Article:

Koster, P, Peer, S and Dekker, T (2015) Memory, expectation formation and scheduling choices. Economics of Transportation, 4 (4). pp. 256-265. ISSN 2212-0122

https://doi.org/10.1016/j.ecotra.2015.09.001

Memory, expectation formation and scheduling choices☆

Paul Koster[a,b,*], Stefanie Peer[c], Thijs Dekker[d]

[a]*Department of Spatial Economics, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands*
[b]*Tinbergen Institute, Gustav Mahlerplein 117, 1082 MS Amsterdam, The Netherlands*
[c]*Vienna University of Business and Economics, Welthandelsplatz 1, 1020 Vienna, Austria*
[d]*University of Leeds, Institute for Transport Studies, 36-40 University Road, Leeds, LS2 9JT, UK.*

Abstract

Limited memory capacity, retrieval constraints and anchoring are central to expectation formation processes. We develop a model of adaptive expectations where individuals are able to store only a finite number of past experiences of a stochastic state variable. Retrieval of these experiences is probabilistic and subject to error. We apply the model to scheduling choices of commuters and demonstrate that memory constraints lead to sub-optimal choices. We analytically and numerically show how memory-based adaptive expectations may substantially increase commuters' willingness-to-pay for reductions in travel time variability, relative to the rational expectations outcome.

Article published as: Koster, P., Peer, S. and Dekker,T. (2015). Memory, expectation formation and scheduling choices. Economics of Transportation. http://dx.doi.org/10.1016/j.ecotra.2015.09.001

Keywords: Memory, Transience, Expectation formation, Adaptive expectations, Retrieval Accuracy, Scheduling, Value of Reliability

☆We like to thank two anonymous referees for very helpful suggestions for improvements. The paper also benefited from suggestions and comments of Alexander Muermann, Hans Koster, Peter Nijkamp, participants of the 2014 conference of the European Association for Research in Transportation (hEART) in Leeds, participants of the 2014 conference of the International Transportation Economic Association (ITEA) in Toulouse, participants of a TLO seminar at Delft University. Paul Koster gratefully acknowledges the financial support of the ERC (Advanced Grant OPTION # 246969).

*Corresponding author. Fax +31 20 5986004, phone +31 20 5984847.

Email addresses: `p.r.koster@vu.nl` (Paul Koster), `speer@wu.ac.at` (Stefanie Peer), `t.dekker@leeds.ac.uk` (Thijs Dekker)

1. Introduction

Imperfect knowledge regarding the true distribution of stochastic state variables, like product quality or travel times, induces individuals to form expectations based on personal experiences and external sources of information. Memory processes are known to influence expectation formation processes (e.g. Hirshleifer and Welch, 2002; Mullainathan, 2002; Wilson, 2003; Sarafidis, 2007) and anchoring constitutes a persistent phenomenon in human behaviour (Wilson et al., 1996; Strack and Mussweiler, 1997; Furnham and Boo, 2011).[1]

This paper develops an adaptive expectations model which explicitly accounts for limited cognitive abilities of decision makers. Expectation formation in our model has the following properties. First, decision makers are assumed to have limited memory, such that only a fixed number of past experiences can be stored. Second, retrieving experiences from memory is probabilistic and decision makers experience difficulty in retrieving more distant experiences; a phenomenon often referred to as transience (Horowitz, 1984; Barucci, 1999, 2000; Schacter, 2002). Third, retrieval may be inaccurate, meaning that retrieved experiences may not correspond to the original experiences. Transience and retrieval inaccuracy are both forms of memory decay. Fourth, decision makers prime their expectations using exogenous anchors. The inclusion of past experiences, limited cognitive abilities and anchoring in the expectation formation model provides a significant deviation of the rational expectations model.

We apply the model to scheduling decisions of commuters facing stochastic daily travel times. Commuters experience dis-utility from travel time variability, as it induces them to depart and/or arrive earlier or later than preferred (e.g. Vickrey, 1969; Small, 1982, 1992; Noland and Small, 1995). The developed model provides a better understanding of empirical findings that hint at the presence of adaptive expectations and anchors in the context of travel related scheduling decisions. For example, Bogers et al. (2007) and Ben-Elia and Shiftan (2010) provide evidence that recently experienced travel times have an over-proportionally large influence on travel decisions. Peer et al. (2015) find that commuters take into account the long-run travel time average as well as day-specific traffic information in their scheduling decisions.

The value commuters attach to a marginal reduction in travel time variability is referred to as the value of (travel time) reliability and can be inferred from observed scheduling choices (Fosgerau and Karlström, 2010; Fosgerau and Engelson, 2011). Typically, the value of reliability is derived using the presumption that commuters have rational expectations and an infinite memory. We find that with adaptive travel time expectations this value of reliability is higher, because sub-optimal scheduling decisions are made. Therefore, improvements in reliability are associated with larger benefits, because they make commuters depart and arrive closer to the times they prefer and decrease the variability in departure times. Empirical revealed preference studies using reduced-form utility functions are likely to already capture these behavioural biases in the coefficient that is estimated for travel

[1] Often anchors corresponds to the information that is obtained first, which is then used as a 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.

time variation. Our results are therefore mainly important for current stated preference practice that ignores the process of expectation formation: our numerical illustration shows that these values of reliability can underestimate our bounded rationality value of reliability by up to 45%, suggesting that the welfare effects of memory biases may be substantial.

Underestimation of the value of reliability may have significant implications for cost-benefit assessments of transport policies. Namely, the benefits from improvements in travel time reliability in road-related transport projects amount to ca. 25% 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% of total user benefits in transport appraisals (Hensher, 2001).

While we apply our model to scheduling choices of commuters, it may very well be relevant to other fields of economics, such as for the study of the effects of heterogeneous expectation formation on (dis)equilibrium in dynamic economic systems (see Hommes (2013)) or for the analysis of repetitive consumer choices with uncertain product quality. Note that bounded rationality in our model is exclusively caused by limited cognitive abilities rather than judgement errors due to selective memory (Gennaioli and Shleifer, 2010) or probability weighting. Therefore this paper stands apart from works modelling bounded rationality as a result of satisficing (Simon, 1955; Caplin et al., 2011), self-deception (Bénabou and Tirole, 2002), or optimal belief formation when the decision utility is affected by anticipatory emotions (Brunnermeier and Parker, 2005; Bernheim and Thomadsen, 2005) as well as by (ex-post) disappointment (Gollier and Muermann, 2010).

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. 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 description of the model

Consider a decision-maker who decides on x_0, where the subscript 0 indicates that the decision is made for the time period to come. Outcome utility $U(x_0, s_0)$ is assumed to be continuous and strictly concave in x_0, and depends on the stochastic state s_0. Let $f(s_0|\omega_0)$ be the probability density function of s_0, where ω_0 is a vector of parameters that characterizes $f(.)$. Expected outcome utility is then defined as:

$$\mathbb{E}(U(x_0, s_0)) = \int U(x_0, s_0) f(s_0|\omega_0) ds_0 \tag{1}$$

With rational expectations, the decision maker knows the distribution $f(s_0|\omega_0)$ and maximizes Equation 1 to decide on x_0. In what follows, we denote x_0^{re} as the optimal choice under rational expectations, and $\mathbb{E}(U_{re}) \equiv \mathbb{E}(U(x_0^{re}, s_0))$ as the corresponding maximal expected utility. Deviations from rational expectations are introduced by assuming that the decision maker has imperfect knowledge regarding $f(s_0|\omega_0)$. In our model, she forms adaptive expectations regarding s_0, using past experiences in combination with primed expectations. Past experiences are denoted by past stochastic realisations of s_k, which are draws from

$f(s_k|\omega_k)$. A higher value of the index k refers to a more distant experience. Primed expectations enter the model in the form of an anchor state s_A. In contrast to the states stored in the decision maker's memory and the corresponding retrievals, the anchor is assumed to be non-stochastic and is a stable element in the expectation formation process.

The decision maker is assumed to have limited cognitive abilities. First, it is assumed that she has a *limited* memory, meaning that only K past experiences $s_1 ... s_K$ are stored in memory. Second, it is assumed that the realisation of s_k is correctly stored in memory, but a stored state can only be retrieved with a probability $\rho_k > 0$. Following Schacter (2002), this allows us to assume that more recent experiences can be retrieved more easily, i.e. $\rho_1 > \rho_2 > ... > \rho_K$. We refer to this phenomena as *transience*. Third, retrieval of the stored states $s_1 ... s_K$ may be *inaccurate*. Instead of $s_1 ... s_K$, the decision maker retrieves $\bar{s}_1 ... \bar{s}_K$ from her memory. Let $g_k(\bar{s}_k|s_k, \phi_k)$ be the retrieval density function, with ϕ_k and s_k as its characterizing parameters. Fourth, *anchoring* is present. The anchor reflects an exogenous, stable belief concerning travel time that is independent of new experiences and the current traffic situation. While we do not model the origin of the anchor explicitly in order to keep the model generic, the anchor could for example be driven by stable publicly available information.

Equation 2 defines the expected decision utility as the weighted average of utilities across the anchor and the set of retrieved states:

$$U^d(.) = \tau U(x_0, s_A) + (1-\tau) \sum_{k=1}^{K} \rho_k U(x_0, \bar{s}_k), \tag{2}$$

where $\sum_{k=1}^{K} \rho_k = 1$. In this equation, τ is the weight assigned to the anchor. When $\tau = 0$, expectations are fully adaptive and when $\tau = 1$, the decision maker ignores her earlier experiences and expected decision utility is solely based on the anchor s_A and the choice of x_0. Equation 2 mimics Equation 1 when $\tau \to 0$, $\rho_k = 1/K$, $\bar{s_k} = s_k$ and $K \to \infty$. Rational expectations are therefore a special case of our model. The decision maker maximizes Equation 2 with respect to x_0. Denote this optimal x_0 by x_0^{ae}, where the ae superscript refers to the fact that the decision maker uses adaptive expectations.[2] Decisions on x_0 are sub-optimal whenever $x_0^{ae} \neq x_0^{re}$. Nevertheless, the situation could arise where $x_0^{ae} = x_0^{re}$, i.e. the decision maker 'coincidentally' makes the optimal choice.

Suppose that we need to make a prediction of the expected outcome utility of the decision maker. This prediction has to 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 take the expected value over all possible combinations of experienced and retrieved states. Mathematically this is tedious, since it involves a $2K+1$ dimensional integral over all possible values of the $K+1$ realised states $s_0 ... s_K$, and the K

[2] A unique solution for x_0^{ae} exists since Equation 2 is a weighted average of strictly concave functions.

possible values of retrieved states $\bar{s}_1 ... \bar{s}_K$:

$$\begin{aligned}\mathbb{E}(U_{ae}) &\equiv \mathbb{E}\left(U(x_0^{ae}, s_0)\right) \\ &= \int ... \int \left(\int ... \int U(x_0^{ae}, s_0) \prod_{k=1}^{K} g_k\left(\bar{s}_k | s_k, \phi_k\right) d\bar{s}_1 ... \bar{s}_K \right) \prod_{k=0}^{K} f(s_k, \omega_k) ds_0 ... ds_K.\end{aligned} \tag{3}$$

This equation obviously has the disadvantage that it is less parsimonious than its rational expectations counterpart, i.e. Equation 1 with x_0^{re}. Nevertheless, this generic set-up helps to structure our thoughts about how earlier experiences and retrieval inaccuracy affect predictions of the expected outcome utility. The next section makes analytical progress by putting more structure on the utility function $U(.)$ and derives an analytical representation of the predicted expected outcome utility $\mathbb{E}(U_{ae})$ for the case of commuters choosing departure times when travel times are stochastic.

3. Memory and the value of travel time reliability

We apply our memory-based adaptive expectation formation model to commuters' scheduling behaviour with stochastic travel times. Commuters face scheduling costs of travel time variability due to departing and/or arriving earlier or later than desired. Noland and Small (1995) were the first to extend the scheduling model of Vickrey (1969) and Small (1982) to expected utility maximization. Their model was recently extended by Fosgerau and Karlström (2010) and Fosgerau and Engelson (2011) who proved that the optimal expected outcome utility depends linearly on some measure of travel time reliability. Here, we extend the results of Fosgerau and Engelson (2011) by showing that this result carries over to the case when memory biases and anchoring are present and the travel time distribution is stable over time. Existing literature on travel time expectation formation typically focuses on learning and perception updating mechanisms in route choice but often ignores the psychological foundation of the adaptation of expectations (e.g. Jha et al., 1998; Arentze and Timmermans, 2003; Chen and Mahmassani, 2004; Avineri and Prashker, 2005; Arentze and Timmermans, 2005; Bogers et al., 2007; Ben-Elia and Shiftan, 2010). Moreover, most existing studies do not quantify behavioural and valuation biases, even when they find that travel time expectations are adaptive. Therefore it is unclear if choice models assuming rational expectations can be viewed as a good approximation of individual choice behaviour. This paper explicitly focuses on the origins of adaptive travel time expectations and characterizes the resulting behavioural and valuation biases.

3.1. Rational expectations

We assume commuters derive utility from being at home, for instance by spending more time with the family, sleeping or having a longer breakfast. Departing earlier or later than preferred therefore reduces utility. Similarly, utility at work is derived from productive work time, which is reduced by arriving later than preferred. An increase in travel time therefore reduces utility on either end. This specification of utility was first introduced by Vickrey (1973) and later used by Fosgerau and Engelson (2011) to derive the value of reductions in travel time variance. Tseng and Verhoef (2008) were the first to find empirical evidence for such scheduling preferences.

Equation 4 describes outcome utility for a given departure time d_0 and a realisation of travel time T_0, where $H'(v)$ is the marginal utility for being at home and $W'(v)$ is the marginal utility for being at work as functions of clock time v. The first part of Equation 4 shows the utility from time spent at home where being at home starts at v_h and ends when the travellers departs at d_0. The second integral gives the utility for being at work, where being at work starts at arrival time $d_0 + T_0$ and ends at v_w. This implies that v_h and v_m span the range of possible departure and arrival times (Börjesson et al. (2012)).[3]

$$V(d_0|T_0) = \int_{v_h}^{d_0} H'(v)dv + \int_{d_0+T_0}^{v_w} W'(v)dv = H(d_0) - H(v_h) + W(v_w) - W(d_0 + T_0). \quad (4)$$

For the remainder of the paper we assume simple linear functional forms for the marginal utilities. Using the normalisation of Börjesson et al. (2012) we have:[4]

$$U(d_0|T_0) = -\int_{d_0}^{0} (\beta_0 + \beta_1 v)dv - \int_{0}^{d_0+T_0} (\beta_0 + \gamma_1 v)dv. \quad (5)$$

For a trip to occur it must hold that $\gamma_1 > \beta_1$. Usually the marginal utility of being at home is decreasing in v, implying $\beta_1 < 0$, whereas the marginal utility of being at work is increasing ($\gamma_1 > 0$). With rational expectations, commuters know the distribution of travel times which is defined by $f(T_0|\mu, \sigma^2)$, where μ is the mean travel time and σ^2 the travel time variance. Accordingly, the expected outcome utility is defined by:

$$\mathbb{E}(U(d_0|T_0)) = \int U(d_0|T_0) f(T_0|\mu, \sigma^2)dT_0. \quad (6)$$

Fosgerau and Engelson (2011) show that when the departure time is optimally chosen, the commuter departs at:

$$d_0^{re} = -\frac{\gamma_1}{\gamma_1 - \beta_1}\mu, \quad (7)$$

resulting in optimal expected outcome utility:

$$\mathbb{E}(U_{re}) \equiv \mathbb{E}(U(d_0^{re}|T_0)) = -\beta_0\mu + \frac{1}{2}\frac{\beta_1\gamma_1}{\gamma_1 - \beta_1}\mu^2 - \frac{1}{2}\gamma_1\sigma^2. \quad (8)$$

[3]For a graphical representation we refer to Tseng and Verhoef (2008), Fosgerau and Engelson (2011) and Börjesson et al. (2012).

[4]Following Börjesson et al. (2012), we normalise utility relative to $V(0|0)$, by defining $U(d_0|T_0) = V(d_0|T_0) - V(0|0)$. This allows us to evaluate $U(d_0|T_0)$ in terms of bounds at 0 rather than at v_h and v_m:

$$U(d_0|T_0) = -\int_{d_0}^{0} H'(v)dv - \int_{0}^{d_0+T_0} W'(v)dv = H(d_0) - H(0) + W(0) - W(d_0 + T_0)$$

Furthermore, we normalise $W(0) - H(0)$ to 0, because this part of utility is independent of departure time d_0 and travel time T_0. We do so by assuming that $H'(v)$ and $W'(v)$ have the same intercept.

This optimal expected outcome utility is a simple function of the mean delay and the travel time variance. Equation 8 does not require any distributional assumptions on the travel time distribution (except that μ and σ^2 are finite). We define the value of reliability (VOR) as the value attached to a marginal decrease in the travel time variance:[5]

$$\text{VOR}_{re} = -\frac{\partial \mathbb{E}(U_{re})}{\partial \sigma^2} = \frac{1}{2}\gamma_1. \tag{9}$$

3.2. Adaptive expectations

Adaptive expectations on the distribution of travel times are based on past travel times stored in memory $T_1 ... T_K$ and the retrievals of these past states $\bar{T}_1 ... \bar{T}_K$. Every retrieval is assumed to be an additive function of the retrieval error and the realised travel time: $\bar{T}_k = T_k + \epsilon_k$, with $\mathbb{E}(\epsilon_k) = 0$, meaning that retrieval is on average correct. The travel times in memory are realizations from $f(T_k|\mu, \sigma^2)$. The accuracy of retrieval ϵ_k is governed by the probability density function $g(\bar{T}_k|T_k, \nu_k^2)$ where $\bar{T}_k$ has mean $\mathbb{E}(\bar{T}_k) = T_k$ and conditional variance $\mathbb{VAR}(\bar{T}_k|T_k) = \nu_k^2$. Using the law of total variance, the unconditional variance of $\bar{T}_k$ is given by: $\mathbb{VAR}(\bar{T}_k) = \sigma^2 + \nu_k^2$.[6] For every day t the commuter has to decide on the departure time and creates a new set of recalled memories from the stored set of past experiences. For large K, the set of past travel time experiences stored in memory for days t and $t+1$ is nearly identical as the experienced travel time at t only replaces a single experience previously stored in memory. The similarity in available memories in combination with transience introduces correlation in the recalled sets, but the process of the recollection and accuracy of these recollections are completely independent between t and $t+1$.

The commuter 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 μ. With adaptive expectations, commuters choose their optimal departure using decision utility

$$U^d(.) = \tau U(d_0, T_A) + (1-\tau) \sum_{k=1}^{K} \rho_k U(d_0, \bar{T}_k), \tag{10}$$

where $\sum_{k=1}^{K} \rho_k = 1$. Solving the first-order condition $\frac{\partial U^d(.)}{\partial d_0} = 0$ gives:

$$d_0^{ae} = -\tau \frac{\gamma_1}{\gamma_1 - \beta_1} T_A - (1-\tau) \frac{\gamma_1}{\gamma_1 - \beta_1} \sum_{k=1}^{K} \rho_k \bar{T}_k. \tag{11}$$

---

[5]For plausibility of the model, additional restrictions may be imposed 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 increases in mean travel time would lead to a higher expected utility.

[6] We assume $\mathbb{COV}(T_k, T_l) = 0$, $\mathbb{COV}(T_k, \epsilon_l) = 0$, and $\mathbb{COV}(\epsilon_k, \epsilon_l) = 0$, $\forall k \neq l$. Together with the assumption that $\mathbb{E}(\epsilon_k) = 0, \forall k$, this results in $\mathbb{COV}(\bar{T}_k, \bar{T}_l) = \mathbb{E}(\bar{T}_k \bar{T}_l) - \mathbb{E}(\bar{T}_k)\mathbb{E}(\bar{T}_l) = \mathbb{E}(T_k T_l) + \mathbb{E}(T_k \epsilon_l) + \mathbb{E}(T_l \epsilon_k) + \mathbb{E}(\epsilon_k \epsilon_l) - \mathbb{E}(T_k)\mathbb{E}(T_l) = 0, \forall k \neq l$. Relaxing these assumptions is an interesting avenue for future research.

The effect of the anchor on departure time choice is captured by the first term, and the effect of limited memory by the second term. A higher K indicates that the commuter is able to store more past travel times. Stored travel times are retrieved with probability ρ_k. Retrieval accuracy enters the departure time choice via the retrieved travel times $\bar{T}_k$. The mean departure time is given by:

$$\mathbb{E}\left(d_0^{ae}\right) = -\tau\frac{\gamma_1}{\gamma_1-\beta_1}T_A - (1-\tau)\frac{\gamma_1}{\gamma_1-\beta_1}\mu, \tag{12}$$

which reduces to d_0^{re} for $T_A = \mu$ $(a = 0)$ (see Equation 7). The variability in departure time choices over time periods is influenced by the variance of travel times and the variance of retrieval inaccuracy. A higher variance in travel times and a higher retrieval inaccuracy result in more variable departure times:[7]

$$\mathbb{VAR}(d_0^{ae}) = (1-\tau)^2\left(\frac{\gamma_1}{\gamma_1-\beta_1}\right)^2\left(\sigma^2\sum_{k=1}^{K}\rho_k^2 + \sum_{k=1}^{K}\rho_k^2\nu_k^2\right). \tag{13}$$

Our model therefore predicts that the variability in departure times increases for increasing variances of travel time and retrieval. This effect is multiplied with the quadratic retrieval probabilities. When transience is stronger, retrieval probabilities will be more unequal, resulting in more volatile behaviour. A higher anchor parameter τ results in less variable departure times because memory biases count less heavily in the decision utility function.
The prediction of the expected outcome utility can be found by integrating over all possible combinations of $T_0, T_1 ... T_K$ and the corresponding stochastic retrievals $\bar{T}_1 ... \bar{T}_K$ (i.e. in a similar way as Equation 3). In Appendix A we show that the predicted expected outcome utility $\mathbb{E}(U_{ae})$ can be written as:[8]

$$\begin{aligned}\mathbb{E}(U_{ae}) &= \mathbb{E}(U_{re}) - \frac{1}{2}\frac{\gamma_1^2}{\gamma_1-\beta_1}\tau^2a^2 - \frac{1}{2}(\gamma_1-\beta_1)\mathbb{VAR}(d_0^{ae}) \\ &= \mathbb{E}(U_{re}) - \frac{1}{2}\frac{\gamma_1^2}{\gamma_1-\beta_1}\tau^2a^2 - \frac{1}{2}(1-\tau)^2\frac{\gamma_1^2}{\gamma_1-\beta_1}\left(\sigma^2\sum_{k=1}^{K}\rho_k^2 + \sum_{k=1}^{K}\rho_k^2\nu_k^2\right).\end{aligned} \tag{14}$$

The first term in Equation 14 is the optimal expected utility with rational expectations (Equation 8). The second term reflects a penalty for relying on an anchor when deciding on the optimal departure time. This penalty only arises when $a \neq 0$, and increases quadratically in the value of a and the anchor parameter τ. When $\tau \to 1$ and $T_A = \mu$, the optimal departure time Equation 11 is equal to the departure time with rational expectations and Equation 14 reduces to Equation 8.

[7]Here we use the assumptions on covariances in footnote 6. When relaxing these assumption additional covariance terms would enter this equation.

[8]The model developed in Appendix A is more general because it assumes a k-specific travel time distribution. We leave this out in the discussion because it gives rise to another type of memory bias related to variations of travel time distributions over subsequent days.

Adaptive expectations are associated with an additional term that is proportionally decreasing in the variance of departure time $\mathbb{VAR}(d_0^{ae})$. More volatile behaviour therefore decreases the predicted expected outcome utility. First, an increase in the travel time variance results in additional dis-utility because of limited memory. Second, K enters the summation over all retrieval variances ν_k. Accordingly, the retrieval variances of the last K time periods decrease expected outcome utility. The negative effect of transience and inaccurate retrieval becomes stronger when commuters' expectations become more adaptive (i.e. when τ decreases).

Equation 14 is derived for general values of retrieval probabilities, travel time variance and retrieval variances. When imposing transience, we have $\rho_1 > \rho_2 > ... > \rho_K$, such that more recent travel time and retrieval variances play a larger role than more distant travel time and retrieval variances in Equation 14. Because retrieval probabilities enter quadratically, transience always reduces utility when retrieval is accurate. However, when retrieval variances are high for more distant memories (i.e. higher values of k), transience may reduce the bias of inaccurate retrieval.

The anchor parameter τ has two roles in Equation 14. Given $T_A \neq \mu$, an increase in τ is associated with a decrease in expected utility due to the sub-optimal choice of T_A. On the other hand, an increase in τ may lead to a decrease in the bias related to transience and retrieval inaccuracy, since past experiences have a lower effect on travel time expectations.

3.3. The value of travel time reliability

The VOR with adaptive expectation is given by:

$$\text{VOR}_{ae} = -\frac{\partial \mathbb{E}(U_{ae})}{\partial \sigma^2} = \frac{1}{2}\gamma_1 + \frac{1}{2}\frac{\gamma_1^2}{\gamma_1 - \beta_1}(1-\tau)^2 \sum_{k=1}^{K} \rho_k^2, \tag{15}$$

where it is assumed that τ is exogenous. As with rational expectations (see 9), the expected outcome utility is linearly decreasing in the travel time variance. The convenient result of Fosgerau and Engelson (2011) thus carries over to the case of adaptive expectations. While the VOR is not affected by retrieval inaccuracy, it is affected by transience and the anchor weight τ. Regarding τ, it is easy to see that the VOR increases as the weight attached to the anchor decreases and expectations thus become more adaptive. Clearly, if $\tau \to 1$ (and hence only the anchor counts), the VOR is not any longer affected by the transience parameter ρ_k.

It is useful to parametrize the retrieval probabilities. These probabilities need to sum up to 1 for any chosen value of $K = 1...\infty$, and for transience to apply, the probabilities need to be decreasing in k, because more recent travel times will then have a higher likelihood of being remembered. A functional form that satisfies these conditions is given by:

$$\rho_k = \frac{r-1}{r(r^K - 1)} r^k, \tag{16}$$

where $0 < r < 1$. 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 limiting case, because $\lim_{r\to 1} \rho_k = \frac{1}{K}$. If we assume that retrieval probabilities are defined by Equation 16, the VOR_{ae} is given by:

$$\text{VOR}_{ae,r} = \frac{1}{2}\gamma_1 + \frac{1}{2}\frac{\gamma_1^2}{\gamma_1 - \beta_1}(1-\tau)^2 \frac{(1-r)(1+r^K)}{(1+r)(1-r^K)}, \tag{17}$$

which is decreasing in the transience parameter r, meaning that more unequal retrieval probabilities increase the value attached to reliable travel times. The limiting case $r \to 1$ gives retrieval probabilities equal to $1/K$. The value of travel time variance is then given by

$$\lim_{r\to 1} \text{VOR}_{ae,r} = \frac{1}{2}\gamma_1 + \frac{1}{2}\frac{\gamma_1^2}{\gamma_1 - \beta_1}(1-\tau)^2 \frac{1}{K}, \tag{18}$$

and therefore in the absence of transience the additional effect of limited memory on the VOR_{ae} is proportional to $\frac{1}{K}$. As expected, the behavioural bias due to limited memory then vanishes when $K \to \infty$, and $\text{VOR}_{ae,r}$ reduces to 9.

3.4. The value of retrieval accuracy

The expected outcome utility (see Equation 14) shows that it is valuable for commuters to have a higher accuracy of retrieval. The value of retrieval accuracy (VORA) for retrieval m is defined as the first derivative of Equation 14 with respect to ν_m^2, multiplied by (-1):

$$\text{VORA}_m = -\frac{\partial \mathbb{E}(U_{ae})}{\partial \nu_m^2} = \frac{1}{2}\frac{\gamma_1^2}{\gamma_1 - \beta_1}(1-\tau)^2 \rho_m^2, \forall m = 1...K. \tag{19}$$

The VORA increases when expectations become more adaptive ($\tau \to 0$), and when that particular experience has a larger impact on the expected outcome utility. A more explicit expression can be derived by replacing ρ_m in Equation 19 by Equation 16:

$$\text{VORA}_{m,r} = \frac{1}{2}\frac{\gamma_1^2}{\gamma_1 - \beta_1}(1-\tau)^2 \left(\frac{r-1}{r(r^K-1)} r^m\right)^2, \forall m = 1...K. \tag{20}$$

In line with intuition, when transience is present, the VORA is lower for more distant memories (higher m). This illustrates the interesting interplay between transience and the effects of retrieval inaccuracy.

3.5. The limiting case of $K \to \infty$ and $\nu_k^2 \to \bar{\nu}k$

This subsection develops a limiting case that may be useful for practical applications. It is assumed 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 with slope $\bar{\nu}$:

$$\nu_k^2 = \bar{\nu}k. \tag{21}$$

If we substitute Equations 16 and 21 in Equation 14 we obtain a parsimonious expression for the expected outcome utility under infinite memory as a function of the transience parameter r and retrieval variance parameter $\bar{\nu}$[9]:

$$\begin{aligned}\mathbb{E}(U_{ae}) = \mathbb{E}(U_{re}) &- \frac{1}{2}\frac{\gamma_1^2}{\gamma_1-\beta_1}\tau^2 a^2 - \frac{1}{2}\frac{\gamma_1^2}{\gamma_1-\beta_1}(1-\tau)^2\frac{1-r}{1+r}\sigma^2 \\ &- \frac{1}{2}\frac{\gamma_1^2}{\gamma_1-\beta_1}(1-\tau)^2\frac{\bar{\nu}}{(1+r)^2}.\end{aligned} \tag{22}$$

The biases due to transience and inaccurate recall do not vanish when memory capacity is unlimited. Equation 22 does show 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 retrieval inaccuracy does not. Accordingly, infinite memory is not a sufficient assumption for rational expectations.

3.6. Endogenous choice of τ

This section generalizes the model to allow for the choice of the anchor parameter, which corresponds to the situation where the decision-maker is aware of her memory limitations. For $a \neq 0$, the decision-maker will trade-off the bias related to the anchor with the memory biases. The change in expected utility for a marginal change in τ is given by:

$$\frac{\partial \mathbb{E}(U_{ae})}{\partial \tau} = -\frac{\gamma_1^2}{\gamma_1-\beta_1}\tau a^2 + \frac{\gamma_1^2}{\gamma_1-\beta_1}(1-\tau)\left(\sigma^2\sum_{k=1}^{K}\rho_k^2 + \sum_{k=1}^{K}\rho_k^2\nu_k^2\right). \tag{23}$$

Solving the first-order condition $\frac{\partial \mathbb{E}(U_{ae})}{\partial \tau} = 0$ results in:

$$\tau^* = \frac{\sum_{k=1}^{K}\rho_k^2\sigma^2 + \sum_{k=1}^{K}\rho_k^2\nu_k^2}{\sum_{k=1}^{K}\rho_k^2\sigma^2 + \sum_{k=1}^{K}\rho_k^2\nu_k^2 + a^2} \tag{24}$$

which is equal to 1 if $a = 0$, and independent of scheduling preferences. The optimal anchor parameter weighs the dis-utility related to imprecision due to anchoring with the dis-utility related to memory biases. Incorporating an anchor (i.e. $\tau^* \neq 0$) is therefore a rational response to cope with imprecision in knowledge about the true distribution. An increase in σ^2 will result in an increase in τ^*, because – as a consequence of the severity of the memory biases – the decision-maker will rely more on her anchor:

$$\frac{\partial \tau^*}{\partial \sigma^2} = \frac{a^2\sum_{k=1}^{K}\rho_k^2}{\left(\sum_{k=1}^{K}\rho_k^2\sigma^2 + \sum_{k=1}^{K}\rho_k^2\nu_k^2 + a^2\right)^2} > 0. \tag{25}$$

[9]For the limiting case of $K \to \infty$ we have: $\sum_{k=1}^{\infty}\rho_k^2 = \frac{1-r}{1+r}$ and $\sum_{k=1}^{\infty}\rho_k^2\bar{\nu}k = \frac{\bar{\nu}}{(1+r)^2}$, resulting in a retrieval inaccuracy bias that is increasing in r. When retrieval inaccuracy increases more rapidly in k, the utilitarian bias resulting from retrieval inaccuracy might be decreasing in r.

A marginal upward change in τ^* results in a higher bias due to anchoring and a lower bias due to transience and retrieval inaccuracy. In Appendix B we show that these marginal changes cancel each other out, resulting in a value of reliability of:

$$\begin{aligned}\mathrm{VOR}_{ae,\tau^*} &= \frac{1}{2}\gamma_1 + \frac{1}{2}\frac{\gamma_1^2}{\gamma_1-\beta_1}(1-\tau^*)^2\sum_{k=1}^{K}\rho_k^2 \\ &= \frac{1}{2}\gamma_1 + \frac{1}{2}\frac{\gamma_1^2}{\gamma_1-\beta_1}\frac{a^4\sum_{k=1}^{K}\rho_k^2}{\left(\sigma^2\sum_{k=1}^{K}\rho_k^2+\sum_{k=1}^{K}\rho_k^2\nu_k^2+a^2\right)^2}.\end{aligned} \tag{26}$$

The last step uses Equation 24. The second positive term captures the combined dis-utility for anchoring and memory biases (i.e. transience and retrieval inaccuracy). For $a = 0$, the VOR reduces to the rational expectations case of Equation 9 because the decision-maker then fully relies on the anchor (see Equation 24). This results in a departure time choice that coincides with the rational expectations case (see Equations 7 and 11). The VOR is now decreasing in the travel time variance, because the second term is lower for higher values of σ^2. The convenient result that expected utility is linear in travel time variance therefore does not hold any longer when the anchor parameter is endogenous.

4. Numerical illustration

This section provides a numerical illustration to investigate the quantitative impact of memory biases. We present results for four parameters, namely r, τ, T_A, and $\bar{\nu}$, and trace their impacts on optimal departure times, expected utility and the value of travel time reliability. The rational expectation levels of these measures (as defined in Section 4.1) are used as the point of reference. Section 4.2 analyses the effect of transience by reducing the value of r such that expectations are increasingly based on recent experiences. At this stage, retrieval is assumed to be accurate. Section 4.3 maintains this assumption but introduces anchoring by increasing the value of τ and varying the value $T_A = \mu + a$. Section 4.4 completes the numerical analysis by introducing inaccurate retrieval. Finally, Section 4.5 summarizes the results of the numerical analysis.

4.1. Parameter assumptions and the rational expectations outcome

The values for the coefficients defining the rational expectations outcomes of the model are based on Tseng and Verhoef (2008) and Fosgerau and Lindsey (2013). Accordingly, β_0, β_1 and γ_1 take the following values: $\beta_0 =$ €40 (p/hour), $\beta_1 =$ €8.86 (p/hour) and $\gamma_1 =$ €25.42 (p/hour). We use the empirical estimates of Peer et al. (2012) to parametrize the distribution of travel times. 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 $\mathbb{VAR}(T_0) = \sigma^2 = \frac{1}{16}$, i.e. a standard deviation of 15 minutes.[10] For this particular set of coefficients, the optimal departure time

[10]The shape parameter of the log-normal distribution is then defined by $\delta_1 = \sqrt{\ln\left(1+\frac{(1/16)}{(1/3)^2}\right)}$, whereas the scale parameter is defined as $\delta_2 = \ln\left(\frac{1}{3}\right) - \frac{\delta_1}{2}$.

with rational expectations is given by $d_0^{re} = -0.51$ (see Equation 7). Using Equation 8, it can be shown that the expected utility under rational expectations equals €-13.37 and the VOR equals €12.71 per hour of variance.

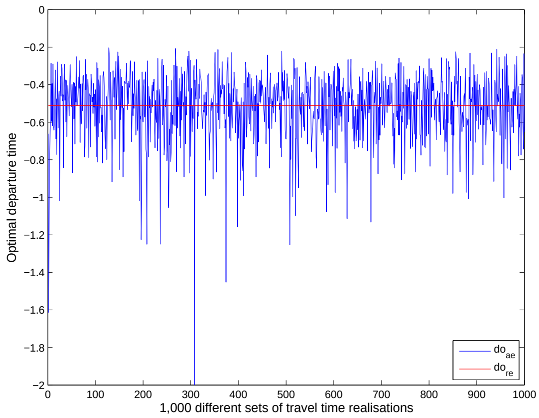
4.2. Accurate retrieval and transience

The first deviation introduced from the rational expectations outcome is transience. Individuals are assumed to store only K past travel time experiences, where K is set to either $K = 5$ or $K = 100$. We systematically change the importance of each of these past experiences in forming expectations by changing the transience parameter r (see Equation 16). Increasing values of r result in a more equal distribution of weights attached across all memories, whereas smaller values assign more weight to more recent periods. We vary r between its upper bound of $r = 1$ (equal weights for all K experienced travel times) and $r = 0.5$ at which the most recent period receives a weight of approximately 50% (i.e. $\rho_1 \approx 0.5$) for both levels of K. Moreover, we assume that the past realisations of T_k are all accurately retrieved, such that $\bar{T}_k = T_k$ and $\nu_k^2 = 0$, $\forall k$. And for the moment, we ignore anchoring by setting $\tau = 0$. We generate 1,000 different sets of K travel time realisations and depict the optimal departure times in Figure 1. A comparison between Figures 1a and 1b highlights that limited storage capacity ($K = 5$ instead of $K = 100$) increases the variance of the optimal departure time considerably. This is a direct consequence of adaptive expectations being formed by a smaller number of travel time realisations. Figures 1c and 1d illustrate that for smaller values of r, the size of K becomes less relevant for the variance of optimal departure times. By definition, reducing r shifts attention towards more recent periods such that more distant travel time realisations have a negligible impact on the optimal departure time.

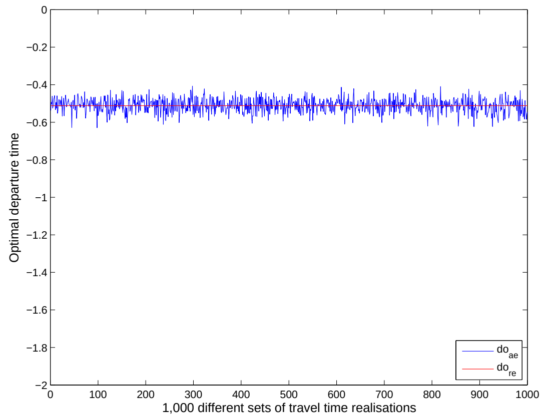
Transience has direct implications on the level of expected outcome utility as illustrated by Figure 2a. Even when $r \to 1$, expected outcome utility falls below EU_{re} for $K < \infty$, because commuters have limited memory capacity to form rational expectations. A decrease in r results in a further deviation of EU_{ae} from EU_{re} because more weight is given to more recent periods. A similar insight is found for the VOR in Figure 2b. Limited memory increases the value of reliability and the penalty is amplified for higher degrees of transience. Equations 9 and 18 indeed confirm that the distance between VOR_{re} and VOR_{ae} decreases for increasing K. The maximum distance between these two lines is in our case $\frac{1}{2}\frac{\gamma_1 2}{(\gamma_1 - \beta_1)}$ = €19.51 per hour for $K = 1$. The latter results in a maximum VOR of €32.22 per hour, which is about 2.5 times higher than the rational expectations outcome. Cantarella (2013) suggests values of 40-80% for the weight of the most recent experience in the expectation of the current traffic situations. When we assume $r = 0.5$ and $K = 5$, the most recent travel time experience determines about 50% of the travel time expectation and the value of travel time reliability is €19.63, which is 45% higher than with rational expectations.

4.3. Accurate retrieval and anchored expectations

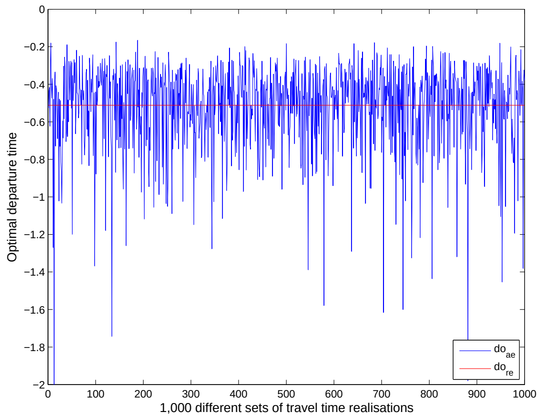
So far we have neglected the presence of an anchor. Equation 11 shows that commuters depart earlier when they have a higher anchor value T_A. Moreover, an increase in τ reduces the influence of past travel time realizations on the optimal departure time and therefore by



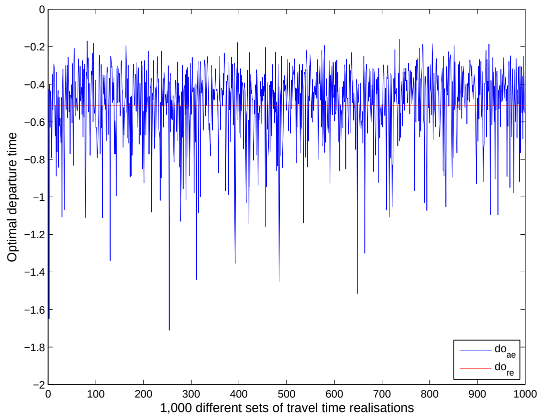
(a) K=5, r=1



(b) K=100, r=1

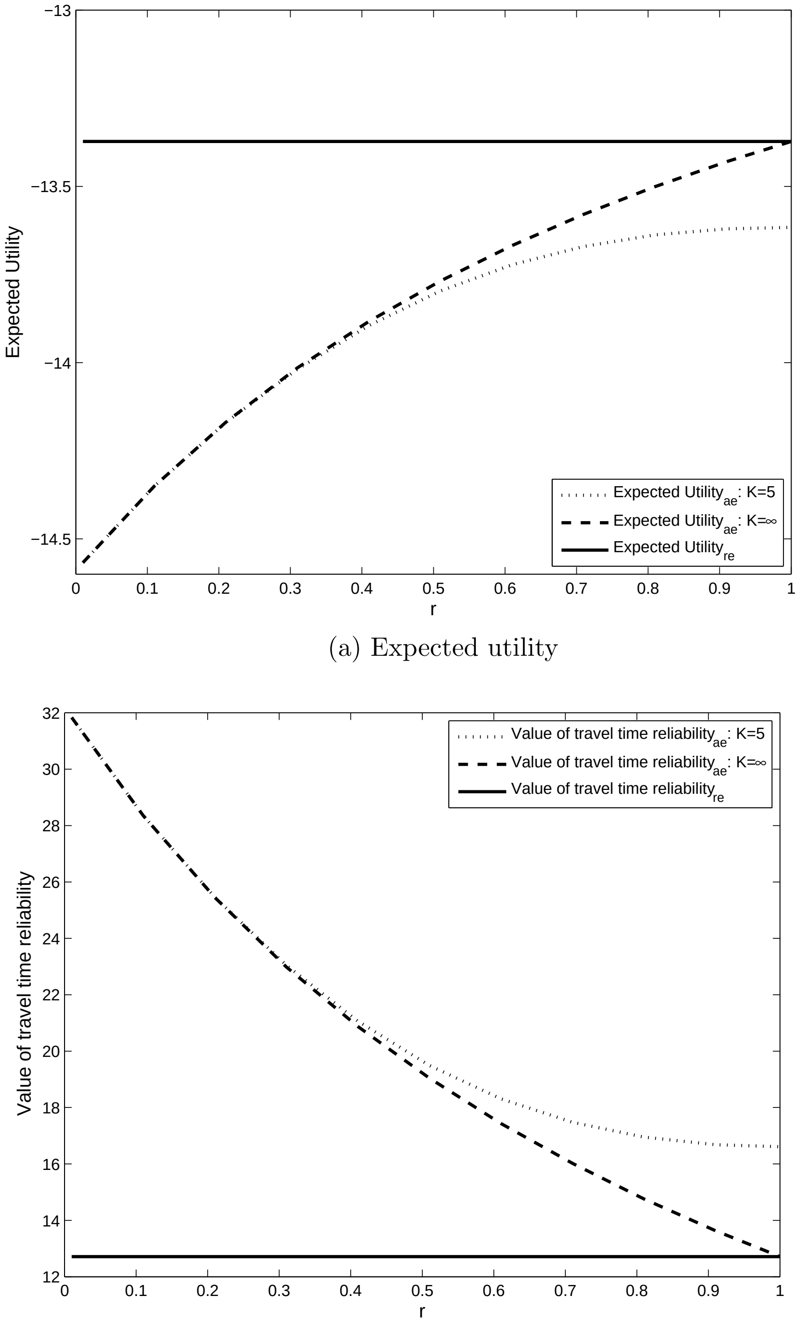


(c) K=5, r=0.5



(d) K=100, r=0.5

Figure 1: Variations in optimal departure time given K and r



(a) Expected utility

(b) Value of travel time reliability

Figure 2: Alternative values of r

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 at $\tau = 1$. The effects of the anchor point on expected utility and the value of travel time reliability are of more interest here. For this exercise we assume $K = 5$ and $r = 1$, resulting in $\rho_k = \frac{1}{K}$.

Figure 3a illustrates that a quadratic penalty applies for deviations $T_A \neq \mu$ when $\tau = 1$. For $\tau = 1$ memory limitations do not play a role (since only the anchor counts), meaning that for $a = 0$ (i.e. when $T_A = \mu$) the expected utility is equal to the rational expectations outcome. For values of τ between 0 and 1, however, an additional deviation from rational expectations due to the limited memory becomes present (even when $a = 0$). For this reason, the dotted horizontal line in Figure 3a falls below the rational expectations utility level even when the anchor has no impact ($\tau = 0$).[11] Moreover the penalty for using a sub-optimal anchor point decreases for lower values of τ. The latter is illustrated by the curve at $\tau = 0.5$.

Figure 3b plots the VOR as a function of τ. It follows directly from Equation 15 that the VOR reduces to the rational expectations outcome for $\tau = 1$, since there is no penalty for forming adaptive expectations. Reducing the value of τ result in a larger effect of recent experiences on the formation of expectations, which increases the value of travel time reliability (see Figure 2). In other words, τ controls the distance between the two horizontal lines in Figure 2*b*. As expected, the effect is strongest for small values of τ. The presence of an anchor point reduces the maximum deviation between the value of travel time reliability with rational expectations and adaptive expectations for any degree of transience.[12]

4.4. Inaccurate retrieval and transience

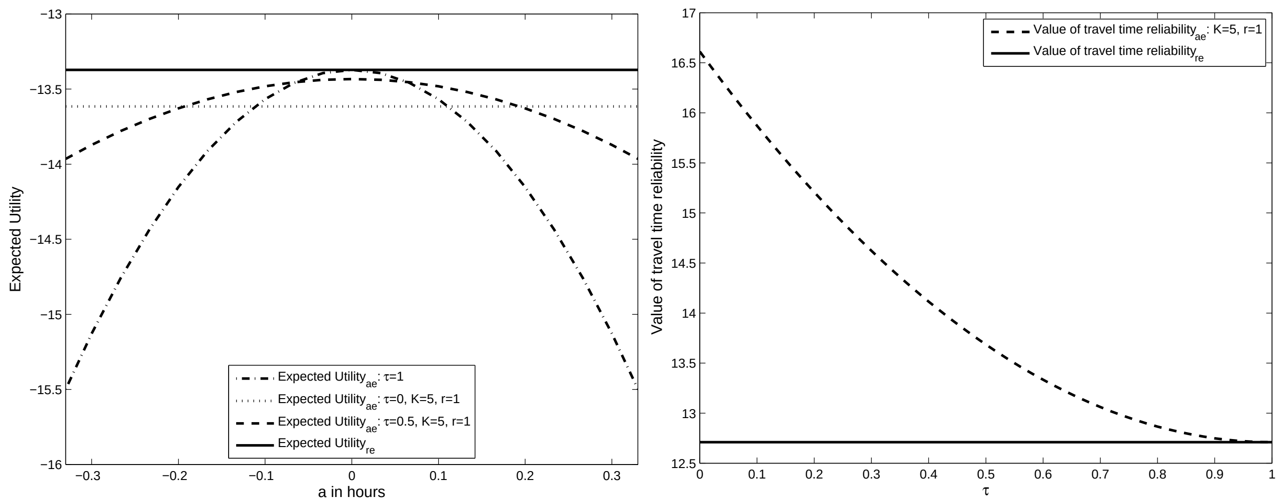
Finally, we illustrate the implications of inaccurate retrieval on expected utility, where we allow $\bar{\nu}$ to vary between zero and σ^2 (see Equation 21). For plausibility, we set this upper bound on $\bar{\nu}$ such that deviations from actual realizations do not fall too much outside of the scale of $f(\cdot)$. In accordance with Equation 22, Figure 4 shows that the penalty for inaccurate retrieval is linear in $\bar{\nu}$, where more inaccuracy reduces expected utility (but does not affect the VOR). This effect is further amplified for smaller values of r.

4.5. Summary of numerical results

We find that r and τ are the most important determinants of differences between adaptive and rational expectations in terms of optimal departure times and the value of travel time reliability. Based on Figure 2b we can conclude that the VOR may be underestimated by up to 45% if the VOR is computed under the assumption that the scheduling decisions are guided by rational expectations, whereas in reality they are guided by adaptive expectations and anchoring. We hereby interpret $K = 5$ and $r = 0.5$ as realistic lower boundaries for the memory storage capacity and the transience parameter, respectively (see Cantarella (2013)). Inaccuracy of retrievals leaves the VOR unaffected, but may induce additional dis-utility, although this effect seems to be relatively small (see Figure 4). The presence of an anchor

[11] For smaller values of r the distance between the two horizontal lines increases.

[12] When τ is endogenously chosen, the VOR will depend on a and $\nu_1 ... \nu_K$ as well (see Equation 26).



(a) Expected utility

(b) Value of travel time reliability as a function of τ (for $a = 0$)

Figure 3: Alternative values of a and τ

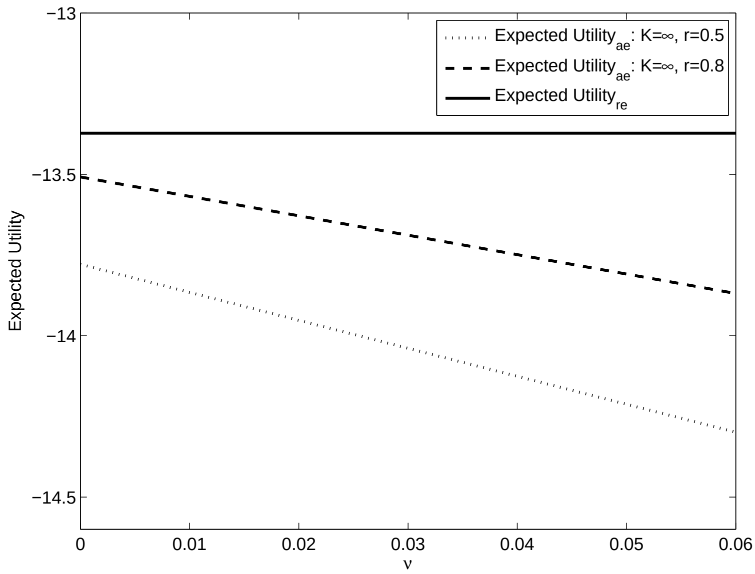


Figure 4: Expected utility for alternative values of $\bar{\nu}$

($\tau \neq 0$) may under certain conditions increase the expected utility. When τ is exogenous, an increase in τ always decreases the bias in the VOR, which is in turn independent of the anchor itself.

5. Conclusions

We developed a model in which adaptive expectations are formed on the basis of past experiences and anchoring. Limited memory storage capacity, transience, inaccurate retrieval and anchoring result in sub-optimal decisions, and thereby translate into reductions in utility relative to the rational expectations outcome. We apply our model to scheduling choices of commuters during the morning commute, where travel times are stochastic. We show that the value of travel time reliability may be underestimated by up to 45% if rational expectations are assumed, while the true expectation formation process is adaptive. The benefits from reliability improvements thus tend to be significantly larger if travel time expectation formation is guided by limited memory, adaptive expectations and anchoring. Revealed preference studies that use a reduced-form utility function probably already capture the biases formulated in this paper. Our results are therefore mainly important for stated preference analyses that ignore the process of expectation formation.

Our functional form assumptions on the utility function allowed us to derive a simple closed-form expression for the memory adjusted value of reliability. The analytical result has the potential to be incorporated in existing static transport network models. Equations 8 and 14 show that trip travel cost functions of the structure $C = b_1 + b_2\mu + b_3\mu^2 + b_4\sigma^2$ are able to capture memory biases in an adequate way. Here, the parameters b_1, b_2, b_3 and b_4 are functions of the underlying behavioural parameters related to scheduling (β_0, β_1 and γ_1), anchoring (τ and a), transience ($\rho_1, ..., \rho_K$) and retrieval inaccuracy ($\nu_1, ..., \nu_K$), and μ and σ^2 are functions of the number of travellers on the links that constitute the trip. For more general forms of the utility function this structure unfortunately breaks down and numerical analysis is needed.

Our dynamic memory model stands apart from static behavioural models where individuals treat probabilities in a non-rational way, since it predicts that commuters are sometimes optimistic and sometimes pessimistic, depending on their most recent experiences and corresponding retrieval probabilities. This is in contrast to rank-dependent utility models that assume that optimism and pessimism are exogenously given and therefore unrelated to earlier experiences (see Koster and Verhoef (2012) and Xiao and Fukuda (2015) for transport applications).

Our approach may serve as an input for the modelling of dynamic systems, both in transport as well as in other fields of economics. In such models adaptive expectations often play a central role but are usually based on simple decision rules (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). Incorporating dynamic learning mechanisms in the model is a fruitful area for further investigation.

Although our model is fairly general, we made several restrictive assumptions in order to keep it analytically tractable. Some of the assumptions can be adjusted in order to arrive at more general results. In the next paragraphs, several possible generalizations are discussed.

First, we assume for simplicity that travel time distributions are independent of departure time, whereas in reality travel time distributions usually vary by time of day. In Appendix A we show how to generalize the resulting expressions for travel time distributions that are changing over subsequent days. This results in additional biases related to *variations in travel time distributions* over subsequent time periods.

Second, we assume that the decision-maker has a fixed anchor $T_A = \mu + a$, whereas in reality this may well be a noisy belief, implying that a is random. The implications of randomness in the anchor can be discussed by looking at the impact on the variance and the mean departure time with adaptive expectations. Equation 12 then will include the mean anchor, whereas Equation 13 would have an additional variance term relating to the variation in the anchor. Because the variance of departure time will increase with a higher variance in the anchor, this will result in additional losses in expected utility (see Equation A.3).

Third, for the main analysis (except for Section 3.6, where we discuss the endogenous choice of the relative weight attached to the anchor, τ), we made the assumption that decision makers are not aware of their memory limitations (Piccione and Rubinstein, 1997). For decisions where the stakes are not so high, this may be a reasonable assumption. When the utilitarian effects of sub-optimal choice are high, the decision maker may take a more reflective attitude and may optimise her anchor or collect additional information in order to reduce behavioural biases.

Fourth, we assumed that retrieval probabilities are independent of the values of the experienced states, meaning that negative experiences do not impact expectations more than positive ones, or *vice versa*. Furthermore, we assumed that the experience of a new state does not affect the memory of the already stored states. Future research should aim at relaxing these assumptions.

Further useful generalizations could be implemented with respect to the specification of the scheduling preferences as well as by including information. One might for instance employ more general scheduling preferences and then use Taylor approximations to arrive at more general results (see Engelson (2011)). It also seems a fruitful direction for future research to extend the model by the possibility to obtain information about future travel times. The quality of the information could then in turn depend on when the information is collected, or on how much one is willing to pay for it.

Finally, we consider the empirical testing of our decision model using laboratory and revealed preference data as a fruitful direction for further research. However, the data requirements will be demanding: high-quality panel data with a substantial sample size will be necessary to estimate the parameters of our model in such a way that the four main components of the model (limited memory capacity, transience, retrieval accuracy and the anchor) can be unambiguously disentangled. For initial applications, it may therefore be useful to disregard one of the components, or to make functional form assumptions that reduce the number of parameters to be estimated.

Bibliography

Arentze, T. A. and Timmermans, H. J. (2003). Modeling learning and adaptation processes in activity-travel choice a framework and numerical experiment. *Transportation*, 30(1):37–62.

Arentze, T. A. and Timmermans, H. J. (2005). Representing mental maps and cognitive learning in micro-simulation models of activity-travel choice dynamics. *Transportation*, 32(4):321–340.

Ariely, D., Loewenstein, G., and Prelec, D. (2003). "Coherent arbitrariness": Stable demand curves without stable preferences. *The Quarterly Journal of Economics*, 118(1):73–106.

Avineri, E. and Prashker, J. N. (2005). Sensitivity to travel time variability: Travelers learning perspective. *Transportation Research Part C: Emerging Technologies*, 13(2):157–183.

Barucci, E. (1999). Heterogeneous beliefs and learning in forward looking economic models. *Journal of Evolutionary Economics*, 9(4):453–464.

Barucci, E. (2000). Exponentially fading memory learning in forward-looking economic models. *Journal of Economic Dynamics and Control*, 24(5):1027–1046.

Ben-Elia, E. and Shiftan, Y. (2010). Which road do I take? a learning-based model of route-choice behavior with real-time information. *Transportation Research Part A: Policy and Practice*, 44(4):249 – 264.

Bénabou, R. and Tirole, J. (2002). Self-confidence and personal motivation. *The Quarterly Journal of Economics*, 117(3):871–915.

Bernheim, D. B. and Thomadsen, R. (2005). Memory and anticipation. *The Economic Journal*, 115(503):271–304.

Bogers, E. A. I., Bierlaire, M., and Hoogendoorn, S. P. (2007). Modeling learning in route choice. *Transportation Research Record*, 2014:1–8.

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

Brunnermeier, M. K. and Parker, J. A. (2005). Optimal expectations. *American Economic Review*, 4(95):1092 – 1118.

Cantarella, G. E. (2013). Day-to-day dynamic models for intelligent transportation systems design and appraisal. *Transportation Research Part C: Emerging Technologies*, 29:117–130.

Caplin, A., Dean, M., and Martin, D. (2011). Search and satisficing. *American Economic Review*, 101(7):2899–2922.

Chen, R. and Mahmassani, H. (2004). Travel time perception and learning mechanisms in traffic networks. *Transportation Research Record*, 1894(1):209–221.

Engelson, L. (2011). Properties of Expected Cost Function with Uncertain Travel Time. *Transportation Research Record*, 2254:151–159.

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

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

Fosgerau, M. and Lindsey, R. (2013). Trip-timing decisions with traffic incidents. *Regional Science and Urban Economics*, 43(5):764–782.

Furnham, A. and Boo, H. C. (2011). A literature review of the anchoring effect. *The Journal of Socio-Economics*, 40(1):35 – 42.

Gennaioli, N. and Shleifer, A. (2010). What comes to mind. *The Quarterly Journal of Economics*, 125(4):1399–1433.

Gollier, C. and Muermann, A. (2010). Optimal choice and beliefs with ex ante savoring and ex post disappointment. *Management Science*, 56(8):1272–1284.

Hensher, D. (2001). The valuation of commuter travel time savings for car drivers: evaluating alternative model specifications. *Transportation*, 28(2):101–118.

Hirshleifer, D. and Welch, I. (2002). An economic approach to the psychology of change: Amnesia, inertia, and impulsiveness. *Journal of Economics & Management Strategy*, 11(3):379–421.

Hommes, C. (2013). *Behavioral rationality and heterogeneous expectations in complex economic systems*. Cambridge University Press.

Horowitz, J. L. (1984). The stability of stochastic equilibrium in a two-link transportation network. *Transportation Research Part B: Methodological*, 18(1):13–28.

Jha, M., Madanat, S., and Peeta, S. (1998). Perception updating and day-to-day travel choice dynamics in traffic networks with information provision. *Transportation Research Part C: Emerging Technologies*, 6(3):189–212.

Koster, P. and Verhoef, E. T. (2012). A rank-dependent scheduling model. *Journal of Transport Economics and Policy*, 46(1):123–138.

Mullainathan, S. (2002). A memory-based model of bounded rationality. *The Quarterly Journal of Economics*, 117(3):735–774.

Noland, R. B. and Small, K. A. (1995). Travel-time Uncertainty, Departure Time Choice, and the Cost of Morning Commutes. *Transportation Research Record*, 1493:150–158.

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

Peer, S., Verhoef, E. T., Knockaert, J., Koster, P., and Tseng, Y.-Y. (2015). Long-run vs. short-run perspectives on consumer scheduling: Evidence from a revealed-preference experiment among peak-hour road commuters. *International Economic Review*, 56(1):303–323.

Piccione, M. and Rubinstein, A. (1997). On the interpretation of decision problems with imperfect recall. *Games and Economic Behavior*, 20(1):3 – 24.

Sarafidis, Y. (2007). What Have you Done for me Lately? Release of Information and Strategic Manipulation of Memories. *The Economic Journal*, 117(518):307–326.

Schacter, D. L. (2002). *The seven sins of memory: How the mind forgets and remembers.* Houghton-Mifflin, New York.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1):99–118.

Small, K. A. (1982). The Scheduling of Consumer Activities : Work Trips. *The American Economic Review*, 72:467–479.

Small, K. A. (1992). Trip scheduling in urban transportation analysis. *The American Economic Review*, 82(2):482–486.

Strack, F. and Mussweiler, T. (1997). Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility. *Journal of Personality and Social Psychology*, 73(3):437.

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

Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157):1124–1131.

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

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

Watling, D. P. and Cantarella, G. E. (2013). Modelling sources of variation in transportation systems: theoretical foundations of day-to-day dynamic models. *Transportmetrica B: Transport Dynamics*, 1(1):3–32.

Wilson, A. (2003). Bounded memory and biases in information processing. *Econometrica*, 6(82):2257–2294.

Wilson, T. D., Houston, C. E., Etling, K. M., and Brekke, N. (1996). A new look at anchoring effects: basic anchoring and its antecedents. *Journal of Experimental Psychology: General*, 125(4):387–402.

Xiao, Y. and Fukuda, D. (2015). On the cost of misperceived travel time variability. *Transportation Research Part A*, (75):96–112.

Appendix A. Proof section 3.3.

In this Appendix we derive the predicted expected outcome utility (Equation 14). The proof is for general travel distributions with k-dependent means and variances and k-dependent anchoring. Assume that travel time distributions have mean μ_k and variance σ_k^2 and probability density $f(T_k|\mu_k, \sigma_k^2)$. The predicted expected outcome utility is given by:

$$\begin{aligned}
&\mathbb{E}(U_{ae}) \equiv \mathbb{E}\left(U(d_0^{ae}, T_0)\right) \\
&= \int ... \int \left(\int ... \int U(d_0^{ae}, T_0) \prod_{k=1}^{K} g\left(\bar{T}_k | T_k, \nu_k^2\right) d\bar{T}_1 ... \bar{T}_K \right) \prod_{k=0}^{K} f(T_k|\mu_k, \sigma_k^2) dT_0 ... dT_K,
\end{aligned} \tag{A.1}$$

The expectation over all values of T_0 is given by:

$$\begin{aligned}
\mathbb{E}_{T_0}(U(d_0^{ae}, T_0)) &= \int \left(- \int_{d_0^{ae}}^{0} (\beta_0 + \beta_1 v) dv - \int_0^{d_0^{ae}+T_0} (\beta_0 + \gamma_1 v) dv \right) f(T_0|\mu_0, \sigma_0^2) dT_0 \\
&= -\beta_0 \mu_0 - \frac{1}{2}(\gamma_1 - \beta_1)(d_0^{ae})^2 - \gamma_1 \mu_0 d_0^{ae} - \frac{1}{2}\gamma_1(\mu_0^2 + \sigma_0^2),
\end{aligned} \tag{A.2}$$

where we use $\mathbb{E}_{T_0}$ to emphasize that the expectation is only over values of T_0. The predicted expected outcome utility with adaptive expectation can be found by taking the expected value over all possible values of the departure time d_0^{ae}:

$$\begin{aligned}
\mathbb{E}(U_{ae}) &= \mathbb{E}\left(-\beta_0 \mu_0 - \frac{1}{2}(\gamma_1 - \beta_1)(d_0^{ae})^2 - \gamma_1 \mu_0 d_0^{ae} - \frac{1}{2}\gamma_1(\mu_0^2 + \sigma_0^2) \right) \\
&= -\beta_0 \mu_0 - \frac{1}{2}(\gamma_1 - \beta_1)\mathbb{E}\left((d_0^{ae})^2\right) - \gamma_1 \mu_0 \mathbb{E}(d_0^{ae}) - \frac{1}{2}\gamma_1(\mu_0^2 + \sigma_0^2) \\
&= -\beta_0 \mu_0 - \frac{1}{2}(\gamma_1 - \beta_1)\left((\mathbb{E}(d_0^{ae}))^2 + \mathbb{VAR}(d_0^{ae})\right) - \gamma_1 \mu_0 \mathbb{E}(d_0^{ae}) - \frac{1}{2}\gamma_1(\mu_0^2 + \sigma_0^2).
\end{aligned} \tag{A.3}$$

This shows that $\mathbb{E}(U_{ae})$ can be written as a function of the mean departure time and the variance of departure time. When travel time distributions depend on k, the departure time with adaptive expectations is given by 11. The mean departure time is given by:

$$\mathbb{E}\left(d_0^{ae}\right) = -\tau \frac{\gamma_1}{\gamma_1 - \beta_1} T_A - (1 - \tau) \frac{\gamma_1}{\gamma_1 - \beta_1} \sum_{k=1}^{K} \rho_k \mu_k, \tag{A.4}$$

which reduces to 12 for $\mu_k = \mu$. The variance of the departure time is given by (here we use the assumptions of footnote 6):

$$\mathbb{VAR}(d_0^{ae}) = (1 - \tau)^2 \left(\frac{\gamma_1}{\gamma_1 - \beta_1} \right)^2 \sum_{k=1}^{K} \rho_k^2 \left(\sigma_k^2 + \nu_k^2\right), \tag{A.5}$$

which reduces to 13, for $\sigma_k^2 = \sigma^2$. Substituting Equation A.4 and A.5 in Equation A.3 gives the result for k-dependent travel time distributions. For the remainder of this Appendix we assume $\mu_k = \mu_0 = \mu$ and $\sigma_k^2 = \sigma_0^2 = \sigma^2$ in order to arrive at the results that are discussed in the main body of the paper. Substituting 12 in A.3 gives:

$$\begin{aligned}
\mathbb{E}(U_{ae}) &= -\beta_0\mu - \frac{1}{2}(\gamma_1 - \beta_1)\left(\left(\frac{\gamma_1}{\gamma_1-\beta_1}\right)^2\left(\mu^2 + 2\mu\tau a + \tau^2 a^2\right) + \mathbb{VAR}(d_0^{ae})\right) \\
&\quad + \frac{\gamma_1^2}{\gamma_1-\beta_1}\left(\mu^2 + \mu\tau a\right) - \frac{1}{2}\gamma_1(\mu^2+\sigma^2) \\
&= -\beta_0\mu + \frac{1}{2}\frac{\gamma_1^2}{\gamma_1-\beta_1} - \frac{1}{2}\gamma_1\mu^2 - \frac{1}{2}\gamma_1\sigma^2 - \frac{1}{2}\frac{\gamma_1^2}{\gamma_1-\beta_1}\tau^2a^2 - \frac{1}{2}(\gamma_1-\beta_1)\mathbb{VAR}(d_0^{ae}) \\
&= -\beta_0\mu + \frac{1}{2}\frac{\beta_1\gamma_1}{\gamma_1-\beta_1}\mu^2 - \frac{1}{2}\gamma_1\sigma^2 - \frac{1}{2}\frac{\gamma_1^2}{\gamma_1-\beta_1}\tau^2a^2 - \frac{1}{2}(\gamma_1-\beta_1)\mathbb{VAR}(d_0^{ae}) \\
&= \mathbb{E}(U_{re}) - \frac{1}{2}\frac{\gamma_1^2}{\gamma_1-\beta_1}\tau^2a^2 - \frac{1}{2}(\gamma_1-\beta_1)\mathbb{VAR}(d_0^{ae}).
\end{aligned} \tag{A.6}$$

Substituting Equation 13 gives the desired result. This concludes the proof.

Appendix B. Proof section 3.6.

We start with 14 where we include the optimal anchor parameter τ^* which depends on the variance of travel time (see 24). Then differentiate expected utility with respect to σ^2 to obtain:

$$\begin{aligned}
\mathrm{VOR}_{\tau^*} &= \frac{1}{2}\gamma_1 + \frac{1}{2}\frac{\gamma_1^2}{\gamma_1-\beta_1}(1-\tau^*)^2\sum_{k=1}^K \rho_k^2 + \frac{\gamma_1^2}{\gamma_1-\beta_1}\tau^*\frac{\partial\tau^*}{\partial\sigma^2}a^2 \\
&\quad - \frac{\gamma_1^2}{\gamma_1-\beta_1}(1-\tau^*)\frac{\partial\tau^*}{\partial\sigma^2}\left(\sigma^2\sum_{k=1}^K\rho_k^2 + \sum_{k=1}^K\rho_k^2\nu_k^2\right) \\
&= \frac{1}{2}\gamma_1 + \frac{1}{2}\frac{\gamma_1^2}{\gamma_1-\beta_1}(1-\tau^*)^2\sum_{k=1}^K\rho_k^2 \\
&\quad + \frac{\gamma_1^2}{\gamma_1-\beta_1}\left(\tau^*\frac{\partial\tau^*}{\partial\sigma^2}a^2 - (1-\tau^*)\frac{\partial\tau^*}{\partial\sigma^2}\left(\sigma^2\sum_{k=1}^K\rho_k^2 + \sum_{k=1}^K\rho_k^2\nu_k^2\right)\right) \\
&= \frac{1}{2}\gamma_1 + \frac{1}{2}\frac{\gamma_1^2}{\gamma_1-\beta_1}(1-\tau^*)^2\sum_{k=1}^K\rho_k^2 \\
&\quad + \frac{\gamma_1^2}{\gamma_1-\beta_1}\frac{\partial\tau^*}{\partial\sigma^2}\left(\tau^*\left(a^2 + \sigma^2\sum_{k=1}^K\rho_k^2 + \sum_{k=1}^K\rho_k^2\nu_k^2\right) - \left(\sigma^2\sum_{k=1}^K\rho_k^2 + \sum_{k=1}^K\rho_k^2\nu_k^2\right)\right) \\
&= \frac{1}{2}\gamma_1 + \frac{1}{2}\frac{\gamma_1^2}{\gamma_1-\beta_1}(1-\tau^*)^2\sum_{k=1}^K\rho_k^2,
\end{aligned} \tag{B.1}$$

Substituting 24 gives:

$$\begin{aligned}
\mathrm{VOR}_{\tau^*} &= \frac{1}{2}\gamma_1 + \frac{1}{2}\frac{\gamma_1^2}{\gamma_1 - \beta_1}\left(\frac{a^2}{\sigma^2\sum_{k=1}^{K}\rho_k^2 + \sum_{k=1}^{K}\rho_k^2\nu_k^2 + a^2}\right)^2 \sum_{k=1}^{K}\rho_k^2 \\
&= \frac{1}{2}\gamma_1 + \frac{1}{2}\frac{\gamma_1^2}{\gamma_1 - \beta_1}\frac{a^4\sum_{k=1}^{K}\rho_k^2}{\left(\sigma^2\sum_{k=1}^{K}\rho_k^2 + \sum_{k=1}^{K}\rho_k^2\nu_k^2 + a^2\right)^2},
\end{aligned} \tag{B.2}$$