Demand model estimation and validation

Daniel McFadden and Antti Talvitie

Main objective

- Estimate demand for Bay Area Rapid Transit (BART) public transport services in San Francisco ***prior*** to the introduction of BART by using disaggregate pre-BART data
- Model estimation and model validation based on surveys of individual commuters before and after the introduction of BART

Model of choice

- Multinomial logit model (MNL)
  - Why MNL? Satisfies Independence from Irrelevant Alternatives (IIA) → demand for transit modes that did not exist pre-BART can be predicted relatively easy
- Several model specifications are estimated by maximum likelihood

Travel modes pre-BART

- Auto alone
- Bus-with-walk-access (walk to bus station)
- Bus-with-auto-access (drive to bus station)
- Carpool

Additional travel modes after introduction of BART

- BART-with-walk-access
- BART-with-bus-access
- BART-with-auto-access

Two validation tasks

1. Compare actual transit mode shares in post-BART sample with transit mode shares predicted by pre-BART sample
2. Compare parameters estimated by post-BART sample with parameters estimated by pre-BART sample

Metrics for the comparison of modal shares

- Percent correct
  - Ratio of the diagonal elements in Figure 1 and the corresponding total column sum, i.e. the predicted probability of person n to choose a given alternative, summed over all individuals who actually chose this alternative is divided by the sum over all individuals of the probability of choosing this alternative
  - High total percent correct (i.e. sum of diagonal elements in Figure 1 divided by total number of individuals in the sample) indicates high accuracy of predictions
- Root mean squared error (RMSE)

Figure 1: Prediction success table

Actual Alternatives	Predicted Alternatives							Row Total	Actual Share (%)
	(1) Auto Alone	(2) Bus/Walk	(3) Bus/Auto	(4) BART/Walk	(5) BART/Bus	(6) BART/Auto	(7) Carpool		
(1) Auto Alone	246.5	21.11	5.984	17.63	1.188	10.91	74.71	378.0	59.53
(2) Bus/Walk	10.97	33.20	2.799	6.689	1.104	.8186	12.43	68.00	10.71
(3) Bus/Auto	1.108	2.367	.5929	1.571	.0057	1.041	2.314	9.000	1.417
(4) BART/Walk	.4660	.2832	.0758	1.307	.1829	.6247	1.060	4.000	0.6299
(5) BART/Bus	.7687	1.526	.0963	1.287	1.040	.1813	1.101	6.000	0.9449
(6) BART/Auto	7.350	2.302	1.365	7.709	.5612	6.201	7.512	33.00	5.197
(7) Carpool	70.65	11.39	3.071	11.62	1.149	5.228	33.89	137.0	21.57
Column Total	337.8	72.17	13.98	47.81	5.230	25.01	133.0	635.0	
Predicted Share (%)	53.19	11.37	2.202	7.529	0.8236	3.938	20.95		
Percent Correct	72.97	45.99	4.240	2.735	19.88	24.79	25.48		

Total Percent Correct: 50.82

Root Mean Squared Error: 9.53

Figure 2: Model comparison

TABLE 16 Predictions Based on Logit Models with Simple Specifications

	(1) Actual Share	(2) Predicted Share based on Model 10 (Final Model)	(3) Predicted Share based on Model 8 *	(4) Predicted Share based on Model 7 *	Predicted Share based on Model 11 ("naive" model)
Auto Alone	59.53	53.19	50.74	47.30	44.68
Bus/Walk	10.71	11.37	13.24	13.77	14.08
Bus/Auto	1.417	2.202	2.617	2.934	3.185
BART/Walk	0.630	7.529	7.587	8.335	10.60
BART/Bus	0.945	0.824	1.950	1.844	1.308
BART/Auto	5.197	3.938	3.020	3.133	4.073
Carpool	21.57	20.95	20.85	22.68	22.08
Root Mean Squared Error		9.53	11.82	15.06	18.33
Total Percent Correctly Predicted		50.82	48.37	44.12	42.33
	(n = 635)				

*The model was reestimated without the variable "Length of Residency in Community"

Main differences between predicted and actual demand

- Demand for BART-with-walk-access substantially over-predicted
- Share of auto alone commute underpredicted

Three possible drivers of inconsistencies

1. IIA assumption not satisfied
    - Two non-MNL models (i.e. no IIA assumption) are estimated on pre-BART data via two-step procedure
        1. Initial choice between auto alone, transit and carpool
        2. Choice among transit alternatives (e.g. bus-with-walk-access, BART-with-walk-access, ...)
    - Both predict somewhat better than MNL but not enough to explain the overprediction of MNL → failure of IIA not primary cause of overprediction

2. Non-genericity of attributes of BART-with-walk-access

   - Significant alternative-specific dummy variable in post-BART model → significant non-genericity
   - Indicates that BART-with-walk-access exhibits some important attributes that none of the pre-BART modes have

3. Incorrect data for walk times in the post-BART sample

Appendix

References

McFadden, D., & A. Talvitie (1977). *Demand model estimation and validation*. Urban Travel Demand Forecasting Project Phase 1 Final Report Series, Vol. V.