

TRANSPORT FINDINGS

Revealed-Preference Evidence Against Valuing Reliability

Jaime Soza-Parra, PhD¹, Emily Moylan, PhD²¹ Department of Human Geography and Spatial Planning, Universiteit Utrecht, ² School of Civil Engineering, The University of Sydney

Keywords: Journey time reliability, Revealed preference, Attribute non-attendance, Hypothetical bias

<https://doi.org/10.32866/001c.172729>

Findings

The literature shows that, when asked to state their preference for a route, travellers will select more reliable ones *ceteris paribus*. But this assumes that semantic information about reliability is available to travellers and acted upon. This finding compares revealed preference route choices with a variety of reliability measures. In contrast to previously-published stated-preference results, real choices revealed by the same respondents do not show the same significance on variables measuring reliability. This implies that the widely accepted value of reliability may be inflated by hypothetical bias and/or attribute non-attendance.

1. Questions

When presented with two alternative routes in a stated preference design, choosers will select the one advertised as more reliable. For realised choices, travellers must both know and care about journey time reliability in order to observe a significant coefficient on reliability. Does the preference for reliability persist when the same people respond to revealed and stated preference route choice questions?

2. Methods

This finding uses revealed preferences for route choices for 689 regular car commuters in NSW, Australia in 2019. Respondents were asked to identify, from a map interface, their usual route to work from a set of either two (28.7%) or three (71.3%) alternatives. Information about the distribution of journey times on each of the alternatives was subsequently collected using *duration in traffic* queried from the Google Maps API over 100 days. More details about the sample and the attributes of the alternatives are provided in Supplement A. A set of reliability measures was calculated from each distribution based on the indicators in the reports on Reliability from the US Strategic Highway Research Program 2 (2013) and previous work using the same dataset (Moylan et al. 2022). The measures are standard deviation, coefficient of variation, buffer index, skewness statistic and rare-event reliability, which capture, respectively, the typical variation, the typical variation as a percent of the average travel time, the amount of extra time needed to be on-time 80% of the time, the importance of the tail of the travel time distribution and the portion of the distribution that is longer than a respondent-identified threshold of lateness. Definitions are provided in Supplement B.

We tested choice model specifications with each reliability measure (r), in addition to median travel time, t (in minutes), travel cost due to tolls, c (in Australian Dollars), and circuitry, u (defined as the ratio between the length of the route and the straight line distance (Levinson and El-Geneidy 2009)). Combinations of reliability measures did not yield improved significance of the findings. In order to account for route overlap correlation, we estimated a series of C-logit models (Equation 1) that modify the systematic utility V_{ij} (Equation 2) of a multinomial logit model for individual i and alternative j by incorporating a commonality factor CF_{ij} (Equation 3) (Cascetta 2013):

$$\tilde{V}_{ij} = V_{ij} - CF_{ij} \quad (1)$$

$$V_{ij} = \beta_t \cdot t_{ij} + \beta_c \cdot c_{ij} + \beta_u \cdot u_{ij} + \beta_r \cdot r_{ij} \quad (2)$$

$$CF_{ij} = \beta_0 \cdot \ln \left[1 + \sum_{\substack{k \in C_i \\ k \neq j}} \left(\frac{L_{jk}}{\sqrt{L_j \cdot L_k}} \right) \left(\frac{L_j - L_{jk}}{L_k - L_{jk}} \right) \right] \quad (3)$$

Here, β_0 corresponds to an additional parameter to be estimated compared to standard logit, and CF_{ij} for individual i and route j is calculated by looking at all routes k that are part of the choice set of the individual (C_i), L_j , which is the length of the specific route j , L_k , which is the length of each other route, and L_{jk} , which is the length of the overlap between routes j and k .

The results from these models are interpreted in the context of previously published results (Moylan et al. 2022) from a stated-preference experiment with the same participants which found statistically significant coefficients on standard deviation of travel time and rare event reliability. Further details in Supplement C.

3. Findings

In five of the six models tested, we see stable coefficients for travel time, travel cost, and circuitry. The Value of Time (VoT) sits around 42.5 AUD/hr (β_{TT}/β_{Cost} \$/hr, [6.1 - 77.6 90% CI]), which falls between regular and business VoT for Australia (Rose 2025). However, the reliability measures tested in these models returned statistically insignificant (albeit intuitively signed for the most part) coefficients. The sixth model is an exception—rare event unreliability has a statistically significant and intuitively signed coefficient, but the median travel time is not significant. This interaction should be expected since this metric is constructed from how often the travel time exceeds a user-specific threshold, so slower alternatives will have higher unreliability by this measure. In contrast to this set of revealed-preference models, the same respondents gave stated preferences that indicated statistical significance with intuitive signs on both standard deviation and rare-event unreliability. This suggests that the accepted value on travel time reliability is overstated.

Table 1. Revealed Preference Models

Variable	Base		Std. Dev.		CV		BI 80%		Skewness		Rare	
	Est.	t-ratio	Est.	t-ratio	Est.	t-ratio	Est.	t-ratio	Est.	t-ratio	Est.	t-ratio
β_{Median}	-1.982	-2.183	-1.899	-2.050	-2.047	-2.248	-2.076	-2.267	-1.949	-2.120	-1.030	-1.087
β_{Cost}	-0.047	-2.436	-0.047	-2.406	-0.047	-2.407	-0.048	-2.476	-0.047	-2.442	-0.047	-2.425
$\beta_{\text{Circuitry}}$	-1.282	-4.295	-1.273	-4.255	-1.257	-4.196	-1.248	-4.150	-1.292	-4.278	-1.253	-4.161
$\beta_{\text{st.dev.}}$	-	-	-0.026	-0.435	-	-	-	-	-	-	-	-
β_{CV}	-	-	-	-	-2.017	-1.034	-	-	-	-	-	-
$\beta_{\text{BI 80}}$	-	-	-	-	-	-	-1.122	-0.800	-	-	-	-
β_{Skewness}	-	-	-	-	-	-	-	-	0.008	0.228	-	-
$\beta_{\text{Rare event unreliability}}$	-	-	-	-	-	-	-	-	-	-	-1.117	-2.509
$\beta_{0,\text{Clogit}}$	1.398	3.274	1.397	3.268	1.404	3.282	1.412	3.302	1.399	3.274	1.302	3.045
Final Log Likelihood	-627.842		-627.747		-627.300		-627.509		-627.817		-622.297	
Adj. ρ_C^2	0.058		0.056		0.057		0.057		0.056		0.065	
AIC	1263.685		1265.494		1264.600		1265.019		1265.633		1254.594	
BIC	1281.814		1288.156		1287.262		1287.680		1288.295		1277.256	

The finding can be interpreted in several ways. A key challenge in revealed preference design is to accurately represent the choice set. In this experiment, the choice set is made explicit by presenting three options (where available) for the nominated origin and destination addresses, but it does not follow that the travellers were aware of and considered these choices before taking the survey. Moreover, the travellers may value reliability but not be aware of the reliability of their alternatives. As a frequently made journey, it can be assumed that the respondents do have semantic knowledge of the travel time distribution of their selected alternative but not necessarily of the routes they don't usually use. Finally, it may be that the travellers do know about the reliability characteristics of all the alternatives, but they do not value them. All three of these explanations (or a combination) would result in the observed outcome of insignificant coefficients. Future work could separate the explanations by collecting information on the traveller's awareness of the alternatives and testing the accuracy of their knowledge of their reliability attributes.

To contextualise these findings, common sense says that travellers probably do value reliability in their route choice decisions, but not in a way that is well captured by the standard deviation or even how often they are subjectively late. There is substantial evidence that recalled travel times are biased (Stopher et al. 2007; Parthasarathi et al. 2013; Tenenboim and Shiftan 2018), which makes sense in the context of the Peak-End Rule (Fredrickson and Kahneman 1993), which implies that travellers should be swayed by their worst or most recently experienced travel time (episodic knowledge) rather than attributes of the distribution (semantic knowledge). This distinction is important for transport system planners tempted to assign user benefits to projects that fractionally reduce the standard deviation of travel times—travellers may say that is attractive in a hypothetical context, but the evidence is weak that it will influence their behaviours.

Acknowledgements

The authors gratefully acknowledge Michiel Bliemer, Taha Hossein Rashidi and Karen Cong for their important roles in the creation of this dataset. In addition, we thank Tom Adelmeijer and Hamza El Idrissi who made use of the NSW toll data available through the NSW Open Data Hub to calculate toll costs for our analysed routes (<https://opendata.transport.nsw.gov.au/>).

Submitted: May 23, 2026 AEDT. Accepted: October 15, 2026 AEDT. Published: October 30, 2026 AEDT.



This is an open-access article distributed under the terms of the Creative Commons Attribution-ShareAlike 4.0 International License (CCBY-SA-4.0). View this license's legal deed at <https://creativecommons.org/licenses/by-sa/4.0> and legal code at <https://creativecommons.org/licenses/by-sa/4.0/legalcode> for more information.

REFERENCES

- Cascetta, Ennio. 2013. *Transportation Systems Engineering: Theory and Methods*. Vol. 49. Springer Science & Business Media.
- Fredrickson, Barbara L., and Daniel Kahneman. 1993. "Duration Neglect in Retrospective Evaluations of Affective Episodes." *Journal of Personality and Social Psychology* 65 (1): 45.
- Levinson, David, and Ahmed El-Geneidy. 2009. "The Minimum Circuitry Frontier and the Journey to Work." Article. *Regional Science and Urban Economics* 39 (6): 732–38. <https://doi.org/10.1016/j.regsciurbeco.2009.07.003>.
- Moylan, Emily K. M., Michiel C. J. Bliemer, and Taha Hossein Rashidi. 2022. "Travellers' Perceptions of Travel Time Reliability in the Presence of Rare Events." *Transportation* 49 (4): 1157–81.
- Parthasarathi, Pavithra, David Levinson, and Hartwig Hochmair. 2013. "Network Structure and Travel Time Perception." *PloS One* 8 (10): e77718.
- Rose, John. 2025. "What Is the Value of Time, and Does It Actually Exist?" Preprint, The University of Sydney Business School News, January. <https://www.sydney.edu.au/business/news-and-events/news/2025/01/06/what-is-the-value-of-time--and-does-it-actually-exist-.html>.
- Stopher, Peter, Camden FitzGerald, and Min Xu. 2007. "Assessing the Accuracy of the Sydney Household Travel Survey with GPS." *Transportation* 34 (6): 723–41.
- Tenenboim, Einat, and Yoram Shiftan. 2018. "Accuracy and Bias of Subjective Travel Time Estimates." *Transportation* 45 (3): 945–69.
- Transportation Research Board and National Academies of Sciences, Engineering, and Medicine. 2013. *Incorporating Reliability Performance Measures into the Transportation Planning and Programming Processes*. The National Academies Press. <https://doi.org/10.17226/22596>.

Appendices



Figure 1. Spatial distribution of participants across the state of New South Wales with concentration in Sydney.

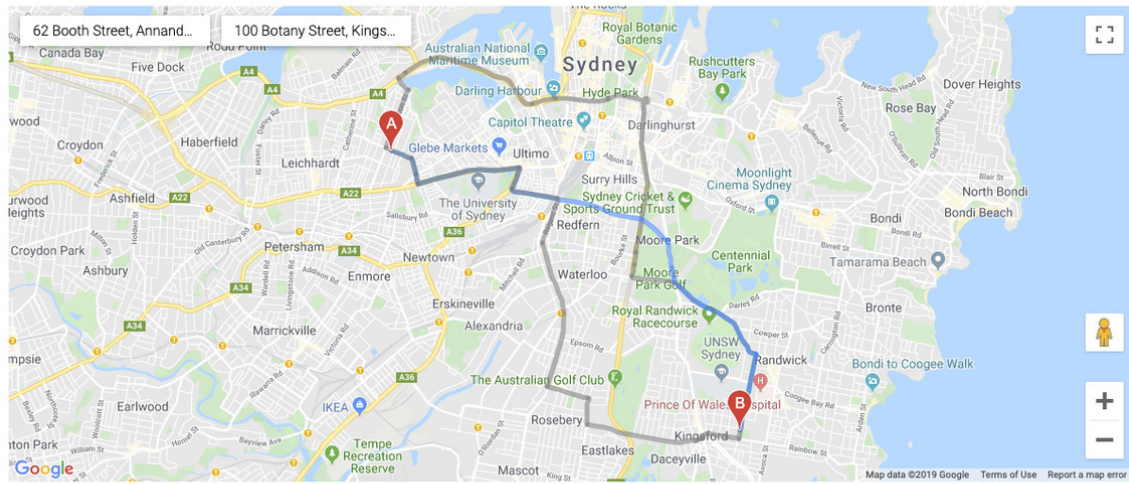
Description of survey participants and choice set creation

This study makes use of a survey combining revealed and stated preference components. It was reviewed and approved by the University of Sydney Human Research Ethics Committee (Application 2019/437). The study is described in detail by Moylan et al. (2022).

The participants who regularly drive to work were recruited in 2019 through a global market research firm. 1001 responses reduced to 914 after cleaning unrealistic responses and speeders. Of these, 689 reported routes with at least two distinct alternatives with queryable travel times. The sample resembles the New South Wales population in distribution of age, gender and household income (further details in Moylan et al. (2022)) and comes from across the state as shown in Figure 1.

The market research firm linked participants to a custom-designed interface. In the first part, participants enter the origin and destination addresses of their usual work commute, which triggered a query to the Google Maps Directions API— three (or two if only two were generated) resulting route choices were displayed on a map and users click to select their usual route

Consider your daily commute. Please enter the origin and destination addresses. Route options will appear on the map. Click to select your typical route on the map below:



If your exact route is not shown, tick this box and choose the closest option.

When do you typically make this trip?

Check all that apply.

☒ Monday ☒ Tuesday ☒ Wednesday ☒ Thursday ☐ Friday ☐ Saturday ☐ Sunday

Time of day for departure:

Hour	Minute	AM/PM
7	40	AM

How much do you pay in tolls to make this trip?

\$ 0

Figure 2. Interface, reproduced from Moylan et al. (2022), offering 3 alternatives for a nominated OD pair. It also collects information on the days of the week, the time of day and the tolls for this trip.

as shown in Figure 2. The interface would be familiar to any user of Google Maps applications. Participants had the option to tick a box if their exact route was not shown.

These choices are used as the revealed preference choice set. Attributes of the travel time distributions of the alternatives were generated through 100 queries to the Google Maps Directions API requesting *duration in traffic* at the participant-provided times of day and days of the week.

Calculation of reliability metrics

The revealed preference models test five reliability metrics:

- Standard Deviation: $\sigma = \sqrt{\frac{\sum (t - \bar{t})^2}{N-1}}$
- Coefficient of Variation: $CV = \sqrt{\frac{\sum (t - \bar{t})^2}{N-1}} / \bar{t}$

- Buffer Index at 80%: $BI_{80} = \frac{t^{80\%} - t^{50\%}}{t^{50\%}}$
- Skewness: $s = \frac{t^{90\%} - t^{50\%}}{t^{50\%} - t^{10\%}}$
- Rare event reliability : $r = P(t \geq \tau)$

where t represents the individual travel times from the N *duration in traffic* queries for each route, where generally $N=100$, $t^{x\%}$ represents the x^{th} percentile travel time from each sample, and τ is a participant-identified threshold of lateness (more details available from Moylan et al. (2022)). Each of these reliability measures is calculated for each alternative available to each participant, and τ is the same for all alternatives for each participant.

Stated preference comparison

After revealing their true route choice and their personal threshold for lateness, the same respondents completed a stated preference experiment offering paired alternatives pivoting off the respondent's reported commute. The experiment compared three presentations of the travel time distribution, and each presentation was used to estimate a route choice model. All models yielded intuitively-signed (negative) and statistically significant coefficients for the central tendency and standard deviation of travel time. Additionally, the coefficient on the risk of extreme delay was negative and statistically significant in the models using the two presentation styles where this information was provided. The results are summarised in Table 1 of Moylan et al. (2022) which concluded that rare event reliability is valued distinctly from day-to-day reliability.