Abstract

We develop a mathematical model of decision makers who use retrievals of earlier experiences from memory to form their expectations. Our model includes several behavioural properties related to memory. First, we assume that memory is limited, and that individuals can only store a finite number of earlier experiences. Second, we assume that retrieval is probabilistic, meaning that earlier experiences can only be retrieved with a certain probability. This allows us to account for the phenomenon that retrieval probability is higher for more recent observations, which is usually referred to as transience. The third property of our model is that retrieval may be inaccurate and subject to stochastic error. Finally, we allow for the fact that individuals use an anchor in their expectations, which results in less volatile behavioural responses than with fully adaptive expectations. The model is applied to optimal scheduling choices of morning commuters, whose travel time expectations are governed by the memory and retrieval of earlier travel experiences and the anchor component. We derive a convenient closed-form expression for the value of travel time reliability that characterizes the utilitarian effects of memory biases for fairly general memory properties. Our numerical results show that ignoring memory biases results in an underestimation of the value of travel time reliability up to 45%.

Keywords: Transience, Retrieval Accuracy, Expectation formation, Adaptive expectations, Scheduling, Value of reliability, Memory
1. Introduction

This paper presents a stylized model for choice under risk, when expectations are formed on the basis of earlier experiences. The model explicitly accounts for memory constraints, both in terms of storage as well as in terms of retrieval of earlier experiences. These constraints lead to violations of the traditional assumptions of rational expectations and infinite memory. We investigate to which extent such violations lead to sub-optimal choices of individuals, and characterize the resulting behavioral biases.

Our paper is part of a recently emerging body of literature on the representation of memory processes in economic modeling. Research on memory goes back to the late 19th century, starting with the influential contribution by Ebbinghaus (1885). During the subsequent decades, research on memory remained mainly in the hands of psychologists. Only with the recent advent of behavioral economics (e.g. Sent, 2004) and neuro-economics (e.g. Camerer and Loewenstein, 2004; Polezzi et al., 2008), also economists started incorporating memory processes in their models, acknowledging their role for expectation formation and decision-making (e.g. Hirshleifer and Welch, 2002; Mullainathan, 2002; Wilson, 2003).

The model developed in this paper assumes that storage capacity of the memory is limited and only a fixed number of past experiences can be stored in memory. We further assume that the retrieval of the stored experiences is inaccurate. Specifically, we make the assumption that the probability of a specific experience being retrievable from memory is higher for more recent experiences. This phenomenon is often referred to as transience (Schacter, 2002) and has been applied earlier in models of evolutionary learning by Barucci (1999, 2000). Additionally, we assume that individuals can remember more distant experiences less precisely. The retrieval accuracy associated with distant memories is thus lower than for more recent ones. Both the lower retrieval probability and the lower retrieval accuracy for more distant memories can be interpreted as forms of memory decay.

In a stochastic environment, the memory properties of our model render expectations adaptive: expectations and hence behavioural reactions adapt when new experiences become available. This effect, however, is counterbalanced by the assumed presence of an anchor in the utility function of decision makers. Anchors can be described as stable information elements that are not updated when new information becomes available, leading thus to less volatile behavioural responses. It has experimentally been proven that they constitute a persistent phenomenon in human behaviour (Wilson et al., 1996; Strack and Mussweiler, 1997; Furnham and Boo, 2011).

1 In his experiments, Ebbinghaus did not only take on the role of the experimenter, but, remarkably for today’s scientific standards, he was also the only participant. Despite this unusual setup, he influenced much of the ensuing psychological research on memory-related topics. His famous curve of forgetting, which describes memory retention as a function of time, resembles an exponential decay function, and is still present in many of today’s psychology textbooks.

2 A more sophisticated formulation is for instance adopted by Dow (1991), who introduces limited memory properties in a search model and investigates the optimal storage structure of information when memory capacity is limited. We assume that the probability that a specific experience is stored in memory is independent of the state associated with that experience.

3 Often anchors corresponds to the information that is obtained first, and which is then used as a
Compared to existing memory and expectation formation models in the economics literature, our model has the advantage of being generic as well as comprehensive. It takes into account both the process of memory storage and of memory retrieval, while the existing literature commonly focuses only on one of them. Moreover, we introduce a rather general model of retrieval, allowing for transience as well as retrieval inaccuracy, whereas most existing studies lack a clear model of retrieval or make very specific assumptions, such as Gennaioli and Shleifer (2010) (on selective retrievals) and Piccione and Rubinstein (1997) (on ‘absent-mindedness’). Most importantly, the model provides a psychological underpinning to adaptive expectations, making explicit the interaction between memory limitations and anchoring. Commonly, adaptive expectations are simply assumed to be present, without explaining the underlying behavioural processes.

To focus on the role of expectations, our model assumes that bounded rationality is exclusively caused by limited cognitive abilities. This paper thus stands apart from works that model bounded rationality as a result of satisficing (Simon, 1955; Caplin et al., 2011), self-deception (Bénabou and Tirole, 2002), anticipatory emotions (Bernheim and Thomadsen, 2005), or other optimality considerations (Brunnermeier and Parker, 2004; Gollier and Muermann, 2010). Neither do we emphasize judgement errors, which may for instance arise from selective memory (Gennaioli and Shleifer, 2010) or from probability weighting, although our model could accommodate for them quite easily. And unlike for instance Piccione and Rubinstein (1997), we also assume that individuals make their choices as if they are not aware of their memory constraints. This seems to be a reasonable assumption for (repeated) decisions where the stakes are not so high, and decision makers therefore take a less reflective attitude towards their own decision processes.

A repetitively occurring decision situation where the stakes are typically not very high is the scheduling choice concerning the morning commute. We apply our generic model of memory and expectation formation to this specific choice situation. Consistent with the generic model, we assume that commuters are able to store a limited number of earlier travel experiences in their memory, that the retrieval of these memorized experiences is subject to transience and inaccuracy, and that travel time expectations are influenced by an anchor. Commuters then optimize their departure time, taking into account that travel times are variable. Because their expectations adapt with new experiences, their departure times vary over time, even if the travel time distribution remains constant.

Commuters experience dis-utility from travel time variability, since it causes them to arrive earlier or later than preferred (e.g. Noland and Small, 1995). The value drivers attach to reducing the costs associated with travel time variability is commonly referred to as value of (travel time) reliability. It can be inferred from their scheduling choices using closed-form expressions that depend on the mean travel time and some measure of travel time reliability (Fosgerau et al., 2010; Fosgerau and Engelson, 2011). Typically the value of reliability is derived based on the presumption that commuters have rational expectations

reference point in subsequent decisions (Tversky and Kahneman, 1974). Ariely et al. (2003), for instance, demonstrated that individuals can be primed to anchors that are as random as the last two digits of their social security number.
and an infinite memory. If, however, in reality their choices are guided by finite memory, transience, inaccurate retrieval and anchoring, decisions will be sub-optimal and the value of reliability based on rational expectations may be substantially biased.

To our best knowledge, this is the first paper that analytically characterizes memory biases in the context of scheduling choices. The results obtained in this paper can guide future empirical research on the interaction between scheduling decisions, the value of reliability and travel time expectation formation. Our analysis has also potentially large implications for the cost-benefit assessments of transport policies, in particular because the benefits from improvements in travel time reliability tend to be large and imprecise estimates thereof may thus have a substantial impact on the outcome of the assessment. For road-related transport projects, the benefits from improvements in travel time reliability amount to ca. 25 percent of the benefits related to travel time gains (Peer et al., 2012); benefits from travel time gains, in turn, are estimated to constitute on average 60 percent of total user benefits in transport appraisals (Hensher, 2001). We chose to apply our generic model to scheduling choices concerning the morning commute. However, it may be well applicable also in other fields of economics where formation of expectations plays a central role in the dynamic economic systems (see for example Hommes (2013) for a recent overview).

The remainder of the paper is structured as follows. Section 2 describes the general setup of the model, Section 3 applies that model to the specific case of scheduling decisions and derives closed-form expressions for the expected utility. Section 4 provides numerical estimates of the biases that may result from memory limitations and anchoring. Finally, Section 5 discusses the modelling assumptions and concludes.

2. General setup of the model

Consider a decision-maker who decides on a variable \( x_0 \), where the subscript 0 indicates that her decision on \( x_0 \) is made for the time period to come. She has a decision utility function which assigns a value to the choice of \( x_0 \), and an outcome utility function that depends on the decision on \( x_0 \) and the stochastic state of nature \( s_0 \). For rational expectations, the decision utility and outcome utility coincide, whereas for adaptive expectations the two utility functions are different, potentially leading to sub-optimal decision making. We define the outcome utility as \( U(x_0, s_0) \), which depends on the state of nature \( s_0 \). Furthermore, the outcome utility is assumed to be continuous and strictly concave in \( x_0 \). The state of nature is stochastic and distributed with a continuous probability density function \( f(s_0|\omega_0) \), where \( \omega_0 \) is a parameter vector that describes this probability density function. For analytical simplicity, it is assumed that the choice of \( x_0 \) does not affect the density function \( f(s_0|\omega_0) \). The expected outcome utility for any chosen \( x_0 \) in time period 0 is given by the expected value over all possible values of \( s_0 \) in time period 0:

\[
E_{s_0}(U(x_0, s_0)) = \int U^0(x_0, s_0)f(s_0|\omega_0)ds_0 \quad (1)
\]

With rational expectations and choice under risk, the decision maker perfectly knows the distribution \( f(s_0|\omega_0) \) and uses Equation 1 to decide on \( x_0 \). We define the optimal choice of
with rational expectations as \( x_0^{re} \), where the superscript \( re \) refers to rational expectations. The corresponding optimal expected utility is then given by \( \mathbb{E}(U_{re}) \equiv \mathbb{E}_0(U_0^{re}(x_0^{re}, s_0)) \).

The decision-maker forms adaptive expectations on the state of nature \( s_0 \), and use these expectations to decide on \( x_0 \). First, it is assumed that she has a limited memory, meaning that only \( K \) experiences can be stored in memory. More precisely, she stores the states \( s_1 \ldots s_K \) in her memory. Here, the subscript refers to the past, meaning that a lower number refers to a more recent experience. The states \( s_1 \ldots s_K \) are draws from the distribution \( f(s_k|\omega_k) \) which depends on \( k \), meaning that the distribution can be changing over subsequent time-periods. For simplicity we assume that the \( K \) stored realised states are equally spaced in time.

A stored memory can only be retrieved with probability. Each state receives a positive weight \( \rho_k > 0 \) in the expectation of the decision maker, where we have the normalisation \( \sum_{k=1}^{K} \rho_k \equiv 1 \) to ensure that retrieval probabilities sum up to 1. It is assumed that the realisation of \( s_k \) is correctly stored in memory, but when \( \rho_k = 0 \), the decision maker cannot retrieve the experience from her memory. For analytical simplicity it is assumed that retrieval probabilities are independent of \( s_k \). Usually more recent experiences can be retrieved more easily. Following Schacter (2002) we refer to this phenomena as transience. Formally transience is defined as \( \rho_1 > \rho_2 > \ldots > \rho_K \), meaning that the retrieval probability decreases for more distant memories. A specific functional form of the retrieval probabilities will be given in Section 3.

Retrieval of the states \( s_1 \ldots s_K \) in memory may be inaccurate, meaning that retrieval is subject to stochastic error. For a sequence of states \( s_1 \ldots s_K \) in memory, the decision maker retrieves a sequence of retrieved states \( \tilde{s}_1 \ldots \tilde{s}_K \) from her memory. With perfectly accurate recall, these retrieved states are equal to the realised states \( s_1 \ldots s_K \). When retrieval is inaccurate, every retrieved state \( \tilde{s}_k \) is a draw from the retrieval density function \( g_k(\tilde{s}_k|s_k, \phi_k) \), where \( \phi_k \) is a parameter vector that describes the probability density function. It is plausible that the distribution of retrieved states bears some relationship with the experienced state \( s_k \). For example, it may be that on average the decision maker remembers the state \( s_k \) correctly, and therefore \( \mathbb{E}(\tilde{s}_k) = s_k \). Another behavioural plausible assumption would be that the conditional retrieval variability is increasing for more distant memories and therefore \( \frac{\partial \text{VAR}(\tilde{s}_k|s_k)}{\partial k} > 0 \). The decision maker may therefore have the idea that more distant memories are more alike, because the retrieval variance around these states is larger. The results in Section 3 are derived for general retrieval densities of retrieval meaning that we do not assume a particular shape for \( g_k(\cdot) \).

The decision maker not only uses earlier experiences to determine her expectations, but also has primed expectations on an anchor state \( s_A \). In contrast to the states in memory and the corresponding retrievals, the anchor is not stochastic and is a stable element in the expectation formation of the decision maker. For example, this anchor can be the mean or the mode of the distribution of experienced states. If the decision maker would only use the anchor, expectations in our model would not be adaptive and there will be no effects of limited memory, transience or inaccurate retrieval. The weight that the decision maker assigns to the anchor therefore plays a crucial role in how adaptive expectations are.

The decision utility conditional on the anchor and the retrieval of earlier experiences is
then given by:

\[ U^d(x_0, \tau, s_A, K, \rho_1, ..., \rho_K, \bar{s}_0, ..., \bar{s}_K) = \tau U(x_0, s_A) + (1 - \tau) \sum_{k=1}^{K} \rho_k U(x_0, \bar{s}_k), \]

where the parameter \( \tau \) is assumed to be independent of the choice of \( x_0 \). This parameter indicates how strongly the individual weights the anchor state in her expectations. The decision maker maximizes the decision utility by calculation of the weighted average of the utility outcome for each retrieved state \( \bar{s}_k \). The weights are governed by \( \tau \) and the retrieval probabilities. When \( \tau = 0 \), expectations are fully adaptive and the decision utility (Equation 2) is solely based on the retrieval probabilities, the retrieved states and the choice of \( x_0 \). When \( \tau = 1 \), the decision maker ignores her earlier experiences and the decision utility is only based on \( s_A \) and the choice of \( x_0 \). The decision utility mimics the outcome utility function (Equation 1) when \( \tau \to 0 \), \( \rho_k = 1/K \), \( E(\bar{s}_k) = s_k \), \( \text{VAR}(\bar{s}_k|s_k) = 0 \) and \( K \to \infty \). Rational expectations can therefore be viewed as a special case of the anchored adaptive expectations.

The decision maker maximizes Equation 2 with respect to \( x_0 \), where a unique solution exists, because Equation 2 is a weighted average of strictly concave function. Denote this optimal \( x_0 \) by \( x_0^{ae} \), the superscript refers to the fact that the decision maker uses adaptive expectations to determine the optimal \( x_0 \). Because decision utility and outcome utility in general do not coincide, decisions on \( x_0 \) are suboptimal whenever \( x_0^{ae} \neq x_0^{re} \). This suboptimal choice is costly in terms of expected outcome utility. When \( x_0^{ae} = x_0^{re} \), the decision maker makes ‘coincidentally’ the optimal choice, but the underlying choice model is not based on rational expectations. This would be the economists’ ideal since the rational expectations model then results in the same optimal expected utility and the more parsimonious presentation of behaviour will not affect the expected outcome utility. Obviously, when the expected outcome utility (Equation 1) is more concave in \( x_0 \), sub-optimal decisions will be more costly than when it is less concave.

The expected outcome utility for a given sequence of states in memory \( s_1...s_K \) and corresponding retrieved states \( \bar{s}_1...\bar{s}_K \) is given by the expected value over all possible states of \( s_0 \): \( E_{s_0} U(x_0^{ae}, s_0) \). This means that adaptive expectations only enter the expected outcome utility via the choice of \( x_0^{ae} \). Since every retrieved state \( \bar{s}_k \) is a draw from the distribution \( g(\bar{s}_k|s_k, \phi_k) \), the decision \( x_0^{ae} \) and the corresponding expected outcome utility also depend on the (stochastic) sequence of states \( s_1...s_K \) in memory.

If we would have observations of the states stored in memory we could simulate the retrieval states using draws from the distributions \( g_1(.)...g_K(.) \) and calculate the corresponding \( x_0^{ae} \) for each set of draws using Equation 2. Now suppose that we need to make a prediction of the expected outcome utility of the decision maker. A good measure of expected utility
will account for the fact that the state in time period 0, the states in memory and the corresponding retrievals of these states are stochastic. To obtain the predicted expected outcome utility, we therefore take the expected value over all possible combinations of experienced states and retrieved states. Mathematically this is a bit tedious, because this involves a $2K + 1$ dimensional integral over all possible values of the $K + 1$ realised states $s_0...s_K$, and the $K$ possible values of retrieved states $\bar{s}_1...\bar{s}_K$:

$$\mathbb{E}(U_{ae}) \equiv \mathbb{E}_{s_0...s_K} \left( \mathbb{E}_{\bar{s}_1...\bar{s}_K} \left( U(x_{0e}^s, s_0) \right) \right) = \int ... \int \left( \int ... \int U(x_{0e}^s, s_0) \prod_{k=1}^{K} g_k (\bar{s}_k|s_k, \phi_k) d\bar{s}_1...d\bar{s}_K \right) \prod_{k=0}^{K} f(s_k, \omega_k) ds_0...ds_K. \quad (3)$$

This equation obviously has the disadvantage that it is less parsimonious then when using Equation 1 with $x_0^{re}$. However, the general set-up helps to structure our thoughts about how earlier experiences and retrieval inaccuracy affect the expected outcome utility of a decision maker. The next section derives a parsimonious representation of expected utility $\mathbb{E}(U_{ae})$ that includes the effects of memory biases in an intuitive way. The remainder of the paper is devoted to work out an analytical example for scheduling decisions of travellers with random travel times.

3. Memory and the value of travel time reliability

3.1. Introduction

We apply the generic memory and expectation formation model of the previous section to travellers’ scheduling behaviour when travel times are stochastic. In such situations, travellers face costs of travel time variability, since they may arrive earlier or later than desired, and they take these costs into account in choosing their optimal departure time. Noland and Small (1995) were the first to show how expected utility maximization can be applied to scheduling choices. Their model was extended by Fosgerau et al. (2010) and Fosgerau and Engelson (2011) who prove that the optimal expected utility depends linearly on the mean travel time and some measure of travel time reliability such as the standard deviation or variance of travel times. Assuming time-of-day-dependent scheduling preferences as in Fosgerau and Engelson (2011), we show that this convenient result carries over to the case when memory biases and anchoring are present and distributions of travel times are stable over time. This allows us to characterize biases in the value of reliability that may result as a consequence of memory limitations and anchoring.

Our model introduces a new layer of analysis to models concerned with travel time expectation formation – by explicitly characterizing the memory processes that underpin the

\[\text{Here we use } \mathbb{E}_{s_0...s_K} \text{ to denote a repeated expectation over all stochastic variables } s_0...s_K.\]

\[\text{Some analytical progress can be made by using strictly concave polynomial utility functions, because the expected outcome utility can then be written as a function of } x_0^{re}, \text{ the (statistical) moments of the distributions of the states and the moments of the retrieval distributions. Instead of proceeding along this line we consider it more insightful to work out an example.}\]
actual expectation formation processes. Unlike existing literature on travel time expectations\textsuperscript{7}, it provides a psychological foundation for the adaptation of expectations when new travel experiences become available. Moreover, existing studies typically do not characterize potential behavioural biases, even when they find that travel time expectations are adaptive. Neither do most of them focus on whether behavioural measures, which can be derived from observed behaviour (such as the value attached to travel time or travel time reliability), may be biased if the actual expectation formation process deviates from the usual assumptions of rationality and infinite memory.\textsuperscript{8}

3.2. Setup of the model

3.2.1. Rational expectations

We assume that travellers derive utility for being at home and at work. An increase in travel time reduces utility because this time could otherwise been used at home by spending more time with the family, sleeping or having a longer breakfast. At work the time could have been used as productive work time. This specification of utility was first introduced by Vickrey (1973) and later used by Fosgerau and Engelson (2011) for departure time choice with rational expectations. Tseng and Verhoef (2008) find empirical evidence for these scheduling preferences.

In line with Fosgerau and Engelson (2011), we assume that the marginal utilities for being at home and at work are linear in the time of the day $v$, and cross at the preferred arrival time $v^*$ which is normalised without loss of generality to 0. Preferences are assumed to be stable over subsequent time periods and therefore independent of $k$. The outcome

\textsuperscript{7}Research on travel time expectations has attracted increasing attention in recent years, especially due to the advent of real-time traffic information. Several studies emphasize the modelling of learning and perception updating: Arentze and Timmermans (2003) use a reinforcement learning approach, Arentze and Timmermans (2005) develop a model based on mental maps, and Jha et al. (1998) and Chen and Mahmassani (2004) draw on Bayesian updating approaches. Some studies investigate day-to-day dynamics in transport networks, (Mahmassani and Liu, 1999; Watling and Cantarella, 2013), however, most of these lack a strong behavioural underpinning of the process of updating expectations. Empirical evidence has mainly been derived from laboratory experiments. Bogers et al. (2007) and Ben-Elia and Shiftan (2010), for instance, find evidence that more recently experienced travel times have an over-proportionally large influence on decision behaviour, suggesting that transience is indeed a behavioural characteristic of commuters. Peer et al. (2013) find that travellers do not only take into account the long-run travel time average in their scheduling decisions (which may be interpreted as an anchor), but also day-specific information.

\textsuperscript{8}Such discussions can almost exclusively be found in those studies that focus explicitly on deviations from rational expectation formation. In the transportation research domain, most studies belonging to this category make use of (cumulative) prospect theory. For example, Borger and Fosgerau (2008) find that reference dependence and loss aversion strongly affect the estimated value of travel time, using stated preference data. They demonstrate that a reference-free value of time can be derived. Koster and Verhoef (2012) analyse travellers’ costs of probability weighting using a rank-dependent scheduling model. They conclude that in reality the resulting bias in travel costs is rather small because both relatively short and long travel times receive higher probability weights resulting in a departure time choice that is on average close to optimal. Analyses on the impact of probability weighting at the network level can for instance be found in Connors and Sumalee (2009) and Gao et al. (2010).
utility for a given departure time $d_0$ and a realisation of travel time $T_0$ is then given by:

$$U(d_0, T_0) = -\int_{d_0}^{0} (\beta_0 + \beta_1 v) dv - \int_{0}^{d_0 + T_0} (\beta_0 + \gamma_1 v) dv,$$

(4)

where it is assumed that $\gamma_1 > \beta_1$. The second derivative of this utility function with respect to the departure time $d_0$ is given by $\beta_1 - \gamma_1$ implying that Equation 4 is strictly concave in $d_0$. With rational expectations, travellers know the distribution of travel times perfectly well and the expected outcome utility is given by:

$$E_{T_0}(U(d_0, T_0)) = \int \left(-\int_{d_0}^{0} (\beta_0 + \beta_1 v) dv - \int_{0}^{d_0 + T_0} (\beta_0 + \gamma_1 v) dv\right) f(T_0|\mu, \sigma_0^2) dT_0.$$

(5)

Fosgerau and Engelson (2011) show that when the departure time is optimally chosen, the traveller departs at $d_{re} = -\frac{\gamma_1}{\gamma_1 - \beta_1} \mu$. Substituting in Equation 5 gives the optimal expected utility with rational expectations:

$$E(U_{re}) \equiv E_{T_0}(U(d_{re}, T_0)) = -\beta_0 \mu + \frac{1}{2} \frac{\beta_1 \gamma_1}{\gamma_1 - \beta_1} \mu^2 - \frac{1}{2} \frac{\gamma_1 \sigma_0^2}{\gamma_1 - \beta_1},$$

(6)

which has the convenient property that the optimal expected utility is a simple function of the mean delay, and is linearly decreasing in the variance of travel time $\sigma_0^2$. Furthermore, Equation 6 does not require any distributional assumptions on the travel time distribution (except that $\sigma_0^2$ is finite).

3.2.2. Adaptive expectations

Next, we introduce the model for adaptive expectations. The states in memory are given by travel times $T_1 \ldots T_K$ and the retrievals of these past travel times are given by $\bar{T}_1 \ldots \bar{T}_K$. To focus on the role of travel time reliability, the mean travel time is assumed to be the same over subsequent days, whereas the variance of travel time is assumed to be dependent on $k$. The travel times in memory are therefore draws from a travel time distribution $f(T_k|\mu, \sigma_k^2)$, with mean $\mu$ and variance $\sigma_k^2$. The traveller decides on her departure time $d_0$ in time period 0. The accuracy of retrieval is governed by the probability density function $g(\bar{T}_k|T_k, \nu_k^2)$ which has mean $E(\bar{T}_k) = T_k$ and conditional variance $\text{VAR}(\bar{T}_k|T_k) = \nu_k^2$. Using the law of total variance, these assumptions together lead to an unconditional variance $\text{VAR}(\bar{T}_k) = \nu_k^2 + \sigma_k^2$.

Furthermore, the traveller has an anchor $T_A$ which is defined as: $T_A = \mu + a$, where $a$ is a parameter that indicates how far the anchor is from the mean travel time $\mu$. The utility of

---

9Here we follow Börjesson et al. (2012) in notation. The subscripts of $\beta$ and $\gamma$ bear no relationship with $k$.

10For plausibility of the model, some additional restrictions on the preference parameters are required, because for some combinations of preference parameters the marginal utility for changes in the mean delay $-\beta_0 + \frac{\beta_1 \gamma_1}{\gamma_1 - \beta_1} \mu$ may be positive, implying that travellers would value increases in travel times positively. This condition is satisfied for our numerical examples.
a realisation of travel time $T_0$ in time period 0 is given by $4$. Decision utility in time period 0 is given by:

$$U^d(d_0, \bar{T}_1, \ldots, \bar{T}_K, T_A, \rho_1, \ldots, \rho_K, \tau, K) = \tau U(d_0, T_A) + (1 - \tau) \sum_{k=1}^{K} \rho_k U(d_0, \bar{T}_k), \quad (7)$$

With adaptive expectations, travellers choose their optimal departure time using Equation 7. The first order condition is given by:

$$\frac{\partial U^d(.)}{\partial d_0} = -(\gamma_1 - \beta_1)d_0 - \tau \gamma_1 T_A - (1 - \tau) \frac{\gamma_1}{\gamma_1 - \beta_1} \sum_{k=1}^{K} \rho_k \bar{T}_k = 0. \quad (8)$$

Solving this first-order condition for $d_0$ gives:

$$d^u_0 = -\tau \frac{\gamma_1}{\gamma_1 - \beta_1} T_A - (1 - \tau) \frac{\gamma_1}{\gamma_1 - \beta_1} \sum_{k=1}^{K} \rho_k \bar{T}_k. \quad (9)$$

The limited memory effect on departure time choice is captured by the last term. If $K$ increases, the traveller has more past travel times stored in memory. These travel times are retrieved with probability $\rho_k$. Retrieval accuracy enters the departure time choice via the retrieved travel times $\bar{T}_k$. If we assume that retrieval is unbiased on average, $E(\bar{T}_k) = T_k$, and the mean departure time is given by:

$$E(d^u_0) = -\tau \frac{\gamma_1}{\gamma_1 - \beta_1} T_A - (1 - \tau) \frac{\gamma_1}{\gamma_1 - \beta_1} \mu, \quad (10)$$

which reduces to $d^u_0$ for $T_A = \mu$ ($a = 0$). The traveller also includes the anchor where the strength of the anchor in the expectation formation is governed by the anchor parameter $\tau$. If the traveller ignores earlier experiences $\tau = 1$, and departure time choice is only governed by the anchor $T_A$ and the parameters of the utility function.

The variability in departure times over time periods is influenced by the variance of travel times and the variance of retrieval. A higher variance in travel times and a higher retrieval inaccuracy result in more variable departure times. This becomes more evident if we derive the variance of departure time:

$$\text{VAR}(d^u_0) = (1 - \tau)^2 \left( \frac{\gamma_1}{\gamma_1 - \beta_1} \right)^2 \sum_{k=1}^{K} \rho_k^2 \left( \sigma_k^2 + \nu_k^2 \right). \quad (11)$$

Our model therefore predicts that the variability in departure times increases when the variance of travel time of the last $K$ experiences increases, where this effect is multiplied with the quadratic retrieval probabilities. Furthermore, the variability in behaviour increases if the variance of retrieval accuracy increases. A higher anchor parameter results in less variable departure time decisions because memory biases count less heavily in the decision utility function. Finally, the variance in departure times increases linearly in $\left( \frac{\gamma_1}{\gamma_1 - \beta_1} \right)^2$.

\[\text{Here we use the rules } \text{VAR}(cx) = c^2 \text{VAR}(x) \text{ and the rule that the variance of the sum of independent random variables is the sum of the variances of these random variables.}\]
3.2.3. Derivation of the expected outcome utility

The expected outcome utility can be found by integrating over the travel time $T_0$, all possible combinations of stochastic experiences $T_1 \ldots T_K$ and the corresponding stochastic retrievals $\bar{T}_1 \ldots \bar{T}_K$:

$$\mathbb{E}(U_{ae}) \equiv \mathbb{E}_{T_0 \ldots T_K} \left( \mathbb{E}_{T_1 \ldots T_K} (U(d_{0e}^{ae}, T_0)) \right) = \int \ldots \int \left( \int \ldots \int U(d_{0e}^{ae}, T_0) \prod_{k=1}^{K} g(T_k | T_0, \nu_k^2) d\bar{T}_1 \ldots d\bar{T}_K \right) \prod_{k=0}^{K} f(T_k | \mu, \sigma_k^2) dT_0 \ldots dT_K,$$

where $\mathbb{E}(U_{ae})$ is the expected utility in time period 0 with adaptive expectations.

**MAIN RESULT**

For retrieval probabilities $\rho_1 \ldots \rho_K$, retrieval variances $\nu_1^2 \ldots \nu_K^2$, anchor $T_A = \mu + a$, and a travel time distribution with mean $\mu$ and variances $\sigma_0^2 \ldots \sigma_K^2$, the optimal expected outcome utility is given by:

$$\mathbb{E}(U_{ae}) = \mathbb{E}(U_{re}) - \frac{1}{2} \frac{\gamma_1}{\gamma_1 - \beta_1} \tau^2 a^2 - \frac{1}{2} \frac{\gamma_1}{\gamma_1 - \beta_1} (1 - \tau)^2 \left( \sum_{k=1}^{K} \rho_k^2 \sigma_k^2 + \sum_{k=1}^{K} \rho_k^2 \nu_k^2 \right).$$

The proof is given in Appendix A. The first term in Equation 13 is the optimal expected utility with rational expectations (Equation 6). There are three additional negative terms related to memory.

Limited memory enters the optimal expected utility in two ways. First, it enters the summation over all travel time variances multiplied with the retrieval probabilities. Therefore the variances of the last $K$ time periods enter the optimal expected utility. This is different from rational expectations where only the variance of the current time period plays a role. Second, $K$ enters the summation over all retrieval variances multiplied with the squared retrieval probabilities. This means that the retrieval variances of the last $K$ time periods decrease expected outcome utility.

The retrieval probabilities in Equation 13 enter the expected outcome utility in a squared manner. If retrieval probabilities for a chosen time period $k$ are high, the negative effect of the travel time variance and the retrieval variance is higher. The negative effect of probabilistic retrieval becomes stronger if travellers’ expectations become more adaptive (e.g. when $\tau$ decreases). The resulting bias is linearly increasing in $\frac{\gamma_1}{\gamma_1 - \beta_1}$. The main result is derived for a general sequence of retrieval probabilities and does not assume transience beforehand. With transience we have $\rho_1 > \rho_2 > \ldots > \rho_K$. Because retrieval probabilities enter Equation 13 in a squared manner, more recent travel time variance play a larger role than more distant travel time variances when transience applies.

The last summation term in Equation 13 is related to the retrieval variances $\nu_k^2$, and drops out if there recall is perfectly accurate. This term is also increasing when adaptive expectations become more important. Furthermore, the term interacts with the squared retrieval probabilities. This is intuitive because for small retrieval probabilities the bias in
recall has a small effect on the formation of expectations and therefore on the departure time choice. The bias due to sub-optimal departure time choices will therefore be smaller.

The second term in Equation 13 is related to the bias because of anchored expectations. For the traveller it is optimal to have $T_A = \mu$. This is because when $\tau \rightarrow 1$, and $T_A = \mu$, the optimal departure time Equation 9 is equal to the departure time with rational expectations: $d_{0e} = -\frac{\gamma_1}{\gamma_1 - \beta_1} \mu$. For any $a \neq 0$, there is an additional penalty in the expected outcome utility. This penalty is quadratically increasing in the value of $a$ and the anchor parameter $\tau$. This means that when a traveller uses the mode or median travel time as an anchor, this results in additional dis-utility. When the anchor is not included in the expectation formation, $\tau \rightarrow 0$, and the second term in Equation 13 drops out. Deviations of the anchor therefore count heavier for individuals that have a higher $\tau$. Again, the resulting bias is linearly increasing in $\frac{\gamma_1^2}{\gamma_1 - \beta_1}$. The marginal change in expected outcome utility due to a change in $\tau$ is given by:

$$\frac{\partial E(U_{ae})}{\partial \tau} = -\frac{\gamma_1^2}{\gamma_1 - \beta_1} \tau a^2 + \frac{\gamma_1^2}{\gamma_1 - \beta_1} (1 - \tau) \left( \sum_{k=1}^{K} \rho_k^2 \sigma_k^2 + \sum_{k=1}^{K} \rho_k^2 \nu_k^2 \right)$$

(14)

When $a \neq 0$, the anchor parameter $\tau$ plays a double role in Equation 13. An increase in $\tau$ leads to a decrease in the expected utility due to a sub-optimal chosen $T_A \neq \mu$, whereas the bias related to probabilistic retrieval and retrieval inaccuracy decreases, because past experiences have a lower effect on travel time expectations. The $\tau$ that weights these effects optimally is independent of scheduling preferences and given by $\tau = \frac{\sum_{k=1}^{K} \rho_k^2 \sigma_k^2 + \sum_{k=1}^{K} \rho_k^2 \nu_k^2}{\sum_{k=1}^{K} \rho_k^2 \sigma_k^2 + \sum_{k=1}^{K} \rho_k^2 \nu_k^2 + \sigma^2}$ which is equal to 1 if $a = 0$.

### 3.3. The value of travel time variance

If we assume that retrieval variances are independent of the variance of travel time, and $\sigma_k^2 \equiv \sigma^2$, the value of travel time variance is given by:

$$\frac{\partial E(U_{ae})}{\partial \sigma^2} = -\frac{1}{2} \gamma_1 - \frac{1}{2} \frac{\gamma_1^2}{\gamma_1 - \beta_1} (1 - \tau)^2 \sum_{k=1}^{K} \rho_k^2,$$

(15)

implying that the expected outcome utility is linearly increasing in the travel time variance. Therefore the convenient result of Fosgerau and Engelson (2011) still applies when travellers have adaptive expectations and travel time variances are stable over time periods. The retrieval probabilities enter the value of travel time variance quadratically leading to a stronger effect for more recent experiences when transience applies. In order to study the effect of transience, it is useful to be more explicit about these retrieval probabilities. Two conditions need to be met for behaviourally plausible retrieval probabilities. First, the probabilities need to sum up to 1 for any chosen value of $K = 1...\infty$. Second, the probabilities

---

12Strictly speaking we could assume that the variance of travel times of the last $K$ time periods is independent of $k$ to arrive at the same expression. This would lead to a value of travel time variance that is still dependent on time (with discrete jumps)
need to be decreasing in $k$, because more recent travel times have a higher likelihood of remembering. A functional form that satisfies these conditions is given by:

$$
\rho_k = \frac{r - 1}{r(r^k - 1)} r^k,
$$

(16)

where we have $0 < r < 1$. In this equation, the parameter $r$ is the transience parameter. A lower value of $r$ indicates more transience, meaning that more recent travel times receive a higher retrieval probability. An increase in $r$ results in more equal weights where equal weights $\frac{1}{K}$ are a special case, because $\lim_{r \to 1} \rho_k = \frac{1}{K}$.

If we assume that retrieval probabilities are given by Equation 16, the value of travel time variance is given by:

$$
\frac{\partial \mathbb{E}(U_{ae})}{\partial \sigma^2} = \left(-\frac{1}{2} \gamma_1 - \frac{1}{2} \gamma_1 \left(1 - \tau\right)^2 \frac{(1 - r)(1 + r^K)}{(1 + r)(1 - r^K)}\right),
$$

(17)

which is increasing in the transience parameter $r$, meaning that if retrieval probabilities are more equal, the value of travel time variance will be lower. The limiting case $r \to 1$ gives retrieval probabilities equal to $\frac{1}{K}$. The value of travel time variance is then given by:

$$
\frac{\partial \mathbb{E}(U_{ae})}{\partial \sigma^2} = \left(-\frac{1}{2} \gamma_1 - \frac{1}{2} \gamma_1 \left(1 - \tau\right)^2 \frac{1}{K}\right),
$$

(18)

and therefore in the absence of transience the additional effect of limited memory on the value of travel time variance is proportional to $\frac{1}{K}$. Intuitively, this effect becomes 0 for infinite memory ($K \to \infty$).

3.4. The value of retrieval inaccuracy

The expected outcome utility (see Equation 13) clearly shows that it is valuable for travellers to have a higher accuracy of retrieval. An increase in the retrieval variance of the $l$th stored memory leads to additional dis-utility. The value of retrieval inaccuracy is given by the first derivative of Equation 13 with respect to $\nu^2_l$:

$$
\frac{\partial \mathbb{E}(U_{ae})}{\partial \nu^2_l} = -\frac{1}{2} \gamma_1 \left(1 - \nu^2_l\right),
$$

(19)

which is increasing if expectations become more adaptive ($\tau \to 0$). Furthermore, the value of retrieval inaccuracy for memory $l$ increases in the squared of the corresponding retrieval probability. A more explicit expression could be derived by substituting 16:

$$
\frac{\partial \mathbb{E}(U_{ae})}{\partial \nu^2_l} = -\frac{1}{2} \left(1 - \frac{\gamma_1^2 (1 - \tau)^2}{\left(\frac{r}{r^K - 1}\right)^l}\right).
$$

(20)

When transience applies the value of retrieval inaccuracy is lower for more distant memories (higher $l$). Therefore there is an interplay between transience and the effects of retrieval inaccuracy.
3.5. The limiting case of \( K \to \infty \)

This subsection assumes that memory is unlimited and that an infinite number of experiences are stored. Furthermore, it is assumed that the retrieval variance is linearly increasing in \( k \):

\[ \nu^2_k = \bar{\nu}k. \]  
(21)

If we substitute Equations 16 and 21 in Equation 13 we obtain an explicit expression for the expected outcome utility as a function of the transience parameter and retrieval variance parameter \( \bar{\nu} \).\(^{13}\) This gives the expected outcome utility for infinite memory:

\[ E(U_{ae}) = E(U_{re}) - \frac{1}{2} \frac{\gamma_1^2}{\gamma_1 - \beta_1} \tau^2 a^2 - \frac{1}{2} \frac{\gamma_1^2}{\gamma_1 - \beta_1} (1-\tau)^2 \frac{1-r}{1+r} \sigma^2 - \frac{1}{2} \frac{\gamma_1^2}{\gamma_1 - \beta_1} (1-\tau)^2 \frac{\bar{\nu}}{(1+r)^2}, \]  
(22)

and therefore the biases due to transience and inaccurate recall do not vanish when memory is unlimited. Furthermore, 22 shows that when the retrieval probabilities are all equal \((r \to 1)\), the third term drops out, and the transience bias vanishes, but the bias due to inaccuracy does not. The rational expectations model can therefore be viewed as a limiting case, where expectations are fully adaptive, retrieval variances are 0, and travellers have infinite memory.

4. Numerical illustration

This section provides a numerical illustration of the model discussed in the previous section. We study variations in and interplay between four of the structural model parameters, namely \( r, T_a, \tau \) and \( \bar{\nu} \), and trace their impacts on optimal departure times, expected utility and the value of travel time reliability. The rational expectations levels of these measures are used as the base for comparison.

In section 4.1 we assign values to the remaining model coefficients and define the rational expectations outcomes. Section 4.2 analyses the quantitative effect of transience by reducing the value of \( r \) such that adaptive expectations become increasingly based on recent periods. At this stage, individuals are still assumed to not be affected by inaccurate retrievals. Section 4.3 maintains the assumption of accurate retrievals but introduces an anchor point by increasing the value of \( \tau \) and varying the value of the anchor point \( T_a \). Section 4.4 completes the numerical analysis by allowing for inaccurate retrievals.

4.1. Defining coefficients and the rational expectations outcome

The values for the coefficients defining the rational expectations outcomes of the model are based upon Fosgerau and Lindsey (2013) who use a linear approximation of the mixed logit estimates of Tseng and Verhoef (2008). Accordingly, \( \beta_0, \beta_1 \) and \( \gamma_1 \) assume the following values: \( \beta_0 = \EUR 40 \) (p/hour), \( \beta_1 = \EUR 8.86 \) (p/hour) and \( \gamma_1 = \EUR 25.42 \) (p/hour). Additionally, we make assumptions on the distribution of travel times using the empirical estimates of Peer

\(^{13}\)For the limiting case of \( K \to \infty \) we have: \( \sum_{k=1}^{\infty} \rho_k^2 = \frac{1-r}{1+r} \) and \( \sum_{k=1}^{\infty} \rho_k^2 \bar{\nu}k = \frac{\bar{\nu}}{(1+r)^2} \).
et al. (2012). They estimate the relationship between the mean and the standard deviation of travel time. Consistent with their results, we assume that with every unit increase in the mean delay the standard deviation will rise with 0.75 units. We assume that \( f(T_k|\mu, \sigma^2) \) is log-normally distributed with an expected travel time of \( \mathbb{E}(T_0) = \mu = \frac{1}{3} \), i.e. 20 minutes and \( \text{VAR}(T_0) = \sigma^2 = \frac{1}{16} \), i.e. a standard deviation 15 minutes.\(^{14}\)

For this particular set of coefficients, the optimal departure time for rational expectations is given by \( d^c_{re} = -0.51 \) hours before the preferred arrival time. Therefore, a fully rational individual departs 31 minutes prior to her desired time of arrival even though her expected travel time is only 20 minutes. Expected utility under rational expectations is \( \mathcal{E} - 13.37 \) and the value of travel time reliability is \( \mathcal{E} - 12.71 \).\(^{15}\) In what follows we study the deviations of this particular rational expectations outcome.

### 4.2. Accurate retrieval and transience

The first deviation introduced from the rational expectations outcome is transience. With transience the \( K \) most recent realisations of travel times determine the optimal departure time for the upcoming time-period where more recent periods have a larger impact on optimal departure time. We generate 1,000 different sets of travel time realisations for \( K = 5 \) and \( K = 100 \), and systematically change the importance of each of these periods in forming expectations by changing \( r \). We vary \( r \) between its upper bound of \( r = 1 \) (equal weights to all \( K \) realisations) and \( r = 0.5 \) where the most recent period receives a weight \( \rho_1 \) of approximately 50\% for both levels of \( K \). Increasing values of \( r \) provides more weight to more distant memories. Additionally, we assume the conditional variances of retrieval \( \nu^2_1 \ldots \nu^2_K \) are all 0. The past realisations of \( T_k \) are then all accurately retrieved, i.e. \( \bar{T}_k = T_k \). For the moment, we ignore anchoring and assume \( \tau = 0 \). Comparing Figures 1a and 1b clearly shows that having a more limited memory (\( K = 5 \) instead of \( K = 100 \)) increases the variance of the optimal departure time substantially. This is a direct consequence of adaptive expectations being formed by a smaller number of travel time realisations. Figures 1c and 1d illustrate that the size of \( K \) becomes less relevant for smaller values of \( r \). Reducing \( r \) focuses attention towards more recent periods such that more distant travel time realisations have a negligible impact on the optimal departure time. In fact, the variance is comparable between the two alternative values of \( K \) in this case since only the last period is of influence in this illustration.

Transience also has direct implications on the level of expected utility as illustrated by Figure 2a. Even when \( r \) is approaching 1, expected utility falls below \( EU_{re} \) for \( K < \infty \), because travellers do not have sufficient memory capacity to form rational expectations about \( f(\cdot) \). A decrease \( r \) results in further of \( EU_{ae} \) from \( EU_{re} \) by giving more weight to more recent periods. A similar insight is found for the value of travel time reliability in Figure 2b. Limited memory increases the willingness to pay for reductions in travel time variability and this penalty is amplified for increasing degrees of transience. Equation 18

---

\(^{14}\)The parameters of the underlying normal distribution are then defined by \( \delta_1 = \sqrt{\ln \left( 1 + \frac{1/16}{(1/3)^2} \right)} \) as the variance and \( \delta_2 = \ln \left( \frac{3}{4} \right) - \frac{\delta_1^2}{2} \) represents the mean.

\(^{15}\)The negative sign of implies that the individual is willing to pay for a reduction in travel time variability.
Figure 1: Variations in optimal departure time given K and r
indeed confirms that the distance between the two horizontal lines in the right panel will decrease for increasing $K$. The maximum distance between these two lines is in our case $\mathcal{E}_{\frac{\gamma_1^2}{2(\gamma_1 - \psi_1)}} = 19.51$ for $K = 1$. The latter would set a maximum value of $\mathcal{E}32.22$ on the value of travel time reliability, which is about 2.5 times higher than the rational expectations outcome. This outcome is somewhat extreme, but even at our lower bound $r = 0.5$ where the most recent period has about 50% impact the value of travel time reliability is already $\mathcal{E}19.63$, which is already 54% higher than the rational expectations prediction.

4.3. Accurate retrieval and anchored expectations

Transience has a clear impact on the parameters of interest, but so far we have neglected the presence of an anchor. It should already be clear from the formulas that people depart earlier when they have a higher anchor value (results not plotted). Moreover, an increase in $\tau$, reduces the impact of past travel time realizations on the optimal departure time and therefore by definition reduces the variance in the latter measure. Naturally, the deviation between the anchor point and mean travel time defines whether optimal departure time coincides with rational expectations for $\tau = 1$. The implications of the anchor point on expected utility and the value of travel time reliability are of more interest here. For this exercise we assume $r = 1$, resulting in $\rho_k = \frac{1}{K}$. Furthermore we assume $K = 5$.

Figure 3a illustrates that expected utility is quadratically decreasing in the distance between $\mu$ and the anchor point. The rational expectations and adaptive expectations utility measure only coincide at $T_a = \mu$ (i.e. 20 minutes). Reducing $\tau$ introduces a deviation from rational expectations due to allowing for transience. This effect is illustrated by the lower horizontal line, where the anchor point has no influence ($\tau = 0$). At any in-between level for $\tau$, the maximum level of expected utility approaches the rational expectations outcome, but still a quadratic penalty applies (but limited) for deviations $T_a \neq \mu$. The latter is illustrated by the curve at $\tau = 0.5$.

Since the value of travel time reliability is only influenced by $\tau$, Figure 3b plots $\tau$ on the horizontal axis. It follows directly from Equation 15 that the value of travel time variability reduces to the rational expectations outcome for $\tau = 1$, since there is no penalty for forming adaptive expectations. Reducing the value of $\tau$ enables the individual to form adaptive expectations, which increases the value of travel time reliability as a result of transience (see Figure 2). In other words, $\tau$ controls the distance between the two horizontal lines in Figure 2b. The effect is strongest for small values of $\tau$. The presence of an anchor point automatically implies that the maximum deviation between the value of travel time reliability in rational expectations and adaptive expectations reduces for any value of $r$.

4.4. Inaccurate retrieval and transience

Finally, we illustrate the implications of inaccurate retrieval on expected utility, where we allow $\bar{\nu}$ to vary between zero and $\sigma^2$. We set this upper bound on $\bar{\nu}$ given that deviations from actual realizations should not fall too much outside of the scale of $f(\cdot)$. Inaccurate
Figure 2: Alternative values of $r$
Figure 3: Alternative values of $T_a$
retrieval introduces an additional penalty on expected utility. Figure 4 points out that this penalty is linear in $\bar{\nu}$, i.e., more inaccuracy reduces expected utility. This effect is further amplified for smaller values of $r$, implying that adaptive expectations are formed only by more recent periods and thereby they prevent over- and under-recalled travel time realizations to cancel out against each other.

We can thus conclude that $r$ and $\tau$ play the most important role in deviations between adaptive and rational expectations in terms of optimal departure time, expected utility, and the value of travel time reliability. Based on Figure 2b we can also conclude that the value of variance may be underestimated by up to 45%, considering that $K = 5$ and $r = 0.6$ are realistic lower boundaries for $K$ and $r$, respectively. It may be further underestimated if retrievals are inaccurate, however, this effect seems to be relatively small (see Figure 4). In contrast, the presence of an anchor equal to $\mu$ would have the opposite effect: It would bring the value of reliability closer to the one derived if expectations are rational (especially for high levels of $\tau$: see Figure 3b).

5. Conclusions

This paper developed a model of decisions under risk, where expectations are formed on the basis of earlier stochastic experiences. We show how memory limitations carry over to additional dis-utility because of sub-optimal decision making. Our model shows that adaptive expectations arise in a stochastic environment as a result of limited memory, transience and retrieval inaccuracy. We applied our model to scheduling choices of travellers, where
travel times are stochastic and show that the value of reliability may be underestimated by up to 45%.

Because it is likely that memory properties of individuals are heterogeneous, our model may serve as an input for the modelling of complex dynamic systems, both in transport as well as in other fields of economics. In these models adaptive expectations often play a central role (see for example Watling and Cantarella (2013) for an overview of day-to-day dynamic transport systems and Hommes (2013) for an overview of adaptive expectations in financial markets).

Our dynamic memory model is stands apart from static behavioural models where individuals treat probabilities in a non-rational way. Our model predicts that travellers are sometime optimistic and sometimes pessimistic depending on their most recent experiences and corresponding retrieval probabilities. This in contrast to rank-dependent utility models that assume that optimism and pessimism are unrelated to earlier experiences (see Koster and Verhoef (2012) for a transport application).

A key assumption of our model is that the decision makers are not aware of their memory limitations. For decisions where the stakes are not so high, this may be a reasonable assumption. When the stakes are higher the decision maker may take a more reflective attitude and may well be choosing his anchor and anchor weight.

Although our model is fairly general, we made several restrictive assumptions in order to keep it analytically tractable. For example, we assumed that retrieval probabilities are independent of the values of states meaning that for example worse experiences do not count heavier than good experiences in the formation of expectation. Furthermore, we assumed that the experience of a new state does not affect the memory of the states in memory.

Finally, we consider the empirical testing of our decision model using laboratory and revealed preference data as a fruitful direction for further research. This may lead to more precise estimates of the biases related to memory limitations both in transport as well in other fields of economics.

Bibliography


**Appendix A.**

The outcome expected utility is given by Equation 5, which can be rewritten in closed-form using the expressions for the first and second moment:

$$
\int T_k f(T_k | \mu, \sigma_k^2) dT_k \equiv \mu,
$$

(A.1)

and

$$
\int T_k^2 f(T_k | \mu, \sigma_k^2) dT_k \equiv \mu^2 + \sigma_k^2.
$$

(A.2)

The expectation over all realisations of the travel time $T_0$ is then given by:

$$
\mathbb{E}_{T_0}(U(d_0, T_0)) = \int \left( - \int_{d_0}^{0} (\beta_0 + \beta_1 v) dv - \int_{0}^{d_0 + T_0} (\beta_0 + \gamma_1 v) dv \right) f(T_0 | \mu, \sigma_0^2) dT_0 = 
\frac{-\beta_0 \mu - \frac{1}{2} d_0^2 (\gamma_1 - \beta_1) - d_0 \gamma_1 \mu - \frac{1}{2} \gamma_1 (\mu^2 + \sigma_0^2)}{23}.
$$

(A.3)
Substituting the solution for the departure time (Equation 9) gives:

\[
\mathbb{E}_{T_0}(U(d_{0e}^e, T_0)) = -\beta_0 \mu - \frac{1}{2} \left( -\tau \frac{\gamma_1}{\gamma_1 - \beta_1} T_A - (1 - \tau) \frac{\gamma_1}{\gamma_1 - \beta_1} \sum_{k=1}^{K} \rho_k \bar{T}_k \right)^2 (\gamma_1 - \beta_1) \tag{A.4}
\]

\[- \left( -\tau \frac{\gamma_1}{\gamma_1 - \beta_1} T_A - (1 - \tau) \frac{\gamma_1}{\gamma_1 - \beta_1} \sum_{k=1}^{K} \rho_k \bar{T}_k \right) \gamma_1 \mu - \frac{1}{2} \gamma_1 (\mu^2 + \sigma_0^2).\]

First, we rewrite the second term in Equation A.4 with \((d_{0e}^e)^2\):

\[- \frac{1}{2} \left( -\tau \frac{\gamma_1}{\gamma_1 - \beta_1} T_A - (1 - \tau) \frac{\gamma_1}{\gamma_1 - \beta_1} \sum_{k=1}^{K} \rho_k \bar{T}_k \right)^2 (\gamma_1 - \beta_1) = \] \[- \frac{1}{2} \frac{\gamma_1^2}{\gamma_1 - \beta_1} \left( -\tau T_A - (1 - \tau) \sum_{k=1}^{K} \rho_k \bar{T}_k \right)^2 = \tag{A.5}
\]

\[- \frac{1}{2} \frac{\gamma_1^2}{\gamma_1 - \beta_1} \left( \tau^2 T_A^2 + 2\tau (1 - \tau) T_A \sum_{k=1}^{K} \rho_k \bar{T}_k + (1 - \tau)^2 \left( \sum_{k=1}^{K} \rho_k \bar{T}_k \right)^2 \right).\]

The last term of this expression is given by:

\[
\left( \sum_{k=1}^{K} \rho_k \bar{T}_k \right)^2 = \sum_{k=1}^{K} \rho_k^2 \bar{T}_k^2 + \sum_{m=1}^{K} \sum_{l=1}^{K} \rho_l \rho_m \bar{T}_l \bar{T}_m. \tag{A.6}
\]

Substituting in Equation A.5 gives the second term of Equation A.4:

\[- \frac{1}{2} \frac{\gamma_1^2}{\gamma_1 - \beta_1} \left( \tau^2 T_A^2 + 2\tau (1 - \tau) T_A \sum_{k=1}^{K} \rho_k \bar{T}_k + (1 - \tau)^2 \left( \sum_{k=1}^{K} \rho_k \bar{T}_k \right)^2 \right). \tag{A.7}
\]

Substituting this term in Equation A.4 gives:

\[
\mathbb{E}_{T_0}(U(d_0, T_0)) = -\beta_0 \mu - \frac{1}{2} \frac{\gamma_1^2}{\gamma_1 - \beta_1} \left( \tau^2 T_A^2 + 2\tau (1 - \tau) T_A \sum_{k=1}^{K} \rho_k \bar{T}_k \right) + \] \[- \frac{1}{2} \frac{\gamma_1^2}{\gamma_1 - \beta_1} \left( (1 - \tau)^2 \left( \sum_{k=1}^{K} \rho_k^2 \bar{T}_k^2 + \sum_{m=1}^{K} \sum_{l=1}^{K} \rho_l \rho_m \bar{T}_l \bar{T}_m \right) \right) \tag{A.8}
\]

\[- \left( -\tau \frac{\gamma_1}{\gamma_1 - \beta_1} T_A - (1 - \tau) \frac{\gamma_1}{\gamma_1 - \beta_1} \sum_{k=1}^{K} \rho_k \bar{T}_k \right) \gamma_1 \mu - \frac{1}{2} \gamma_1 (\mu^2 + \sigma_0^2).\]
To obtain $\mathbb{E}(U_{ae})$ we need to integrate Equation A.8 over all possible values of $\bar{T}_i$ and $\bar{T}_m$. It is assumed that the mean retrieval is unbiased, and therefore:

$$ \int \bar{T}_k g(\bar{T}_k | T_k, \nu_k^2) d\bar{T}_k \equiv T_k. \quad (A.9) $$

and furthermore we have:

$$ \int \bar{T}_k^2 g(\bar{T}_k | T_k, \nu_k^2) d\bar{T}_k \equiv T_k^2 + \nu_k^2. \quad (A.10) $$

Because the distributions $g(\cdot)$ are independent, the following identities can be used to get rid of the integrals:

$$ \int \ldots \int \left( \sum_{k=1}^{K} \rho_k \bar{T}_k \right) \prod_{k=1}^{K} g(\bar{T}_k | T_k, \nu_k^2) \ d\bar{T}_1 \ldots d\bar{T}_K = \sum_{k=1}^{K} \rho_k T_k. \quad (A.11) $$

$$ \int \ldots \int \left( \sum_{k=1}^{K} \rho_k^2 \bar{T}_k^2 \right) \prod_{k=1}^{K} g(\bar{T}_k | T_k, \nu_k^2) \ d\bar{T}_1 \ldots d\bar{T}_K = \sum_{k=1}^{K} \rho_k^2 (T_k^2 + \nu_k^2). \quad (A.12) $$

$$ \int \ldots \int \ldots \int \left( \sum_{m=1}^{K} \sum_{l=1}^{K} \rho_l \rho_m \bar{T}_l \bar{T}_m \right) \prod_{l=1}^{K} g(\bar{T}_l | T_l, \nu_l^2) \ d\bar{T}_1 \ldots d\bar{T}_K = \sum_{l=1}^{K} \sum_{m=1}^{K} \rho_l \rho_m \mu^2. \quad (A.13) $$

Using Equations A.11, A.12 and A.13 to integrate A.8 over all possible values of $\bar{T}_1 \ldots \bar{T}_K$, we obtain:

$$ \mathbb{E}_{\bar{T}_1 \ldots \bar{T}_K}(U_{ae}) = -\beta_0 \mu - \frac{1}{2} \frac{\gamma_1^2}{\gamma_1 - \beta_1} \left( \tau^2 T_A^2 + 2\tau (1-\tau) T_A \sum_{k=1}^{K} \rho_k T_k \right) + $$

$$ -\frac{1}{2} \frac{\gamma_1^2}{\gamma_1 - \beta_1} \left( (1-\tau)^2 \left( \sum_{k=1}^{K} \rho_k^2 (T_k^2 + \nu_k^2) + \sum_{l=1}^{K} \sum_{m=1 \neq l}^{K} \rho_l \rho_m \mu^2 \right) \right) \quad (A.14) $$

$$ - \left( -\tau \frac{\gamma_1}{\gamma_1 - \beta_1} T_A - (1-\tau) \frac{\gamma_1}{\gamma_1 - \beta_1} \sum_{k=1}^{K} \rho_k T_k \right) \gamma_1 \mu - \frac{1}{2} \gamma_1 (\mu^2 + \sigma_0^2). $$

The next step is to integrate over all possible realisations of travel times. Here we can use the identities as in Equations A.11, A.12 and A.13, with $\nu_k^2$ replaced by $\sigma_k^2$, and $T_k$ replaced
by $\mu$. This results in:
\[
\mathbb{E}(U_{ae}) \equiv \mathbb{E}_{T_1 \ldots T_K}(\mathbb{E}_{T_1 \ldots T_K}(U_{ae})) = -\beta_0 \mu - \frac{1}{2} \frac{\gamma_1^2}{\gamma_1 - \beta_1} \left( \tau^2 T_A^2 + 2\tau(1 - \tau)T_A \sum_{k=1}^{K} \rho_k \mu \right) + \\
- \frac{1}{2} \frac{\gamma_1^2}{\gamma_1 - \beta_1} \left( (1 - \tau)^2 \left( \sum_{k=1}^{K} \rho_k^2 (\sigma_k^2 + \mu^2 + \nu_k^2) \right) + \sum_{l=1}^{K} \sum_{m=1}^{K} \rho_l \rho_m \mu^2 \right) \\
- \left( -\tau \frac{\gamma_1}{\gamma_1 - \beta_1} T_A - (1 - \tau) \frac{\gamma_1}{\gamma_1 - \beta_1} \sum_{k=1}^{K} \rho_k \mu \right) \gamma_1 \mu - \frac{1}{2} \gamma_1 (\mu^2 + \sigma_0^2). 
\]  
(A.15)

Next we use:
\[
\sum_{k=1}^{K} \rho_k^2 (\sigma_k^2 + \mu^2 + \nu_k^2) + \sum_{l=1}^{K} \sum_{m=1}^{K} \rho_l \rho_m \mu^2 = \\
\sum_{k=1}^{K} \rho_k^2 \sigma_k^2 + \sum_{k=1}^{K} \rho_k^2 \nu_k^2 + \sum_{k=1}^{K} \rho_k^2 \mu^2 + \sum_{l=1}^{K} \sum_{m=1}^{K} \rho_l \rho_m \mu^2 - \sum_{l=1}^{K} \rho_l \mu^2 = \\
\sum_{k=1}^{K} \rho_k^2 \sigma_k^2 + \sum_{k=1}^{K} \rho_k^2 \nu_k^2 + \sum_{l=1}^{K} \sum_{m=1}^{K} \rho_l \rho_m \mu^2 = \\
\sum_{k=1}^{K} \rho_k^2 \sigma_k^2 + \sum_{k=1}^{K} \rho_k^2 \nu_k^2 + \mu^2. 
\]  
(A.16)

Substituting in Equation A.15 and simplifying gives:
\[
\mathbb{E}(U_{ae}) = -\beta_0 \mu - \frac{1}{2} \frac{\gamma_1^2}{\gamma_1 - \beta_1} \left( \tau^2 T_A^2 + 2\tau(1 - \tau)T_A \mu \right) + \\
- \frac{1}{2} \frac{\gamma_1^2}{\gamma_1 - \beta_1} \left( (1 - \tau)^2 \left( \sum_{k=1}^{K} \rho_k^2 \sigma_k^2 + \sum_{k=1}^{K} \rho_k^2 \nu_k^2 + \mu^2 \right) \right) \\
- \frac{\gamma_1^2}{\gamma_1 - \beta_1} \left( -\tau T_A - (1 - \tau) \mu \right) \mu - \frac{1}{2} \gamma_1 (\mu^2 + \sigma_0^2). 
\]  
(A.17)

Collecting terms gives:
\[
\mathbb{E}(U_{ae}) = -\beta_0 \mu - \frac{1}{2} \frac{\gamma_1^2}{\gamma_1 - \beta_1} \left( \tau^2 T_A^2 + 2\tau(1 - \tau)T_A \mu - 2\tau T_A \mu - 2(1 - \tau) \mu^2 + (1 - \tau)^2 \mu^2 \right) \\
- \frac{1}{2} \frac{\gamma_1^2}{\gamma_1 - \beta_1} (1 - \tau)^2 \left( \sum_{k=1}^{K} \rho_k^2 \sigma_k^2 + \sum_{k=1}^{K} \rho_k^2 \nu_k^2 \right) - \frac{1}{2} \gamma_1 (\mu^2 + \sigma_0^2). 
\]  
(A.18)
Collecting terms around $T_A$, $\mu$ and $\mu^2$ gives:

\[
\mathbb{E}(U_{ae}) = -\beta_0\mu - \frac{1}{2} \frac{\gamma_1^2}{\gamma_1 - \beta_1} \tau^2 T_A^2 + \frac{\gamma_1^2}{\gamma_1 - \beta_1} \tau^2 T_A \mu \\
+ \frac{1}{2} \frac{\beta_1 \gamma_1}{\gamma_1 - \beta_1} \mu^2 - \frac{1}{2} \frac{\gamma_1^2}{\gamma_1 - \beta_1} \tau^2 \mu^2 \\
- \frac{1}{2} \frac{\gamma_1^2}{\gamma_1 - \beta_1} (1 - \tau)^2 \left( \sum_{k=1}^{K} \rho_k^2 (\sigma_k^2 + \nu_k^2) \right) - \frac{1}{2} \gamma_1 \sigma_0^2
\]  

(A.19)

Substituting $T_A = \mu + a$ and simplifying gives:

\[
\mathbb{E}(U_{ae}) = -\beta_0\mu - \frac{1}{2} \frac{\beta_1 \gamma_1}{\gamma_1 - \beta_1} \mu^2 - \frac{1}{2} \gamma_1 \sigma_0^2 - \frac{1}{2} \frac{\gamma_1^2}{\gamma_1 - \beta_1} \tau^2 a^2 \\
- \frac{1}{2} \frac{\gamma_1^2}{\gamma_1 - \beta_1} (1 - \tau)^2 \left( \sum_{k=1}^{K} \rho_k^2 (\sigma_k^2 + \nu_k^2) \right)
\]  

(A.20)

The first part of Equation A.20 is given by the optimal expected utility with rational expectations (Equation 6), and therefore we have:

\[
\mathbb{E}(U_{ae}) = \mathbb{E}(U_{re}) - \frac{1}{2} \frac{\gamma_1^2}{\gamma_1 - \beta_1} \tau^2 a^2 - \frac{1}{2} \frac{\gamma_1^2}{\gamma_1 - \beta_1} (1 - \tau)^2 \left( \sum_{k=1}^{K} \rho_k^2 (\sigma_k^2 + \nu_k^2) \right)
\]  

(A.21)

This concludes the proof.