Estimating Values of Time on National Travel Survey data.

Juan Manuel Lorenzo Varela[1], Maria Börjesson, Andrew Daly*

The Value of Travel Time (VTT) is fundamental in transport economics. Since 1984 (MVA et al., 1984) best practice for VTT estimation has been to use Stated Choice (SC) data. However, there is now plenty of evidence of reference dependence and gain-loss asymmetry in SC data, implying that such data do not reveal long-term stable preferences. This is a serious problem since the value of time is often applied in welfare analyses, where long-term stability of the preferences is a key assumption. A potential reason for the strong reference dependence found in SC data is the emphasis on a short-term reference point often used in SC data to reduce hypothetical bias. In the long-run there is no stable reference point. Also, the use of Stated Choice data always raises the issue of the credibility of hypothetical responses.

Furthermore, the standard way of estimating the VTT using SC experiments is to estimate car VTT on car drivers, bus VTT on bus users, and so forth. With this approach, all marginal utilities will vary across modes because of self-selection, and because the direct utility of travel time varies between modes (due to differences in comfort, attractiveness, etc.). The self-selection will impact mode differences because travellers with low marginal utility of income and high resource value of time for the trip under study will tend to choose faster but more expensive trips and vice versa. Because of self-selection train travellers (faster but more expensive) will often have higher values of time than bus users (slower and cheaper), despite the usual preference for train over bus. This is found in the Swedish Value of Time study (Börjesson and Eliasson, 2014). However, self-selection can also arise from individual variation in the direct utility of travel time for any given mode.

An alternative to SC data is to use revealed preference (RP) data and a mode choice model to estimate the VTT. Observed behaviour has adapted to the (more stable) travel conditions and should thus be ruled by more long-term preferences. Many countries collect NTS (national travel survey) data and spend considerable resources on making them representative, which is an argument for using them for VTT estimation. However, a key problem in the use of NTS data for VTT estimation is measurement errors in the travel time and travel cost variables. Time and cost in NTS data is either self-reported or derived from a network assignment model.

In this paper we estimate the distribution of the VTT whilst controlling for errors in the self-reported and model computed time and cost variables. We achieve this by treating travel time and travel cost as latent variables in the estimation of a mode choice model. We use Swedish NTS data, and a Transcad network to simulate travel time and cost with the state-of-practice method in large-scale modelling. We test how different assumptions regarding the errors in input variables, latent variables and VTT distributions change the estimated VTT.

[1] KTH Royal Institute of Technology, Department of Transport Science, Stockholm, Sweden
* Contact email address: jmlv@kth.se

We admit that we face a potential identification problem, i.e. that the random error in the choice model cannot be separated from error in the latent variables, or the variability of VTT, without implementing strong assumptions. Assumptions of the error structure in the choice model influence the estimated errors in the time and cost variables, which influence the latent variables and these influence the estimated VTT distributions. We explore this issue by comparing different model formulations, to understand whether parameter estimates are robust against different modelling assumptions. We use maximum likelihood measurement-error models, focussing on their application to mode choice models. To our knowledge, no previous study on large-scale transport models has explored the impacts of different model assumptions in VTT estimation based on NTS data in this way.

Previous studies have shown that regression models - including discrete choice models - are sensitive to errors in variables if they are not accounted for; parameter estimates are biased towards zero, an effect known as regression dilution. Perception, reporting (e.g. rounding) and modelling errors are just a few of the errors transport models face; therefore, models that can account for these errors are of paramount importance to prevent biased estimates.

We start our work by analysing the range of trade-off values between time and cost of different alternatives, in order to understand for what range of bids the analysis reveals the VTT distribution. Then, we estimate a series of increasingly flexible models in willingness-to-pay space. We start by estimating a multinomial logit assuming no error in the input variables. Subsequent models use latent variables to control for errors in the time and cost variables. We estimate different model specifications, some of which exploit specific model assumptions of the distributions for the latent variables, and show how this influences results. Moreover, we test how the results depend on the assumed distribution of the input errors and the VTT distributions.

We apply the 2005/06 Swedish National Survey for the Greater Stockholm Region. The final sample consists of 3777 observations, of which 41% used Public Transport (PT), 48% car (driver and passenger) and 10% used active modes (walk and cycle)

Results indicate that time and cost variables normally used in mode choice models, whether reported or derived from networks, carry errors with them; hence, parameter estimates are diluted, and therefore biased. Estimated error distributions show that residuals for cost variables exhibit much larger variance than do time variables. This suggests that cost parameters incur larger errors than time parameters.

Moreover, when estimating a distributed VTT, we observe a reduction of its variance, when compared against an identical model specification that does not control for measurement errors in the input variables (time and cost). We also find that means of the estimated VTT distributions are reasonably robust to the modelling assumptions.

The resulting VTT from the different models are reported. Models not accounting for measurement errors yield higher values of time – between two and four times the values currently used in appraisal – than the models with latent variables. Results thus suggest that when the model do not account for errors in the travel cost variables, cost parameter estimates are diluted resulting in a too high VTT. Furthermore, VTT estimates

from the specifications with latent values yield lower estimates than current VTT from SC data used in appraisal. We believe that the results we have obtained are at the least promising for further research."

References

Börjesson, M. and Eliasson, J. 2014. Experiences from the Swedish Value of Time study. Transportation Research A, 59, 144–158.

MVA, ITS Leeds, and TSU Oxford (1987): 'The value of travel time savings', Policy Journals, Newbury, Berkshire.