Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 00 (2018) 000–000

Transportation Research Procedia

www.elsevier.com/locate/procedia

XIII Conference on Transport Engineering, CIT2018

Parameter bias in misspecified Hybrid Choice Models: an empirical study.

"Juan Manuel Lorenzo Varela" *

"KTH Royal Institute of Technology, 114 28 Stockholm ,Sweden"

Abstract

Model misspecification is likely to occur when working with real datasets. However, previous studies showing the advantages of hybrid choice models have mostly used models where structural and measurement equations match the functions employed in the data generating process, especially when parameter biases were discussed. The aim of this study is to investigate the extent of parameter bias in misspecified hybrid choice models, and assess if different modelling assumptions impact the parameter estimates of the choice model. For this task, a mode choice model is estimated on synthetic data with efforts focus on mimicking the conditions present in real datasets, where the postulated structural and measurement equations are less flexible than the functions used to generate the data. Results show that hybrid choice models, even if misspecified, manage to recover better parameter estimates than a multinomial logit. However, hybrid choice models are not unbeatable, as results also indicate that misspecified hybrid choice models might still yield biased parameter estimates. Moreover, results suggest that hybrid choice models successfully isolate the source of model bias, preventing its propagation to other parameter estimates. Results also show that parameter estimates from hybrid choice models are sensible to modelling assumptions, and that parameter estimates of the utility function are robust given that errors are modelled.

Keywords: Hybrid Choice Models (HCM); Integrated Choice and Latent Variable models (ICLV); Mode choice; Latent variables; Model misspecification; Parameter bias; Synthetic dataset

* Corresponding author. Tel.: +46-8-790-66-95.
E-mail address: jmlv@kth.se

2352-1465 © 2018 The Authors. Published by Elsevier Ltd.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/)
Selection and peer-review under responsibility of the scientific committee of the XIII Conference on Transport Engineering, CIT2018.

1. Introduction

The use of Hybrid Choice Models (HCM) has grown exponentially during the last decade. These models are motivated by findings in the social sciences, where evidence supports that latent variables (attitudes, norms, perceptions, affects and/or beliefs) can often override the influence of observable variables on disaggregate behaviour. Making the HCM operative requires several modelling assumptions, which include the specification of structural and measurement equations. Through these assumptions, researchers postulate theories in the form of statistical models and test how well the theories model the observed data. Frequently, and especially when working with real datasets, the postulated theories may not be accurate and result in misspecified models. Hence, we should ask ourselves how the estimation results might be affected by model misspecification. Nearly all models are to some degree structurally misspecified, (Browne and Cudeck, 1992). As a result, most of the ideal properties of the maximum likelihood estimator need not hold in the real world, (Bollen et al., 2007). Following this line of reasoning, Kolenikov (2011) dealt with the problem of quantifying the degree to which parameter estimates in a structural equation model (SEM) can be biased when structural relationships were not specified correctly, and proposed a framework to assess the degree to which the parameter estimates may be biased.

Although considerably research has been devoted to structural misspecification bias in SEM, rather less attention has been paid to its effects in HCM. For instance, Walker et al. (2010) studied how to estimate travel demand models when the underlying quality of level of service data is poor. In this study, the authors used synthetic data to test the capabilities of the HCM framework to correct measurement errors in explanatory variables; and found that the HCM framework was able to accurately estimate the true value of the parameters without knowing the true travel time. Nevertheless, the estimated HCM was correctly specified and matched the formulation used for the data generating process. A more recent study (Vij and Walker, 2016) systematically evaluated the benefits of the HCM framework in comparison with a more traditional choice model without latent variables. In this study, the authors carried out a detail analysis of goodness-of-fit and bias of the parameter estimates through different Monte Carlo experiments with synthetic data. Among these experiments, the HCM in experiment III had a misspecified utility function, whilst the HCM in experiment IV had misspecified structural equations. Unfortunately, the focus of those experiments was the goodness-of-fit of the HCM compared to a reduced form mixed logit, and parameter estimates were not discussed, neither provided. Furthermore, in their discussion of measurement error bias, results from another Monte Carlo experiment were given, but the structural and measurement equations of the HCM tested were not only correctly specified, but also the structural equation included an extra parameter. In other words, the model tested was more flexible than the one used for the data generation. In their analysis, the hypothesis that the mean parameter estimates were equal to the true values could not be rejected.

Whilst these exercises provided useful insights into the HCM capabilities to correct for measurement errors, I argue that misspecification effects expected from the complexity of real datasets, where the specified HCM is not flexible enough to model the data generation process, might not have been captured. The aim of this paper is to investigate the extent of parameter bias in misspecified HCM. Efforts are focus on mimicking the conditions present when real datasets are used, and the assumed structural and measurement equations are less flexible than the functions used for the data generating process. Furthermore, the paper compares two different HCM formulations to explore how modelling assumptions needed to estimate the HCM affect parameter estimates of the choice model. Finally, biases in parameters are discussed in relation to goodness-of-fit measurements frequently used in testing models with latent variables; and I show that, contrary to what intuition and common practices indicate, a better fitting model does not always guarantee better parameter estimates.

The paper is structured as follows: Section 2 describes the synthetic data generation process. Section 3 gives a description of the HCM framework as well as the models and goodness-of-fit measurements used in this study. Section 4 presents the results, and Section 5 concludes.

2. Data

For this study, 10000 synthetic observations for a multinomial logit (MNL) mode choice model were generated. Synthetic datasets are especially useful for bias quantification because the functions and true parameters underlying the data generating process are known. The MNL model includes 3 alternatives and explanatory variables of time and cost, along with alternative specific constants. The utilities are specified as:

$$U_1 = \beta_{t1} \cdot t_1 + \beta_c \cdot c_1 + asc_1 + \varepsilon_1, \quad (1)$$
$$U_2 = \beta_{t2} \cdot t_2 + \beta_c \cdot c_2 + asc_2 + \varepsilon_2, \quad (2)$$

$$U_3 = \beta_{t3} \cdot t_3 + \beta_c \cdot c_3 + asc_3 + \varepsilon_3, \tag{3}$$

where, U_i is the utility for alternative i={1,2,3}, β_{ti} the parameter for travel time, β_c the parameter for travel cost, t_i the travel time, c_i the travel cost, asc_i the alternative specific constant, and ε_i are independent and identical distributed (iid) extreme value errors. Time and cost variables are drawn independently from a mixture of gaussian and lognormal distributions described in *Table 1*, with mixing proportions 50-50%; parameters β_{ti} and β_c are set to the values reported in *Table 3* and ε_i are drawn independently from the standard Gumbel distribution.

Table 1. Distributions used to generate the synthetic variables.

		Time1	*Time2*	*Time3*	*Cost1*	*Cost2*	*Cost3*	*Time1_measured*
Normal distribution								
	μ	55	40	25	10	25	15	0
	σ	30	20	15	5	10	5	5
Lognormal distribution								
	$log(\mu)$	3.5	3.15	2.95	2.2	2.5	3.1	0
	$log(\sigma)$	0.3	0.4	0.6	0.46	0.55	0.39	1

In addition, to better mimic the conditions of real datasets, an imperfect measurement for t_1 ($t_{1\,measured}$) was generated. Measurement errors introduced in $t_{1\,measured}$ consist of two components; one multiplicative and other additive. These components represent typical sources of measurement error in our models. For instance, multiplicative disturbances (γ) might be expected in level of service time variables from assignment models when the assumed average speed is incorrect; and additive measurement errors (α) might represent the aggregation effect into zone centroids, or rounding effects when using self-reported variables. Below I present the generating function for this variable,

$$t_{1\,measured} = time_1 * \gamma + \alpha \tag{4}$$

where γ is lognormally distributed with parameters $log(\mu)$ = 0 and $log(\sigma)$ = 1; and, α is normally distributed with zero mean and $\sigma = 5$. I acknowledged that other measurement error definitions are possible, for instance psychological research has found evidence suggesting that perceived time follows a power function of the clock time (see Roeckelein, 2000), and that the ability of a particular HCM formulation to recover good parameter estimates might depend on the type of measurement errors. Therefore, the adopted function in equation (4) for $t_{1\,measured}$ is not expected to represent all the errors present on all real datasets, but it is expected to share some of their characteristics, such as the existence of imperfect measurements, guarantying that the HCM we estimate in this study are misspecified.

Finally, observations where time and/or cost variables have negative values were excluded from the synthetic dataset used for estimation, leaving a total of 9719 observations. This was done to maintain theoretical consistency with the types of variables being modelled, as times and cost cannot yield negative values. Using these observations, utilities are calculated for each of the synthesized observations, and the 'chosen' alternative is set to be the one with the highest utility. Note that removing observations with negative time and/or cost variables will not bias the estimation results, as these observations were removed prior to the calculation of the utilities and setting of the chosen alternatives. Distributions of the synthetic variables are shown in *Figure 1* below. Here we can observe how measurement errors modify the shape of the variable *time1*, where the resulting variable with measurement errors has thicker tales and lower mean.

3. Methodology

3.1. Hybrid choice model framework

The HCM framework has two main components: a discrete choice model and a latent variable model. According to the components included, this structure has also been referred to as the integrated choice and latent variable (ICLV) model. A HCM allows us to introduce observable variables in the utility functions, as well as unobservable ones. The key concept is that while the value of a latent variable is unknown, an approximation is available in the value measured with error, the indicator. The approximated and true values are then connected via a measurement relationship. The model schematics are shown by *Figure 2*, where observed variables, indicators and choices are represented by rectangular boxes, whilst unobserved variables such as utilities and latent variables are represented by ellipses. In addition, structural equations are represented by continuous lines and measurement equations by dashed lines.

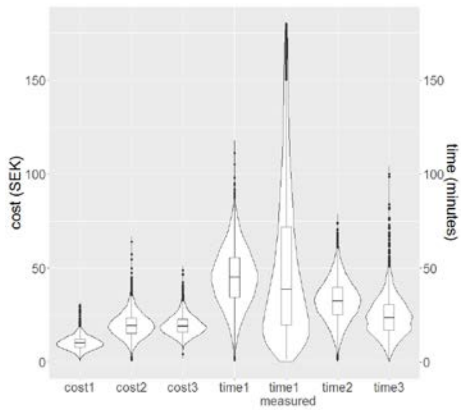

Fig. 1. Violin plots of synthetic variables

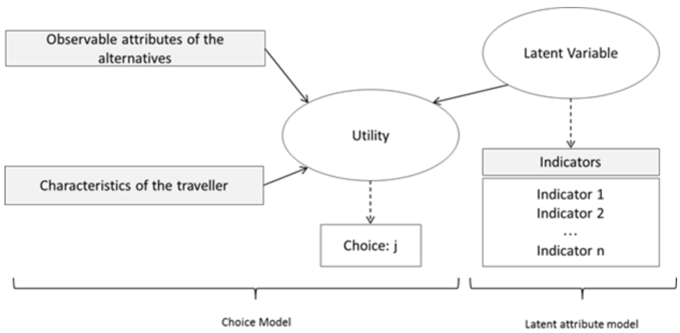

Fig. 2. Hybrid Choice Model framework

Equations for the HCM framework were presented in Walker et al. (2010) as follows; First, we choose a variable to be treated as a latent variable (X), known only to a distribution f_X and a set of estimated parameters θ:

$$f_X(X;\theta) \quad (5)$$

In terms of the mode choice model, if the explanatory variable were known, we would have a typical mode choice model that depicts the probability of an individual choosing a mode *i* conditional on a set of estimated parameters β and explanatory variables (which includes the true value of *X*):

$$P(i|\beta,X) \quad (6)$$

However, since X is unknown, it is necessary to integrate the conditional choice probability over the distribution of X:

$$P(i|\beta,\theta) = \int P(i|\beta,X)\, f_X(X;\theta)dX \quad (7)$$

The measurement equation now comes into play to incorporate the measured value as an indicator of X. For this, the distribution of the indicator (I) conditional on X and a set of estimated parameters λ is necessary:

$$f_M(I|X;\lambda) \quad (8)$$

Assembling all elements together, the likelihood function of the entire framework shown in Figure 1 is:

$$L(i,I|\beta,\theta,\lambda) = \int P(i|\beta,X)\, f_X(X;\theta) f_M(I|X;\lambda)dX \quad (9)$$

Finally, the unknown parameters (β, θ, λ) can be estimated using maximum likelihood estimation from observed modal choices.

3.2. Models and goodness-of-fit measurements.

Making the HCM framework operative requires assumptions on the functional form of the choice model, equation (6), the prior distribution of latent variables, equation (5), and the definition of the measurement models, equation (8); hence, models with different assumptions will typically result in different specifications. In this paper we define two HCM as described below.

3.2.1. HCM1: Latent variable with normal prior and additive error model.

This is how most of the HCM are specified in practice. First, the latent variable (*LV*) is assumed to be normally distributed, $LV \sim N(\mu\ \ , \sigma^2)$. Second, the measurement equation is modelled as additive, and residuals are assumed to follow a normal distribution with zero mean. This formulation implies that indicators are centred on the true value of the latent variable; that errors are introduced by the measuring device; and that the magnitude of the error is independent from the value being measured. Mathematically, this model is expressed by equations (6)-(11).

$$U_1 = \beta_{t1} \cdot LVt_1 + \beta_c \cdot c_1 + asc_1 + \varepsilon_1 \quad (6)$$
$$LVt_1 = \mu + \sigma \cdot \phi \qquad with\ \phi \sim N(0, 1^2) \quad (7)$$
$$t_{1\,measured} = LVt_1 + \eta \quad (8)$$
$$\eta = \sigma_{error} \cdot \phi' \qquad with\ \phi' \sim N(0,1^2) \quad (9)$$
$$U_2 = \beta_{t2} \cdot t_2 + \beta_c \cdot c_2 + asc_2 + \varepsilon_2 \quad (10)$$
$$U_3 = \beta_{t3} \cdot t_3 + \beta_c \cdot c_3 + asc_3 + \varepsilon_3 \quad (11)$$

where U_i is the utility for alternative i={1,2,3}; β_{ti} the parameter for travel time, β_c the parameter for travel cost; LVt_1 the Latent Variable modelling t_1 with mean μ, standard deviation σ, and indicator $t_{1\,\mathrm{measured}}$; t_i the travel time; c_i the travel cost; asc_i the alternative specific constant; ε_i are iid errors from a standard Gumbel distribution; η the measurement error with parameter σ_{error} ; and, ϕ and ϕ' are draws from a standard normal distribution.

3.2.2. HCM2: Latent variable with lognormal prior and multiplicative error model.

Due to the nature of the latent variable being modelled (travel time), we define a second specification that exploits the fact that time variables must be positive. For this, we select the lognormal distribution, which has support on the required interval (0,+∞), and the additive measurement equation is replaced by a multiplicative one. This formulation implies that measurement errors are proportional to the values being measured, and can be interpreted as scaling factors; hence, it is critical to guarantee that the error term will be always positive. Again, this is achieved by modelling the error term with a lognormal distribution.

Finally, the formulation assumes that measurement errors will have an expected value of 1. This condition is similar to the additive errors having mean zero, and it is necessary to make the model identifiable. In our particular case, the expected value of a lognormally distributed variable equals one as long as its parameters fulfil the following condition: $\mu_{error} = -0.5 \cdot \sigma_{error}{}^2$. This model is expressed by equations (12)-(17).

$$U_1 = \beta_{t1} \cdot LVt_1 + \beta_c \cdot c_1 + asc_1 + \varepsilon_1 \quad (12)$$
$$LVt_1 = e^{(\mu + \sigma \cdot \phi)} \qquad with\ \phi \sim N(0, 1^2) \quad (13)$$
$$t_{1\,measured} = LVt_1 * \eta \quad (14)$$
$$\eta = e^{(-0.5 \cdot \sigma_{error}{}^2 + \sigma_{error} \cdot \phi')} \qquad with\ \phi' \sim N(0,1^2) \quad (15)$$
$$U_2 = \beta_{t2} \cdot t_2 + \beta_c \cdot c_2 + asc_2 + \varepsilon_2 \quad (16)$$
$$U_3 = \beta_{t3} \cdot t_3 + \beta_c \cdot c_3 + asc_3 + \varepsilon_3 \quad (17)$$

where again U_i is the utility for alternative i={1,2,3}; β_{ti} the parameter for travel time, β_c the parameter for travel cost; LVt_1 the Latent Variable modelling t_1 with mean μ, standard deviation σ, and indicator $t_{1\,\mathrm{measured}}$; t_i the travel time; c_i the travel cost; asc_i the alternative specific constant; ε_i are iid errors from a standard Gumbel distribution; η the measurement error with parameter σ_{error}; and, ϕ and ϕ' are draws from a standard normal distribution.

3.2.3. Model selection and goodness-of-fit.

Despite the increasing popularity of the HCM framework, discussion of the accuracy of measurement of latent variables in the HMC framework is largely absent in transportation research (Motoaki and Daziano, 2015). Unfortunately, traditional goodness-of-fit measurements of transport demand models, (e.g. likelihood ratio test and ρ^2) cannot be used to assess model fit, reliability, and validity of HCM; hence, researchers face difficulties to evaluate the accuracy of measurement of the latent variables, as there still is no consensus about how to test the HCM goodness-of-fit. A common practice to evaluate how well the data supports the HCM modelling assumptions is to test the normality assumption of the model residuals. This task is frequently carried out by Quantile-Quantile plots (QQplots). A QQplot is a graphical technique for determining if two data sets come from populations with a common distribution, and consists in plotting the quantiles of the first data set against the quantiles of the second data set. If the two sets come from a population with the same distribution, the points should follow a straight line. Other current practices to evaluate goodness-of-fit in HCM come from studies in structural equation modelling, where there are established goodness-of-fit measurements such as the chi-square statistic, the root mean square error of approximation (RMSEA), and fit indexes such as the Bayesian Information Criterion (BIC), and the Akaike Information Criterion (AIC). Hauber et al. (2016), presents this two indexes as:

$$BIC = -2LL + ln[sample\ size]K, \tag{18}$$

$$AIC = -2LL + 2K, \tag{19}$$

where *LL* is the log-likelihood of the full model, and *K* is the number of parameter estimates corresponding to the number of explanatory variables in the model. These are comparative measures of the relative quality of statistical models for a given dataset, and they evaluate the plausibility of the models focusing on minimizing information loss. An advantage of using these indexes is that comparing multiple models becomes trivial, as the model with the lowest value is preferred. On the other hand, these goodness-of-fit measurements are based on the model's final likelihood; hence, when applied to HCM with different number of latent variables and or different measurement equation formulations, the goodness-of-fit measurements might provide counterintuitive results. Some authors have expressed their concerns about these goodness-of-fit measurements. For instance, Barrett (2007) and Ropovik (2015), state that fit indexes add nothing to the analysis, and Hayduk et al. (2007) argue that fit index thresholds can be misleading and subject to misuse. Moreover, Motoaki & Daziano (2015), showed through a Monte Carlo experiment that the behaviour of SEM fit assessment tools did not work as expected for the HCM. Nevertheless, is common to find fit indexes reported in current literature despite of these criticisms. In this study, the BIC and AIC goodness-of-fit indexes are reported to inform further discussions.

4. Results

In this study, three different models are estimated using pythonBiogeme, (Bierlaire, 2016). These models include one multinomial logit (MNL) model and the two different Hybrid Choice Models (HCM) defined in Section 3. All three models assume that the time variable for alternative 1 suffer from measurement errors. Hybrid choice models have been estimated using 5000 draws. Parameter estimates, log-likelihood and goodness-of-fit measurements for the models are reported in *Table 2*.

Table 2. Estimation results.

Parameter	True value	*MNL*		*HCM1*		*HCM2*	
		Estimate	t-test	Estimate	t-test	Estimate	t-test
$asc2$	-1.00	3.940	18.96	-2.29	-1.65*	-1.96	-1.34*
$asc3$	-2.00	3.560	18.06	-2.89	-2.04	-2.63	-1.75
β_{cost}	-1.00	-0.615	-41.64	-0.991	-12.45	-0.999	-12.35
β_{t1}	-0.25	-0.006	-10.63	-0.143	-5.16	-0.135	-4.73
β_{t2}	-0.25	-0.149	-24.09	-0.236	-11.60	-0.238	-11.46
β_{t3}	-0.30	-0.195	-19.90	-0.305	-11.32	-0.304	-11.26
Structural equations							
μ	Refer to Table 1	-	-	63.5	89.43	4.15	284.79
$log(\sigma\ \ )$		-	-	3.24	31.54	-0.976	-8.85
Measurement equations							
$log(\sigma_{error})$	Refer to Table 1	-	-	4.21	429.30	-0.0134	-34.21
Goodness-of-fit measurements							
Log-likelihood	-	-2123		-69480		-28408	
BIC	-	4301		139043		56899	
AIC	-	4258		138978		56834	

* Parameter not statistically different from zero at 95% confidence level

4.1. MNL Model estimation results

The first set of estimation results in Table 2 shows how measurement errors bias the MNL parameter estimates. As can be seen, the largest bias can be found on β_{t1}, where the estimated parameter *(-0.006)* is heavily diluted when compared with its true value *(-0.25)*. However, parameter bias not only appears in the coefficients of the variable suffering from measurement errors, *time 1*, but also spreads to all other parameter estimates. As a consequence, these biased parameters seriously underestimate the value of time for alternative 1, the ratio of the time and cost parameters, with a value of *-0.009*, as oppose to the true ratio of *-0.25*.

4.2. HCM1 Model estimation results

When working with HCMs, it is good practice to look at how well the modelling assumptions fit the observed data. For this task, a QQplot is commonly used as a graphical diagnostic tool of the normality assumption of the measurement equation error terms. Remember, as mentioned in Section 3.2, that in a QQplot if the two sets come from a population with the same distribution, the points should follow a straight line. *Figure 3* shows the QQplot of the simulated residuals for the additive measurement error formulation. It is evident, that the plotted line departs from a straight line, suggesting that the normality assumption of the residuals from the additive measurement error formulation might not be accurate. Nevertheless, *HCM1* manages to recover parameter estimates that are closer to the true values than the ones provided by the MNL. For instance, results in *Table 2* show that *HCM1* provides a better estimation of β_{t1} *(-0.143)* than the MNL *(-0.006)*, when compared with its true value *(-0.25)*. Furthermore, it seems that the HCM formulation is successful at isolating the source of the measurement error, preventing the propagation of the bias to other parameter estimates. Accordingly, this model provides a value of time for alternative 1 of *-0.14*; which is much closer to the true value of *-0.25*.

4.3. HCM2 Model estimation results

In the case of the multiplicative measurement error model, residuals were assumed to be lognormally distributed; hence, their logarithmic transformation should follow a normal distribution. Thus, a QQplot can be used to assess the normality assumption of the measurement equation error residuals. As can be seen in *Figure 4*, the multiplicative error formulation provides a better fit to the observed data than the additive formulation, as simulated residuals for the multiplicative error formulation show a smaller departure from a straight line than simulated residuals from the additive measurement error model. Furthermore, results from additional goodness-of-fit measurements (*BIC* and *AIC* indexes) also favour the use of a multiplicative error structure.

Results show that the modelling assumptions implemented in this model *(HCM2)* are also capable of isolating the source of the measurement error, preventing the propagation of the bias to other parameter estimates. However, despite of all this evidence in favour of the multiplicative measurement error model, parameter estimates from *HCM2* are not better than the ones provided by *HCM1*. For instance, results in *Table 2* show that β_{t1} from *HCM2 (-0.135)* is slightly more underestimated than the parameter from *HCM1 (-0.143)*, and that both underestimate its true value *(-0.25).* Nevertheless, modelling assumptions might impact the model scale; hence, we should look at the ratio of parameters which is scale independent. Comparing the values of time, we observe that *HCM2* yields a value of time for alternative 1 of *-0.13*, whilst *HCM1* yields a value of *-0.14*. Looking at the parameter estimates of β_{t1} from models *HCM1* (normal priors and additive error model) and *HCM2* (lognormal priors and multiplicative error model), we can observe that the two parameter values differ by less of one standard deviation: Hence, there is no statistical evidence to support that this values are different from each other. This suggests that parameter estimates of the choice model are reasonably robust to these modelling assumptions.

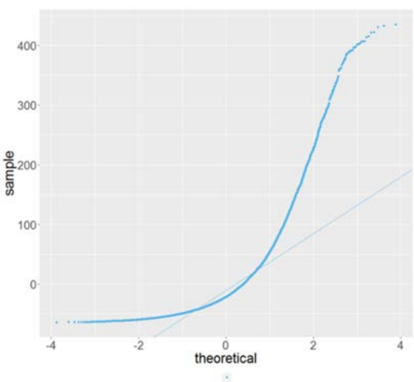

Figure 3. QQplot of the time variable residuals following an additive error formulation

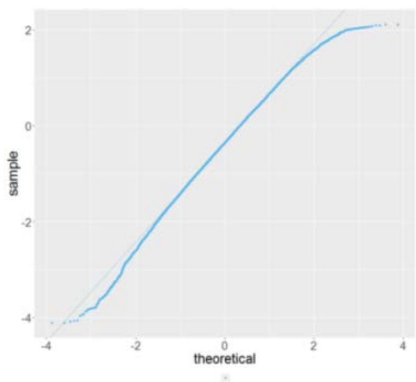

Figure 4. QQplot of the time variable residuals following a multiplicative error formulation

5. Conclusions

Model misspecification is likely to occur when working with real datasets. Despite of this fact, previous empirical studies showing the advantages of the HCMs have mostly used correctly specified models, especially when parameter estimate biases are discussed. However, the paper at hand explores the extent of parameter bias in misspecified hybrid choice models, through Monte Carlo experiments using synthetic data. Efforts are made on mimicking the conditions present when real datasets are used, and the assumed structural and measurement equations are less flexible than the functions used for generating the data. The main conclusions can be summarized as follows:

- Parameter estimates of the choice model seem robust to modelling assumptions required to estimate the HCM.
- The HCM framework seems to be able to isolate the source of the measurement error, and prevent the propagation of the bias to other parameter estimates.
- Whilst HCMs are not unbeatable, both HCM specifications provide better parameter estimates than the MNL.
- The multiplicative error formulation provides a better fit to the observed data. This empirical finding is particular to this dataset, and a priori it is not possible to know for any given dataset whether the multiplicative error formulation will provide a better fit.
- A better fit of the modelling assumptions to the observed data does not guarantee better parameter estimates.

These results advocate for the use of advancing modelling techniques such as the HCM framework. However, there is evidence that model misspecification might still cause the parameters of these models to be incorrectly estimated; hence, it is important that practitioners and decision makers keep a critical attitude towards parameter estimates from hybrid choice models. These are empirical findings for this particular dataset, and whether if they can be generalised to other datasets is still an open question; therefore, a more theoretical analysis should be carried out. Moreover, different measurement errors might modify the ability of the HCM to recover good parameter estimates; hence, it might be the case that some modelling assumptions (assumptions regarding the structural and measurement equations of the HCM) perform better in the majority of the situations. This is an interesting question for future research. Moreover, it is also necessary to further understand the implications of these types of models on policy analysis. For instance, after accounting for measurement errors in travel times, users appear to be more sensitive to travel time changes, and have higher values of time. Similar findings have been reported by Walker et al. (2010), and Varotto et al. (2017). However, these findings should be handle with care, as all these studies considered that a single variable suffered from measurement errors; hence, it is unclear what the changes in parameters will be if all variables were tested for measurement errors. Finally, this paper draws attention into current practices used on testing hybrid choice models, where there is lack of a rigorous evaluation of its results, and it is not unusual to find selective reporting of fit indices: hence, the development of a common framework to assess the performance of HCM should be prioritised.

6. References

Bierlaire, M., 2016. pythonbiogeme: a short introduction. report transp-or 160706, series on biogeme. transport and mobility laboratory, school of architecture, civil and environmental engineering, ecole polytechnique fédérale de lausanne, switzerland.

Bollen, K. A., Kirby, J. B., Curran, P. J., Paxton, P. M., and Chen, f., 2007. Latent variable models under misspecification two-stage least squares (2sls) and maximum likelihood (ml) estimators. sociological methods and research, 36(1), 48-86. doi: 10.1177/0049124107301947.

Browne, M., and Cudeck, R., 1992. Alternative ways of assessing model fit. Sociological methods and research, vol. 21, no. 2, pp. 230–258.

Barrett, P., 2007. Structural equation modeling: adjudging model fit. pers. individ. dif. 42, 815–824.doi:10.1016/j.paid.2006.09.018.

Hauber AB, González JM, Groothuis-Oudshoorn CG, Prior T, Marshall DA, Cunningham C, et al. 2016. Statistical methods for the analysis of discrete choice experiments: a report of the ISPOR Conjoint Analysis Good Research Practices Task Force. Value Health.

Hayduk, L.A., Cummings, G., Boadu, K., Pazderka-Robinson, H., and Boulianne, S., 2007. Testing! testing! one, two, three – testing the theory in structural equation models! pers. individ. dif. 42, 841–850.doi:10.1016/j.paid. 2006.10.001.

Kolenikov, S., 2011. Biases of parameter estimates in misspecified structural equation models. Sociological methodology, 41(1), 119–157. doi:10.1111/j.1467-9531.2011.01236.x

Motoaki, Y. and Daziano, R.A., 2015. Assessing goodness of fit of hybrid choice models. transp. res. rec. j. Transp. res. board. 2495, 131–141.

Roeckelein, J.E., 2000. The Concept of Time in Psychology: A Resource Book and Annotated Bibliography. Greenwood Press, Westport.

Ropovik, I., 2015. A cautionary note on testing latent variable models. Front. psychol.6:1715. doi:10.3389/fpsyg.2015.01715.

Varotto, S. F., Glerum, A., Stathopoulos, A., Bierlaire, M., and Longo, G., 2017. Mitigating the impact of errors in travel time reporting on mode choice modelling, in journal of transport geography, volume 62, 2017, pages 236-246, issn 0966-6923, https://doi.org/10.1016/j.jtrangeo.2017.05.016.

Vij, A. and Walker, J., 2016. How, when and why integrated choice and latent variable models are latently useful. TR part B. 90, 192–217.

Walker, J., Li, J., Srinivasan, S., and Bolduc, D., 2010. Travel demand models in the developing world: correcting for measurement errors. transportation letters 2, 231–243.