

*Mode choice endogeneity in value of travel  
time estimation*

STEFAN L. MABIT AND MOGENS FOSGERAU

DTU Transport  
Technical University of Denmark

**Abstract**

It is often found that the value of travel time (VTT) is higher for car drivers than for public transport passengers. As there is no theory to support this finding this paper examines a possible explanation. The difference could be due to a selection effect. The result is an inability to measure the effect of a mode difference, e.g., comfort, between transport modes. We specify a model that captures the mode difference through a mode dummy and use econometric techniques that allow treatment of the mode dummy as the result of individual choice and hence endogenous. Using first a standard logit model we find a large and significant difference between the VTT for bus and car. When we allow for endogeneity, the mode parameter becomes smaller and just significant, meaning that the problem has been reduced. The difference in the sample mean VTT between modes is no longer significant.

# 1 Introduction

We are often concerned with the estimation of the value of travel time (VTT) as it plays an important role in demand models and appraisal. It is essential to recognise that the VTT is heterogeneous in the population. Some variation in the VTT may be explained by observed variables such as income and mode. In the stylised case of car versus bus, we would expect the mode effect to cause the VTT to be higher in bus, since riding a bus is mostly less comfortable than driving a car. On the other hand, we expect individuals with a high VTT, perhaps explainable by higher income, to be more likely to choose car. For this reason we may observe a higher VTT among car drivers than among bus passengers.

As mentioned by Wardman (2004) it is difficult to separate such mode and user-type effects. In many empirical studies the combined impact of the two effects is that individuals in car are seen to have higher VTT than for example in bus. This may be an adequate representation of the actual VTT in car and bus, but for many purposes this state of affairs is not satisfactory. For example, a traffic model that includes a higher VTT in car than in bus would predict a shift in the mode share towards bus if travel times were increased equally in the two modes. Such a result clearly does not make sense.

The typical way to address this issue is to specify separate models for each mode. These models are by their nature good at capturing heterogeneity within a specific mode. But they suffer a fundamental problem since they are estimated conditional on mode. This makes them good at assessment of specific projects where the mode-specific population may be reasonably assumed to represent the population of interest. On the other hand strong assumptions are necessary to apply mode-specific models to a population of mixed travellers where a improvement of the transport system may motivate some travellers to change modes.

Many applications have shown a gap between the VTT in car and the VTT in public transport. Discussions of how self selection may explain the differences found in VTT among modes are found in Mabit & Fosgerau (2006) and Fosgerau et al. (2007). In Mabit & Fosgerau (2006) they use a Heckman-type selection approach to model self-selection into modes. In Fosgerau et al. (2007), VTT is estimated both in a reference mode and an alternative mode to investigate whether differences among modes may be explained by strategic choices or self-selection. Their results favour the self-selection explanation.

This is the motivation behind the model presented in this paper. We present a model for the VTT in a population where a mode dummy are included to separate user effects from mode effects. In our estimation we then allow for the possible endogeneity of the mode dummy. Endogeneity means that the mode dummy may be correlated with the errors in the model, since the errors reflect unobserved characteristics of travellers which may also affect mode choice. Such correlation violates the assumptions underlying the standard estimation procedures and may cause bias.

Our approach for the estimation of models with endogenous explanatory

variables is taken from Lewbel (2004). The principle is to transform the discrete choice into a linear model using a special regressor and then utilise methods dealing with endogenous variables in linear models. Our model is simple, e.g., it does not take the panel dimension of data into account. This is the price we have to pay with the present econometric state of the art for being able to take endogeneity into account. Future research may be able to address the shortcomings of the methodology that we apply. We believe it is of interest to start working in the direction of being better able to deal with possible endogeneity.

The remainder of the paper is organised as follows. In section 2 we present our model for VTT together with an approach from Lewbel (2004) for the estimation of models with endogenous explanatory variables. Section 3.1 presents the data and section 3.2 discusses the model specification. Section 3.3 presents the estimation results. The final section 4 contains some concluding remarks.

## 2 Model formulation

### 2.1 VTT model

To model VTT we use the framework of discrete choice models based on random utility maximisation, see Train (2003). Here we restrict ourselves to binary choices. Suppose that individual  $n$  faces a choice between two alternatives and that each alternative has just two attributes: cost and time. We assume that the fast alternative is chosen if and only if

$$VTT_n > -\Delta C_n / \Delta T_n, \quad (1)$$

where  $VTT_n$  is the unknown VTT of individual  $n$  while  $\Delta C_n$  and  $\Delta T_n$  refer to the difference in the cost attributes and time attributes between the two alternatives. If we assume positive VTT, we may use a log transformation, and add a random error representing measurement and specification errors. The equation above then becomes

$$\ln(VTT_n) - \ln(-\Delta C_n / \Delta T_n) + \varepsilon_n > 0, \quad (2)$$

where  $\varepsilon_n$  is a random error term. This approach to model VTT is described in more detail in Fosgerau (2007).

To make the model empirically tractable we assume a linear parameterisation of  $\ln(VTT)$ . As discussed by Wardman (2004) we may split heterogeneity in VTT into mode and user-type effects. Here we include mode effects through a mode dummy while the remaining variables describe the individual and the choice context, i.e.,

$$\ln(VTT_n) = \beta' x_n = \beta' x_{1,n} + \gamma' y_n, \quad (3)$$

where  $y_n$  is the mode dummy. The final model describing whether an individual chooses the fast alternative,  $D_n = 1$ , becomes

$$D_n = I(\beta' x_{1,n} + \gamma' y_n - \ln(-\Delta C_n / \Delta T_n) + \varepsilon_n > 0), \quad (4)$$

where  $I(\cdot)$  is an indicator function equal to one if the expression in parenthesis is true and zero if it is false. Our interest lies in the estimation of  $\gamma$ .

## 2.2 Estimation

Suppose first that we are willing to make the standard strong assumption behind the logit model. That is we assume the model in equation 4 together with an assumption that  $\varepsilon$  is independent of  $x$  and identically logistically distributed. The model may then be estimated using maximum likelihood estimation.

Included in the logit model is the standard but strong assumption that  $x$  is exogenous. To allow for the possibility that  $y$  and  $\varepsilon$  are dependent we use an estimation procedure presented in Lewbel (2004). Suppose the model  $D = I(\beta'x_1 + \gamma'y + v + \varepsilon > 0)$  where  $v$  and  $x_1$  are exogenous variables,  $y$  is an endogenous variable, and that we have instruments  $z_1$  for  $y$ , i.e., additional variables that are correlated with  $y$  but not with  $\varepsilon$ .<sup>1</sup> Let  $s = (y', x'_1, z'_1)$  denote all variables except  $v$  and  $x = (y', x'_1)$ . Define  $v$  to be a very exogenous regressor (VER) if

1.  $v = g(\nu, s)$  for some function  $g$ ,  $g$  is differentiable, strictly monotonically increasing in its first element,  $\nu \perp s, \varepsilon$ , and  $\nu$  is continuously distributed.
2.  $\text{supp}(-[\beta'x + \varepsilon]) \subset \text{supp}(v|s)$ .

The first condition tells us that  $v$  depends on  $s$  and an independent error. This error term  $\nu$  is assumed independent of other explanatory variables  $s$  and  $\varepsilon$ . Lewbel notes that this is slightly stronger than ordinary exogeneity. The second condition tells us that the support of  $v$  is large. This is used to identify the  $\beta$  parameter. We will discuss both of these conditions later when we discuss our specific VER and the assumed distribution of  $\nu$ . In Lewbel (2004), the following theorem is proved using the notation  $z = (z'_1, x'_1)$  to denote the exogenous variables.

**Theorem I** Suppose we have the model  $D = I(\beta'x + v + \varepsilon > 0)$  and instrument,  $z$ , such that  $E(z\varepsilon) = 0$ . Assume that  $v$  is a VER, i.e.,  $\text{supp}(\beta'x + \varepsilon) \subset \text{supp}(-v|s)$ ,  $v = g(\nu, s)$ ,  $g$  is differentiable, strictly monotonically increasing in its first element,  $\nu \perp s, \varepsilon$ , and  $\nu$  is continuously distributed with density function  $f(\nu)$ . Define  $T$  and  $e$  by

$$T = \frac{D - I(v \geq 0)}{f(\nu)} \frac{\partial g(\nu, s)}{\partial \nu} \text{ and } e = T - \beta'x.$$

Then  $E(ze) = 0$ .

The theorem tells us that we may estimate the parameters  $\beta$  using two stage least squares regressing  $T$  on  $x$  using  $z$  as instruments.

---

<sup>1</sup>The normalisation of  $v$ 's coefficient to unity is not a restriction as long as we do not assume anything about the variance of  $\varepsilon$

Table 1: DESCRIPTIVE STATISTICS, 0 – 1 INDICATORS, SHARE= 1 IN PERCENT

variable	All	Bus	Car	description
$x_{area12}$	42.5	50.2	37.6	residence in Copenhagen
$x_{cars}$	17.6	10.3	22.2	more than one car in household
$x_{carsin}$	12.0	5.3	16.3	single with car
$x_{gender}$	46.4	53.3	42.0	female
$x_{grp2}$	27.2	13.2	36.1	travel w. household member
$x_{wage0}$	7.9	4.2	10.2	no wage earner in household

Table 2: MORE DESCRIPTIVE STATISTICS

variable	mean	std.dev.	min	max	mean(bus)	mean(car)
$x_{age}$ (years)	3.72	0.37	2.77	4.43	3.59	3.80
$x_{dist}$ (km)	3.08	1.16	0	6.18	2.49	3.45
$x_{inc}$ (scale)	1.15	0.56	0	2.20	1.05	1.20
$x_{jcost}$ (DKK)	3.04	1.15	0	6.62	2.68	3.27
$x_{jtime}$ (min)	3.35	0.79	1.39	6.17	3.18	3.46
$v$ (DKK/min)	-0.47	0.86	-3.00	1.72	-0.54	-0.43
$x_{\Delta T_{max}}$ (min)	8.52	4.25	0	14	8.87	8.30

In this article we use the special case assuming that the VER is normally distributed conditional on the other exogenous variables, i.e., we apply the assumption  $v = b's + \nu$ ,  $\nu \perp s, \varepsilon$ , and  $\nu \sim N(0, \sigma^2)$  with density function  $f(\nu)$ . Given this both of the conditions for a VER are fulfilled. As an example of a suitable  $v$  Lewbel mentions a bid determined by experimental design.

Neither of the procedures discussed above take random effects into account.<sup>2</sup> This could be a problem if data come as a panel with several choices for each individual. Our approach may seem simplistic in that we do not take any panel dimension into account. In this respect we have chosen to make one restriction to be able to relax another.

### 3 Data and Estimation

#### 3.1 Data

The data are from the 2004 Danish VTT study known as DATIV, see Fosgerau et al. (2006). Each individual made 8 stated preference (SP) choices in an unlabelled experiment referring to a current trip, i.e., car users only made car choices. Every choice was a binary choice where the alternatives were only described by travel time and cost. There was 1050 individuals of which 416 used bus and 634 car. This gave 3109 bus observations and 4867 car observations.

Table 1 summarises the 0 – 1 dummies used as explanatory variables. The attributes are summarised in Table 2 together with the continuous variables.

<sup>2</sup>Lewbel's approach does allow for heteroscedastic errors.

Here  $v$  denotes  $-\ln(-\Delta C/\Delta T)$  where  $\Delta T$  is the time attribute of the first alternative minus the time attribute of the second, and likewise  $\Delta C$  is the cost difference. The alternatives have been arranged such that alternative 1 is the fastest. The bid  $v$  is chosen as our VER. The variable was generated within the design of the SP experiment. This makes it plausible that the independence assumption holds concerning the remaining variables and the error term in the model. Furthermore the experimental design allowed for a wide range of bids. We discuss this issue further in section 3.3 as part of the estimation.

The variable  $x_{age}$  is the log of age,  $x_{dist}$  is the log of the distance in kilometres between origin and destination. If distance was zero,  $x_{dist}$  is set to zero. The variable  $x_{inc}$  is the log of gross personal income.<sup>3</sup> The variable  $x_{jtime}$  is the log of reported travel time. The variable  $x_{jcost}$  is the log of reported travel cost. The last variable is  $x_{\Delta Tmax}$ . This is defined as  $x_{\Delta Tmax} = \max(15 + \Delta T, 0)$ . This allows us to see how VTT varies for different time differences up to a difference of 15 min.<sup>4</sup>

## 3.2 Model specification

In the specification we have to decide which variables to use as  $x$  in equation 3. We use variables that describe cost and time as restrictions on these cause VTT to exist. In addition we use other design variables as they describe the choice context together with age and gender as the causal relation between VTT and these variables is clear. Based on preliminary testing we ended up with the specification below. Gender was excluded as it turned out to be insignificant. The testing considered both which variables to include and whether to use a linear or logarithmic parameterisation.

$$\begin{aligned} \ln(VTT) &= \beta_0 + \beta_{age}x_{age} + \beta_{jtime}x_{jtime} + \beta_{jcost}x_{jcost} \\ &+ \beta_{inc}x_{inc} + \eta_t S(t) + \eta_c S(c) \\ &+ \beta_{\Delta Tmax}x_{\Delta Tmax} + \beta_{mode}x_{mode}. \end{aligned}$$

The remaining variables  $S(t)$  and  $S(c)$  indicate how the cost and time attributes in the SP choice relate to the reference values. We define  $S(t)$  as the sign of  $(T_1 + T_2)/2 - jtime$  and likewise for  $S(c)$ . A positive coefficient on either corresponds to loss aversion concerning that attribute. We may also use these two variables to divide the choice situation into one of four types depending on the reference situation: equivalent gain, equivalent loss, willingness to pay, and willingness to accept. These concepts are discussed in detail in De Borger & Fosgerau (2007). They show that it is relevant to include these in the estimation of VTT since VTT differs significantly among the choice situations.

In the specification of the VER approach we suspect that the mode dummy is endogenous. We retain the assumption that the other variables are exogenous. We then have to find suitable instruments for the mode dummy. These

<sup>3</sup>Income is a discrete variable with 9 levels where level 1 is income below 100,000 DKK, level 2 is income between 100,000 DKK and 200,000 DKK, etc., until level 9 which is income above 800,000 DKK. One Euro is 7.5 DKK (050908).

<sup>4</sup>Beyond 15 min VTT was tested to be constant with respect to the time difference.

Table 3: LOGIT ESTIMATION

Explanatory variable	Estimated coefficient	Lower bound	Upper bound
Constant	0.61	-0.02	1.22
LN( age )	-0.63	-0.79	-0.45
LN( journey time )	-0.35	-0.45	-0.25
LN( journey cost )	0.44	0.37	0.52
LN( income )	0.60	0.51	0.71
S(t)	0.19	0.15	0.25
S(c)	0.13	0.08	0.18
Time difference less than 15	-0.04	-0.06	-0.03
Car dummy	0.59	0.47	0.72
Scale	1.02	0.95	1.09

Number of parameters: 10  
Log-likelihood at convergence: -4541  
Adjusted  $\rho^2$ : 0.18

should affect the mode choice and at the same time they should be redundant in the VTT equation conditional on the other variables. An ideal instrument would be a variable having a large influence on selection of mode but being uncorrelated with VTT conditional on mode. As instruments we use  $x_{area12}$ ,  $x_{cars}$ ,  $x_{carsin}$ ,  $x_{gender}$ ,  $x_{grp2}$ ,  $x_{wage0}$ , and  $x_{logdis}$ . Their explanatory power concerning mode choice is supported by visual inspection of tables 1 and 2. It was also supported by a binary logit estimation (not reported) of mode choice depending on the socioeconomic variables .

### 3.3 Model estimation

The estimation is performed by a program written in Ox, see Doornik (2001). The results for the logit model are given in Table 3.

The estimates are found using maximum likelihood estimation while the confidence limits are found by bootstrapping<sup>5</sup>, i.e., the model is estimated 400 times. The lower and upper bounds correspond to the boundaries on the 95% confidence interval for each coefficient. The estimation shows that all the parameters are significantly different from zero. The VTT decreases with age and increases with income which is reasonable. Furthermore the income parameter is a direct estimate of the income elasticity and 0.60 is a reasonable estimate. VTT decreases with journey time and increases with journey cost in contrast to some former studies. Many studies show increasing VTT with journey time which also has been supported theoretically in Jara-Diaz (1998). In our case we obtained a positive sign on the journey time parameter in preliminary estimations without journey cost. But as soon as journey cost was entered the

<sup>5</sup>Other procedures, e.g., Jackknife, were not considered. We chose bootstrapping since this was suggested by Lewbel.

Table 4: ESTIMATION USING A VER

Explanatory variable	Estimated coefficient	Lower bound	Upper bound
Constant	-0.02	-1.74	1.31
LN( age )	-0.44	-0.70	-0.13
LN( journey time )	-0.52	-1.04	-0.07
LN( journey cost )	0.43	0.11	0.86
LN( income )	0.54	0.34	0.79
S(t)	0.16	-0.02	0.34
S(c)	0.22	0.07	0.42
Time difference less than 15	0.03	-0.03	0.13
Car dummy	0.41	0.06	0.77

sign changed. This is not in contrast with the empirical findings supporting the positive sign as these do not include journey cost. The two sign variables relating to the framing of the choice experiment show the expected pattern, see De Borger & Fosgerau (2007). We have loss aversion concerning both time and cost with the aversion being largest for time. The time difference parameter shows a rising VTT with the time difference up to a difference of 15 min (when  $\Delta T$  increases then  $x_{\Delta Tmax}$  decreases).

The mode dummy parameter is the parameter of interest. If the logit model is correct we get a significant positive dummy. This dummy indicates that individuals who chose car on their reference trip have a higher VTT than those who chose bus, even when controlling for both socioeconomic characteristics and design variables. The final parameter is the scale on the logistic error. It is possible to estimate the scale as we restrict the coefficient on the bid  $v$  to unity. The scale is not significantly different from 1 but we keep it in the model to make the model more comparable to the VER approach.

Next we consider the estimation using the VER approach. The first step in the estimation procedure using the bid as a VER is to regress the bid  $v$  against all other variables. This is just a way to obtain a linear projection of  $v$  on the other variables and the regression coefficients have no interpretation. The only concern is whether the normality assumption is appropriate. In our case we regressed  $v$  against the other variables and investigated the normality assumption using a QQ plot. This check did not indicate problems with the normality assumption.

The estimation results are shown in Table 4. The upper and lower bounds refer to the 95% confidence interval found by bootstrapping. The parameters for age, journey time, journey cost, and income are significant and similar to the logit estimates. Concerning the design related variables, the parameter for  $S(t)$  is now insignificant, i.e., loss aversion has become insignificant for time. The car dummy becomes smaller but it is still positive and significant. But the estimate of 0.41 is clearly outside the 95% confidence interval of the logit estimates which has a lower bound of 0.47. This indicates a possible endogeneity problem in the

Table 5: EVALUATION OF MEAN VTT (DKK/HOUR)

	Bus		Car	
Logit	27.9	(24.6-33.0)	50.1	(44.9-58.0)
VER approach	16.8	( 5.9-30.7)	25.4	( 9.3-44.9)

mode dummy as the logit estimate is biased. Finally it has to be noted that this procedure does not estimate a scale coefficient. Therefore there is one coefficient less in Table 4.

Table 5 shows the mean VTT for both the logit estimates and the VER estimates. Confidence intervals are obtained through bootstrapping. The logit estimates indicate that the VTT in bus is only about 60 percent of the VTT in car. The mean VTT's are quite precisely estimated and the confidence intervals do not overlap. Even though we had a significant mode dummy we see that the mode-specific VTT's are no longer significantly different using the VER approach. The values are fairly small but if they are compared to values based on mixed logit estimates it should be noted that the present values corresponds to medians which are in general smaller.

## 4 Summary and conclusions

We have formulated a model where the chosen mode affects the VTT through a dummy. The model is estimated both as an ordinary logit model assuming that the mode dummy is exogenous and using a VER approach to account for the likely endogeneity of the mode dummy.

Results indicate that parameter estimates change significantly. The mode dummy estimate using the VER approach is 0.41. This falls outside the confidence interval between 0.47 and 0.72 from the logit estimation. This is an indication that the independence assumption of the logit estimation fails and that there is in fact some dependence between the mode dummy and the error term.

With logit estimates, the mode dummy is large and very significant. Using the VER approach, the mode dummy decreases, but it is still positive and (just) significant. So it would seem we were able to reduce the problem but that we did not succeed in accounting for all user-type effects. In the end, this is a question of having the right instruments. We seek variables that have a strong influence on mode choice, but which do not have a direct causal effect on the VTT.

An interesting and relevant extension would be to investigate whether the VER could be developed for the mixed logit model and whether it could account for panel data.<sup>6</sup> This would make the VER approach applicable to more of the models that are used in practice.

<sup>6</sup>In Lewbel (2004) a fixed effects approach is presented which is suitable for panel data. Unfortunately this does not allow for the estimation of a mode dummy.

Finally, we note that the VER approach does not assume a specific distribution for the error term  $\varepsilon$  in (2), whereas the logit model does. A better comparison for the VER approach would therefore be a model that does not assume a specific distribution for the error term. Such a model may be estimated using, e.g., the Klein & Spady (1993) estimator. This is an issue we will take up in future versions of this paper.

## References

- De Borger, B. & Fosgerau, M. (2007), Discrete choices and the trade-off between money and time: A test of the theory of reference-dependent preferences. working paper, [mpira.ub.uni-muenchen.de/3904/](http://mpira.ub.uni-muenchen.de/3904/).
- Doornik, J. (2001), *Ox: An Object-Oriented Matrix Language*, Timberlake Consultants Press, London.
- Fosgerau, M. (2007), 'Using nonparametrics to specify a model to measure the value of travel time', *Transportation Research Part A* **41**, 842–856.
- Fosgerau, M., Hjorth, K. & Lyk-Jensen, S. V. (2006), *An integrated approach to the estimation of the value of travel time*, European Transport Conference, Strasbourg, 2006.
- Fosgerau, M., Hjorth, K. & Lyk-Jensen, S. V. (2007), *Between mode differences in the value of travel time: Self-selection or strategic behaviour?*, European Transport Conference, Leiden, 2007.
- Jara-Diaz, S. R. (1998), Time and income in travel choice: Towards a microeconomic activity-based theoretical framework, in T. Gärling, T. Laitia & K. Westin, eds, 'Theoretical Foundations of Travel Choice Modelling', Pergamon, Oxford, pp. 51–73.
- Klein, R. W. & Spady, R. H. (1993), 'An efficient semiparametric estimator for binary response models', *Econometrica* **61** (2), pp. 387–421.
- Lewbel, A. (2004), Simple estimators for hard problems: Endogeneity in discrete choice related models. working paper.
- Mabit, S. L. & Fosgerau, M. (2006), *Controlling for sample selection in the estimation of the value of travel time*, paper presented at the 11<sup>th</sup> International Conference of Travel Behaviour Research, Kyoto, Japan.
- Train, K. (2003), *Discrete Choice Methods with Simulation*, Cambridge University Press, New York, NY, USA.
- Wardman, M. (2004), 'Public transport values of time', *Transport policy* **11**, 363–377.