Forecasting with Strategic Transport Models Corrected for Endogeneity

Thomas E. Guerrero[1], C. Angelo Guevara[2], Elisabetta Cherchi[3] and Juan de Dios Ortúzar[4]

[1] Department of Civil Engineering, Universidad Francisco de Paula Santander Ocaña, Sede el Algodonal Ocaña, Colombia.
[2] Department of Civil Engineering, Universidad de Chile, Blanco Encalada 2002, Santiago, Chile.
[3] School of Engineering, Newcastle University, Newcastle upon Tyne, NE1 7RU, United Kingdom.
[4] Department of Transport Engineering and Logistics, Pontificia Universidad Católica de Chile, Avda. Vicuña Mackenna 4860, Macul, Santiago, Chile.

SHORT SUMMARY

We propose a variation of the classical *Control Function* (*CF*) method, called *Control Function Updated* (*CFU*), which considers updating the endogeneity correction using information from the future equilibria. The proposed method is assessed using Monte Carlo simulation for a strategic transport model affected by three endogeneity sources, examining the equilibrium results for various future scenarios. We compare the *CFU* method by doing nothing and with the classical *CF* approach. The forecasts are evaluated in terms of recovering the true (simulated) travel times and two indices of fit. Results show that the endogenous (do nothing) model produces large biases in simulated travel times and poor goodness-of-fit measures that steeply worsen with time in future scenarios. The corrected models perform much better and, in particular, the new *CFU* approach shows statistically significant improvements over the classical approach in all scenarios tested.

Keywords: Endogeneity, discrete choice models, forecasting, control function, strategic transport models.

1. INTRODUCTION AND MOTIVATION

The correction of endogeneity in strategic mode choice models using the *CF* approach has been addressed recently by Guerrero et al. (2020). However, in this paper, we address the important question of how to make forecasts with such models in the case of strategic supply-demand equilibration settings. Guevara and Ben-Akiva (2012) proposed an approach to forecasting with models that have been corrected for endogeneity using the *CF* method. However, their approach has the limitation of assuming exogenous shifts to analyse future scenarios, making it unsuitable for a supply-demand equilibration context where, for example, travel time is the result of a future equilibrium. To address this limitation, we suggest a new approach, the *Control Function Updated* (*CFU*) method. It allows correcting for endogeneity when models are used to forecast demand after changes in level-of-service variables determined at equilibrium for future scenarios (e.g. 10 to 40 years ahead from the calibration year).

To test our proposed approach, we used simulated data. We considered three typical sources of endogeneity that may occur in strategic transport models: (*i*) measurement error, (*ii*) omitted variables and (*iii*) the simultaneous estimation of key variables in a supply-demand equilibration mechanism. We also considered six transport modes, in an attempt to emulate the application of a strategic transport model in an urban case. Future scenarios were assessed considering exogenous changes in the explanatory variables.

In this setting, we compared three different approaches: (*i*) do nothing (i.e., no endogeneity correction), (*ii*) the *CF* approach of Guevara and Ben-Akiva (2012) and (*iii*) our proposed *CFU* procedure. The forecasts were evaluated in terms of the recovery of the *true* (simulated) travel time, the logarithm of the likelihood expected value - $E(l(\theta))$, and the Akaike Information Criteria's expected value - $E(AIC)$ (Akaike 1974) for the future scenarios 10 to 40 years ahead.

2. ADDRESSING ENDOGENEITY USING THE CF APPROACH AND THE CFU METHOD

The CF approach

The *CF* approach's idea is attributed to Heckman (1978), who first developed it to correct endogeneity in a simultaneous equation problem. The *CF* classical version can be applied following a two-stage procedure or simultaneously (Train 2009).
Guevara and Ben-Akiva (2012) considered how to forecast and simulate in practice using the *CF* method under exogenous shifts. Formally, the estimation of the probability P_{in}^1 given Gumbel errors corresponds to the following multinomial logit expression:

$$P_{in}^1 = \frac{e^{\widehat{ASC}_i + \hat{\beta}_t \tilde{t}_{in}^1 + \hat{\beta}_c \tilde{c}_{in}^1 + \hat{\beta}_{\hat{\delta}^t} \hat{\delta}_{in}^t + \hat{\beta}_{\hat{\delta}^c} \hat{\delta}_{in}^c}}{\sum_{j \in A} e^{\widehat{ASC}_j + \hat{\beta}_t \tilde{t}_{jn}^1 + \hat{\beta}_c \tilde{c}_{jn}^1 + \hat{\beta}_{\hat{\delta}^t} \hat{\delta}_{jn}^t + \hat{\beta}_{\hat{\delta}^c} \hat{\delta}_{jn}^c}} \quad (1)$$

where the parameters $\widehat{ASC}_i$ and $\hat{\beta}$ in (1) are estimators obtained following the two-stage *CF* approach shown previously. Here, the superscript 1 is used to highlight that the model attributes vary in the forecasting phase; however, $\hat{\delta}_{in}$ is fixed. This last comment is crucial because unlike Guevara and Ben-Akiva (2012), we hypothesize that in future scenarios (where new supply-demand equilibria are reached) the *CF* approach's first stage must be updated. Namely, the instruments and level-of-service attributes might endogenously change in future scenarios. If this is not corrected, it may cause biases as we will show later. On the other hand, if the *CF* approach's simultaneous (one-stage) procedure is preferred, forecasting would follow (2):

$$P_{in}^1 = \frac{e^{\widehat{ASC}_i + \hat{\beta}_t \tilde{t}_{in}^1 + \hat{\beta}_c \tilde{c}_{in}^1 + \hat{\beta}_{\hat{\delta}^t}(\tilde{t}_{in} - \hat{\alpha}_t - \hat{\gamma}_z^t z_{in}^0 - \hat{\gamma}_{fft}^t fft_{in}^0) + \hat{\beta}_{\hat{\delta}^c}(\tilde{c}_{in} - \hat{\alpha}_c - \hat{\gamma}_z^c z_{in}^0 - \hat{\gamma}_{fft}^c fft_{in}^0)}}{\sum_{j \in A} e^{\widehat{ASC}_j + \hat{\beta}_t \tilde{t}_{jn}^1 + \hat{\beta}_c \tilde{c}_{jn}^1 + \hat{\beta}_{\hat{\delta}^t}(\tilde{t}_{jn} - \hat{\alpha}_t - \hat{\gamma}_z^t z_{jn}^0 - \hat{\gamma}_{fft}^t fft_{jn}^0) + \hat{\beta}_{\hat{\delta}^c}(\tilde{c}_{jn} - \hat{\alpha}_c - \hat{\gamma}_z^c z_{jn}^0 - \hat{\gamma}_{fft}^c fft_{jn}^0)}} \quad (2)$$

where, superscript 0 indicates data coming from the sample used for estimation and superscript 1 indicates attributes that vary in the forecasting phase (Guevara and Ben-Akiva 2012).

The CFU method

For explanatory purposes, we continue using the superscript 0 to indicate data coming from the sample used for estimation and the superscript 1 to indicate attributes that vary in the forecasting phase. The proposed *CFU* method for forecasting consists in considering that the residuals from the first stage of the *CF* approach for the future scenarios must be obtained using the value of the instruments for the year of forecasting (z_{in}^1 and fft_{in}^1) instead of z_{in}^0 and fft_{in}^0 in (8) and (9). Therefore, these new residuals are identified as $\hat{\delta}_{in}^{1t}$ and $\hat{\delta}_{in}^{1c}$. They are obtained applying an OLS regression of $\tilde{c}_{in}^1$ and $\tilde{t}_{in}^1$ on the exogenous variables in the DCM and the instruments. In this way, the CFU approach's first stage is represented in (3) and (4):

$$\tilde{t}_{in}^1 = \alpha_t + \gamma_z^t z_{in}^1 + \gamma_{fft}^t fft_{in}^1 + \delta_{in}^{1t} \quad \xrightarrow{OLS} \quad \hat{\delta}_{in}^{1t} = \tilde{t}_{in}^1 - t_{in}^1 \qquad (3)$$

$$\tilde{c}_{in}^{1} = \alpha_c + \gamma_z^c z_{in}^{1} + \gamma_{fft}^{c} fft_{in}^{1} + \delta_{in}^{1c} \quad \overset{OLS}{\longrightarrow} \quad \hat{\delta}_{in}^{1c} = \tilde{c}_{in}^{1} - c_{in}^{1} \qquad (4)$$

where, α is the intercept of the regression model, γ_z and γ_{fft} are parameters to be estimated for the exogenous attributes z_{in}^1 and fft_{in}^1, respectively. Finally, δ_{in}^1 is the error term of the regression model.

For the second stage of the CFU approach, $\hat{\delta}_{in}^{1c}$ and $\hat{\delta}_{in}^{1t}$ are added as explanatory variables within the utility function. However, the choices are unknown; therefore, we use P_{in}^1 coming from (1) or (2) as a proxy of the choice. Furthermore, it is known that forecasting in transport planning and social evaluation projects requires the parameters estimated for the calibration year (i.e. $\widehat{ASC}_i$, $\hat{\beta}_t$ and $\hat{\beta}_c$). The attribute vectors ($\tilde{t}_{in}^1$, $\tilde{c}_{in}^1$, $\hat{\delta}_{in}^{1c}$ and $\hat{\delta}_{in}^{1t}$) are known for the modeller, therefore the only unknown elements are $\hat{\beta}_{\hat{\delta}^{1c}}$ and $\hat{\beta}_{\hat{\delta}^{1t}}$. These can be re-estimated using the linear regression in (5), which can be seen as an application of the Berkson-Theil transformation procedure (Ortúzar and Willumsen 2011).

$$Ln\frac{P_{in}^1}{P_{jn}^1} = \left(\widehat{ASC}_i - \widehat{ASC}_j\right) + \hat{\beta}_t\left(\tilde{t}_{in}^1 - \tilde{t}_{jn}^1\right) + \hat{\beta}_c\left(\tilde{c}_{in}^1 - \tilde{c}_{jn}^1\right) + \beta_{\hat{\delta}^{1c}}\left(\hat{\delta}_{in}^{1c} - \hat{\delta}_{jn}^{1c}\right) + \beta_{\hat{\delta}^{1t}}\left(\hat{\delta}_{in}^{1t} - \hat{\delta}_{jn}^{1t}\right) + \left(\tilde{\tilde{\varepsilon}}_{in} - \tilde{\tilde{\varepsilon}}_{jn}\right) \quad (5)$$

where the left-hand side of (5) is estimated based on Guevara and Ben-Akiva (2012) approach and it acts as the dependent variable; whereas $\left(\hat{\delta}_{in}^{1c} - \hat{\delta}_{jn}^{1c}\right)$ and $\left(\hat{\delta}_{in}^{1t} - \hat{\delta}_{jn}^{1t}\right)$ act as the independent variables. Then, the sum of $\hat{\beta}_t\left(\tilde{t}_{in}^1 - \tilde{t}_{jn}^1\right)$, $\hat{\beta}_c\left(\tilde{c}_{in}^1 - \tilde{c}_{jn}^1\right)$ and $\left(\widehat{ASC}_i - \widehat{ASC}_j\right)$ is the intercept. The calculation of the probability $P_{in}^{1\_CFU}$ after updating the residuals is given by (6) for the two-stage *CF* approach. Note that the parameters used are $\widehat{ASC}_i$, $\hat{\beta}_t$ and $\hat{\beta}_c$ and come from the calibration year model, however $\hat{\beta}_{\hat{\delta}^{1c}}$ and $\hat{\beta}_{\hat{\delta}^{1t}}$ are used instead $\beta_{\hat{\delta}^c}$ and $\beta_{\hat{\delta}^t}$. On the other hand, if the one-stage estimation is used, then it must follow (7):

$$P_{in}^{1\_CFU} = \frac{e^{\widehat{ASC}_i + \hat{\beta}_t \tilde{t}_{in}^1 + \hat{\beta}_c \tilde{c}_{in}^1 + \hat{\beta}_{\hat{\delta}^{1t}} \hat{\delta}_{in}^{1t} + \hat{\beta}_{\hat{\delta}^{1c}} \hat{\delta}_{in}^{1c}}}{\sum_{j \in A} e^{\widehat{ASC}_j + \hat{\beta}_t t_{jn}^1 + \hat{\beta}_c c_{jn}^1 + \hat{\beta}_{\hat{\delta}^{1t}} \hat{\delta}_{jn}^{1t} + \hat{\beta}_{\hat{\delta}^{c1}} \hat{\delta}_{jn}^{1c}}} \quad (6)$$

$$P_{in}^{1\_CFU} = \frac{e^{\widehat{ASC}_i + \hat{\beta}_t \tilde{t}_{in}^1 + \hat{\beta}_c \tilde{c}_{in}^1 + \hat{\beta}_{\hat{\delta}^{1t}}(\tilde{t}_{in}^1 - \hat{\alpha}_t - \hat{\gamma}_z^t z_{in}^1 - \hat{\gamma}_{fft}^t fft_{in}^1) + \hat{\beta}_{\hat{\delta}^{1c}}(\tilde{c}_{in}^1 - \hat{\alpha}_c - \hat{\gamma}_z^c z_{in}^1 - \hat{\gamma}_{fft}^c fft_{in}^1)}}{\sum_{j \in A} e^{\widehat{ASC}_j + \hat{\beta}_t \tilde{t}_{jn}^1 + \hat{\beta}_c \tilde{c}_{jn}^1 + \hat{\beta}_{\hat{\delta}^{1t}}(\tilde{t}_{jn}^1 - \hat{\alpha}_t - \hat{\gamma}_z^t z_{jn}^0 - \hat{\gamma}_{fft}^t fft_{jn}^0) + \hat{\beta}_{\hat{\delta}^{1c}}(\tilde{c}_{jn}^1 - \hat{\alpha}_c - \hat{\gamma}_z^c z_{jn}^0 - \hat{\gamma}_{fft}^c fft_{jn}^0)}} \quad (7)$$

3. RESULTS

Base year assessment

As the correction for endogeneity in a DCM produces a change of scale in the parameter estimates, we checked the ratios among parameters. Let β_t^{PT} be the parameter of travel time for the public transport modes (Bus, Train and Share Taxi), β_t^{CM} for the car modes (Car driver and Share car), β_t^W for Walking and β_c the parameter of cost (which was considered generic). Using the parameters shown in Table 1, we can see that the *true* population ratios are $\frac{\beta_t^{PT}}{\beta_c} = \frac{-0.03}{-0.01} = 3$, $\frac{\beta_t^W}{\beta_c} = \frac{-0.05}{-0.01} = 5$ and $\frac{\beta_t^{CM}}{\beta_c} = \frac{-0.04}{-0.01} = 4$. Our results for the benchmark (*true* model), endogenous and corrected models are reported in Table 1, which shows the ratios' mean $\frac{\beta_t^{PT}}{\beta_c}$, $\frac{\beta_t^W}{\beta_c}$ and $\frac{\beta_t^{CM}}{\beta_c}$, the t-test against the *true* ratios and the bias (in percentage) for each case. For the base year, only the classical *CF* approach is used to correct for endogeneity and to recover the parameters. Our approach (*CFU*) is applied later for forecasting.

Table 1 Statistics for benchmark, corrected and endogenous model

Model	$\frac{\beta_t^{PT}}{\beta_c}$ (t-test)	$\frac{\beta_t^{W}}{\beta_c}$ (t-test)	$\frac{\beta_t^{CM}}{\beta_c}$ (t-test)	% Bias $\frac{\beta_t^{PT}}{\beta_c}$	% Bias $\frac{\beta_t^{W}}{\beta_c}$	% Bias $\frac{\beta_t^{CM}}{\beta_c}$
True	3.00	5.00	4.00	-	-	-
Endogenous	1.47 (26.89)	10.19 (13.78)	3.67 (3.36)	51.1%	103.8%	8.3%
Corrected	2.97 (0.49)	4.92 (0.58)	3.94 (1.03)	0.9%	1.6%	1.6%

The biases for the endogenous ratios are large, varying from 8.3% (i.e. $\left[\frac{4.00-3.67}{4.00}\right] * 100$) to 103.8% ($\left[\frac{10.19-5.00}{5.00}\right] * 100$). In this case too, the t-test[1] against the null hypotheses that $\frac{\beta_t^{PT}}{\beta_c} = 3$, $\frac{\beta_t^{W}}{\beta_c} = 5$ and $\frac{\beta_t^{CM}}{\beta_c} = 4$ can be easily rejected and it can be concluded that the parameter ratios in the endogenous model are significantly different from the *true* values for a one-sided test at the 95% confidence level[2]. On the other hand, the t-test for the parameter ratios in the corrected model are all accepted for a one-sided test at the 95% confidence level, meaning that the corrected ratios are not significantly different from the *true* ratios. The biases for the corrected ratios are small, varying from 0.9% to 1.6%.

Future scenarios assessment

Exogenous changes impact individual choices. This happens because the exogenous changes affect the explanatory variables of the model (travel time and cost), which in turn modify the supply-demand equilibration in future scenarios. With our Monte Carlo simulation, we estimate several measures (travel times at equilibrium, the logarithm of the likelihood expected value, and the expected value of the Akaike Information Criteria) and compare them with those obtained for the *true* model.

We assessed and compared the new CFU method with (1) *No endogeneity correction* and (2) the *Guevara and Ben-Akiva (2012) CF* approach for the eight scenarios shown in Table 2. In each case, the model forecasts were evaluated in terms of their ability to recover the *true* (simulated) travel times at equilibrium (TTE), the logarithm of the likelihood expected value - $E(l(\theta))$ for the *true* model, and the expected value of the Akaike Information Criteria[3] – $E(AIC)$ for future scenarios in 10, 20, 30 and 40 years ahead. The *true* model is the benchmark and $E(l(\theta))$ and $E(AIC)$ were calculated as shown in (8) and (9), respectively:

$$E\left(l(\theta)_I^{m,t}\right) = \sum_n \sum_{i \epsilon A(n)} ln\left(P_{in}^{m,t} * P_{in}^{True,t}\right) \quad (8)$$

$$E(AIC_I^{m,t}) = 2k - 2E(l(\theta)_I^{m,t}) \quad (9)$$

[1] The t-test statistic has the form $t = \frac{\bar{X}-\mu}{sd/\sqrt{n}}$ where $\bar{X}$ is the sample mean from a sample $X_1, X_2, \ldots, X_n$, of size *n*, *sd* is the estimate of the standard deviation of the population, and μ is the population mean.

[2] When the sign of the parameter is known a one-sided test should be applied; the critical value of *t* is 1.64 for a one-sided test at the 95% confidence level.

[3] AIC is a measure of the relative quality of a statistical model. It provides a trade-off between the goodness of fit of the model and its complexity (Akaike, 1974).

where $P_{in}^{True,t}$ represents the choice probability of alternative *i* for individual *n* in the *true* model for year *t*; $P_{in}^{m,t}$ is the choice probability of alternative *i* for individual *n* in approach *m* (*No endogeneity correction*, *Guevara and Ben-Akiva (2012)* and *CFU*) for year *t*, and the subscript (*I*) indicates the number of repetitions. In our case, in (9) we used $E(l(\theta)_I^{m,t})$ to estimate $E(AIC_I^{m,t})$ for approach *m* and year *t*, where *k* corresponds to the number of model parameters.
Table 2 summarizes the average for all repetitions of the free flow time in the base year and with the exogenous change (see Table 2), as well as the travel times achieved for the future equilibria in the simulations for each scenario.

Table 2 Average of the free flow time and travel times in equilibrium (TTE) for 100 replications

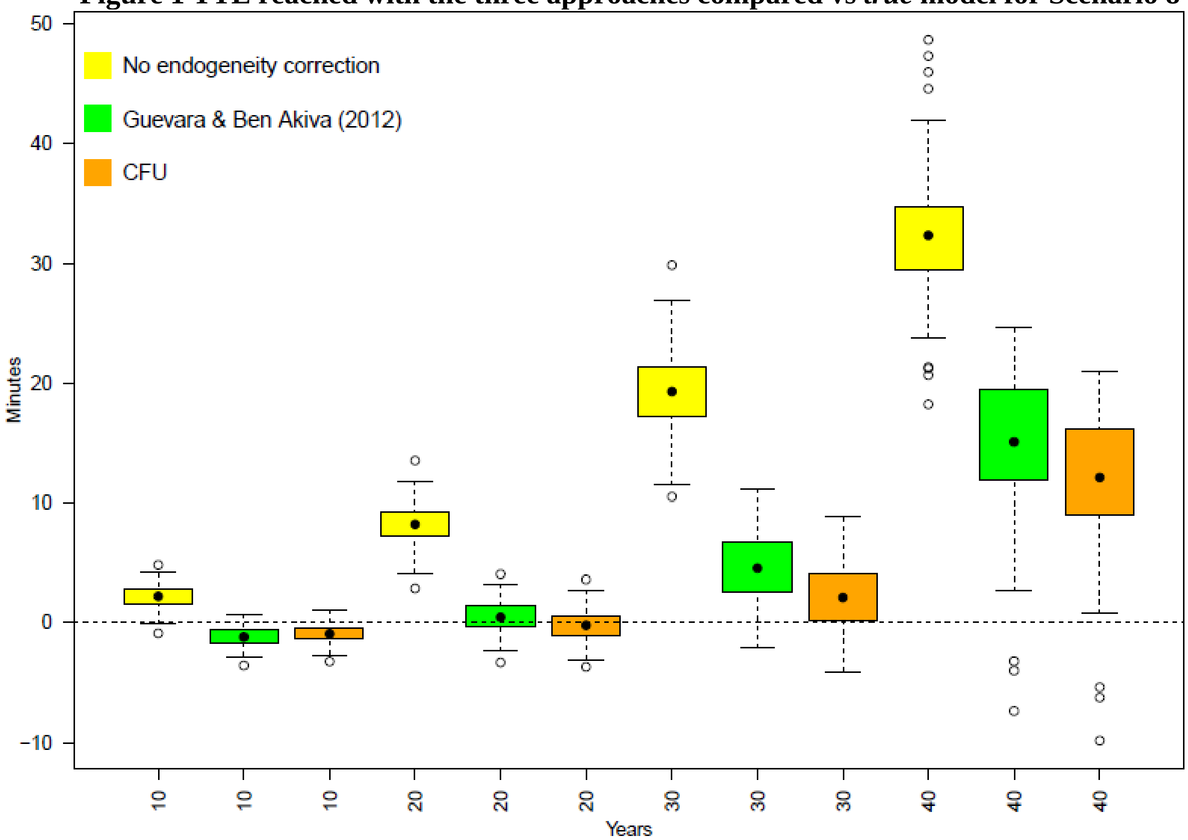
Year	Approach	Scenario 1			Scenario 2			Scenario 3			Scenario 4		
		Car modes	Public modes	Walking mode	Car modes	Public modes	Walking mode	Car modes	Public modes	Walking mode	Car modes	Public modes	Walking mode
Calibration	Free flow time	45.03	49.53	35	45.03	49.53	35	45.03	49.53	35	45.03	49.53	35
	TTE	64.04	64.1	35	64.04	64.1	35	64.04	64.1	35	64.04	64.1	35
10	TTE True	72.28	62.91	35	73.2	66.62	35	68.74	67.7	35	73.29	71.19	35
	TTE No endogeneity correction	75.61	65.47	35	75.89	68.68	35	69.27	68.11	35	75.52	72.9	35
	TTE Guevara and Ben-Akiva (2012)	71.79	62.54	35	72.66	66.21	35	67.79	66.97	35	72.68	70.72	35
	TTE *CFU*	71.51	62.33	35	72.52	66.1	35	68.06	67.18	35	72.76	70.78	35
20	TTE True	87.78	74.8	35	89	78.73	35	83.91	79.33	35	89.19	83.38	35
	TTE No endogeneity correction	97.5	82.25	35	97.95	85.59	35	89.68	83.75	35	97.51	89.75	35
	TTE Guevara and Ben-Akiva (2012)	89.3	75.96	35	90.54	79.91	35	84.89	80.08	35	90.62	84.47	35
	TTE *CFU*	87.98	74.95	35	89.35	79	35	84.26	79.6	35	89.69	83.76	35
30	TTE True	115.63	96.15	35	117.15	100.32	35	109.68	99.09	35	117.5	105.08	35
	TTE No endogeneity correction	135.35	111.27	35	136.04	114.8	35	124.79	110.67	35	135.7	119.04	35
	TTE Guevara and Ben-Akiva (2012)	122.39	101.33	35	124.02	105.58	35	115.22	103.34	35	124.31	110.3	35
	TTE *CFU*	119.97	99.48	35	121.64	103.75	35	113.12	101.72	35	122.04	108.56	35
40	TTE True	170.61	138.3	35	172.41	142.68	35	159.6	137.36	35	173.05	147.67	35
	TTE No endogeneity correction	200.12	160.92	35	201.11	164.68	35	184.73	156.63	35	201.11	169.18	35
	TTE Guevara and Ben-Akiva (2012)	185.76	149.92	35	187.74	154.44	35	173.91	148.33	35	188.5	159.51	35
	TTE *CFU*	183.38	148.09	35	185.35	152.6	35	171.48	146.47	35	186.11	157.68	35
Free flow time with exogenous change		45.03	42.03	35	45.03	45.03	35	40.53	46.07	35	45.03	49.53	35

Year	Approach	Scenario 5			Scenario 6			Scenario 7			Scenario 8		
		Car modes	Public modes	Walking mode	Car modes	Public modes	Walking mode	Car modes	Public modes	Walking mode	Car modes	Public modes	Walking mode
Calibration	Free flow time	45.03	49.53	35	45.03	49.53	35	45.03	49.53	35	45.03	49.53	35
	TTE	64.04	64.1	35	64.04	64.1	35	64.04	64.1	35	64.04	64.1	35
10	TTE True	60.66	54.01	35	69.97	61.14	35	66.20	61.26	35	58.44	52.30	35
	TTE No endogeneity correction	61.57	54.7	35	74.10	64.31	35	68.14	62.74	35	60.21	53.66	35
	TTE Guevara and Ben-Akiva (2012)	59.59	53.19	35	69.51	60.79	35	65.40	60.64	35	57.48	51.57	35
	TTE *CFU*	59.78	53.33	35	69.29	60.63	35	65.50	60.72	35	57.69	51.73	35
20	TTE True	74.03	64.26	35	84.78	72.50	35	80.55	72.26	35	71.11	62.01	35
	TTE No endogeneity correction	79.69	68.59	35	95.35	80.60	35	87.98	77.95	35	77.68	67.05	35
	TTE Guevara and Ben-Akiva (2012)	74.52	64.63	35	86.08	73.49	35	81.46	72.96	35	71.48	62.30	35
	TTE *CFU*	73.88	64.14	35	84.86	72.56	35	80.66	72.34	35	70.93	61.88	35
30	TTE True	96.34	81.36	35	112.05	93.40	35	105.55	91.42	35	92.76	78.62	35
	TTE No endogeneity correction	110.83	92.47	35	132.68	109.22	35	122.46	104.39	35	108.2	90.45	35
	TTE Guevara and Ben-Akiva (2012)	100.37	84.45	35	118.42	98.29	35	110.77	95.43	35	96.39	81.40	35
	TTE *CFU*	98.34	82.9	35	116.07	96.49	35	108.58	93.75	35	94.45	79.91	35
40	TTE True	139.51	114.45	35	166.77	135.36	35	154.85	129.22	35	135.52	111.4	35
	TTE No endogeneity correction	164.36	133.51	35	197.21	158.69	35	181.82	149.90	35	161.39	131.23	35
	TTE Guevara and Ben-Akiva (2012)	151.9	123.95	35	181.64	146.76	35	168.82	139.93	35	147.61	120.67	35
	TTE *CFU*	149.44	122.07	35	179.31	144.97	35	166.42	138.09	35	145.22	118.83	35
Free flow time with exogenous change		36.03	35.12	35	45.03	42.03	35	40.53	41.57	35	40.53	36.03	35

Results are shown for the base year and the 10 to 40 forecasting years for the three approaches compared (*No endogeneity correction*, *Guevara and Ben-Akiva (2012)* and *CFU*) as well as for the benchmark (*true* model). As can be seen, for the base year, the TTE are the same in all scenarios because the estimates replicate the values. Note that the free flow time and the TTE do not change for the Walking mode, because walking does not share infrastructure with other modes; therefore, it is not affected by congestion. The TTE (in the base year) are higher than the free flow time because of congestion[4].

On the other hand, both the *CFU* and *Guevara and Ben-Akiva (2012)* results are better than those of the endogenous model, but they are still different from the *true* model. This can be attributed to a simulation error. Notwithstanding, the TTE reached with the *CFU* approach are closer to the values of the *true model* than the TTE reached with the *Guevara and Ben-Akiva (2012)* approach. We applied the t^*-statistics (Ortúzar and Willumsen 2011, 341-342) to test the null hypothesis that the mean difference of TTE for the 100 repetitions between *CFU* and *Guevara and Ben-Akiva (2012)* was zero. The t^*-statistics applied for 20, 30 and 40 years show that these values are superior to the critical value (i.e., 1.96 for a two-sided test at the 95% confidence level); therefore, the null hypothesis is rejected, and there is a significant difference between the means. Consequently, we can conclude that the *CFU* approach is better than the approach proposed by *Guevara and Ben-Akiva (2012)*. This happened in all the scenarios analysed. Figure 1 focuses only on the car modes and shows the boxplot for the TTE reached by each of the three approaches.

Figure 1 TTE reached with the three approaches compared vs *true* model for Scenario 8



---

[4] We simulated a case with endogeneity only in travel time due to the equilibrium conditions (i.e., no endogeneity due to measurement error and omitted variables), so the cost was not endogenous. We found that the effect was similar but smaller than in the case with three endogeneity sources, an expected result; as there is less bias, the TTE will tend to be closer to those in the *true* model. So, there appears to be an additive effect regarding endogeneity sources; that is, if the endogeneity sources increase, the bias also does.

Again, the *true* model's travel times are the benchmark. These are represented by the dashed line drawn at zero. The points that belong to the boxplots are the difference between the TTE reached from any other approaches and the *true* model. We show only the case corresponding to Scenario 8 for explanatory purposes because it is one of two scenarios affected by the most significant number of exogenous shifts. Besides, for the other scenarios, the performance was very similar. As can be seen, "*No endogeneity correction*" is the worst in recovering the TTE. Figure 1 also shows that the endogenous model overestimates congestion. This makes sense given that, as shown in Table 1, the parameter ratios show large bias for the endogenous estimations.

To assess the relative quality of each approach, we used again $E(l(\theta))$ and $E(AIC)$. A summary of their estimates for each approach and year are given in Table 3 for the scenarios shown in Table 2, and all the repetitions run in the simulation. Given that "*No endogeneity correction*" uses a model (endogenous) that is a restricted version of the model used in both *Guevara and Ben-Akiva (2012)* and in the *CFU* approach, it is possible to apply the likelihood ratio (LR) test (Ortúzar and Willumsen 2011, 281) to compare them. LR is asymptotically distributed χ_r^2 with *r* degrees of freedom, where *r* is the number of linear restrictions required to transform the more general model into the restricted version. The null hypothesis is that the two models compared are equivalent; rejecting it, implies that the restricted model is erroneous. In our case, *r* = 2 (because the restrictions are that both $\hat{\beta}_{\hat{\delta}\tilde{t}}$ and $\hat{\beta}_{\hat{\delta}\tilde{c}}$ are zero). The LR test for the results of scenario 8 and year 10[5] comparing "*No endogeneity correction*" against *Guevara and Ben-Akiva (2012)* is $LR = -2(-5459.9 + 5427.4) = 65$, and comparing "*No endogeneity correction*" against *CFU* is $LR = -2(-5459.9 + -5398.5) = 122.8$. These values must be compared with the critical value for two degrees of freedom at the 95% confidence level ($\chi_2^2 = 5.99$). As LR > χ_2^2 (for both cases), the null hypothesis is confidently rejected, and we can conclude that the corrected version models are superior. Given that the *CFU* approach has $E(l(\theta))$ better than the *Guevara and Ben-Akiva (2012)* approach, we can conclude that the *CFU* approach performs best. Given that $E(AIC)$ values are calculated using the $E(l(\theta))$ and *k* (see expression 41), then it is expected that the $E(AIC)$ from *CFU* approach will also be better than those of the other methods (see Table 3).

[5] For the other scenarios and years, the performance is similar. We highlighted in **bold** the values from Table 5 used to apply the LR test.

Table 3 Average estimates of $E(l(\theta))$ and $E(AIC)$ for 100 replications

Year	Approach	Scenario 1		Scenario 2		Scenario 3		Scenario 4		Scenario 5		Scenario 6		Scenario 7		Scenario 8	
		$E(l(\theta))$	$E(AIC)$	$E(l(\theta))$	$E(AIC)$	$E(l(\theta))$	$E(AIC)$	$E(l(\theta))$	$E(AIC)$	$E(l(\theta))$	$E(AIC)$	$E(l(\theta))$	$E(AIC)$	$E(l(\theta))$	$E(AIC)$	$E(l(\theta))$	$E(AIC)$
10	No endogeneity correction	-5303.4	10624.8	-5296.3	10610.6	-5378.5	10775.0	-5259.3	10536.7	-5508.2	11034.4	-5265.8	10549.6	-5375.7	10769.5	**-5459.9**	10937.9
	Guevara and Ben-Akiva (2012)	-5194.3	10410.5	-5212.2	10446.4	-5336.7	10695.4	-5183.2	10388.4	-5482.2	10986.3	-5157.6	10337.2	-5329.3	10680.6	**-5427.4**	10876.9
	CFU	-5136.0	10293.9	-5157.9	10337.9	-5298.2	10618.4	-5135.1	10292.2	-5449.2	10920.4	-5102.9	10227.8	-5288.3	10598.7	**-5398.5**	10818.9
20	No endogeneity correction	-5052.8	10123.6	-5046.4	10110.8	-5149.0	10316.0	-5024.1	10066.1	-5292.5	10603.0	-5037.7	10093.4	-5154.8	10327.6	-5266.0	10549.9
	Guevara and Ben-Akiva (2012)	-4871.9	9765.7	-4892.2	9806.4	-5056.0	10134.0	-4890.0	9802.0	-5215.5	10453.0	-4869.2	9760.5	-5055.5	10133.0	-5195.3	10412.6
	CFU	-4785.4	9592.9	-4807.2	9636.5	-4978.9	9979.7	-4807.7	9637.4	-5141.1	10304.3	-4783.7	9589.5	-4976.5	9975.0	-5123.9	10269.9
30	No endogeneity correction	-4694.9	9407.8	-4687.8	9393.6	-4814.0	9646.0	-4679.9	9377.8	-4980.3	9978.6	-4706.1	9430.2	-4832.4	9682.9	-4980.8	9979.6
	Guevara and Ben-Akiva (2012)	-4402.0	8826.0	-4417.0	8856.0	-4606.5	9235.1	-4430.8	8883.5	-4797.7	9617.3	-4434.3	8890.5	-4625.2	9272.4	-4816.8	9655.6
	CFU	-4319.1	8660.3	-4333.0	8687.9	-4517.3	9056.5	-4343.8	8709.6	-4706.6	9435.3	-4348.0	8718.1	-4534.9	9091.9	-4724.2	9470.4
40	No endogeneity correction	-4191.0	8400.0	-4180.7	8379.3	-4327.8	8673.5	-4180.4	8378.7	-4529.2	9076.5	-4222.9	8463.7	-4361.6	8741.2	-4556.4	9130.7
	Guevara and Ben-Akiva (2012)	-3793.1	7608.2	-3796.3	7614.6	-3982.1	7986.1	-3808.1	7638.2	-4213.3	8448.5	-3847.5	7717.0	-4027.3	8076.7	-4264.8	8551.6
	CFU	-3741.0	7504.0	-3742.6	7507.1	-3918.1	7858.3	-3749.9	7521.8	-4143.7	8309.3	-3791.0	7603.9	-3963.0	7947.9	-4190.5	8403.1

4. CONCLUSIONS

We emulated a complex transport modelling process affected by three different endogeneity sources that are common in strategic studies (measurement error, omitted variables and simultaneous estimation in a supply-demand equilibration context). In this setting, we compared three different approaches: (1) No endogeneity correction (which has been, so far, the only method used in practice), (2) the *CF* approach proposed by Guevara and Ben-Akiva (2012) and (3) our new proposal, the *CFU* approach. Forecasts were evaluated in terms of recovery of the *true* (simulated) travel times, and two goodness-of-fit indices, $E(l(\theta))$ and the $E(AIC)$, for future scenarios in 10 to 40 years ahead. This was a challenge because this type of modelling is usually done using commercial software packages, where the level-of-service variables are estimated from complex supply-demand equilibria processes considering multiple user classes, several transport modes, complex networks and many other aspects.

Monte Carlo simulation helped us to demonstrate that under fairly reasonable conditions, consistent with observed data, the new *CFU* proposal performs better than both other approaches; in particular, it performs much better than the "*doing nothing*" approach and marginally (but significantly) better than the more classical *CF* approach of Guevara and Ben-Akiva (2012). We also show that the adverse effects of endogeneity increase over the years, severely impacting future scenarios' forecasts.

Our methodological findings suggest two important recommendations for practice. The first is to avoid forecasting with endogenous models as the problem is severe in the case studied. The second is that even when correcting for endogeneity, in forecasting the residuals from the first stage of the *CF* approach for the future scenarios should always be updated. Thus, our new CFU approach is especially recommended to correct for endogeneity when discrete choice models are used to forecast strategic scenarios involving supply-demand equilibration.

REFERENCES

Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**, 716-723.

Guerrero, T.E., Guevara, C.A., Cherchi, E. and Ortúzar, J. de D. 2020. Addressing endogeneity in strategic urban mode choice models. *Transportation*, https://doi.org/10.1007/s11116-020-10122-y.

Guevara, C.A. and Ben-Akiva, M.E. 2012. Change of scale and forecasting with the control-function method in logit models. *Transportation Science*, **46**, 425-437.

Heckman, J.J. 1978. Dummy endogenous variables in a simultaneous equation system. *Econometrica*, **46**, 931–959.

Kono, T., Kishi, A., Seita, E. and Yokoi, T. 2018. Limitations of using generalized transport costs to estimate changes in trip demand: a bias caused by the endogenous value of time. *Transportmetrica A: Transport Science*, **14**, 192-209.

Ortúzar, J. de D. and Willumsen, L.G. 2011. *Modelling Transport* (4th ed). John Wiley & Sons, Chichester.

Train, K.E. 2009. *Discrete Choice Methods with Simulation* (2nd ed). Cambridge University Press, Cambridge.