Quantifying uncertainty in travel demand forecasts

Engineering New Zealand – Transportation Group

Annual conference Wellington, 9 March 2026

Stuart Donovan, Motu Research

Pete Clark, UrbanUmber

Motu

economic & public policy research

Outline

- Background
- Approach & results
- Final comments

Background

We want to plan for the future ...

... but the future is inherently uncertain.

Uncertainty in travel demand forecasts

- Transport planning and investment processes often rely on travel demand forecasts that extend many decades into the future.

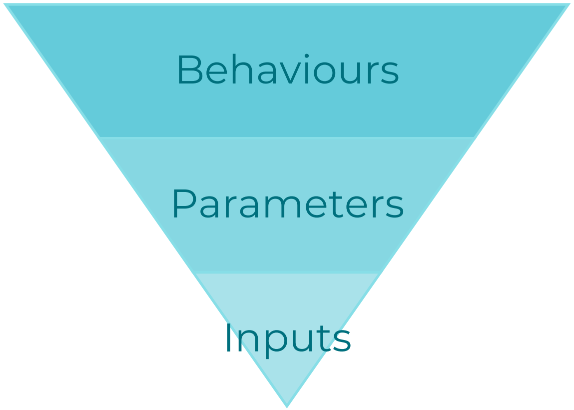
- The future is *inherently* uncertain, as are the models we use to generate our forecasts (whether four-step models or activity-based)

- **Research question:** How might we quantify uncertainty in travel demand forecasts?

Typologies of uncertainty

Uncertainty arises from multiple sources:

- *Behavioural uncertainty*, which arises from uncertainty in how we model individual behaviour (people and firms)
- *Parameter uncertainty*, which arises from uncertainty in the values of the parameters that are used within our behavioural models
- *Input uncertainty,* which arises from uncertainty in the inputs that we feed into models, e.g. assumptions on the future state of the world.



Different types of uncertainty might be best addressed at different times, e.g. address parameter uncertainty during model development and input uncertainty during model application.

In our research, we focus on *input uncertainty.*

Approach & results

Approach

1. *Identify the effects of inputs on travel demand forecasts.*
   - Identify input variables most relevant to uncertainty (e.g., based on evidence and judgement).
   - Vary chosen inputs and re-run the transport model to quantify changes in travel demands.
   - Use a regression model to identify the effects of inputs on travel demand forecasts.

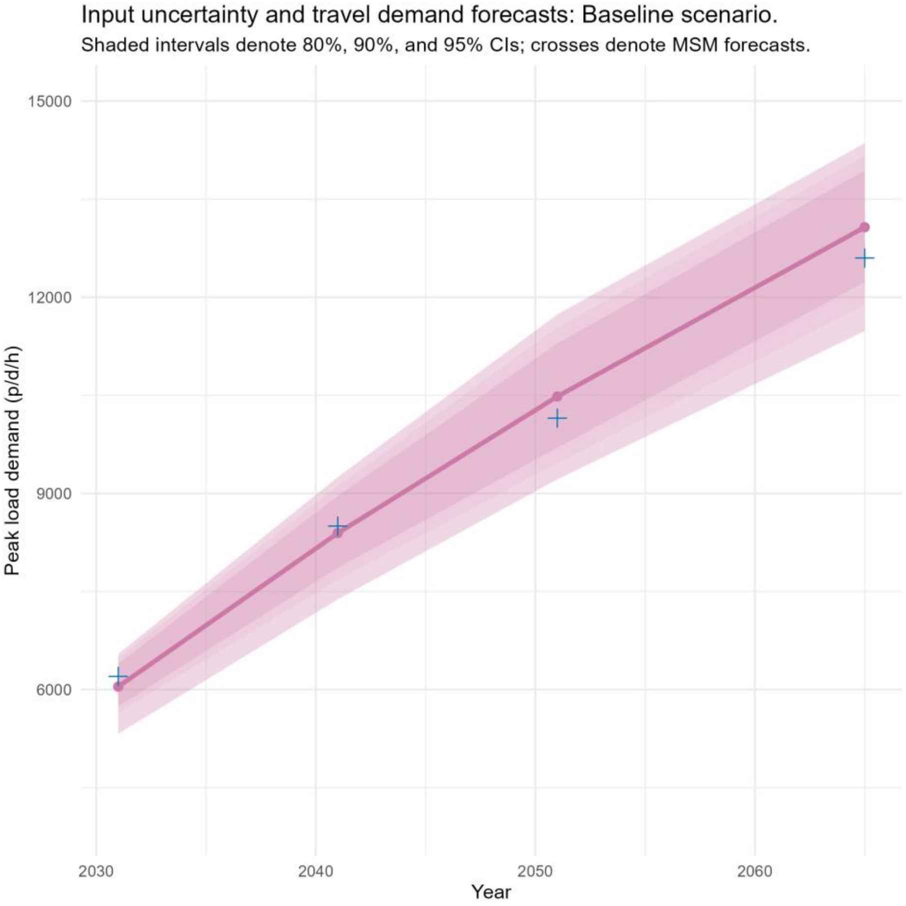
2. *Use simulations to quantify uncertainty.*
   - Assume probability distributions (and associated parameters) for each input of interest.
   - Run Monte Carlo simulation, i.e. repeatedly sample inputs from the assumed distributions.
   - Combine sampled inputs with random draws from parameters estimated in step 1.

This produces a **distribution** of travel demand forecasts for each future year.

Case study: Auckland Light Rail

143 model runs were undertaken during the development of the business case

Results


Input uncertainty and travel demand forecasts: Baseline scenario.
Shaded intervals denote 80%, 90%, and 95% CIs; crosses denote MSM forecasts.
15000
12000
9000
6000
Peak load demand (p/d/h)
2030
2040
2050
2060
Year

Final comments

Summary

- Quantify the effects of uncertainty in inputs on travel demand forecasts
- In our case study, variation in five input assumptions → uncertainty of approximately +/-10% (95% CIs)
- Further research:
    - Use systematic approaches ("orthogonal designs") to reduce run requirements
    - Ground assumptions used in the simulation in empirical analyses
    - Model correlations between inputs, e.g. population and employment.

Thank you & questions?

Step 1: Identify the effects of inputs

$$\log D_i = \beta_0 + \beta_1 \log P_i + \beta_2 \log E_i + \beta_3 \log I_i + \beta_4 \log C_i + \beta_5 \log F_i + \beta_6 W_i + \beta_7 \delta_i^C + \beta_8 \delta_i^P$$

Where for each of the 147 MSM runs denoted by i:

- $\log D_i$ denotes the natural log of passengers per hour per direction (p/h/d) at the peak load point (PLP) in the AM peak
- $\log P_i$ denotes the natural log of population in the corridor (000s)
- $\log E_i$ denotes the natural log of employment in the city centre (000s)
- $\log I_i$ denotes the natural log of in-vehicle travel-time by transit from the city centre to the airport, i.e. for the full length of the corridor (mins)
- $\log C_i$ denotes the natural log of total capacity at the peak load point (p/h/d, 000s)
- $\log F_i$ denotes the natural log of frequency (trains per hour)
- W_i denotes work from home uptake (%), which is set to either 7% or 14%
- δ_i^C denotes a dummy for congestion charging (0=No, 1=Yes)
- δ_i^P denotes a dummy for parking prices (0=No, 1=Yes)
- β denote parameters in the regression model to be estimated.

Note: Total capacity, C_i, and frequency, F_i, will be strongly positively correlated, because $C_i = V_i \cdot F_i$, where V_i = the capacity of rolling stock. To remove this correlation, future applications will model V_i and F_i instead of C_i and F_i.

Step 1: Identify the effects of inputs

Parameter	Variable	Interpretation / Measure	Estimate	s.e.	t-stat
β_1	$\log P_i$	Population in corridor (000s)	0.55	0.14	4.01
β_2	$\log E_i$	Employment in city centre (000s)	0.89	0.21	4.33
β_3	$\log I_i$	In-vehicle travel time (mins)	-0.85	0.06	-15.26
β_4	$\log C_i$	Capacity at PLP (p/h/d, 000s)	0.14	0.01	9.43
β_5	$\log F_i$	Frequency (trains per hour)	0.16	0.04	3.81
β_6	W_i	WfH (7%, 14%)	-2.85	0.45	-6.27
β_7	δ_i^C	Congestion charge (0=No, 1=Yes)	0.02	0.02	0.62
β_8	δ_i^P	Parking price increase (0=No; 1=Yes)	0.13	0.02	5.33
				R^2	0.912
				n	147

Step 1: Identify the effects of inputs

- All variables in the regression model have the expected sign and all but one (congestion charging) are estimated precisely (NB: Happy to discuss the details at the end).

- For continuous variables, the magnitudes of the estimated elasticities are plausible, e.g:

  - Population in the corridor, 0.55
  - Employment in the city centre, 0.89
  - In-vehicle time, -0.85
  - Frequency, 0.16
  - Capacity, 0.14

  NB: Elasticities reported in the literature typically don't separate the effects of frequency from capacity. As we control for both attributes, our elasticity for frequency will be smaller than normal.

- Doubling WfH from 7% to 14% is predicted to reduce peak demand for Auckland Light Rail by approximately ~20% (0.07*-2.85).

- See paper for more details.

Step 1: Identify the effects of inputs