure times and travel times are available. We make use of these data to construct long-run travel time expectations, as well as day-specific travel time expectations for the short-run model. This allows us to largely avoid the use of reported data, which are frequently imprecise and often even biased (e.g Small and Verhoef, 2007, p.21). A distinction between long-run and short-run scheduling behavior would probably be much more difficult to achieve using SP data, as the choice experiment would have to be phrased such that participants follow the intention of the researcher in interpreting the choice situations in a long-run sense or in a short-run sense, respectively. A more generic drawback of SP data is that they may lead to hypothetical biases in the estimates. These

[1]Certainly, also alternative definitions of the 'short run' and the 'long run' are feasible. For instance, it is possible to endogenize the structural arrival time preferences and let them depend on job choices as well as residential and work locations choices. Some studies that focus on such choices also determine valuations of travel time (e.g Van Ommeren et al., 2000; Van Ommeren and Fosgerau, 2009). However, the resulting values are usually not directly comparable to the values derived in this paper, not in the last place because they usually do not distinguish between time and scheduling-related costs.

arise if individuals react differently in a laboratory environment than in real life, which has been found to be true also for travel choice behavior (Hensher, 2010).

We use discrete choice models to analyze the participants' choices of routines and departure times. We find that drivers value more permanent (long-run) travel time gains substantially higher than short-run gains, with a factor three to six, presumably as these can be exploited better through the adjustment of routines. After all, an incidental minute of time gain can typically be used less effectively than a structural gain of one minute. This result is consistent with the results obtained by Tseng et al. (2011), who find that the value attached to a permanent one-hour travel time gain is 14.5 Euro/hour, compared to 3 Euro/hour for an incidental travel time gain of one hour. On the other hand, we find a substantially higher value attached to less permanent (short-run) changes of schedule delays, with the difference mounting to an order of two to three. This finding is consistent with the notion that scheduling restrictions are normally more binding in the short run. Börjesson (2009) and Börjesson et al. (2012) speculate along similar lines to explain the differences they find between the valuation of planned and unplanned delays in models that are based on SP data and do not distinguish between short-run and long-run values of time and schedule delays.

Our results strongly suggest that the distinction between short-run and long-run modeling of consumer behavior in the study of valuation of time and scheduling is a very important one to make. It also calls for a reconsideration of equilibrium models of travel behavior, for example as used to assess the social desirability of infrastructure investments or road pricing. Moreover, we believe that the large dispersion of the values of time and schedule delays that is found in the literature may partly be explained by differences in the relative prominence of long-run and short-run definitions of the model attributes.

The outline of the paper is as follows. In Section 2 we introduce the modeling framework, in particular the distinction between the long-run and the short-run model. Section 3 gives an overview of the data, including the experimental setup and variable definitions. The estimation results as well as several robustness checks are presented in Section 4. Section 5 concludes the paper.

2 Modeling Framework

2.1 Main models

In the conventional scheduling model, introduced by Small (1982), the utility of departing at a specific time t, V_t, is a linear function of four attributes. It depends on the monetary attribute, a toll or reward R_t, travel time T_t, and the extent of earliness or lateness with respect to the preferred arrival time, PAT. The last two attributes are generally referred to as schedule delay early, SDE_t, and schedule delay late, SDL_t. The subscript t signifies dependence on the moment of departure, t. More recently, the conventional scheduling model was extended to account for stochastic travel times (e.g. Noland and Small, 1995; Bates et al., 2001; Fosgerau and Karlström, 2010). In the formulation that resembles most closely the deterministic case[2], $R_t,\ T_t,\ \text{SDE}_t,$ and SDL_t represent distributions which arise from day-to-day variations in travel times. The attributes of the utility function are then defined in expected terms, denoted by the operator $E[.]$. In a conventional linear model we may denote the coefficients attached to these attributes by $\beta_R,\ \beta_T,\ \beta_E$ and β_L, respectively:

$$\begin{aligned} E[V_t] &= \beta_R\ E[R_t] + \beta_T\ E[T_t] + \beta_E\ E[\text{SDE}_t] + \beta_L\ E[\text{SDL}_t] \\ \text{where } E[\text{SDE}_t] &= E[\max[\text{PAT} - t - T_t,\ 0]] \\ E[\text{SDL}_t] &= E[\max[t + T_t - \text{PAT},\ 0]] \end{aligned} \tag{1}$$

If R_t is defined as a reward, the term -1/β_R can be interpreted as the marginal utility of income.[3] The value of time (VOT) and the values of schedule delay early (VSDE) and late (VSDL) can then be expressed as follows:

$$\text{VOT} = -\beta_T/\beta_R \qquad \text{VSDE} = -\beta_E/\beta_R \qquad \text{VSDL} = -\beta_L/\beta_R \tag{2}$$

In this paper, we adapt the standard scheduling model such that it allows for the possibility to study differences between long-run and short-run scheduling behavior. There is reason to expect that these are inherently different in at least two senses. First, in the long run, drivers are able to adapt their routines. Specifically, in a world without traffic congestion, a driver's preferred arrival time may for instance be 9:00 in the morning, whereas recurring congestion may induce him to establish a routine with a usual arrival time earlier than 9:00. As a consequence, he may make appointments at work earlier, and given those appointments, the actually preferred arrival time also shifts to an earlier time of the day. That is, if the preferred arrival time under uncongested conditions is given as in the above example, but the driver has organized routines such that he expects to arrive around 8:00 and therefore plans a meeting at 8:15, a delayed arrival at 8:30 brings schedule delay late, not early.

Second, also the availability of information on travel times is likely to differ between the long and the short run. In the long run, only average traffic conditions are known, whereas in the short run, more information on actual travel time realizations becomes available. We expect these differences between long-run and short-run scheduling decisions to be reflected also in the valuation of travel time and schedule delays. In particular, we expect that travel time gains are valued higher in the long run, as these can be exploited better through the re-scheduling of routines. Schedule delays, on the other hand, are likely to be valued higher in the short run, as scheduling restrictions tend to be more binding in the short run.

We will thus distinguish between a long-run and a short-run scheduling model. In the long-run, arrival routines are chosen subject to structural arrival time preferences, whereas in the short run, departure times are chosen subject to the arrival routines chosen in the long-run model. In the terminology that we will use, the structural arrival time preference

[2]Sometimes the probability of being late is added to the utility function, allowing for an extra penalty for lateness. Also measures of dispersion are occasionally added, with the underlying idea that they capture disutility from travel time variability that is not directly related to scheduling, such as stress or anxiety.

[3]In case of R_t being defined as a toll, the marginal utility of income equals $1/\beta_R$.

is referred to as long-run preferred arrival time (LRPAT), and the preferred arrival time given the chosen routines are referred to as the short-run preferred arrival time (SRPAT). Schedule delays in the long-run model are therefore defined as deviations from the LRPAT, while schedule delays in the short-run model are defined as deviations from the SRPAT. The distinction between LRPAT and SRPAT is new to the relevant literature. We are not aware of any study that examines the sensitivity of the coefficients of the scheduling model towards these different PAT definitions, despite their intuitive interpretations. Most studies adopt preferred arrival time definitions that comply with the idea of the LRPAT. The preferred arrival time is often defined in the context of hypothetical behavior in uncongested conditions (e.g. Börjesson, 2009), or as the official work start time (e.g. Noland and Small, 1995; Small, 1982). Only few studies define the preferred arrival time as based on usual behavior, corresponding to the definition of the SRPAT. One such example is Tseng et al. (2011). However, they define a preferred departure rather than arrival time.

The long-run model, which models the choice of arrival routines, differs from the conventional scheduling model with stochastic travel times (Equation 1) in various respects. First, long-run travel time expectations denoted by the superscript LR are used. Second, rather than actual departure times t, the long-run model explains the choice of arrival time a as short-run preferred arrival time, SRPAT. As a consequence, the schedule delays, and hence the deviations from the LRPAT, become deterministic. In order to account for the effect of costs of travel time variability on the choice of the SRPAT, we add a term $C^{LR}[\sigma_a]$ to the equation, which is a function of the standard deviation of those travel times that result in an arrival time at time a, σ_a. This formulation thus uses the standard deviation of travel times for an arrival time exactly at a, ignoring that travelers may adjust actual arrival moments through their short-run decisions.[4] The indirect utility for SRPAT$= a$ can thus be written as:

$$E[V_a]^{LR} = \beta_R\ E[R_a]^{LR} + \beta_T\ E[T_a]^{LR} + \beta_E\ \mathrm{SDE}_a + \beta_L\ \mathrm{SDL}_a + C^{LR}[\sigma_a], \qquad (3)$$
$$\text{where } \mathrm{SDE}_a = \max[\mathrm{LRPAT} - a,\ 0]$$
$$\mathrm{SDL}_a = \max[a - \mathrm{LRPAT},\ 0]$$

The optimal arrival routine, the SRPAT, is therefore a deliberate choice, which maximizes the long-run utility function:

$$\mathrm{SRPAT} = \arg\max_a\ E[V_a]^{LR} \qquad (4)$$

The short-run utility function closely resembles the conventional scheduling model that accounts for stochastic travel times (Equation 1). Arrival routines are fixed in the short run. The SRPAT thus enters the indirect short-run utility function as the relevant anchor point for defining the schedule delays. Moreover, we use short-run travel time expectations denoted by the superscript SR, which account for the availability of more up-to-date information on

[4]Given the correlation between standard deviations at close arrival moments, this may be considered a mild simplification. Moreover, we will not be estimating $C^{LR}[\sigma_a]$ but instead will impute it, for reasons to be outlined shortly.

travel time realizations when the departure moment has to be chosen, than what applies in the long-run model. Note that the short-run model assumes that disutility of travel time variability is entirely captured by its impact on expected schedule delay costs.

$$E[V_t]^{SR} = \beta_R\ E[R_t]^{SR} + \beta_T\ E[T_t]^{SR} + \beta_E\ E[\mathrm{SDE}_t]^{SR} + \beta_L\ E[\mathrm{SDL}_t]^{SR}, \tag{5}$$
$$\text{where } E[\mathrm{SDE}_t] = E[\max[\mathrm{SRPAT} - t - T_t,\ 0]]$$
$$E[\mathrm{SDL}_t] = E[\max[t + T_t - \mathrm{SRPAT},\ 0]]$$

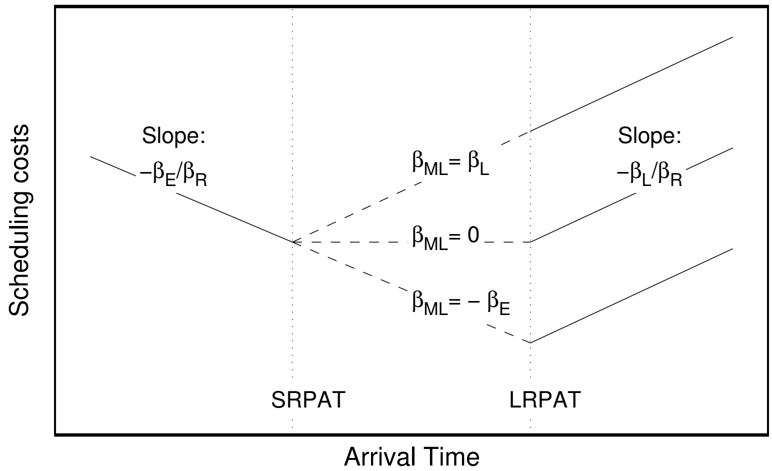
Besides the basic long-run and the short-run models of Equations 3 and 5, we introduce an alternative short-run model, which is used exclusively to test whether it is true that the SRPAT is indeed the relevant anchor point in departure time decisions, as we assumed in Equation 5. If this is true, a move away from the SRPAT towards the LRPAT would increase schedule delay costs. Following our earlier example, an arrival time at 8:30, given a SRPAT at 8:00 and an LRPAT at 9:00, would yield costs of schedule delay late rather than early. We therefore re-formulate the short-run model, changing the definition of the schedule delays such that we can determine whether scheduling costs are at their minimum at the SRPAT or at the LRPAT. To do so, we introduce the adapted schedule delays SDAE_t and SDAL_t, which capture arrival moments that are early or late with respect to both measures of SRPAT and LRPAT. In addition, we define the intermediate domains SDME_t and SDML_t. SDME_t captures arrival moments that are early with respect to the SRPAT but late with respect to the LRPAT, when LRPAT $<$ SRPAT. Likewise, SDML_t captures arrival moments that are late with respect to the SRPAT but early with respect to the LRPAT, when SRPAT $\leq$ LRPAT. The corresponding utility in this "three-domains" model is denoted by $V_t^{\text{SR-3D}}$, and the travel time expectations are defined in the same way as in the short-run model:

$$V_t^{\text{SR-3D}} = \beta_R\ E[R_t]^{SR} + \beta_T\ E[T_t]^{SR} + \beta_E\ E[\mathrm{SDAE}_t]^{SR} + \beta_L\ E[\mathrm{SDAL}_t]^{SR} + \tag{6}$$
$$\beta_{ME}\ E[\mathrm{SDME}_t]^{SR} + \beta_{ML}\ E[\mathrm{SDML}_t]^{SR},$$
$$\text{where}$$

	LRPAT $<$ SRPAT	SRPAT $\leq$ LRPAT
$E[\mathrm{SDAE}_t] =$	$E[\max[\mathrm{LRPAT} - t - T_t,\ 0]]$	$E[\max[\mathrm{SRPAT} - t - T_t,\ 0]]$
$E[\mathrm{SDAL}_t] =$	$E[\max[t + T_t - \mathrm{SRPAT},\ 0]]$	$E[\max[t + T_t - \mathrm{LRPAT},\ 0]]$
$E[\mathrm{SDME}_t] =$	$E[\min[\mathrm{SRPAT} - \mathrm{LRPAT},$ $\max[\mathrm{SRPAT} - t - T_t,\ 0]]]$	0
$E[\mathrm{SDML}_t] =$	0	$E[\min[\mathrm{LRPAT} - \mathrm{SRPAT},$ $\max[t + T_t - \mathrm{SRPAT},\ 0]]]$

Consider the case when SRPAT $\leq$ LRPAT, depicted in Figure I. If β_{ML} would be equal to β_L, the model would suggest that the SRPAT is the relevant anchor to describe short-run behavior, and the LRPAT has no influence whatsoever in the short-run model. In contrast, when $\beta_{ML} = -\beta_E$, the model suggests that the LRPAT is the relevant anchor point, and the SRPAT has no impact on short-run choices. We may expect β_{ML} to be somewhere between these polar cases, and would take a significant negative sign as an indicator of the desirability

Figure I: Diagrammatic exposition: SRPAT $\leq$ LRPAT



of an arrival at the SRPAT over an arrival at the LRPAT. A similar argument can be made for the case where LRPAT $<$ SRPAT. If both β_{ML} and β_{ME} are negative, we thus have the confirmation that deviations from the SRPAT towards the LRPAT indeed increase scheduling costs. The three domains model, therefore, helps identifying the degree to which the SRPAT rather than the LRPAT indeed determines short-run scheduling behavior.

2.2 Empirical Model Specifications

The models discussed in Section 2 can be estimated by employing the additive random-utility model developed by McFadden (1974). It assumes that choices between discrete alternatives are made such that the utility of the decision maker is maximized. Choices are probabilistic as they are not only affected by observed attributes, which are captured in the systematic part of the utility function, but also by unobserved attributes, which are captured by a random term. We will apply this method for the estimation of the long-run as well as the short-run models.

We allow arrival routines to be weekday-specific. Therefore, in the long-run model, each driver $z = 1, \ldots, Z$ chooses between $j = 1, \ldots, J$ discrete arrival routines for each weekday $l = 1, \ldots, 5$. In the short-run model the same drivers choose between $j = 1, \ldots, J$ discrete departure time alternatives on days $k = 1, \ldots, K$. The systematic part of the long-run utility function for arrival routine a, $E[V_a]^{LR}$, is therefore re-written as $E[V_{zlj}]^{LR}$. Similarly, the systematic part of the short-run utility function for departure time t, $E[V_t]^{SR}$, is re-written as $E[V_{zkj}]^{SR}$.[5] The corresponding random terms are denoted by ϵ_{zlj} and ϵ_{zkj}, and the overall utilities, U_{zlj}^{LR} and U_{zkj}^{SR} are thus given by:

$$U_{zlj}^{LR} = E[V_{zlj}]^{LR} + \epsilon_{zlj} \qquad \text{and} \qquad U_{zkj}^{SR} = E[V_{zkj}]^{SR} + \epsilon_{zkj} \tag{7}$$

[5] Clearly, a similar notational specification should be applied to the $SR - 3D$-model.

We apply the most basic discrete-choice specification and assume that the random terms are iid with the extreme-value distribution. McFadden (1974) showed that then the choice probabilities for alternative i, P^{LR}_{zli} and P^{SR}_{zki}, respectively, have the logit form:

$$P^{LR}_{zli} = \frac{\exp(E[V_{zli}]^{LR})}{\sum_{j=1}^{J} \exp(E[V_{zlj}]^{LR})} \qquad \text{and} \qquad P^{SR}_{zki} = \frac{\exp(E[V_{zki}]^{SR})}{\sum_{j=1}^{J} \exp(E[V_{zkj}]^{SR})} \tag{8}$$

From this assumption, the multinomial logit (MNL) model arises. The model parameters are then estimated by maximizing the corresponding log-likelihood function, which is defined as the sum of the logs of the choice probabilities across observations. Due to the fact that arrival routines are weekday-specific, and departure time decisions are made on multiple days, multiple observations per driver are included in the long-run as well as short-run datasets. The MNL model ignores this panel nature of the datasets. This assumption of independent observations is probably less drastic for RP data compared to SP data, as the valuation attached to travel time and schedule delays may differ across days also for a given driver. Still, in order to correct for a possible under-estimation of the standard errors, the panel specification of the sandwich estimator is used (e.g Menard, 2009). Moreover, we will later (in Section 4.2) test whether the results still hold if the models allow for heterogeneity across drivers. For this purpose, we re-estimate the main models using panel latent-class models. These models assume that drivers can be sorted into a set of Q classes, with coefficient estimates being class specific. It is unknown to the analyst to which class a particular driver belongs to. The estimation procedure differs from the estimation of MNL models mainly by the fact that not only choice probabilities for the alternatives $j = 1, \ldots, J$ need to be determined, but also the probabilities of being member of class $q = 1, \ldots, Q$ (see McCutcheon (1987) for a general discussion on latent class analysis and Greene and Hensher (2003) for an application of latent class analysis to panel data and discrete choice problems).

2.3 Travel time expectations

Like the arrival routines, also travel time expectations are defined as weekday-specific. In fact, differences in travel time expectations for different weekdays may be an important reason to also differentiate routines over weekdays. We then define the long-run expectation for attribute $A = \{R,\ T,\ \text{SDE},\ \text{SDL}\}$ on weekday l as the average attribute value on that weekday over a time period $k = 1, \ldots, K$. It is therefore implicitly assumed that the driver has information about future travel times when choosing the optimal routines. While it is clearly impossible to know travel time realizations on single days (in particular in the long run), the average weekday-specific travel time seems to be a good representation of the knowledge a regular commuters may have about average future travel time realizations. We thus define an indicator function $\mathbf{1}_{lk}$, which is equal to 1 if it is true that weekday(k)=l, and 0 otherwise. The long-run expectation for attribute A, $E[A_{zlj}]^{LR}$, is then given by

$$E[A_{zlj}]^{LR} = \sum_{k=1}^{K} \mathbf{1}_{lk} A_{zlj} / \sum_{k=1}^{K} \mathbf{1}_{lk}. \tag{9}$$

In the short run, which means at the moment that the actual departure time decision is taken, more up-to-date information on travel time realizations is available compared to the long run, for instance due to a better knowledge of expected weather conditions. We represent this enhanced information availability by computing a weighted average of the long-run expectation of attribute A and the according attribute value based on the actual travel time realization on the day of travel $\breve{k}$. The weight attached to the actual travel time realizations, relative to the long-run travel time expectations, is denoted by θ and will be estimated.[6] Although it is unrealistic that these actual travel times are known to drivers at the moment of departure, they represent a benchmark for the maximum extent of information they may have at that moment. Again, we define an indicator function $\mathbf{1}_{\breve{k}k}$, which is equal to 1 if it is true that weekday(k)=weekday($\breve{k}$), and 0 otherwise.

$$E[A_{z\breve{k}j}]^{SR} = \theta \; A_{z\breve{k}j} + (1-\theta) \; \sum_{k=1}^{K} \mathbf{1}_{\breve{k}k} A_{zkj} / \sum_{k=1}^{K} \mathbf{1}_{\breve{k}k} \tag{10}$$

3 Data[7]

3.1 Experimental Setting

We use RP data that were gathered during a large-scale "Peak Avoidance" ("Spitsmijden" in Dutch) experiment in the Netherlands. In this experiment, participants were eligible for a monetary reward of 4 Euro per day if they avoided traveling on a specific highway link[8] during the morning peak (6:30–9:30 a.m.). The highway link has a length of 9.21 km and is frequently congested during morning peak hours. We refer to the link as 'C1–C2' as it is defined as the road segment between two cameras 'C'. Rewards could not be earned on weekends or school vacation days, and also not if a driver had already exceeded the maximum number of days per 2-week period for which he could obtain a reward. This maximum is driver-specific, and is based on a driver's reference behavior, which is defined as the average number of trips he had undertaken along the C1–C2 link per 2-week period before the start of the experiment.

The peak avoidance experiment lasted for more than a year, from November 2008 until December 2009. Participants could join as well as leave the experiment during this period. Overall, about 5000 commuters participated. More than 15000 commuters had been invited to participate in the experiment after they were observed passing the C1–C2 link, resulting

[6] We also tested methods (i.e. state-space and semi-parametric estimations) that allowed for the expected travel time to depend on day-specific variables that may affect travel times, such as weather conditions. However, we found that the difference between realized and expected travel times barely decreased as a consequence of adding these variables. This can probably be attributed to the fact that only 75 working days are taken into account here, and that weather conditions were quite stable over that time period.

[7] Additional details on the data used in this study can be found in the web report: `http://personal.vu.nl/e.t.verhoef/datamanual.pdf`

[8] Highway A12, between Gouda and Zoetermeer with driving direction towards The Hague.

in almost 3000 participants. Another 2000 participants were recruited through lease car companies and from an earlier peak avoidance experiment. The selection of the participants is therefore voluntary and not random. From a survey among drivers who were invited to participate but chose not to participate in the experiment, it was found that participants tend to be older, have a higher education and income, and (not surprisingly) are more flexible in choosing their arrival time at work (Knockaert et al., 2012). Moreover, their reference travel behavior differs. On average, participants travel more frequently along the C1–C2 link, and their passage times of this link are less clustered in the middle of the peak (7:30-8:30 a.m.). We believe that selection effects will affect the observed sensitivity to rewards, but are unlikely to have a systematic impact on the distinction between short-run and long-run decision making that we are interested in here.

Travel times and passage times of the participants have been directly measured along the C1–C2 link using cameras capable of number plate detection. For our analysis, we would like to use door-to-door measures of travel time instead. Peer et al. (2011) develop a method to approximate driver, day and time-of-day-specific door-to-door travel times as well as departure times from home and arrival times at work, combining speed measurements from loop detectors along the C1–C2 link and GPS-based speed measurements on the home–C1 and C2–work links. They draw on the common finding that speeds are correlated positively across links (for a given day and time of the day). By using geographically weighted regression (GWR), they take into account that this correlation pattern varies over space and is therefore dependent on the home and work location of a driver. It is the travel time calculations from that model that we will be using in the present analysis.

In the analysis performed in this paper, we focus on drivers who participated in the experiment between September 2009 and December 2009 (K =75 working days), and who were observed to use the C1–C2 link at least occasionally during that time period. From these participants, only those are selected who filled in a survey that included questions about their long-run arrival time preference, and for whom we are able to approximate door-to-door travel times given the limited coverage of the GPS-base GWR model of Peer et al. (2011). These restrictions leave use with $Z = 371$ drivers. Besides reasons of data availability and consistency, another reason for restricting the analysis to a 3-month period is to increase the likelihood that drivers have fixed arrival routines during that period. Moreover, the beginning of the 3-month period coincides with the end of the summer vacation for many Dutch people as well as the start of a new school year, both of which can be considered natural moments for adjusting one's morning commute routine.

3.2 Operationalization of the LRPAT and SRPAT

For each driver, measures of the LRPAT and the SRPAT need to be determined. We derive the LRPAT from reported data and the SRPAT from actual behavior. By nature of its definition, the LRPAT refers to a guaranteed congestion-free situation, and can therefore not be derived from actual behavior. Instead, the LRPAT is obtained from a questionnaire conducted among the participants of the experiment. They were asked to state their preferred arrival time at work if they knew for sure that they would not face any congestion during their commuting trip[9].

We consider the SRPAT to be weekday-specific, as travel times differ substantially across weekdays, and drivers are likely to adapt their arrival routines accordingly (Figure IIa). The SRPAT is defined as the driver and weekday-specific median arrival time at work. For each driver only those weekdays are considered for which at least three passages of the C1–C2 link were observed. Moreover, only those days are taken into account during which a driver was eligible for a reward, to be sure that the monetary incentive indeed applies. If the number of observations per driver and weekday is an even number, we randomly assign one of the two middle values as median. The intuition behind this procedure is that we want to be able to identify those departure time decisions for which the actual arrival time by definition coincides with the corresponding SRPAT. These choices will be excluded from the short-run model, as they are fully endogenous.

Figure IIb shows a scatterplot for all drivers and weekdays for which valid measures of the SRPAT and the LRPAT are available. As expected given that long travel times are unattractive, we find that the distribution of the SRPAT is relatively more dispersed in time than the LRPAT. The figure also shows that drivers with a relatively early LRPAT (the median LRPAT is indicated in the figure) tend to choose a SRPAT that is even earlier than their LRPAT, while drivers with a relatively late LRPAT tend to choose an even later SRPAT. This is confirmed by the histograms of the LRPAT and SRPAT distributions, shown in Figures IIc and IId, respectively. We find that the density of the LRPAT distribution is highest between 8:00 and 9:00, roughly corresponding to the typical work starting times. The distribution of the SRPAT is much flatter. Since rewards can be gained for passing C1–C2 before 6:30, a very pronounced peak can be observed for arrival times at work between 6:30 and 7:00.[10] A relatively small number of participants has a SRPAT that gives them the opportunity to receive a reward after the end of the peak (9:30). These are strong indications that the choice of the SRPAT is indeed the result of a trade-off between deviations from the LRPAT, average congestion patterns, as well as the distribution of the monetary incentive over time of the day.

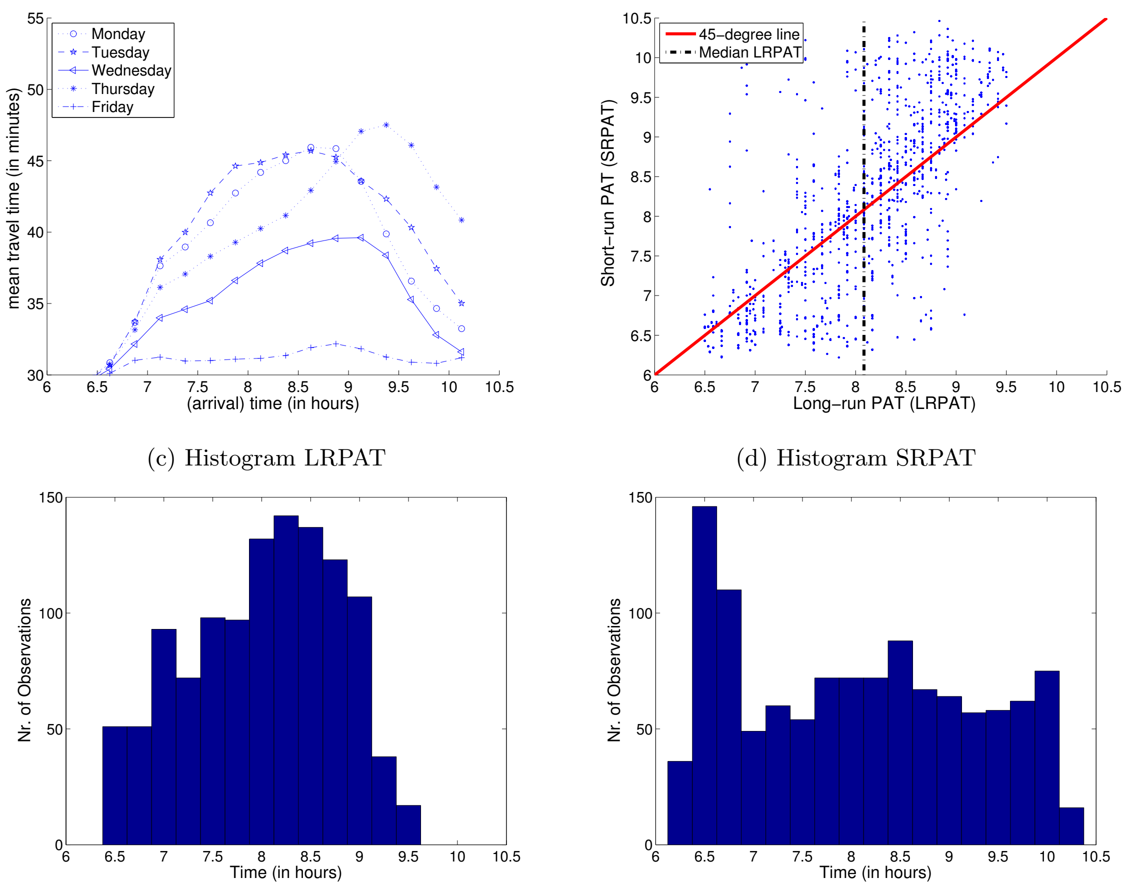
3.3 Selection of Observations and Choice Set Definitions

For both the long-run and the short-run models, the choice set consists of $J = 16$ discrete choice alternatives. Each alternative corresponds to a 15-minute interval. In the long-run model, a driver is able to choose between arrival routines that result in an arrival time at work between 6:15 to 10:00. In the short-run model, the choice set is driver-specific. This is to ensure that despite of differences in home–work distances only the most relevant departure time alternatives are included for each driver, in particular those that yield a trade-off with

[9] Drivers with an LRPAT earlier than 6:30 or later than 9:30 were removed from the dataset as these drivers do not have an incentive to travel during the peak period, so that it is more natural to assume that the scheduling model applies to them.

[10] In one of the sensitivity analyses to be presented below, we will weight observations such to correct for possible biases that may result from this peak.

Figure II: Descriptives LRPAT and SRPAT



(a) Average weekday-specific travel times

(b) Scatter LRPAT vs. SRPAT

(c) Histogram LRPAT

(d) Histogram SRPAT

respect to the reward. Thus, for each driver the choice set contains 12 departure time alternatives that in expected terms result in a passage time of the C1–C2 link during the peak, and therefore do not yield a reward, and 4 alternatives that result, again in expected terms, in an off-peak passage time of the C1–C2 link (2 alternatives before, and 2 after the peak).

Table I provides some descriptive statistics on the datasets used for the estimation of the long-run and the short-run models. The number of observations in the long-run model is determined by the number of driver-weekday combinations for which valid measures of the SRPAT are available. On average, the number of weekdays for which a SRPAT can be specified, is about three. The short-run dataset then consists of all observations that were used in determining the SRPAT, except for the ones for which the corresponding arrival times have been identified as SRPAT (resulting in an exclusion of 1306 observations). We find that the remaining number of departure time choices per driver is on average about

16. Finally, Table I also shows the differences between the actual arrival times that result from the departure time decisions and the SRPAT. Table I shows that more than 60% of the departure time choices result in an arrival time that is less than 15 minutes early or late with respect to the SRPAT. Hence, drivers tend to choose their departure times such that they arrive close to their SRPAT.

Table I: Descriptives datasets

Variables	Value	St. Dev.
General		
Nr. of days (K)	75	–
Nr. of drivers (Z)	371	–
Nr. of choice alternatives (J)	16	–
Long-run dataset		
Nr. of SRPAT choices	1158	–
Avg. nr. of SRPAT choices per driver	3.12	1.42
Short-run dataset		
Total nr. of departure time choices	7271	–
Nr. of excluded choices due to coincidence of SRPAT and arrival time	1306	–
Remaining nr. of departure time choices	5965	–
Avg. remaining nr. of departure time choices per driver	16.08	11.21
Arrival time deviations from the SRPAT		
more than 30 minutes early	7.66%	–
between 15 and 30 minutes early	11.03%	–
between 15 minutes early and 15 minutes late	62.35%	–
between 15 and 30 minutes late	9.64%	–
more than 30 minutes late	9.32%	–

4 Estimation Results

4.1 Main Models

Table II shows the results obtained for the long-run and the short-run model, as well as for the 3-domains model that is used to test whether actual departure time decisions are indeed driven by the SRPAT rather than the LRPAT.

For all models, credible point estimates are obtained. The VOT is between 5.20 and 30.16 Euro/hour, the VSDE between 9.34 and 23.16 Euro/hour and the VSDL between 7.22 and 20.22 Euro/hour.[11] Significant differences between the long-run and the short-run

[11] It is usually found that the VSDL is higher than the VSDE. The reason why we obtain the opposite result is most likely due to the fact that a large number of participants in this study have a rather early SRPAT. These participants may be quite reluctant to switch to an even earlier departure time, as the utility

Table II: Main estimation results

	Long-Run		Short-Run		Short-run: 3 domains	
Coefficient	Value	t-Statistic	Value	t-Statistic	Value	t-Statistic
β_R	0.22	4.87	0.13	5.78	0.17	6.75
β_T	-6.56	-7.31	-0.69	-1.45	-1.75	-3.09
β_E	-2.03	-13.28	-2.89	-18.38	-3.83	-14.52
β_L	-1.57	-13.90	-2.70	-20.34	-2.53	-20.80
β_{ME}	–	–	–	–	-1.45	-8.74
β_{ML}	–	–	–	–	-3.23	-9.88
θ	–	–	0.43	6.25	0.36	5.49
VOT (Euro/h)	30.16	4.14	5.20	1.31	10.56	2.55
VSDE (Euro/h)	9.34	4.70	21.62	5.47	23.16	5.90
VSDL (Euro/h)	7.22	4.72	20.22	5.52	15.27	6.22
VSDME (Euro/h)	–	–	–	–	8.76	5.22
VSDML (Euro/h)	–	–	–	–	19.54	5.57
Nr. Obs.	1158		5965		5965	
LogLik.	-2681		-10550		-10410	
Pseudo R^2	0.17		0.36		0.37	

estimates of the VOT, VSDE and VSDL confirm our main hypothesis that the higher scheduling flexibility in the long run is reflected by the coefficient estimates. We find that travel time is valued six times higher in the long-run model than in the short-run model, indicating that drivers attach a higher value to more permanent (repetitive) travel time changes, supposedly because these can be exploited better through the rescheduling of routines. Schedule delays are valued two to three times higher in the short-run model than in the long-run model. This is likely because of the more binding nature of short-run scheduling constraints, which renders the re-planning of activities more costly. Moreover, the fact that θ is around 0.4 and significantly different from 0, shows that in the short-run drivers indeed have more information on travel time realizations than in the long-run and take this information into account when deciding on their departure time.

Since all scheduling coefficients are negative and significantly different from 0 in the 3-domains model, we conclude that indeed a net scheduling disutility results from moving away from the SRPAT towards the LRPAT. Therefore, the SRPAT is the valid anchor point of the short-run schedule delays. Specifically, we find that β_E is almost three times larger than β_{ME}, indicating that schedule delay early is valued higher if one arrives early with respect to both PAT definitions than in the intermediate domain with earliness with respect to the SRPAT but lateness with respect to the LRPAT. While this finding might indicate that drivers still take into account the LRPAT in their departure time decisions, although to

derived from being at home rather than at work tends to be larger in the early morning, resulting in a high VSDE. See Tseng and Verhoef (2008) for a model that allows scheduling costs to differ by time of the day.

a lesser extent than the SRPAT, it may also reflect non-linearities in the valuation of earliness. Regarding the coefficients for schedule delay late, we find that β_{ML} is approximately 1.3 times as large as β_L. The costs of being late with respect to both PAT definitions are therefore almost equal to the costs resulting from lateness with respect to the SRPAT but earliness with respect to the LRPAT. We also find that the VOT is significantly higher in the 3-domains model compared to the short-run model (10.56 vs. 5.20 Euro, respectively). Allowing for kinks in the schedule delay costs may have the consequence that departure time choices that result in arrival times relatively relatively far away from the SRPAT can be better explained. These choices may be determined to a relatively large extent by travel time.

In the original equation of the long-run model (Equation 3) a term for travel time variability was included ($C^{LR}[\sigma_a]$). However, this term could not be estimated due to the strong correlation between variability and expected travel time. A consequence may be that the relatively high VOT in the long-run model partly reflects the value of variability associated with longer travel times. To check whether the difference between the long-run and the short-run VOT remains if we correct for this, we impute the share of the long-run VOT that may be due to variability. For this purpose, we use the relationship between the valuation of reliability (VOR) and the valuation of schedule delays that was established by Noland and Small (1995) and later generalized by Fosgerau and Karlström (2010). Fosgerau and Karlström (2010) showed that for any travel time distribution that is constant over the time of the day, the VOR attached to the regarding standardized travel time distribution Φ is given by

$$\text{VOR} = (\text{VSDE+VSDL})\ H(\Phi, \frac{\text{VSDE}}{\text{VSDE+VSDL}}), \tag{11}$$

where $H(\Phi, \frac{\text{VSDE}}{\text{VSDE+VSDL}})$ is the so-called mean lateness factor, which is equal to the average lateness conditional on being late. The costs of variability provided that the travel time distribution Φ has standard deviation σ are then given by VOR$*\sigma$. This result remains a good approximation if the mean and the standard deviation of the travel time distribution change over the time of the day while the standardized distribution is constant.

To approximate the long-run VOR, we derive for each driver a standardized travel time distribution Φ_z and mean lateness factor $H(\Phi_z, \frac{\text{VSDE}}{\text{VSDE+VSDL}})$, using the VSDE and VSDL derived in the short-run model.[12] These can be interpreted as the values attached to rather incidental variations of travel time and therefore an upper limit to the valuation of travel time variability in the long run. We compute the costs of variability that arise if travel time changes by one hour, as these can be compared directly to the VOT. Therefore, we multiply the VOR by the derivative $(\partial\sigma/\partial E[T])_z$, which is computed using a simple OLS regression. On average, $(\partial\sigma/\partial E[T])_z$ is found equal to 0.54, confirming earlier research on the positive correlation between mean travel time and travel time variability (see Fosgerau (2010) for a

[12] For each driver we form a distribution that contains all J time-of-day-specific standardized travel time distributions based on the long-run travel time expectations. The assumption that travel time distributions are constant over the time of the day is not perfectly fulfilled. However, it was verified that our results hold also if standardized travel times are based either on peak or off-peak travel time distributions only.

theoretical derivation and Peer et al. (2012) for empirical evidence). We then find that the average cost of variability, across drivers, that results from a one-hour change in travel times is equal to 16.71 Euro. These are fairly similar across drivers, with a standard deviation amounting to 1.00 Euro. We can therefore conclude that about half of the long-run VOT as presented in Table II should be attributed to the costs resulting from travel time variability. Even after removing the costs of variability, the long-run VOT of 13.45 Euro still exceeds the short-run VOT by a factor 2.6. Our results also imply that the reliability ratio, defined as VOR divided by the VOT amounts to 2.33 on average. This is on the higher side of the range of reliability ratios that were found in earlier studies (e.g. Li et al. (2010) cite a range of 0.5–3.3), which does not come at a surprise, given that most of these studies compute the costs of variability using long-run definitions of variability and/or schedule delays, whereas we apply the relatively high short-run values of schedule delay.

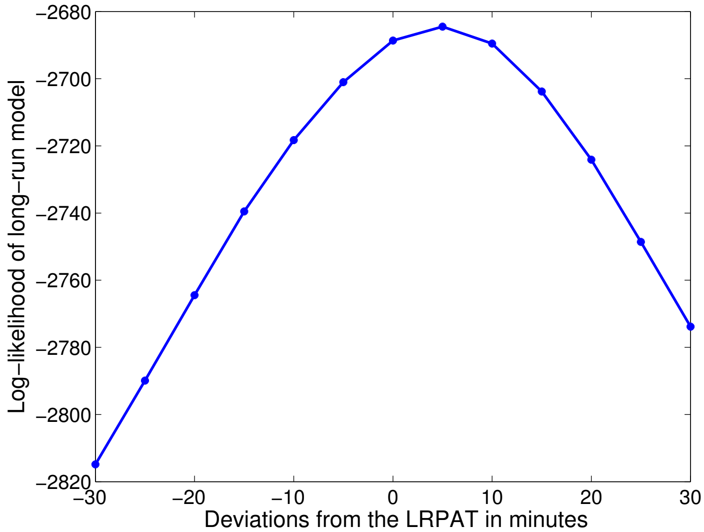
4.2 Robustness Checks

MNL Models

In this section, various robustness tests of the above results are presented, again using MNL models (as opposed to the latent class models considered in the next section). We implement two tests to ensure that the LRPAT provided by the respondents is indeed a valid measure for the preferred arrival time in the long-run choice, and therefore explains SRPATs better than other measures of the LRPAT. In the first test, we alter the LRPAT for all respondents by $[\pm 5, \pm 10, \pm 15, \pm 20, \pm 25, \pm 30]$ minutes. Figure III shows that indeed the log-likelihood peaks very near the model where the LRPAT is left unchanged, confirming that the LRPAT, although drawn from a questionnaire, indeed seems to represent the drivers' most desired arrival time from a long-run perspective. In the second robustness check for the long-run model, we divide the observations into two groups: the earliest 50% and the latest 50% of LRPATs. This is to check that the schedule delay cost valuations in the long-run model are not mainly due to early commuters for schedule delay early, or mainly to late commuters for schedule delay late. Table III shows that lower values of time and schedule delay are obtained for the earlier LRPATs, induced by a higher reward coefficient. This is probably the result of relatively more drivers with an early LRPAT choosing for early, rewarded travel routines than drivers with a late LRPAT choosing for late, rewarded travel routines. Despite of the differences between the two groups, the result of travel time being valued higher in the long run and schedule delays being valued higher in the short run is still true for both.

We also perform two robustness checks for the short-run model. For the first one, we re-estimate the short-run model as presented in Table II, with the only difference that travel time expectations are assumed equal to the actual travel time realizations (hence, $\theta = 1$). If this model yields similar results as the standard short-run model, we can conclude that the results of the standard model are not driven by long-run travel time expectations. If the latter was true, differences between long-run arrival routines and actual arrival times would be due to unobserved factors only, rendering the distinction between the long and the short-run model meaningless. The estimation results in Table III show that similar values

Figure III: LRPAT ± x minutes



of schedule delay as in the standard short-run model are obtained, indicating that actual travel time realizations are able to explain short-run behavior. The finding that the time coefficient decreases can probably be attributed to the volatile, to some extent unpredictable nature of the actual travel times. This provides an argument for including the long-run travel time expectations as part of the short-run expectations (i.e. $\theta < 1$), assuming the role of a stabilizing element in the formation of short-run travel time expectations. But anyway, the main qualitative differences in relative valuations compared to the long-run model survive.

Finally, we test whether the particular shape of the SRPAT distribution over the time of the day, with a large share of drivers departing before the peak (Figure IId), drives the outcome of the short-run model. For this purpose, we re-estimate the short-run model while applying weights to the observations that are defined as inversely proportional to the density of the SRPAT distribution[13]. The corresponding results are presented again in Table III. We find that values of schedule delay resulting from this weighted model are close to those of the standard short-run model, while the VOT decreases. The differences between the short-run and the long-run values of time and schedule delay are therefore not due to the irregular shape of the SRPAT distribution.

Latent Class Models

Finally, we estimate panel latent class models to investigate to which extent the results differ from the standard models in Table II if we account for unobserved heterogeneity

[13]The density of the SRPAT distribution is calculated using a normal kernel function with a bandwidth of 0.25.

Table III: Robustness checks

	Long-run				Short-run			
	Early LRPAT		Late LRPAT		$\theta = 1$		Inverse Density	
Coefficient	Value	t-Stat.	Value	t-Stat.	Value	t-Stat.	Value	t-Stat.
β_R	0.23	3.61	0.15	2.05	0.15	7.36	0.15	6.33
β_T	-5.59	-4.59	-7.18	-5.45	-0.13	-0.50	-0.05	-0.13
β_E	-1.80	-5.72	-2.09	-10.67	-2.80	-19.19	-2.70	-20.84
β_L	-1.57	-10.20	-1.48	-7.06	-2.65	-21.50	-2.69	-19.48
θ	–	–	–	–	1	–	0.47	8.33
VOT (Euro/h)	23.97	2.85	47.10	2.02	0.88	0.37	0.36	0.07
VSDE (Euro/h)	7.76	3.13	13.74	2.13	19.20	6.85	18.30	5.96
VSDL (Euro/h)	6.72	3.47	9.68	2.08	18.14	6.96	18.18	5.93
Nr. Obs.	579		579		5965		5965	
LogLik.	-1269		-1409		-10592		-10871	
Pseudo R^2	0.21		0.12		0.36		0.34	

among drivers. Compared to mixed logit models, which are frequently used to capture unobserved heterogeneity, latent class models have the advantage that they do not require distributional assumptions regarding heterogeneity. Moreover, in a latent class setting it is usually not necessary to assume that certain coefficients do not differ across individuals. This assumption is often present in mixed logit models, especially in cases where the variation in explanatory variables is low and correlation between the variables is high, rendering the identification of heterogeneity more difficult. As these characteristics are also present in the datasets used in this study, panel latent class models are used as a simple yet insightful check for the effects of heterogeneity among drivers.

We re-estimate the standard long- and the short-run model allowing for two latent classes[14] in both models (Table IV). In both cases, the log-likelihood of the models improves significantly, implying that the latent class models are able to capture the unobserved heterogeneity among drivers. For the long-run model, we found a large group of drivers (ca. 65% of all drivers) with a reward coefficient close to zero. We therefore fix the reward coefficient to zero for this class, in order to avoid unrealistically high values of time and schedule delay. While it is impossible to determine the monetary valuations of time and schedule delays of this class, it is reassuring that the time coefficient is higher (in absolute size) than the scheduling coefficients in the long-run model, while the opposite is true for the short-run model, confirming the results of the standard model. For the group of drivers who take into account the reward in the long-run model, the values of time and schedule delays are fairly consistent with the results of the standard long-run model in Table II. The evidence from short-run model regarding heterogeneity among drivers differs substantially

[14] We tested also higher numbers of classes, however, we found the resulting coefficients then to be fairly sensitive to the starting values.

Table IV: Latent class estimation results

	Long-Run				Short-Run			
	Class 1		Class 2		Class 1		Class 2	
Coefficients	Value	t-Stats.	Value	t-Stats.	Value	t-Stats.	Value	t-Stats.
β_R	0.27	5.53	0	–	0.13	6.76	0.14	7.21
β_T	-10.01	-6.02	-8.64	-10.16	-1.01	-1.65	-1.19	-2.36
β_E	-1.47	-6.74	-4.97	-10.66	-5.67	-32.89	-1.60	-25.39
β_L	-4.00	-7.92	-1.36	-15.21	-4.82	-34.24	-1.45	-24.66
θ	–	–	–	–	0.36	7.21	0.36	2.36
Class Prob.	0.35	8.66	0.65	16.16	0.55	17.38	0.45	14.38
VOT (Euro/h)	37.44	3.84	–	–	7.68	1.60	8.62	2.19
VSDE (Euro/h)	5.48	3.82	–	–	43.15	6.24	11.60	6.91
VSDL (Euro/h)	14.97	4.19	–	–	36.70	6.25	10.54	6.89
Nr. Obs.	1158				5965			
LogLik.	-2503				-9915			
Pseudo R^2	0.22				0.40			

from that of the long-run model. We find similar coefficient estimates for the reward and travel time in both classes, while the scheduling coefficients differ significantly between them. We can identify one group of drivers with very high values of schedule delay (55% of drivers), and another group of drivers whose values of schedule delay are in about the same range as the long-run VSDE and VSDL.

We can therefore conclude that also if we account for (unobserved) heterogeneity between drivers, the consequences of increased flexibility in the long run compared to the short run remain robust. So, the VOT is still found substantially higher in the long run than in the short run. The values of schedule delay seem to be similar between long run and short run for one group of drivers, however, on average they are again significantly higher in the short run.

5 Conclusions

We decompose scheduling decisions of commuters into long-run choices of routines and short-run choices of departure times. Data from a large-scale revealed preference experiment are used. Participants of the experiment were able to gain monetary rewards for not using a specific highway link during the morning rush hour. We find that drivers attach different values to travel time and schedule delays in the long run compared to the short run. Travel times are valued higher in the long run and schedule delays are valued higher in the short run. These findings are consistent with the intuitive notion that commuters are typically more flexible in their scheduling decisions in the longer run than in the shorter run. Travel time gains can thus be exploited better in the long run through the adaptation of routines, while scheduling constraints are more binding in the short run when routines are fixed.

Since the short-run and the long-run shadow prices differ by factors ranging from 2 to 5 in our basic model, our results may have substantial impacts on optimal choices for transport policies such as pricing and investment. Our results in particular imply that different values of time and schedule delay should be used in the evaluation of policy options, depending on whether they lead to permanent changes in travel times (e.g. an increase in road capacity) or changes in travel times that occur only under certain circumstances (e.g. incident management). Our results also have implications for the determination of optimal (road) pricing schemes. One question that we consider in follow-up research concerns the optimal pricing schedule for a congestible facility when preferences take on the form suggested by this paper. One particularly interesting aspect of the problem is whether a consistent application of optimal short-run toll schedules, which optimize departure times given fixed routines, lead to a long-run expected toll schedule that optimizes the choice of routines given structural (long-run) arrival time preferences.

Affiliations

VU University Amsterdam and Tinbergen Institute
VU University Amsterdam and Tinbergen Institute
VU University Amsterdam
VU University Amsterdam
VU University Amsterdam

References

Arnott, R., De Palma, A., and Lindsey, R. (1999). Information and time-of-usage decisions in the bottleneck model with stochastic capacity and demand. *European Economic Review*, *43*(3), 525–548.

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*(2-3), 191–229.

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

Börjesson, M. (2008). Joint RP-SP data in a mixed logit analysis of trip timing decisions. *Transportation Research Part E: Logistics and Transportation Review*, *44*(6), 1025–1038.

Börjesson, M. (2009). Modelling the preference for scheduled and unexpected delays. *Journal of Choice Modelling*, *2*(1), 29–50.

Börjesson, M., Eliasson, J., and Franklin, J. (2012). Valuations of travel time variability in scheduling versus mean-variance models. *Transportation Research Part B: Methodological*. (forthcoming)

DeSerpa, A. (1971). A theory of the economics of time. *The Economic Journal*, *81*(324), 828–846.

Fosgerau, M. (2010). On the relation between the mean and variance of delay in dynamic queues with random capacity and demand. *Journal of Economic Dynamics and Control*, *34*(4), 598–603.

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

Greene, W., and Hensher, D. (2003). A latent class model for discrete choice analysis: Contrasts with mixed logit. *Transportation Research Part B: Methodological*, *37*(8), 681–698.

Hensher, D. (2010). Hypothetical bias, choice experiments and willingness to pay. *Transportation Research Part B: Methodological*, *44*(6), 735–752.

Knockaert, J., Ettema, D., Verhoef, E., Koster, P., Peer, S., and Tseng, Y. (2012). *Spitsmijden Gouda-Zoetermeer* [Scientific Report].

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), 231–251.

Li, Z., Hensher, D., and Rose, J. (2010). Willingness to pay for travel time reliability in passenger transport: A review and some new empirical evidence. *Transportation Research Part E: Logistics and Transportation Review*, *46*(3), 384–403.

McCutcheon, A. (1987). *Latent class analysis. Sage University Paper Series on Quantitative Applications in the Social Sciences* (No. 07-064). Newbury Park, CA: Sage Publications.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in econometrics.* New York, US: Academic Press.

Menard, S. (2009). *Logistic regression: From introductory to advanced concepts and applications.* Thousand Oaks, CA: Sage Publications.

Noland, R., and Small, K. (1995). Travel-time uncertainty, departure time choice, and the cost of the morning commute. *Transportation Research Board*(1493).

Peer, S., Knockaert, J., Koster, P., Tseng, Y., and Verhoef, E. (2011). Door-to-door travel times in RP departure time choice models. An approximation method based on GPS data. *Tinbergen Institute Discussion Papers*.

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

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

Small, K., and Verhoef, E. (2007). *The economics of urban transportation.* London, UK: Routledge.

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

Tseng, Y., Knockaert, J., and Verhoef, E. (2011). A revealed-preference study of behavioural impacts of real-time traffic information. *Transportation Research Part C: Emerging Technologies*.

Tseng, Y., and Verhoef, E. (2008). Value of time by time of day: A stated-preference study. *Transportation Research Part B: Methodological*, *42*(7-8), 607–618.

Van Ommeren, J., and Fosgerau, M. (2009). Workers' marginal costs of commuting. *Journal of Urban Economics*, *65*(1), 38–47.

Van Ommeren, J., Van Den Berg, G., and Gorter, C. (2000). Estimating the marginal willingness to pay for commuting. *Journal of Regional Science*, *40*(3), 541–563.

Vickrey, W. (1969). Congestion theory and transport investment. *The American Economic Review*, 251–260.