Transportmetrica A: Transport Science

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ttra21

Forecasting with strategic transport models corrected for endogeneity

Thomas E. Guerrero, C. Angelo Guevara, Elisabetta Cherchi & Juan de Dios Ortúzar

TRANSPORTMETRICA A: TRANSPORT SCIENCE
https://doi.org/10.1080/23249935.2021.1891154

Taylor & Francis
Taylor & Francis Group

Check for updates

Forecasting with strategic transport models corrected for endogeneity

Thomas E. Guerrero [a,b,c], C. Angelo Guevara [d,f], Elisabetta Cherchi [c] and Juan de Dios Ortúzar [b,e,f]

[a]Department of Civil Engineering, Universidad Francisco de Paula Santander Ocaña, Sede el Algodonal Ocaña, Colombia; [b]Department of Transport Engineering and Logistics, Pontificia Universidad Católica de Chile, Santiago, Chile; [c]School of Engineering, Future Mobility Group, Newcastle University, Newcastle upon Tyne, UK; [d]Department of Civil Engineering, Universidad de Chile, Santiago, Chile; [e]BRT+Centre of Excellence, Santiago, Chile; [f]Instituto Sistemas Complejos de Ingeniería (ISCI), Santiago, Chile

ABSTRACT

The correction of endogeneity is a problem in strategic transport modelling; the question remains on how to make appropriate forecasts in this case. 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.

ARTICLE HISTORY
Received 3 August 2020
Accepted 11 February 2021

KEYWORDS
Endogeneity; discrete choice models; forecasting; control function; strategic transport models

Introduction and motivation

The classic strategic transport model has been often used to represent the dynamic and complex urban mobility phenomenon, where travel demand and supply interact. As a result, consistent flows and levels of service are obtained. This conceptual model has been implemented in various modelling suites used worldwide, such as ESTRAUS (De Cea et al. 2005), EMME/2 (INRO 1996) or CUBE (Citilabs 2016), to name a few. Decision-makers use outputs from these modelling packages in urban planning and to optimise supply given certain transport demands. In particular, the mode choice stage (third of the classic transport model), is typically modelled using discrete choice models (DCM) and

CONTACT Thomas E. Guerrero teguerrerob@ufpso.edu.co Department of Civil Engineering, Universidad Francisco de Paula Santander Ocaña, Sede el Algodonal Ocaña, Colombia; Department of Transport Engineering and Logistics, Pontificia Universidad Católica de Chile, Avda. Vicuña Mackenna 4860, Macul, Santiago, Chile; School of Engineering, Future Mobility Group, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK

plays a fundamental role in short and long-term transport planning and policy formulation (Ortúzar and Willumsen 2011).

Endogeneity arises in DCM when the systematic part of the utility function is correlated with the error term. As a result, the estimated parameters of the model may be inconsistent. Typically, this occurs due to measurement or specification errors, omitted variables, simultaneous determination and/or self-selection (Guevara 2015). If a model is affected by endogeneity, any conclusion, analysis or policy derived from the model may be wrong (Guevara and Ben-Akiva 2006). The planning and social evaluation of transport projects require models to deliver forecasts and examine their sensitivity to changes in key variables as required by the analyst (Ortúzar and Willumsen 2011). Quality forecasts are crucial to identify appropriate investments and the feasibility of transport projects in cost-benefit analyses. Changes in explanatory variables, such as travel time, can affect the estimation of benefits (Kato and Uchida 2018). Therefore, it is critical that the DCM used to forecast travel demand are corrected for endogeneity.

Ebbes, Papies, and Van Heerde (2011) found that linear model functions corrected for endogeneity did not necessarily predict better than endogenous models when analysing holdout samples. This fact occurs because if the model has an intercept (or a set of alternative specific constants in the case of DCM), it will necessarily reproduce the sample means (or shares) for the base year situation, regardless of the potentially poor recovery of model coefficients. However, forecasting with strategic transport models corrected for endogeneity implies simulating beyond the base year. As we will show in this research, this affects not only the parameter estimates but also the models' predictive capacity in terms of both modal split and simulated levels-of-service. Furthermore, this effect may be enhanced because the attributes in future scenarios imply exogenous shifts and are also the result of a simulated equilibria. Kono et al. (2018) state that the value of time (VoT) estimated endogenously can generate a significant bias in trip demand forecasts when travel times and costs change.

Endogeneity problems can be addressed from several approaches available in the recent literature. Some of these are the Control Function (*CF*) method (Heckman 1978), Latent Variables – LV (Walker and Ben-Akiva 2002), Maximum Likelihood (Park and Gupta 2009; Train 2009), BLP (Berry, Levinsohn, and Pakes 1995), Proxies (Guevara 2015), and MIS (Guevara and Polanco 2016). The *CF* method is particularly appealing for the correction of endogeneity in strategic mode choice models. The approach consists of adding to the econometric model a new variable (or *control function*), obtained from the regression residuals of the endogenous variables on the exogenous variables. This allows consistent estimation of the model parameters (Guevara and Ben-Akiva 2010). An important challenge here is the practical difficulty of finding valid *instruments* (or *instrumental variables*), which are fundamental to apply the method. In the transport field, the *CF* approach has been successfully used by Guerrero et al. (2020), Wen and Chen (2017), Lurkin et al. (2017) and Mumbower, Garrow, and Higgins (2014).

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). Our approach takes as a point of reference to the *CF* approach of Guevara and Ben-Akiva (2012). This approach is used first to correct for endogeneity in the base year. Then, the correction for future scenarios (where new supply-demand equilibria are reached) consists in updating the *CF* approach used in the first stage, given that the instruments might change in future scenarios (i.e. new equilibria are reached for demand increases and higher flows in the network) and, therefore, affect the control function, which needs to be updated.

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.

We show that the endogenous model may lead to strong biases that steeply grow in time in practical terms. On the other hand, the *CF* method (Guevara and Ben-Akiva 2012) and the new *CFU* approach produce much better results. This confirms that neglecting endogeneity in strategic transport models may severely mislead decision-making about mode choice and level-of-service in future scenarios. Furthermore, when comparing the *CF* and the *CFU* methods, results show that our new approach achieves outcomes that become increasingly better than the former as the time horizon increases.

The rest of the paper is structured as follows. In Section 2, we describe the *CF* method's fundamentals and our approach (*CFU*) to correct for endogeneity in DCM. In Section 3, we describe the Monte Carlo experiment designed to test our methodological proposal (*CFU*) to correct for endogeneity, when forecasting the demand after changes in level-of-service variables determined at equilibrium. The results are shown and analysed in Section 4. Finally, in Section 5, we conclude summarising the main findings and suggesting future lines of research.

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). To explain the method's fundamentals, consider a DCM with a utility function represented by (1):

$$U_{in} = ASC_i + \beta_t t_{in} + \beta_c c_{in} + \beta_q q_{in} + \varepsilon_{in} \tag{1}$$

where U_{in} represents the utility perceived by individual n for alternative A_i belonging to the individual's choice set $A(n)$. ASC_i is an alternative specific constant for alternative A_i. β_t and β_c are parameters associated with the explanatory variables t_{in} and c_{in} (representing, for example, travel time and cost). β_q is a parameter associated with the qualitative attribute q_{in} (e.g. safety, comfort or reliability). Finally, ε_{in} is an exogenous error term for individual n and alternative A_i.

Given the above, let us assume that for illustrative purposes of this explanation the qualitative attribute q_{in} is unknown to the modeller; the modeller's specification will be as in (2), where the new error term $\tilde{\varepsilon}_{in}$ contains both ε_{in} and q_{in}:

$$U_{in} = ASC_i + \beta_t t_{in} + \beta_c c_{in} + \tilde{\varepsilon}_{in} \tag{2}$$

For explanatory purposes, we will also consider that from (2) the explanatory variables c_{in} and t_{in} are endogenous due to three possible sources of endogeneity: measurement errors, omitted variables, simultaneous determination. Guerrero et al. (2020) showed that these variables are usually endogenous in strategic mode choice models. We assume that the endogeneity for c_{in} is due to measurement error and omitted variables, whereas the endogeneity for t_{in} is due to its simultaneous estimation process.

For the explanatory variable c_{in}, we consider that it is correlated with q_{in} as shown in (3). where, θ_c is a constant, z_{in} is an exogenous attribute, which then works as instrument, since it partially explains c_{in}. θ_z is parameter to be estimated. $\tilde{\varphi}_{in}$ is an error term containing both φ_{in} and q_{in}. The correlation between c_{in} and q_{in} in (3) yields that c_{in} and the error term $\tilde{\varepsilon}_{in}$ are also correlated in (2); therefore, by definition, c_{in} is endogenous. As can be seen, in this case, the endogeneity arises due to the omission of the attribute q_{in} in (2):

$$c_{in} = \theta_c + \theta_z z_{in} + \theta_q q_{in} + \varphi_{in} = \theta_c + \theta_z z_{in} + \tilde{\varphi}_{in} \tag{3}$$

On the other hand, to account for the measurement error, let us assume that the modeller observes $\tilde{c}_{in}$ instead of c_{in} as shown in (4), where η_{in} represents a measurement error and $\tilde{c}_{in}$ is the explanatory variable but affected by η_{in}. Therefore, c_{in} can be re-written as $c_{in} = \tilde{c}_{in} - \eta_{in}$ and the new functional form in (2) changes as follows in (5):

$$\tilde{c}_{in} = c_{in} + \eta_{in} \tag{4}$$

$$U_{in} = ASC_i + \beta_t t_{in} + \beta_c(\tilde{c}_{in} - \eta_{in}) + \tilde{\varepsilon}_{in} = ASC_i + \beta_t t_{in} + \beta_c \tilde{c}_{in} - \underbrace{\beta_c \eta_{in} + \tilde{\varepsilon}_{in}}_{\tilde{\tilde{\varepsilon}}_{in}} \tag{5}$$

where the new error term ($\tilde{\tilde{\varepsilon}}_{in}$) contains both $\beta_c \eta_{in}$ and $\tilde{\varepsilon}_{in}$. As can be seen, the endogeneity arises in (5) given that $\tilde{c}_{in}$ is correlated with $\tilde{\tilde{\varepsilon}}_{in}$ through η_{in} and q_{in}. Namely, the expected value of $\tilde{\tilde{\varepsilon}}_{in}$ conditional on $\tilde{c}_{in}$ is not zero ($E(\tilde{\tilde{\varepsilon}}_{in}|\tilde{c}_{in}) \neq 0$), therefore $\tilde{c}_{in}$ is endogenous.

To represent endogeneity in the explanatory variable t_{in}, let us consider the supply-demand equilibrium mechanism that occurs in the classical four-stage transport model. In this case, new equilibria are reached due to changes in demand, which in turn yield changes on network flows (supply). Namely, the error term $\tilde{\tilde{\varepsilon}}_{in}$ affects the probability (P_{in}) in (6), which triggers a change in $\tilde{t}_{in}$. For explanatory purpose, $\tilde{t}_{in}$is expressed through of a BPR function (Bureau of Public Roads 1964) as shown in (7), where fft_{in} represents the free-flow time, and μ and ρ are parameters to be estimated. The values reached as a result of this equilibrium will be known as $\tilde{t}_{in}$, and they are considered endogenous because the expected value of $\tilde{\tilde{\varepsilon}}_{in}$ conditional on $\tilde{t}_{in}$ is not zero ($E(\tilde{\tilde{\varepsilon}}_{in}|\tilde{t}_{in}) \neq 0$).

$$P_{in} = \frac{e^{ASC_i+\beta_t\tilde{t}_{in}+\beta_c\tilde{c}_{in}+\tilde{\tilde{\varepsilon}}_{in}}}{\sum_{j\epsilon A} e^{ASC_j+\beta_t\tilde{t}_{jn}+\beta_c\tilde{c}_{jn}+\tilde{\tilde{\varepsilon}}_{jn}}} \tag{6}$$

$$\tilde{t}_{in} = fft_{in}[1 + \mu(P_{in})^{\rho}] \tag{7}$$

As mentioned above, an essential requirement for using the *CF* method to correct for endogeneity is the availability of proper instruments (Hausman 1978). An instrument can be considered valid when fulfilling two requirements: (*i*) it is correlated with the endogenous variable, and (*ii*) it is independent of the model's error term. The former is known as *relevance condition* and the second as *exogeneity condition*. There are some tests in the literature to determine whether both conditions are fulfilled for the case of DCM. The relevance condition can be checked using the results based on the F test of the CF method's first stage regression, as originally proposed by Guevara and Navarro (2015) and recently advanced by Guerrero et al. (2021). To check the exogeneity condition, Guevara (2018) recently proposed the *Refutability* test and its variation, the *Modified Refutability* test. These can be applied when there is overidentification of the model, which holds when there are more instruments than endogenous variables. If the instruments are relevant and exogenous, the *CF* method can be used to obtain consistent estimators of the model parameters, up to a scale. Nevertheless, it is always a challenge (and even a controversial one, see Bresnahan 1997) to find proper instruments fulfilling both conditions. To apply the *CF* method, at least one instrument is needed for each endogenous variable considered.

In practice, following the two-stage approach of the *CF* method (Train 2009; Wooldridge 2010), the first stage below consists in obtaining the residuals ($\hat{\delta}_{in}$) from an ordinary least square (OLS) regression of $\tilde{t}_{in}$ and $\tilde{c}_{in}$ on the exogenous variables in the DCM and the instruments. One control function must be estimated for each endogenous variable. Therefore, the first stage of the *CF* approach for $\tilde{t}_{in}$ and $\tilde{c}_{in}$ is as follows:

$$\tilde{t}_{in} = \alpha_t + \gamma_z^t z_{in} + \gamma_{fft}^t fft_{in} + \delta_{in}^t \quad \overset{OLS}{\rightarrow} \quad \hat{\delta}_{in}^t = \tilde{t}_{in} - \hat{\tilde{t}}_{in} \tag{8}$$

$$\tilde{c}_{in} = \alpha_c + \gamma_z^c z_{in} + \gamma_{fft}^c fft_{in} + \delta_{in}^c \quad \overset{OLS}{\rightarrow} \quad \hat{\delta}_{in}^c = \tilde{c}_{in} - \hat{\tilde{c}}_{in} \tag{9}$$

where, α_c and α_t are the intercepts of the regression model, γ are parameters to be estimated, and δ are the regression model's error terms. On the other hand, the instruments (i.e. exogeneous variables) considered are z_{in} and fft_{in} given that they are correlated with the endogenous variable, but they do not influence the individuals' choice.

In the second stage, the DCM is estimated considering $\hat{\delta}_{in}^t$ and $\hat{\delta}_{in}^c$ (coming from the first stage) as explanatory variables within the utility function, as shown in (10):

$$U_{in} = ASC_i + \beta_t\tilde{t}_{in} + \beta_c\tilde{c}_{in} + \beta_{\hat{\delta}^t}\hat{\delta}_{in}^t + \beta_{\hat{\delta}^c}\hat{\delta}_{in}^c + \tilde{\tilde{\tilde{\varepsilon}}}_{in} \tag{10}$$

The intuition behind this method is that the residuals $\hat{\delta}_{in}^t$ and $\hat{\delta}_{in}^c$ capture all of $\tilde{t}_{in}$ and $\tilde{c}_{in}$ that is correlated with the error term $\tilde{\tilde{\varepsilon}}_{in}$, effectively controlling for this problem in (10).

However, this approach implies a different error term of the model ($\tilde{\tilde{\varepsilon}}_{in}$ instead of $\tilde{\varepsilon}_{in}$), resulting in a change of scale in the estimated parameters of the DCM (Yatchew and Griliches 1985; Cramer 2007; Guevara and Ben-Akiva 2012). Thereby, it would be wrong to check the recovery of the parameter for $\tilde{t}_{in}$ and $\tilde{c}_{in}$ in (10); rather, its ratio (i.e. β_t/β_c) needs to be checked.

Now, while the two-stage procedure guarantees consistent estimators up to a scale, their standard errors cannot be calculated, in general, directly from the Fisher information matrix due to the estimated regressor problem (Guevara 2015). Therefore, they must be corrected to carry out statistical inference analysis. The only case in which this correction is not needed is for testing that $\beta_{\hat{\delta}t}$ and $\beta_{\hat{\delta}c}$ in (10) are equal to zero, because under such a null assumption, there is no endogeneity. Therefore, this can be used as a direct test to discard the presence of endogeneity (Rivers and Vuong 1988). Other alternatives to correct the standard errors are the Bootstrap method (Petrin and Train 2003) or the formula proposed by Hardin (2002), by Karaca-Mandic and Train (2003), or the one proposed by Terza (2016). If the modeller wants to avoid the standard error correction, the maximum likelihood simultaneous *CF* approach should be applied (Train 2009).

Furthermore, the two-stage estimation approach results in an efficiency loss, which can be avoided when the error terms $\tilde{\tilde{\varepsilon}}_{in}$ and $\tilde{\varphi}_{in}$ are homoscedastic and non-autocorrelated (Rivers and Vuong 1988). For other cases, efficiency may only be regained by following a maximum likelihood approach, as Train (2009) suggested.

Note finally, that the mathematical derivation of the *CF* approach considers that the error terms $\tilde{\tilde{\varepsilon}}_{in}$ and $\tilde{\varphi}_{in}$ distribute bivariate Normal. Therefore, the specification in (1) leads to a Probit model. However, Villas-Boas (2007) showed that, under some considerations, approximating it to a Gumbel distribution (also called Extreme Value Type I) did not involve any issue.

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^1_{in} given Gumbel errors corresponds to the following multinomial logit expression:

$$P^1_{in} = \frac{e^{\widehat{ASC}_i + \hat{\beta}_t \tilde{t}^1_{in} + \hat{\beta}_c \tilde{c}^1_{in} + \hat{\beta}_{\hat{\delta}t} \hat{\delta}^t_{in} + \hat{\beta}_{\hat{\delta}c} \hat{\delta}^c_{in}}}{\sum_{j \in A} e^{\widehat{ASC}_j + \hat{\beta}_t \tilde{t}^1_{jn} + \hat{\beta}_c \tilde{c}^1_{jn} + \hat{\beta}_{\hat{\delta}t} \hat{\delta}^t_{jn} + \hat{\beta}_{\hat{\delta}c} \hat{\delta}^c_{jn}}} \tag{11}$$

where the parameters $\widehat{ASC}$ and $\hat{\beta}$ in (11) 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, $\hat{\delta}$ is fixed. This last comment is crucial because unlike Guevara and Ben-Akiva (2012), we hypothesise 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 (12):

$$P^1_{in} = \frac{e^{\widehat{ASC}_i + \hat{\beta}_t \tilde{t}^1_{in} + \hat{\beta}_c \tilde{c}^1_{in} + \hat{\beta}_{\hat{\delta}t} (\tilde{t}_{in} - \hat{\alpha}_t - \hat{\gamma}^t_z z^0_{in} - \hat{\gamma}^t_{fft} fft^0_{in}) + \hat{\beta}_{\hat{\delta}c} (\tilde{c}_{in} - \hat{\alpha}_c - \hat{\gamma}^c_z z^0_{in} - \hat{\gamma}^c_{fft} fft^0_{in})}}{\sum_{j \in A} e^{\widehat{ASC}_j + \hat{\beta}_t \tilde{t}^1_{jn} + \hat{\beta}_c \tilde{c}^1_{jn} + \hat{\beta}_{\hat{\delta}t} (\tilde{t}_{jn} - \hat{\alpha}_t - \hat{\gamma}^t_z z^0_{jn} - \hat{\gamma}^t_{fft} fft^0_{jn}) + \hat{\beta}_{\hat{\delta}c} (\tilde{c}_{jn} - \hat{\alpha}_c - \hat{\gamma}^c_z z^0_{jn} - \hat{\gamma}^c_{fft} fft^0_{jn})}} \tag{12}$$

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

The approach of Guevara and Ben-Akiva (2012) has a potential limitation in forecasting because it assumes that exogenous shifts do not affect predictions in future scenarios. In strategic transport models, the simulation of future scenarios requires achieving new supply/demand equilibria, which must be considered endogenous as they result from an equilibrium process. This might involve a change, not necessarily linear, of the relationship between the residuals and the endogenous variables, invalidating the use of (11) or (12). Namely, the endogeneity continues arising in the future years affecting the calculation of the probabilities; however, there is no guarantee that the instruments used for the calibration year correct it, given that the correlation between the instruments (z_{in} and fft_{in}) and the endogenous variable ($\tilde{c}_{in}$ and $\tilde{t}_{in}$) changes in (8) and (9). This fact is ignored by Guevara and Ben-Akiva (2012), and all related literature since they are only concern with the analysis of exogenous shifts. In practice, these situations can occur – for example – when the free-flow times change due to improvements in the infrastructure or when the fuel cost changes due to governmental policies. Note that the shifts are exogenous for both cases, but it triggers endogenous changes in the forecasting of future scenarios. We propose to address this limitation by updating the residuals for applying the *CF* method for futures scenarios. We termed this proposed variation of the *CF* method as Control Function Updated (*CFU*), which will be explained below.

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^1_{in} and fft^1_{in}) instead of z^0_{in} and fft^0_{in} in (8) and (9). Therefore, these new residuals are identified as $\hat{\delta}^{1t}_{in}$ and $\hat{\delta}^{1c}_{in}$. They are obtained applying an OLS regression of $\tilde{c}^1_{in}$ and $\tilde{t}^1_{in}$ on the exogenous variables in the DCM and the instruments. In this way, the CFU approach's first stage is represented in (13) and (14):

$$\tilde{t}^1_{in} = \alpha_t + \gamma^t_z z^1_{in} + \gamma^t_{fft} fft^1_{in} + \delta^{1t}_{in} \quad \overset{OLS}{\rightarrow} \quad \hat{\delta}^{1t}_{in} = \tilde{t}^1_{in} - \hat{\tilde{t}}^1_{in} \tag{13}$$

$$\tilde{c}^1_{in} = \alpha_c + \gamma^c_z z^1_{in} + \gamma^c_{fft} fft^1_{in} + \delta^{1c}_{in} \quad \overset{OLS}{\rightarrow} \quad \hat{\delta}^{1c}_{in} = \tilde{c}^1_{in} - \hat{\tilde{c}}^1_{in} \tag{14}$$

Where, α is the intercept of the regression model, γ_z and γ_{fft} are parameters to be estimated for the exogenous attributes z^1_{in} and fft^1_{in}, respectively. Finally, δ^1_{in} is the error term of the regression model.

For the second stage of the CFU approach, $\hat{\delta}^{1c}_{in}$ and $\hat{\delta}^{1t}_{in}$ are added as explanatory variables within the utility function. However, the choices are unknown; therefore, we use P^1_{in} coming from (11) or (12) 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. $\widehat{ASC}$, $\hat{\beta}_t$ and $\hat{\beta}_c$). The attribute vectors ($\tilde{t}^1_{in}$, $\tilde{c}^1_{in}$, $\hat{\delta}^{1c}_{in}$ and $\hat{\delta}^{1t}_{in}$) are known for the modeller, therefore the only unknowns elements are $\hat{\beta}_{\hat{\delta}1c}$ and $\hat{\beta}_{\hat{\delta}1t}$. These can be re-estimated using the linear regression in (15), which can be seen as an application of the Berkson-Theil transformation procedure (Ortúzar and Willumsen 2011).

$$Ln\frac{P^1_{in}}{P^1_{jn}} = (\widehat{ASC}_i - \widehat{ASC}_j) + \hat{\beta}_t(\tilde{t}^1_{in} - \tilde{t}^1_{jn}) + \hat{\beta}_c(\tilde{c}^1_{in} - \tilde{c}^1_{jn}) + \beta_{\hat{\delta}1c}(\hat{\delta}^{1c}_{in} - \hat{\delta}^{1c}_{jn}) + \beta_{\hat{\delta}1t}(\hat{\delta}^{1t}_{in} - \hat{\delta}^{1t}_{jn}) + (\tilde{\tilde{\tilde{\varepsilon}}}_{in} - \tilde{\tilde{\tilde{\varepsilon}}}_{jn}) \tag{15}$$

where the left-hand side of (15) is estimated based on Guevara and Ben-Akiva (2012) approach and it acts as the dependent variable; whereas $(\hat{\delta}^{1c}_{in} - \hat{\delta}^{1c}_{jn})$ and $(\hat{\delta}^{1t}_{in} - \hat{\delta}^{1t}_{jn})$ act as the independent variables. Then, the sum of $\hat{\beta}_t(\tilde{t}^1_{in} - \tilde{t}^1_{jn})$, $\hat{\beta}_c(\tilde{c}^1_{in} - \tilde{c}^1_{jn})$ and $(\widehat{ASC}_i - \widehat{ASC}_j)$ is the intercept.

The calculation of the probability $P^{1\_CFU}_{in}$ after updating the residuals is given by (16) for the two-stage *CF* approach. Note that the parameters used are $\widehat{ASC}$, $\hat{\beta}_t$ 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 (17):

$$P^{1\_CFU}_{in} = \frac{e^{\widehat{ASC}_i + \hat{\beta}_t\tilde{t}^1_{in} + \hat{\beta}_c\tilde{c}^1_{in} + \hat{\beta}_{\hat{\delta}1t}\hat{\delta}^{1t}_{in} + \hat{\beta}_{\hat{\delta}1c}\hat{\delta}^{1c}_{in}}}{\sum_{j\in A} e^{\widehat{ASC}_j + \hat{\beta}_t t^1_{jn} + \hat{\beta}_c c^1_{jn} + \hat{\beta}_{\hat{\delta}1t}\hat{\delta}^{1t}_{jn} + \hat{\beta}_{\hat{\delta}c1}\hat{\delta}^{1c}_{jn}}} \tag{16}$$

$$P^{1\_CFU}_{in} = \frac{e^{\widehat{ASC}_i + \hat{\beta}_t\tilde{t}^1_{in} + \hat{\beta}_c\tilde{c}^1_{in} + \hat{\beta}_{\hat{\delta}1t}(\tilde{t}^1_{in} - \hat{\alpha}_t - \hat{\gamma}^t_z z^1_{in} - \hat{\gamma}^t_{fft} fft^1_{in}) + \hat{\beta}_{\hat{\delta}1c}(\tilde{c}^1_{in} - \hat{\alpha}_c - \hat{\gamma}^c_z z^1_{in} - \hat{\gamma}^c_{fft} fft^1_{in})}}{\sum_{j\in A} e^{\widehat{ASC}_j + \hat{\beta}_t\tilde{t}^1_{jn} + \hat{\beta}_c\tilde{c}^1_{jn} + \hat{\beta}_{\hat{\delta}1t}(\tilde{t}^1_{jn} - \hat{\alpha}_t - \hat{\gamma}^t_z z^0_{jn} - \hat{\gamma}^t_{fft} fft^0_{jn}) + \hat{\beta}_{\hat{\delta}1c}(\tilde{c}^1_{jn} - \hat{\alpha}_c - \hat{\gamma}^c_z z^0_{jn} - \hat{\gamma}^c_{fft} fft^0_{jn})}} \tag{17}$$

A Monte Carlo framework to represent simulation equilibrium and forecasting

To investigate the problem at hand, we designed a Monte Carlo experiment. The simulation process considered a sample size of 5000 individuals and was repeated 100 times, to guarantee randomness in the estimates. All experiment runs were generated using the open-source software R (R Development Core Team 2008). We considered six transport modes: Bus, Car driver, Shared car, Walking, Train and Shared taxi, emulating the typical alternatives considered in the application of a strategic 4-stage urban transport model; we also assumed that the availability of modes could vary among individuals (4 to 6 alternatives available). On the other hand, the network structure (i.e. square grid, monocentric, radial, or other) was not considered. Besides, we did not consider a particular number of OD pairs either. This could be interpreted as each observation being an OD pair isolated from the others and with different free-flow times.

For the choice's simulation process, we specified six utility functions shown in (18)-(23), one for each simulated transport mode:

$$U_{Bus,n} = \beta_c c_{Bus,n} + \beta^{PT}_t t_{Bus,n} + \beta_q q_{Bus,n} + \varepsilon_{Bus,n} \tag{18}$$

$$U_{Car\ driver,n} = \beta_c c_{Car\ driver,n} + \beta^{CM}_t t_{Car\ driver,n} + \beta_q q_{Car\ driver,n} + \varepsilon_{Car\ driver,n} \tag{19}$$

$$U_{Shared\ Car,n} = \beta^{CM}_t t_{Shared\ Car,n} + \varepsilon_{Shared\ Car,n} \tag{20}$$

$$U_{Walking,n} = \beta_t^W t_{Walking,n} + \varepsilon_{Walking,n} \tag{21}$$

$$U_{Train,n} = \beta_c c_{Train,n} + \beta_t^{PT} t_{Train,n} + \beta_q q_{Train,n} + \varepsilon_{Train,n} \tag{22}$$

$$U_{Shared\ taxi,n} = \beta_c c_{Shared\ taxi,n} + \beta_t^{PT} t_{Shared\ taxi,n} + \beta_q q_{Shared\ taxi,n} + \varepsilon_{Shared\ taxi,n} \tag{23}$$

where cost (c_{in}) has a generic parameter (β_c) in the utility function of Bus, Car driver, Train and Shared taxi (i.e. we assumed that Shared Car and Walking had not cost); travel time (t_{in}), has a specific parameter depending on whether the mode was walking (β_t^W), Car driver and Shared car (β_t^{CM}) or public transport, that is Bus, Train and Shared taxi (β_t^{PT}), and the error term ε_{in} was considered independently and identically distributed (IID) Extreme Value Type I (0,1).

We considered that both c_{in} and t_{in}, were endogenous. We assumed that the endogeneity for c_{in} was due to measurement error and omitted variables, whereas the endogeneity for t_{in} was due to its simultaneous estimation in the supply-demand equilibration process. In this way, we covered the three typical sources of endogeneity that may occur in the strategic transport modelling.

The variable c_{in} was constructed using the parameters shown in (24). Here, c_{in} is a function of an intercept term (α_c), q_{in}, the model's error term φ_{in}, and z_{in} which is an instrument required to correct for endogeneity in c_{in}.

$$c_{in} = \alpha_c + \gamma_z^c z_{in} + \gamma_q q_{in} + \varphi_{in} \tag{24}$$

where γ_z^c is the parameter of z_{in}^c that considers the weakness/strength of the instrument in the simulation (Staiger and Stock 1997) and should be much higher than zero. The parameter γ_q allows the correlation between c_{in} and q_{in}. If this value is zero, then endogeneity does not arise as will be explained below. Finally, the intercept term (α_c) is a constant that means a minimal cost.

On the other hand, a measurement error ω_{in} is added to the explanatory variable c_{in} as shown in (25), yielding the variable $\tilde{c}_{in}$ and, therefore, the first source of endogeneity for the simulation. In practice, this measurement error may come from the inaccuracy and/or complexity involved in the on-site data collection process, which is typical in large-scale mobility survey. Besides, the aggregation process during the modelling may also affect the estimation of the explanatory variables such as travel times and travel costs.

$$\tilde{c}_{in} = c_{in} + \omega_{in} \tag{25}$$

The second source of endogeneity in our simulation comes from omitting q_{in} in (18), (19), (22) and (23), when the model is estimated using the simulated data. The inclusion of q_{in} in (24) accounts for this source of endogeneity in the simulation. Note that if q_{in} is omitted in the specification of the utility functions in (18), (19), (22) and (23), then the new error term in those equations would be equal to $\tilde{\varepsilon}_{in} = \beta_q q_{in} + \varepsilon_{in}$. Therefore, $\tilde{c}_{in}$ would be correlated with $\tilde{\varepsilon}_{in}$, and by definition, $\tilde{c}_{in}$ would be endogenous. Note that this only affects the utility function for Bus, Car driver, Train and Shared taxi, because $\tilde{c}_{in}$ is an explanatory variable for these modes. The omission of q_{in} does not affect t_{in} because it is not correlated with q_{in}.

In practice, variables such as safety, comfort and/or reliability are often omitted in strategic transport models. These variables are usually significant in explaining the mode choice stage. Still, they are challenging to measure and can also be correlated with cost and/or travel time, causing additional endogeneity.

To create endogeneity in the variable t_{in}, we simulated a simultaneous equilibrium. For this, we reproduced the estimation process in the strategic transport modelling suite ESTRAUS[1] (De Cea et al. 2005), where the levels-of-service are estimated for the distribution, mode choice and assignment sub-models. First, six BPR (Bureau of Public Roads 1964) type functions were defined, as shown in (26) to (31):

$$t_{Bus,n} = \pi_{Bus} t_{Car\ driver,n} \tag{26}$$

$$t_{Car\ driver,n} = fft_{Car\ driver}\{1 + \mu[demand(\tau_{Bus}P_{Bus} + P_{Car\ driver} + \tau_{Shared\ taxi}P_{Shared\ taxi})^{\rho}]\} \tag{27}$$

$$t_{Shared\ car,n} = t_{Car\ driver,n} \tag{28}$$

$$t_{Walking,n} = fft_{Walking,n} \tag{29}$$

$$t_{Train,n} = fft_{Train,n} \tag{30}$$

$$t_{Shared\ taxi,n} = \pi_{Shared\ taxi} t_{Car\ driver,n} \tag{31}$$

where fft_{in} corresponds to the free-flow time for alternative i, which was considered a positive non-zero uniform random variable. These values are exogenous and come from previous researches reported in the literature using real databanks (Guerrero et al. 2020). In practice, the parameters (μ and ρ) of the BPR function must be calibrated. We considered that Bus, Car driver, Shared car and Shared taxi used the same infrastructure; therefore, the travel times for Bus, Shared car and Shared taxi depend on the Car driver's travel time. In the same way, the Car driver's travel time is only affected by its choice probability ($P_{Cardriver}$) and those of Bus and Shared taxi (P_{Bus} and $P_{Shared\ taxi}$). The parameters τ_{Bus} and $\tau_{Sharedtaxi}$ refer to the car equivalent units. On the other hand, the parameters π_{Bus} and $\pi_{Sharedtaxi}$ emulate increases in the travel time for the modes Bus and Shared taxi in comparison with the travel time for the car driver mode, given that they share the same infrastructure of transport. The variable called demand was added to the BPR function to consider the effect of increased demand over the years; it takes a positive non-zero value in the base year. Finally, given that Walking and Train do not share infrastructure with other modes, their travel times only depend on their free-flow time.

Having defined the travel time functions for each transport mode, the simulated simultaneous equilibrium for the *true* model followed the iterative process shown in Figure 1. The values reached at equilibrium will be known as $\tilde{t}_{in}$ from now on, and they are considered endogenous. First $\tilde{t}_{in}$ is estimated using the functions (26) to (31). Then, using the variable values obtained in (24) and the error term (ε_{in}), the utility functions in (18) to (23) are estimated. Finally, the probabilities obtained are used back again in step one, to find new values for $\tilde{t}_{in}$.

The iterative process must converge to an equilibrium, and the stop criterion is triggered when the tolerance (between $\tilde{t}_{in}$ in two successive iterations) is less than or equal to 0.001[2]. The convergence in our simulation was achieved implementing the method of successive weighted averages (MSWA) proposed by Liu, He, and He (2009). Note that fft_{in} is, by construction, a valid instrument of $\tilde{t}_{in}$; because fft_{in} is exogenous and correlated with $\tilde{t}_{in}$, as it appears in the right-hand side of equations (26) to (31).

The choices for the base year were simulated with the parameters specified in Table 1. The values of the parameters β_t^{PT}, β_t^{CM}, β_t^{W} and β_c are typical of results found in applications

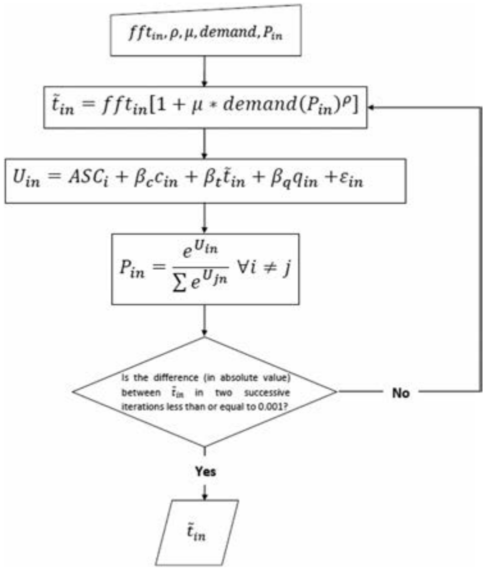

Figure 1. Simulated simultaneous equilibrium process flowchart for the *true* model.

Table 1. Parameters of the Monte Carlo simulation.

Parameter	Value	Parameter	Value
β_t^{PT}	−0.03	π_{Bus}	1.25
β_t^{CM}	−0.04	$\pi_{Sharedtaxi}$	1.05
β_t^{W}	−0.05	μ	0.80
β_c	−0.01	ρ	1.20
α_c	10.0	τ_{Bus}	0.05
γ_z^c	2.0	$\tau_{Sharedtaxi}$	0.40
γ_q	1.0	*demand*	5

of strategic urban mode choice models using DCM (Varela, Börjesson, and Daly 2018; Guerrero et al. 2020). The value of the parameter associated with γ_z^c represents the weakness/strength of the instrument in the simulation (Staiger and Stock 1997). This value guarantees that the instrument is strong enough since it must be higher than zero. Besides, the value of γ_q associated with the variable q_{in} allows the correlation between c_{in} and q_{in}. If this value is zero, endogeneity does not arise as explained above. The intercept term (α_c) is a constant to guarantee that c_{in} will always be positive. The coefficients τ_{Bus} and $\tau_{Sharedtaxi}$ refer to the car equivalent units; for example, 0.05 comes from assuming that 20 people on a bus occupy the same road space as one person in a car (SECTRA 2013). On the other hand, the parameters π_{Bus} and $\pi_{Sharedtaxi}$ mean that the travel times of Bus and Shared taxi are fixed as 25% and 5% higher than the Car driver's travel time, respectively (e.g. additional time spent to stop at stations to pick up passengers). Finally, μ and ρ are the BPR function parameters, which must be $\mu > 0$ and $\rho > 1$ (Márquez, García, and Guarín 2014).

We estimated two different models for the base year: *endogenous* and *corrected*. The last one used the *CF* classical approach (Heckman 1978). The endogenous model includes $\tilde{c}_{in}$ and $\tilde{t}_{in}$ in the specification of the utility function, and the corrected model includes $\tilde{c}_{in}$, $\tilde{t}_{in}$, $\hat{\delta}_{in}^c$ and $\hat{\delta}_{in}^t$, where $\hat{\delta}_{in}^c$ and $\hat{\delta}_{in}^t$ are the residuals from the first stage of the *CF* approach for the endogenous variables $\tilde{c}_{in}$ and $\tilde{t}_{in}$, as shown in (32) and (33):

$$\tilde{t}_{in} = \alpha_t + \gamma_z^t z_{in} + \gamma_{fft}^t fft_{in} + \delta_{in}^t \quad \overset{OLS}{\rightarrow} \quad \hat{\delta}_{in}^t = \tilde{t}_{in} - \hat{\tilde{t}}_{in} \tag{32}$$

$$\tilde{c}_{in} = \alpha_c + \gamma_z^c z_{in} + \gamma_{fft}^c fft_{in} + \delta_{in}^c \quad \overset{OLS}{\rightarrow} \quad \hat{\delta}_{in}^c = \tilde{c}_{in} - \hat{\tilde{c}}_{in} \tag{32}$$

where α_t, α_c, γ_z^t, γ_z^c, γ_{fft}^t, γ_{fft}^c, $\hat{\delta}_{in}^t$ and $\hat{\delta}_{in}^c$ are obtained from an OLS regression of $\tilde{t}_{in}$ and $\tilde{c}_{in}$ on the exogenous variables z_{in} and fft_{in}. This is the first stage of the two-stage *CF* approach proposed by Wooldridge (2010).

Until here, this explains only the process of the choices' simulation and estimation of the endogenous and corrected model parameters. The forecasting stage has not started yet. To answer our research question – 'how correcting for endogeneity would work in forecasting future scenarios (i.e. 10 to 40 years ahead) when the variables that change are endogenous?' – we compared three forecasting approaches: (*i*) do nothing (*No endogeneity correction*); (*ii*) the Guevara and Ben-Akiva (2012) *CF* approach and (*iii*) our proposal (*CFU* method). We considered two critical aspects of modelling transport in future scenarios: first, exogenous changes in the explanatory variables and second, increases in traffic demand (congestion). For the first aspect, we simulated several hypothetical scenarios where the travel time and/or travel cost change (for example, decreasing travel time due to infrastructure improvements or increasing travel cost due to change in fuel price).

In the simulation of future scenarios for each approach, fft_{in} and z_{in} can change in future situations. For example, infrastructure improvements may yield a decrease in free-flow times, identified as fft_{in}^{FUT} for future scenarios. On the other hand, increasing travel cost ($\tilde{c}_{in}^{FUT}$) can be due to the fuel price (represented in z_{in}^{FUT}) and suffer a shift for future scenarios; therefore, $\tilde{c}_{in}^{FUT}$ must be estimated as shown in (34):

$$\tilde{c}_{in}^{FUT} = \underbrace{(\alpha_c + \gamma_z^c z_{in}^{FUT} + \gamma_q q_{in} + \tilde{\delta}_{in})}_{c_{in}^{FUT}} + \omega_{in} \tag{34}$$

In particular, we simulated the effect of eight scenarios shown in Table 2.

Finally, to estimate the future value of the demand (*demand*^{FUT}), we applied a simple model (35):

$$demand^{FUT} = demand(1 + r)^N \tag{35}$$

where r is the demand growth rate (say, 5%), and N is the forecasting year (i.e. 10 to 40 years).

Table 2. Description of the hypothetical future scenarios.

Scenario	Exogenous change
1	50% decrease in free flow time for the train mode
2	30% decrease in free flow time for the train mode
3	10% decrease in free flow time for the car driver, shared car, bus and shared taxi modes
4	10% increase in travel cost for the car mode
5	50% decrease in free flow time for the train mode and 20% decrease in free flow time for the car driver, shared car, bus and shared taxi modes
6	50% decrease in free flow time for the train mode and 20% increase in travel cost for the car mode
7	30% decrease in free flow time for the train mode, 10% decrease in free flow time for the car driver, shared car, bus and shared taxi modes and 10% increase in travel cost for the car mode
8	50% decrease in free flow time for the train mode, 20% decrease in free flow time for the car driver, shared car, bus and shared taxi modes and 20% increase in travel cost for the car mode

In the case of the future scenarios, a new equilibrium must be reached as shown in Figure 2, which is affected by the changes in fft_{in}^{FUT}, $\tilde{c}_{i,n}^{FUT}$ and $demand^{FUT}$. The values reached in this equilibrium process are $\tilde{t}_{in}^{FUT}$, and are considered endogenous because they come from a simultaneous equilibrium process. Note that in the process described, U_{in}^{FUT} corresponds to the model utility corrected using the *CF* approach of Guevara and Ben-Akiva (2012) in the forecasting stage (i.e. (11) or (12)).

Given that the equilibrium for future scenarios is reached in each approach, the specification of U_{in}^{FUT} must change, accordingly, in each approach. Besides, note that as we are forecasting, in general $\widehat{ASC}$, $\hat{\beta}$,$\hat{\alpha}$,$\hat{\gamma}$ and $\hat{\delta}$, the parameters used in U_{in}^{FUT} are those estimated in the base year by the *true*, endogenous and corrected models. However, our new methodological approach (*CFU*) requires updating the residuals $\hat{\delta}$ as follows.

Once the values $\tilde{t}_{in}^{FUT}$ are estimated (Figure 2), we need to update the *CF* for the future scenarios as shown in (36) and (37). The update consists in replacing the instrument fft_{in} and z_{in} with fft_{in}^{FUT} and z_{in}^{FUT} (respectively) in the first stage of the *CFU* approach, so that new values for the residuals are obtained ($\hat{\delta}_{in}^{c\_FUT}$ and $\hat{\delta}_{in}^{t\_FUT}$).

$$\tilde{t}_{in}^{FUT} = \alpha_t + \gamma_z^t z_{in}^{t\_FUT} + \gamma_{fft}^t fft_{in}^{FUT} + \delta_{in}^{\tilde{t}\_FUT} \overset{OLS}{\rightarrow} \hat{\delta}_{in}^{\tilde{t}\_FUT} = \tilde{t}_{in}^{FUT} - \hat{\tilde{t}}_{in}^{FUT} \tag{36}$$

$$\tilde{c}_{in}^{FUT} = \alpha_c + \gamma_z^c z_{in}^{c\_FUT} + \gamma_{fft}^c fft_{in}^{FUT} + \delta_{in}^{\tilde{c}\_FUT} \overset{OLS}{\rightarrow} \hat{\delta}_{in}^{\tilde{c}\_FUT} = \tilde{c}_{in}^{FUT} - \hat{\tilde{c}}_{in}^{FUT} \tag{36}$$

For the second stage of the *CFU* approach, where $\hat{\delta}_{in}^{\tilde{t}\_FUT}$ and $\hat{\delta}_{in}^{\tilde{c}\_FUT}$ are added as explanatory variables within the utility function, we can re-estimate the parameters $\beta_{\hat{\delta}\tilde{t}\_FUT}$ and $\beta_{\hat{\delta}\tilde{c}\_FUT}$ using the Berkson-Theil transformation (Ortúzar and Willumsen 2011) as shown in (38) and (39). This can be done because all the terms (less $\beta_{\hat{\delta}\tilde{t}\_FUT}$ and $\beta_{\hat{\delta}\tilde{c}\_FUT}$) in (38) and (39)

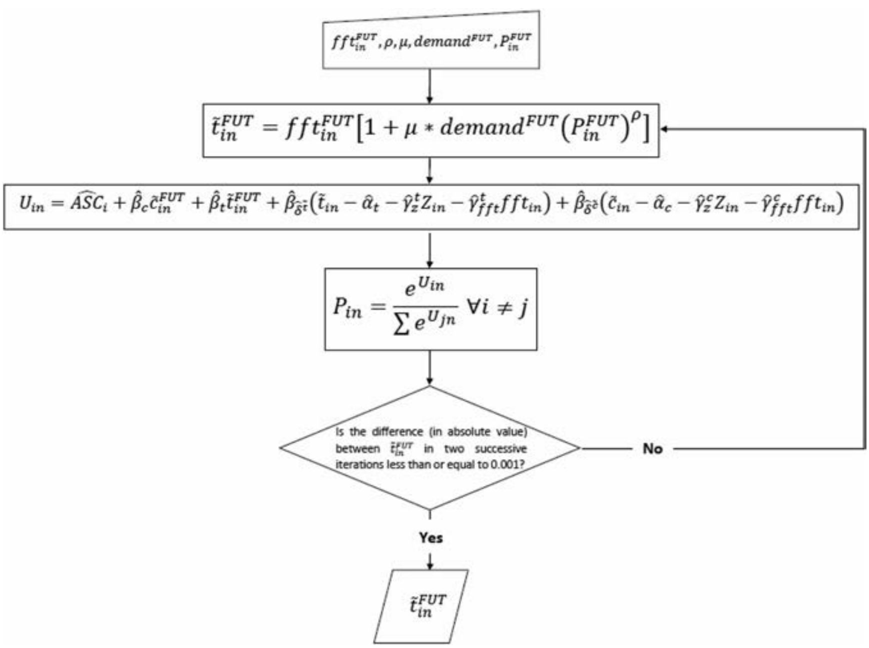

Figure 2. Simultaneous equilibrium process flowchart for future scenarios using the *CF* method.

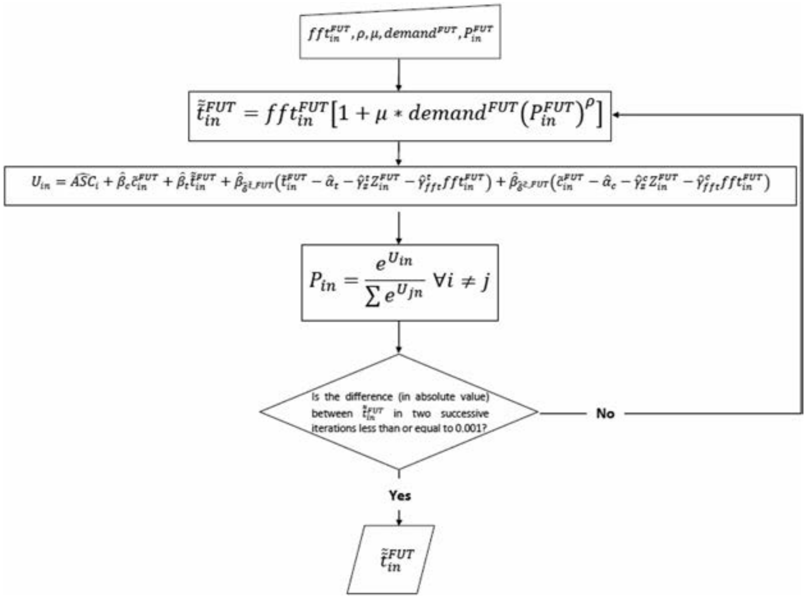

Figure 3. Simulated simultaneous equilibrium flowchart for future scenarios with the *CFU* approach.

are scalars, as the modeller knows them.

$$Ln\frac{P_{in}^{FUT}}{P_{jn}^{FUT}} = (\widehat{ASC}_i - \widehat{ASC}_j) + \hat{\beta}_t(\tilde{t}_{in}^{FUT} - \tilde{t}_{jn}^{FUT}) + \hat{\beta}_c(\tilde{c}_{in}^{FUT} - \tilde{c}_{jn}^{FUT}) + \beta_{\hat{\delta}\tilde{t}\_FUT}(\hat{\delta}_{in}^{\tilde{t}\_FUT} - \hat{\delta}_{jn}^{\tilde{t}\_FUT})$$
$$+ \beta_{\hat{\delta}\tilde{c}\_FUT}(\hat{\delta}_{in}^{\tilde{c}\_FUT} - \hat{\delta}_{jn}^{\tilde{c}\_FUT}) \tag{38}$$

$$Ln\frac{P_{in}^{FUT}}{P_{jn}^{FUT}} - \{(\widehat{ASC}_i - \widehat{ASC}_j) + \hat{\beta}_t(\tilde{t}_{in}^{FUT} - \tilde{t}_{jn}^{FUT}) + \hat{\beta}_c(\tilde{c}_{in}^{FUT} - \tilde{c}_{jn}^{FUT})\} = \beta_{\hat{\delta}\tilde{t}\_FUT}(\hat{\delta}_{in}^{\tilde{t}\_FUT} - \hat{\delta}_{jn}^{\tilde{t}\_FUT})$$
$$+ \beta_{\hat{\delta}\tilde{c}\_FUT}(\hat{\delta}_{in}^{\tilde{c}\_FUT} - \hat{\delta}_{jn}^{\tilde{c}\_FUT}) \tag{39}$$

In this way, $\beta_{\hat{\delta}\tilde{t}\_FUT}$ and $\beta_{\hat{\delta}\tilde{c}\_FUT}$ can be re-estimated (i.e. updated) using an OLS regression, and a new equilibrium point is reached considering the updated parameters $\beta_{\hat{\delta}\tilde{t}\_FUT}$ and $\beta_{\hat{\delta}\tilde{c}\_FUT}$ (Figure 3). The values reached in this new equilibrium are $\tilde{\tilde{t}}_{in}^{FUT}$.

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 3, which shows the ratios'

Table 3. 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%

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.

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[3] 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[4]. 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%.

The boxplots[5] in Figure 4 show the parameter ratios for the endogenous and corrected model using the classical *CF* approach (Heckman 1978). We do not show the median; instead, we show the mean as a black dot depicting the average for the 100 repetitions, as we are interested in the average of all observations. The dashed line represents the *true* ratio value. The variance of the corrected ratios seems lower than those of the endogenous ratios. Also, there are more outliers for the endogenous ratios than for the corrected ratios.

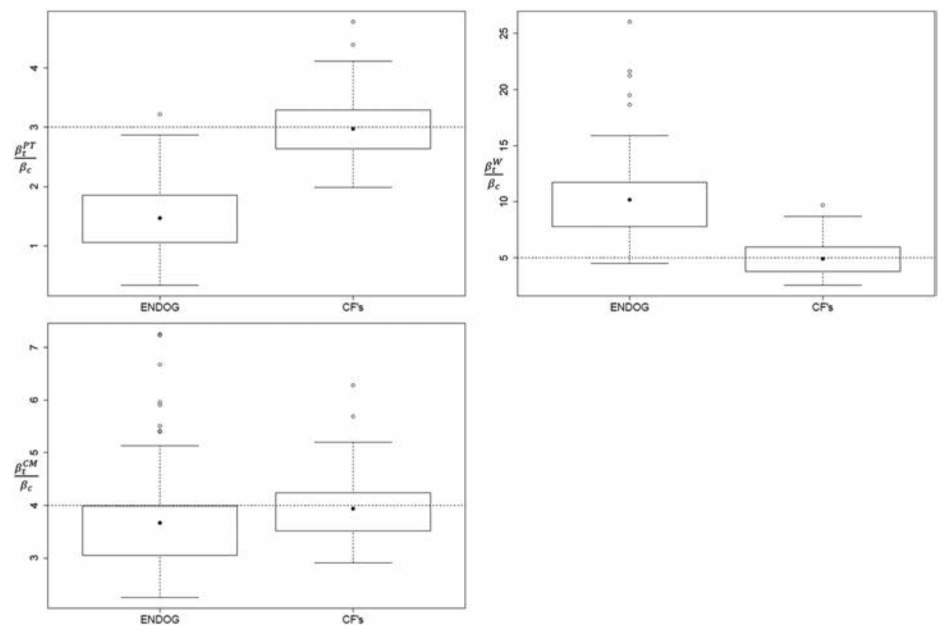

Figure 4. Boxplots of parameter ratios for the endogenous and *CF* corrected model.

The former is far from the *true* ratios, showing the impact of endogeneity in yielding inconsistent parameters. The parameter ratios for the corrected model are close to the *true* ratios, showing the power of the *CF* method to correct for endogeneity.

Note that for all the cases analysed, reported in Table 4, the TTE for the endogenous model were larger than those of the true and corrected models[6]. This does not need to be the case in general. In our case study, the TTE for the endogenous model was larger because, for the specific settings considered, the endogenous model resulted in an underestimation of the value of time of the motorized modes that experienced congestion (see Figure 4). So, smaller values of time implied that the simulated individuals were willing to pay less for reducing travel time (compared to what they would do in the true model), resulting in choices of slower alternatives and, hence, larger TTE[7].

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[8] – $E(AIC)$ 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 (40) and (41), respectively:

$$E(l(\theta)_l^{m,t}) = \sum_n \sum_{i\epsilon A(n)} ln(P_{in}^{m,t} P_{in}^{True,t}) \tag{40}$$

$$E(AIC_l^{m,t}) = 2k - 2E(l(\theta)_l^{m,t}) \tag{41}$$

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 (l) indicates the number of repetitions. In our case, in (41) we used $E(l(\theta)_l^{m,t})$ to estimate $E(AIC_l^{m,t})$ for approach m and year t, where k corresponds to the number of model parameters.

Table 4 summarises 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.

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

Table 4. Average of the free flow time and travel times in equilibrium (TTE) for 100 replications.

		Scenario 1			Scenario 2			Scenario 3			Scenario 4		
Year	Approach	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

(*continued*)

Table 4. Continued.

		Scenario 5			Scenario 6			Scenario 7			Scenario 8		
Year	Approach	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

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[9].

To evaluate the future scenarios in terms of recovering the *true* (simulated) travel times, we used the % bias as shown in (42). This way, we can indicate how well or poorly the TTE are recovered in comparison with the *true* model. The TTE for the 10-to-40-year forecasts with the endogenous model are worse than in the *true* model in all scenarios. Also, the effects of '*No endogeneity correction*' increase with time.

$$\%bias = \frac{(\mathrm{TTE}_{Approach} - \mathrm{TTE}_{true})}{\mathrm{TTE}_{true}} * 100 \tag{42}$$

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 5 focuses only on the car modes and shows the boxplot for the TTE reached by each of the three approaches.

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

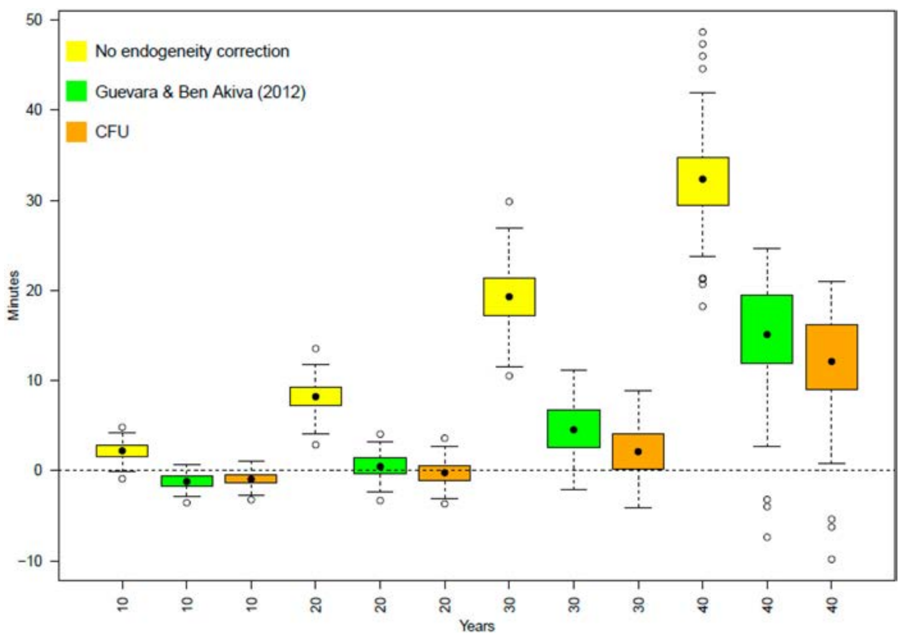

Figure 5. TTE reached with the three approaches compared vs *true* model for Scenario 8.

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 5 also shows that the endogenous model overestimates congestion. This makes sense given that, as shown in Table 3, 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 5 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^2_r 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[10] 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 $> \chi^2_2$ (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

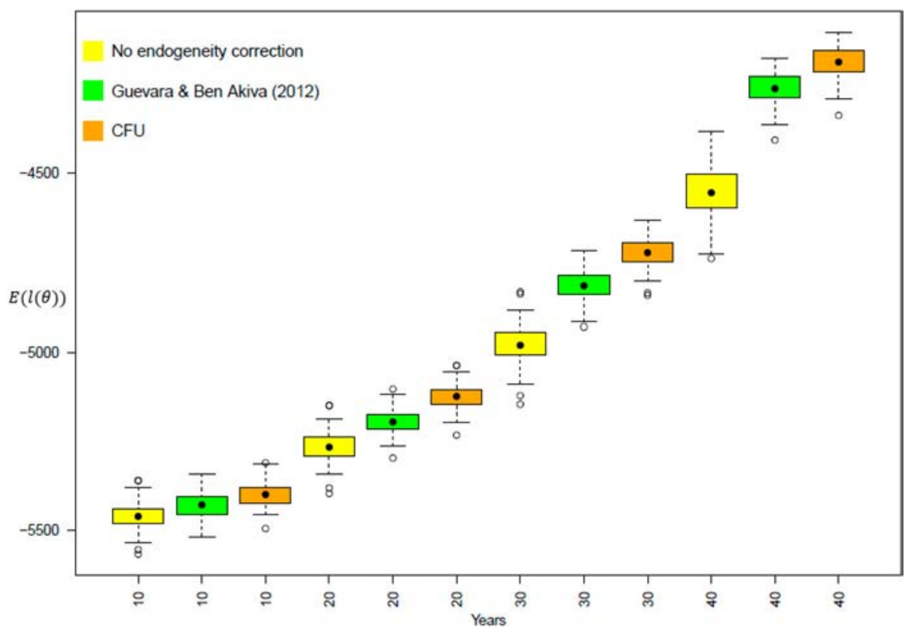

Figure 6. Boxplots of $E(l(\theta))$ for approaches over the years for scenario 8.

Table 5. 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	
		$\boldsymbol{E(l(\theta))}$	$\boldsymbol{E(AIC)}$	$\boldsymbol{E(l(\theta))}$	$\boldsymbol{E(AIC)}$	$\boldsymbol{E(l(\theta))}$	$\boldsymbol{E(AIC)}$	$\boldsymbol{E(l(\theta))}$	$\boldsymbol{E(AIC)}$	$\boldsymbol{E(l(\theta))}$	$\boldsymbol{E(AIC)}$	$\boldsymbol{E(l(\theta))}$	$\boldsymbol{E(AIC)}$	$\boldsymbol{E(l(\theta))}$	$\boldsymbol{E(AIC)}$	$\boldsymbol{E(l(\theta))}$	$\boldsymbol{E(AIC)}$
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

Table 6. Average of the market shares 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	Endogenous	30.7	44.3	25.0	30.7	44.3	25.0	30.7	44.3	25.0	30.7	44.3	25.0
	Corrected	30.7	44.3	25.0	30.7	44.3	25.0	30.7	44.3	25.0	30.7	44.3	25.0
10	True	26.6	48.8	24.6	27.3	47.3	25.4	29.9	44.8	25.3	28.3	45.4	26.3
	No endogeneity correction	28.4	46.4	25.3	28.6	46.1	25.4	30.0	44.9	25.2	28.7	45.6	25.6
	Guevara and Ben-Akiva (2012)	26.7	49.0	24.3	27.4	47.5	25.1	29.9	44.9	25.2	28.4	45.6	26.1
	CFU	26.4	49.1	24.5	27.2	47.5	25.3	29.9	44.7	25.4	28.2	45.5	26.3
20	True	23.9	50.1	26.0	24.5	48.7	26.8	27.0	46.1	26.9	25.5	46.7	27.8
	No endogeneity correction	25.2	48.7	26.1	25.4	48.4	26.3	26.9	47.0	26.0	25.6	47.9	26.5
	Guevara and Ben-Akiva (2012)	23.5	50.7	25.8	24.2	49.3	26.6	26.6	46.6	26.8	25.1	47.3	27.6
	CFU	23.2	50.9	26.0	23.8	49.3	26.8	26.4	46.5	27.1	24.8	47.2	27.9
30	True	20.5	52.0	27.5	21.1	50.5	28.4	23.4	47.9	28.7	22.1	48.5	29.5
	No endogeneity correction	**20.7**	**52.0**	**27.3**	20.8	51.7	27.5	22.4	50.3	27.3	21.0	51.3	27.7
	Guevara and Ben-Akiva (2012)	19.2	53.4	27.5	19.7	52.0	28.4	22.0	49.2	28.8	20.5	50.0	29.5
	CFU	19.1	53.5	27.5	19.6	52.0	28.4	21.9	49.2	29.0	20.5	50.0	29.6
40	True	16.8	54.1	29.1	17.3	52.7	30.0	19.3	50.1	30.6	18.2	50.6	31.2
	No endogeneity correction	**15.6**	**55.5**	**28.9**	15.7	55.2	29.1	17.1	54.0	28.9	15.9	54.8	29.3
	Guevara and Ben-Akiva (2012)	14.3	56.4	29.3	14.7	55.0	30.3	16.6	52.4	31.0	15.4	53.1	31.5
	CFU	14.5	56.4	29.1	14.9	55.0	30.2	16.7	52.4	30.9	15.6	53.0	31.4

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	Endogenous	30.7	44.3	25.0	30.7	44.3	25.0	30.7	44.3	25.0	30.7	44.3	25.0
	Corrected	30.7	44.3	25.0	30.7	44.3	25.0	30.7	44.3	25.0	30.7	44.3	25.0
10	True	29.7	47.7	22.6	26.4	49.1	24.5	28.8	47.0	24.3	29.5	48.0	22.5
	No endogeneity correction	30.6	45.0	24.4	28.2	46.5	25.3	29.6	45.5	25.0	30.4	45.2	24.4
	Guevara and Ben-Akiva (2012)	29.6	47.8	22.6	26.5	49.3	24.2	28.8	47.1	24.1	29.4	48.2	22.5
	CFU	29.5	47.8	22.6	26.2	49.4	24.4	28.6	47.1	24.3	29.3	48.2	22.5
20	True	27.0	48.9	24.1	24.0	50.2	25.8	26.1	48.1	25.8	27.0	49.0	23.9
	No endogeneity correction	27.9	47.0	25.2	25.2	48.7	26.1	26.7	47.5	25.8	27.8	47.0	25.2
	Guevara and Ben-Akiva (2012)	26.7	49.3	24.0	23.6	50.8	25.6	25.7	48.6	25.6	26.7	49.4	23.9
	CFU	26.4	49.4	24.2	23.3	51.0	25.8	25.4	48.7	25.9	26.4	49.6	24.0
30	True	23.6	50.6	25.8	20.8	51.9	27.3	22.8	49.8	27.4	23.9	50.5	25.6
	No endogeneity correction	23.6	50.1	26.3	20.7	52.0	27.3	22.3	50.8	27.0	23.7	50.0	26.3
	Guevara and Ben-Akiva (2012)	22.5	51.7	25.9	19.4	53.4	27.2	21.4	51.2	27.5	22.7	51.7	25.6
	CFU	22.3	51.8	26.0	19.4	53.4	27.2	21.3	51.2	27.6	22.5	51.8	25.7
40	True	19.7	52.8	27.6	17.3	53.9	28.8	18.9	52.0	29.1	20.1	52.6	27.3
	No endogeneity correction	18.3	53.8	27.9	15.8	55.4	28.8	17.1	54.4	28.6	18.5	53.7	27.8
	Guevara and Ben-Akiva (2012)	17.2	54.9	27.9	14.7	56.3	29.0	16.3	54.3	29.5	17.6	54.8	27.6
	CFU	17.3	54.9	27.8	14.8	56.3	28.9	16.4	54.3	29.4	17.7	54.8	27.5

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 5).

Figure 6 deploys the boxplot of $E(l(\theta))$ estimations for the three approaches above, across the 100 repetitions of the simulation for Scenario 8. As can be seen, '*No endogeneity correction*' shows the worst $E(l(\theta))$. These findings reinforce the negative impact of endogeneity in forecasting. Besides, the *CFU* approach shows, once more, a better $E(l(\theta))$ over the years in comparison with the Guevara and Ben-Akiva (2012) approach. These findings confirm the conclusion that suggests *CFU* shows better performance over Guevara and Ben-Akiva (2012), both of which are far from the endogenous model.

Finally, we show the average for 100 replications of the base year's market shares and the future scenarios in Table 6. Note that ignoring the impact of endogeneity does not affect the base year's market shares, despite the significant bias in the model estimates, because the calibration of a strategic transport model always implies replicating the base year results. Note that the market shares estimated from Guevara and Ben-Akiva (2012) and *CFU* are consistently similar. In most cases, the market shares estimated with the 'No endogeneity correction' approach are far worse than those calculated with the Guevara and Ben-Akiva (2012), and CFU approaches, except for two cases (in Scenario 1 for years 30 and 40)[11]. As the results in Table 6 are shown as aggregate probabilities by mode, namely, Public modes (Bus, Train and Shared Taxi), Car modes (Car driver and Shared car), and Walking mode, this could be purely circumstantial. Our primary interest was to ensure that the CFU approach performed better on measures such as TTE, the logarithm of the expected probability value, and the expected value of the Akaike Information Criteria.

Conclusions

This paper provides a framework for correcting endogeneity, and the correction applied at the forecasting stage of supply-demand equilibration models when the endogenous variables change over the years. 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.

Three areas for further research can be identified. First, we believe it is crucial to examine in greater depth how the social evaluation of transport projects may be affected by endogeneity. Especially given our findings regarding changes in the level-of-service variables when forecasting over several years. Second, the functional form used in our simulation was of multinomial logit type. We recommend exploring other functional forms, such as Nested Logit or Mixed Logit, to see if results vary (although the latter is certainly difficult to implement in the context of a large-scale supply-demand equilibration model). Finally, an exciting extension of this research would be to apply our methodological proposal to real data.

Notes

1. ESTRAUS is a simultaneous equilibrium model designed to analyse and evaluate multimodal urban transport systems with multiple user classes, which has been extensively applied in Santiago and other Chilean cities.
2. We made a sensitivity analysis considering changes in the tolerance (0.1, 0.01 and 0.0001). The results showed no significant differences with the tolerance initially contemplated.
3. 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 $X_1, X_2, \dots$, X_n, of size n, sd is the estimate of the standard deviation of the population, and μ is the population mean.
4. 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.
5. Boxplots were introduced by Tukey (1977). They consist of a rectangle with bottom and topside at the 1st and 3rd quartile (i.e. 25th and 75th percentiles). The distance between the 1st and 3rd quartiles is known as the *interquartile range* (IQR). It allows getting an idea of the dispersion (accumulation) of the values drawn. Usually, they have a horizontal line added at the median (2nd quartile or 50th percentile), and other lines known as *whiskers*. They have a length 1.5 times the IQR added at the top and bottom. Observations outside of the whiskers are plotted and considered as *outliers*.
6. Note that the TTE reached with endogenous and corrected models could be 'theoretically' higher or smaller than the true TTE. All this depends on the simulation conditions considered.
7. We are grateful to an anonymous reviewer for having made us look deeper into this finding.
8. 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).
9. 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.
10. 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.
11. Again, these values are highlighted in **bold** in Table 6.

Acknowledgements

We wish to acknowledge the insightful comments of the Editor and four unknown referees, which helped us to improve the paper considerably. Of course, any remaining errors are our fault.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was partially funded by Agencia Nacional de Investigación y Desarrollo (ANID), FONDECYT 1191104 and by the Instituto Sistemas Complejos de Ingeniería (ISCI), through the grant ANID PIA/BASAL AFB180003. We are also grateful for the support received from the BRT+ Centre of Excellence (www.brt.cl), financed by the Volvo Research and Educational Foundations.

ORCID

Thomas E. Guerrero http://orcid.org/0000-0003-3690-256X

References

Akaike, H. 1974. "A new Look at the Statistical Model Identification." *IEEE Transactions on Automatic Control* 19: 716–723.

Berry, S., J. Levinsohn, and A. Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica* 63: 841–890.

Bresnahan, T. 1997. *The Apple-Innamon Cheerios War: Valuing New Goods, Identifying Market Power, and Economic Measurement*. Working Paper, Department of Economics, Stanford University. https://web.stanford.edu/~tbres/Unpublished_Papers/hausman%20recomment.pdf.

Bureau of Public Roads. 1964. *Traffic Assignment Manual*. Washington, DC: Urban Planning Division, US Department of Commerce.

Citilabs. 2016. *Discover Cube 6.4 Tutorial*. Sacramento: Citilabs Inc.

Cramer, J. S. 2007. "Robustness of Logit Analysis: Unobserved Heterogeneity and mis-Specified Disturbances." *Oxford Bulletin of Economics and Statistics* 69: 545–555.

De Cea, J., J. E. Fernández, V. Dekock, and A. Soto. 2005. "Solving Network Equilibrium Problems on Multimodal Urban Transportation Networks with Multiple User Classes." *Transport Reviews* 25: 293–317.

Ebbes, P., D. Papies, and H. J. Van Heerde. 2011. "The Sense and non-Sense of Holdout Sample Validation in the Presence of Endogeneity." *Marketing Science* 30: 1115–1122.

Guerrero, T. E., C. A. Guevara, E. Cherchi, and J. de D. Ortúzar. 2020. "Addressing Endogeneity in Strategic Urban Mode Choice Models." *Transportation*. doi:10.1007/s11116-020-10122-y.

Guerrero, T. E., C. A. Guevara, E. Cherchi, and J. de D. Ortúzar. 2021. *A Monte Carlo Method to Detect Weak Instruments: Application to Linear and Discrete Choice Models*. *Working Paper*.

Guevara, C. A. 2015. "Critical Assessment of Five Methods to Correct for Endogeneity in Discrete-Choice Models." *Transportation Research Part A: Policy and Practice* 82: 240–254.

Guevara, C. A. 2018. "Overidentification Tests for the Exogeneity of Instruments in Discrete Choice Models." *Transportation Research Part B: Methodological* 114: 241–253.

Guevara, C. A., and M. E. Ben-Akiva. 2006. "Endogeneity in Residential Location Choice Models." *Transportation Research Record: Journal of the Transportation Research Board* 1977: 60–66.

Guevara, C. A., and M. E. Ben-Akiva. 2010. "Addressing Endogeneity in Discrete Choice Models: Assessing Control-Function and Latent-Variable Methods." In *Choice Modelling: The State-of-the-Art and The State-of-Practice*, edited by S. Hess and A. Daly, 353–370. Cheltenham: Edward Elgar.

Guevara, C. A., and M. E. Ben-Akiva. 2012. "Change of Scale and Forecasting with the Control-Function Method in Logit Models." *Transportation Science* 46: 425–437.

Guevara, C. A., and P. Navarro. 2015. "Detection of Weak Instruments When Correcting for Endogeneity in Binary Logit Models." Proceedings 14th International Conference on Travel Behaviour Research (IATBR), Windsor.

Guevara, C. A., and D. Polanco. 2016. "Correcting for Endogeneity due to Omitted Attributes in Discrete-Choice Models: the Multiple Indicator Solution." *Transportmetrica A: Transport Science* 12: 458–478.

Hardin, J. W. 2002. "The Robust Variance Estimator for two-Stage Models." *The Stata Journal: Promoting Communications on Statistics and Stata* 2: 253–266.

Hausman, J. A. 1978. "Specification Tests in Econometrics." *Econometrica* 46: 1251–1271.

Heckman, J. J. 1978. "Dummy Endogenous Variables in a Simultaneous Equation System." *Econometrica* 46: 931–959.

INRO. 1996. EMME/2 *User's Manual*. INRO Inc. Montreal.

Karaca-Mandic, P., and K. Train. 2003. "Standard Error Correction in two-Stage Estimation with Nested Samples." *The Econometrics Journal* 6: 401–407.

Kato, T., and K. Uchida. 2018. "A Study on Benefit Estimation That Considers the Values of Travel Time and Travel Time Reliability in Road Networks." *Transportmetrica A: Transport Science* 14: 89–109.

Kono, T., A. Kishi, E. Seita, and T. Yokoi. 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.

Liu, H. X., X. He, and B. He. 2009. "Method of Successive Weighted Averages (MSWA) and Self-Regulated Averaging Schemes for Solving Stochastic User Equilibrium Problem." *Networks and Spatial Economics* 9: 485–503.

Lurkin, V., L. A. Garrow, M. J. Higgins, J. P. Newman, and M. Schyns. 2017. "Accounting for Price Endogeneity in Airline Itinerary Choice Models: an Application to Continental U.S. Markets." *Transportation Research Part A: Policy and Practice* 100: 228–246.

Márquez, L., D. E. García, and L. C. Guarín. 2014. "Funciones Volumen-Demora BPR y Cónica en Vías Multicarriles de Bogotá." *Revista de Ingeniería* 41: 30–39. (in Spanish).

Mumbower, S., L. A. Garrow, and M. J. Higgins. 2014. "Estimating Flight-Level Price Elasticities Using Online Airline Data: a First Step Toward Integrating Pricing, Demand, and Revenue Optimization." *Transportation Research Part A: Policy and Practice* 66: 196–212.

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

Park, S., and S. Gupta. 2009. "Simulated Maximum Likelihood Estimator for the Random Coefficients Logit Model Using Aggregate Data." *Journal of Marketing Research* 46: 531–542.

Petrin, A., and K. Train. 2003. *Omitted Product Attributes in Discrete Choice Models*. NBER Working Paper 9452. Cambridge, MA: National Bureau of Economic Research, Inc.

R Development Core Team. 2008. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna. http://www.Rproject.org.

Rivers, D., and Q. H. Vuong. 1988. "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models." *Journal of Econometrics* 39: 347–366.

SECTRA. 2013. Mediciones de Aforos de Tráfico y Perfiles de Carga en Servicios Troncales en el Gran Santiago. SECTRA, Santiago (in Spanish).

Staiger, D., and J. H. Stock. 1997. "Instrumental Variables Regression with Weak Instruments." *Econometrica* 65: 557–586.

Terza, J. V. 2016. "Simpler Standard Errors for two-Stage Optimization Estimators." *The Stata Journal: Promoting Communications on Statistics and Stata* 16: 368–385.

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

Tukey, J. W. 1977. *Exploratory Data Analysis*. Reading, MA: Addison-Wesle.

Varela, J. M., M. Börjesson, and A. Daly. 2018. "Quantifying Errors in Travel Time and Cost by Latent Variables." *Transportation Research Part B: Methodological* 117: 520–541.

Villas-Boas, J. M. 2007. *A Note on Limited Versus Full Information Estimation in non-Linear Models*. Discussion Paper. Berkeley: University of California.

Walker, J., and M. Ben-Akiva. 2002. "Generalized Random Utility Model." *Mathematical Social Sciences* 43: 303–343.

Wen, C. H., and P. H. Chen. 2017. “Passenger Booking Timing for low-Cost Airlines: a Continuous Logit Approach.” *Journal of Air Transport Management* 64: 91–99.

Wooldridge, J. 2010. *Econometric Analysis of Cross-Section and Panel Data*. 2nd ed. Cambridge, MA: The MIT Press.

Yatchew, A., and Z. Griliches. 1985. “Specification Error in Probit Models.” *The Review of Economics and Statistics* 67: 134–139.