

An artificial neural network based approach to uncover Value-of-Time distributions

Sander van Cranenburgh^a, Marco Kouwenhoven^a

^aTransport and Logistics Group, Delft University of Technology

Motivation

Without a doubt, the Value-of-time (VoT) is one of the most important and most researched notions in transport economics (Abrantes and Wardman 2011; Kouwenhoven et al. 2014; Hess et al. 2017). As travel time savings typically constitute a large part of benefits of transport projects, acquiring an as accurate as possible estimate of the (mean) VoT is considered important by many analysts and transport departments of governments.

Since the 1980s the dominant method to obtain VoT estimates is based on discrete choice theory in combination with Stated Choice (SC) data. However, despite the extensive knowledge that has accumulated over the years on estimating the VoT using discrete choice models, there are still a number of unresolved issues. One particular issue concerns how to uncover the underlying VoT distribution. Most studies that aim to do so use Mixed Logit (ML) models. Typically, the analyst tests a number of competing distributions (e.g. normal, log-normal, triangular, log-uniform) and then accepts the distribution that yields the highest model fit (after accounting for degrees of freedom).

However, a major disadvantage of this approach is that it is parametric: the analyst needs to presuppose the shape of VoT distribution. Whilst the analyst chooses the best fitting distribution, this distribution may overall still poorly represent the true underlying distribution. This, in turn, may result in inaccurate (mean) VoT estimates and in misguided policy recommendations. Therefore, in light of this shortcoming, recently a number of semi- and non-parametric methods to uncover VoT and Willingness-To-Pay (WTP) distributions from choice data have been proposed (e.g. Fosgerau 2007; Bastin et al. 2010; Fosgerau and Mabit 2013; Train 2016; Vij and Krueger 2017).

The purpose of this study is to contribute to the literature on non-parametric methods to uncover VoT and Willingness-To-Pay (WTP) distributions. In this paper we propose a new non-parametric approach, which is based on Artificial Neural Networks (ANNs). ANNs are highly flexible mathematical models which are loosely inspired by the structure and functional aspects of biological neural systems. Specifically, we develop a novel ANN based approach, designed to predict travellers' individual VoTs based on sets of observed choices of individuals. The overall VoT distribution is recovered by combining all individual VoT estimates. This approach is assumption free, in the sense that it does not impose any assumption on the VoT distribution, nor regarding the distribution of the error terms or the underlying decision-rule.

Methodology: an ANN based approach to uncover VoT distributions

The core idea behind our approach is as follows. We train an ANN to learn the functional relation, $f : X \rightarrow Y$, between a set of choices made in previous choice tasks (X) on the one hand and the choice made in the final choice task (Y) on the other hand. By formulating the (supervised) learning task in this way, the network essentially learns to represent a set of data

generating processes.¹ After having trained the ANN, we use the ANN to create, for each individual, a choice set² for which he or she is indifferent (conditional on the set of previous choices made by that individual). We do this for all individuals and infer for each individual his or her VoT based on this choice set. Finally, we recover the distribution of the VoT by combining all individual VoTs.

Monte Carlo experiments

To assess how well our ANN based approach is able to recover the underlying VoT distributions we have conducted Monte Carlo experiments. To do so, we have trained the ANN based on a (generic) training data set, with simulated series of choices. After having tested several different architectures, we find that a sparsely connected architecture consisting of two hidden layers performs best. Our network is able to correctly predict the final choice in approx. 90% of the cases in the hold-out sample. After having trained the network, we have presented it with a number of synthetic data sets with different underlying VoT distributions.

Figures 1 & 2 show the results for two synthetic data sets with respectively a normal and a log-normal true underlying VoT distribution. The four plots show (from left to right): a histogram of the true VoT, a histogram of the recovered VoT, a plot in which both the true and recovered VoT distributions are fitted, and a scatter plot – showing the correlation between the true and recovered VoTs. The figures clearly reveal that the ANN has not only been able to accurately recover both VoT distributions, but that it has also been able to fairly well recover the individual VoTs (see the scatter plots).

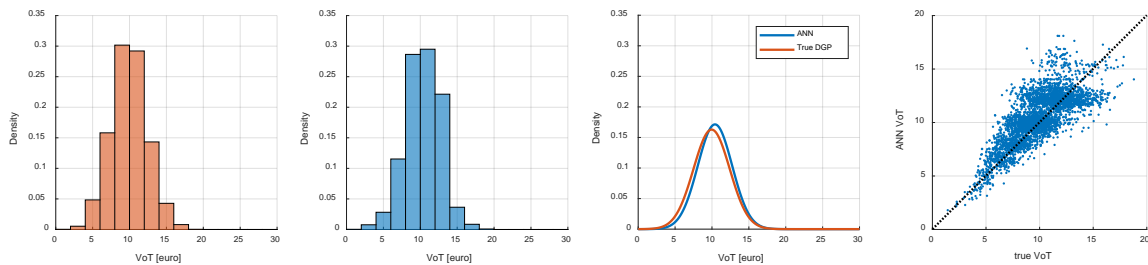


Figure 1: Normal distribution: True DGP (left), ANN (middle-left), Fits (middle-right), Scatter plot (right)

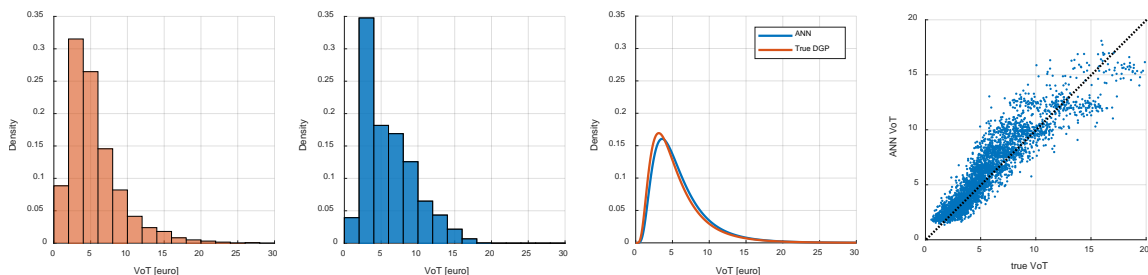


Figure 2: Log-normal distribution: True DGP (left), ANN (middle-left), Fits (middle-right), Scatter plot (right)

Empirical application

Next, we apply our method to empirical data. For this study we analyse SC data from the 1997 Dutch VoT study (HCG 1998). The data consist of about 4,000 respondents. Each respondent was confronted with eleven binary choice tasks, in which each alternative consisted of two attributes:

¹ Note that ANNs are capable of approximating any measurable function, given that sufficient processing neurons are available at the hidden layer and sufficient data is available for training. This property is known as the Universal Approximation Theorem (Cybenko 1989; Hornik et al. 1989).

² The choice sets are created such that they are in consonance with each individual's current travel time

travel cost and travel time. Using these data, we retrained the ANN to predict the eleventh choice based on the previous ten choices. This trained network is able to correctly predict 85% of these choices (in the hold-out sample). Figure 3 shows the recovered VoT distribution. The main observation in Figure 3 is that the shape of the VoT distribution is asymmetric and bounded. This asymmetric shape is behaviourally intuitive and has occasionally been found in previous VoT studies. Therefore, this result provides additional support for the notion that ANNs can be employed to uncover VoT distributions in a non-parametric fashion.

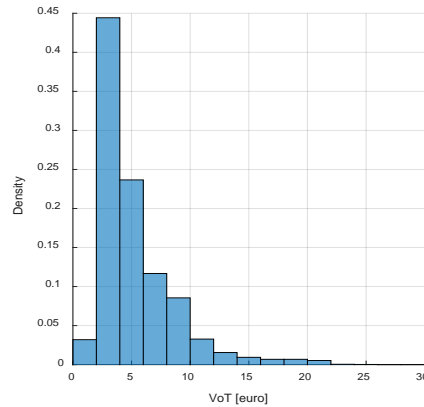


Figure 3: Recovered VoT distribution

Cross-validation

Lastly, we cross-validate our ANN results by comparing the mean VoT obtained using our ANN with those obtained using a series of commonly used discrete choice models (e.g. MNL, Mixed Logit, Latent Class, Random Valuation). We find that the mean VoT derived by the ANN is quite similar to what one finds when analysing these data using discrete choice models.

Conclusions and next steps

We believe the results from our ANN based approach are very encouraging. This suggests that ANNs could complement discrete choice based inference of VoT distributions. Next steps involve getting a better understanding of the sensitivity of our results towards design dimensions, such as the number of nodes, the size of the training and the balancing of the data during training. Furthermore, we aim to make further steps on how to deal with non-traders in the context of ANNs.

References

- Abrantes, P. A. L. & Wardman, M. R. (2011). Meta-analysis of UK values of travel time: An update. *Transportation Research Part A: Policy and Practice*, 45(1), 1-17.
- Bastin, F., Cirillo, C. & Toint, P. L. (2010). Estimating Nonparametric Random Utility Models with an Application to the Value of Time in Heterogeneous Populations. *Transportation Science*, 44(4), 537-549.
- Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems (MCSS)*, 2(4), 303-314.
- Fosgerau, M. (2007). Using nonparametrics to specify a model to measure the value of travel time. *Transportation Research Part A: Policy and Practice*, 41(9), 842-856.
- Fosgerau, M. & Mabit, S. L. (2013). Easy and flexible mixture distributions. *Economics Letters*, 120(2), 206-210.
- HCG (1998) *The second Netherlands' value of time study - final report*. Report 6098-1 The Hague.

- Hess, S., Daly, A., Dekker, T., Cabral, M. O. & Batley, R. (2017). A framework for capturing heterogeneity, heteroskedasticity, non-linearity, reference dependence and design artefacts in value of time research. *Transportation Research Part B: Methodological*, 96, 126-149.
- Hornik, K., Stinchcombe, M. & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural networks*, 2(5), 359-366.
- Kouwenhoven, M., de Jong, G. C., Koster, P., van den Berg, V. A. C., Verhoef, E. T., Bates, J. & Warffemius, P. M. J. (2014). New values of time and reliability in passenger transport in The Netherlands. *Research in Transportation Economics*, 47, 37-49.
- Train, K. (2016). Mixed logit with a flexible mixing distribution. *Journal of Choice Modelling*, 19, 40-53.
- Vij, A. & Krueger, R. (2017). Random taste heterogeneity in discrete choice models: Flexible nonparametric finite mixture distributions. *Transportation Research Part B: Methodological*, 106, 76-101.