Transportation
https://doi.org/10.1007/s11116-020-10122-y

Addressing endogeneity in strategic urban mode choice models

Thomas E. Guerrero[1,2,3] · **C. Angelo Guevara**[4] · **Elisabetta Cherchi**[3] · **Juan de Dios Ortúzar**[5]

Abstract
Endogeneity is a potential anomaly in econometric models, which may cause inconsistent parameter estimates. Transport models are prone to this problem and applications that properly correct for it are scarce. This paper focuses on how to address this issue in the case of strategic urban mode choice models (i.e., the third stage of classic strategic transport models), possibly the main tool for the assessment of costly transport projects. To address this problem, we propose and validate, for the first time, adequate instruments that may be obtained from data that is already available in this context. The proposed method is implemented using the Control Function approach, which we use to detect and correct for endogeneity in a case study in Valparaiso, Chile. The effects arising from the neglected endogeneity in this case study reflect on an overestimation between 26–49% of the subjective value of time and an underestimation of 33–75% of modal elasticities.

Keywords Endogeneity · Discrete choice models · Instrumental variables · Strategic urban mode choice models

Introduction and motivation

The term "*endogeneity*" is used in the literature when there is correlation between one or more of the observed explanatory variables and the error term of an econometric model. As a result, the parameter estimates of these variables may be inconsistent. Endogeneity is considered an unavoidable problem in econometric modelling as it may be caused by omitted attributes, measurement or specification errors, simultaneous determination and/or self-selection (Guevara 2015).

Endogeneity is not a new problem and was initially studied in areas such as marketing, in the context of the simultaneity problem between advertising and sales (Bass 1969). The correction of endogeneity in linear models has been widely addressed (Wooldridge 2010), but it is not possible to model certain phenomena using linear models. Discrete Choice Models (DCM) are an example of non-linear models, much used in econometrics when in

✉ Thomas E. Guerrero
teguerrerob@ufpso.edu.co

Extended author information available on the last page of the article

Published online: 04 June 2020

the phenomenon under study the dependent variable is not continuous but discrete; a classical case is when individuals must choose an alternative belonging to a finite set of options. In transport modelling, DCM play a fundamental role (Ortúzar and Willumsen 2011). If the effects of endogeneity are not considered, the analyst can arrive to wrong forecasts and conclusions (Guevara and Ben-Akiva 2006), leading to potentially faulty decision making.

Classic strategic transport models, such as the 4-stage classic transport model, are possibly the main tool for the assessment of costly transport projects (Ortúzar and Willumsen 2011). As any other econometric model, these tools are susceptible to endogeneity problems because when modelling transport, it is common to have omitted attributes, measurement or specification errors, and/or simultaneous determination. For example, the inaccuracy and/or complexity involved in the on-site data collection process may cause measurement errors in several variables (i.e. costs and travel times, among others) derived from the Origin and Destination survey. On the other hand, variables such as safety, comfort and/or reliability are often omitted in large scale models. These variables are usually significant in explaining this stage of the classic transport model, but they are difficult to measure and can additionally be correlated with cost and/or travel time, causing endogeneity. Another cause of endogeneity is the simultaneity since the level of service that defines the choice is a result of a demand–supply equilibration mechanism.

To the best of our knowledge, no mode choice model included in a classic strategic transport model has been corrected for endogeneity. This is a critical drawback because it implies that almost all strategic urban models are based on inconsistent estimates of the subjective value of time (SVT), modal elasticities and forecast, which are crucial for planning and the social evaluation of transport projects. Therefore, there is a need to develop methods to solve this problem and to quantify the effects of correcting the model parameters in model forecasts.

Having detected this gap in the literature, we intend to contribute to the state of knowledge by solving the following three challenges: (i) How is endogeneity detected in strategic urban mode choice models? This stage includes correcting the model; (ii) Solving the practical difficulty of finding adequate instrumental variables (IV) or instruments (Hausman 1978) to correct for endogeneity in this context (Bresnahan 1997; Guevara 2010; Mumbower et al. 2014); the problem comes from the fact that the instruments must fulfil two conflicting properties: be correlated with the endogenous variable, and be independent of the model error; (iii) Quantifying the impacts of neglecting the problem of endogeneity in the estimation of strategic urban modal choice model's parameters.

The rest of the paper is organised as follows. The "Theoretical framework" section details the methodology used, with a focus on how endogeneity arises in DCM and the importance of defining appropriate instruments. The "Application" section describes the databank used and its general characteristics; we present an endogenous model, its corrected version, the instruments used to correct it and a quantification of the impacts of using the corrected version. In the final section we discuss the main findings and conclusions of our research.

Theoretical framework

Endogeneity and DCM

DCM enjoy high applicability in econometrics (Train 2009; Ortúzar and Willumsen 2011). They are used when the dependent variable is discrete in the phenomenon studied, for example, when individuals must choose an alternative belonging to a finite set of options. The use of DCM is very common in areas such as transport demand (Yáñez et al. 2010; Bass et al. 2011; Jensen et al. 2013; Orozco-Fontalvo et al. 2018), road safety (Rizzi and Ortúzar 2006; Anderson and Hernandez 2017), marketing (Lam et al. 2010), spatial economics (Hurtubia and Bierlaire 2014), tourism (Chou and Chen 2014), urbanism (Torres et al. 2013) and environmental economics (Hess and Beharry-Borg 2012).

Like other econometric models, DCM are not exempt from endogeneity, but methods, tests and effects differ from those observed in linear models. For example, the correction of endogeneity in DCM implies a change of scale in the estimated models, and this is not the case in linear models (Guevara and Ben-Akiva 2012). While the problem has been addressed for many types of DCM, to the best of our knowledge it has not been studied under the framework of strategic urban mode choice models, suggesting a research gap that we want to fulfil with this research.

DCM are based on Random Utility Maximization (RUM), whereby the utility, U_{in}, of a certain alternative i for an individual n, is explained by the analyst as the sum of an observed component (systematic, representative or measurable utility, V_{in}) and a random term (Domencich and McFadden 1975; Williams 1977), ε_{in} as shown in (1):

$$U_{in} = V_{in} + \varepsilon_{in} \tag{1}$$

here, V_{in} is a function of a set of observable and measurable attributes X_{ikn}, where the subscript k denotes the attribute; ε_{in} reflects individual tastes and idiosyncrasies not captured in X_{ikn}, in addition to any measurement errors or attributes omitted by the modeller.

This form allows explaining how two individuals with the same attributes and the same set of alternatives ($\underline{A}$) available, can choose differently, or why an individual does not always select the best alternative (from the modeller's point of view, Ortúzar and Willumsen 2011). Thus, individual n will choose alternative A_i belonging to her set of choices $A(n)$ if and only if (2) is fulfilled:

$$U_{in} \geq U_{jn}, \quad \forall A_i \in A(n) \tag{2}$$

If it is assumed that the errors follow an independent and homoscedastic (IID) Gumbel distribution (also called Extreme Value Type I), the popular Multinomial Logit (MNL) model is obtained (Domencich and McFadden 1975); other assumptions about the nature and characteristics of the error term distribution will allow to define different models.

The control function (CF) method

It consists in identifying an auxiliary variable (or control function), such that when it is added to the systematic part of the DCM's utility function, it makes the error of the model uncorrelated with the observed variables (Guevara and Ben-Akiva 2010). This auxiliary variable or CF is constructed by means of an instrumental variable (IV). The

CF method has been used and reported as a suitable approach for correcting endogeneity (Train 2009; Petrin and Train 2010; Wooldridge 2015). Besides the application of the CF method for correcting endogeneity at the individual level presents the advantage of being easy to apply and requiring low consumption of computational resources (Guevara 2015).

Rivers and Vuong (1988) and Villas-Boas and Winer (1999) among others, show that the IVs needed for the application of the CF method in DCM are valid if they fulfil two properties: (i) be correlated with the endogenous variable, and (ii) be independent of the DCM's error. However, identifying proper IVs in practice is always a difficult and even controversial process (see e.g. the debate in Bresnahan 1997). In particular, the CF method can be hard to apply in the case of strategic urban transport modelling suites, because it is not clear how to obtain proper IV to correct for endogeneity in these models.

For explanatory purposes, we will consider a DCM with endogeneity due to the omission of a certain variable q. Assume that its true linear utility function is represented by (3):

$$U_{in} = ASC_i + \beta_x X_{in} + \beta_q q_{in} + e_{in} \tag{3}$$

where ASC_i is an alternative specific constant for alternative A_i; β_x and β_q are parameters to be estimated, X_{in} and q_{in} are explanatory variables of the model, and e_{in} is the exogenous error term. In particular, we will assume that X_{in} represents a set of known (measurable) attributes while the variable q_{in} is unknown to the modeller.

Given the above, let us assume that the specification proposed by the modeller is as in (4):

$$U_{in} = ASC_i + \beta_x X_{in} + \varepsilon_{in} \tag{4}$$

where the new error term ε_{in} obviously contains both e_{in} and q_{in}. Now, let us consider that one of the elements of the set that makes up for X_{in} (for example, the *kth* term) is correlated with q_{in}, as follows:

$$X_{kin} = \gamma_0 + \gamma_{z_1} z_{1in} + \gamma_{z_2} z_{2in} + \gamma_q q_{in} + \varphi_{in} \tag{5}$$

where φ_{in} is an exogenous error term, z_{1in} and z_{2in} are exogenous attributes, which then work as instruments or IV, since they partially explain X_{kin}, but are at the same time independent from ε_{in}. For the model to be identifiable, there must be (at least) as many IV as endogenous variables in the model (Guevara and Ben-Akiva 2012). Following the assumption that q_{in} is a variable not considered by the modeller, a specification that can be set up to treat potential endogeneity would be:

$$X_{kin} = \gamma_0 + \gamma_{z_1} z_{1in} + \gamma_{z_2} z_{2in} + \delta_{in} \tag{6}$$

where the error term δ_{in} contains both φ_{in} and q_{in}. As it is now clear, endogeneity arises because the error terms ε_{in} (4) and δ_{in} (6) are correlated with each other, as q_{in} was not included in the model specification originally proposed by the modeller.

Thus, if (6) is valid, such that z_{1in} and z_{2in} are truly exogenous, then δ_{in} will capture the entire part of X_{kin} that is endogenous. This way, the DCM corrected by endogeneity using the CF approach would have the functional form shown in (7), which implies using a proper estimator of δ_{in}:

$$U_{in} = ASC_i + \hat{\beta}_x X_{in} + \beta_{\hat{\delta}} \hat{\delta}_{in} + \tilde{e}_{in} \tag{7}$$

Thus, in practice the CF method follows two-stages:

(i) to obtain the residuals $\hat{\delta}_{in}$ by applying an ordinal least squares (OLS) regression to X_{kin} on z_{1in}, z_{2in} and all the exogenous variables in X_{in}.
(ii) to estimate the DCM considering $\hat{\delta}_{in}$ and the X_{in} attributes within the utility function,

This allows obtaining consistent estimators $\hat{\beta}_x$ for the β_x in (7) up to a scale (Guevara and Ben-Akiva 2012), but the CF method can also be estimated simultaneously (Train 2009). Theoretically, the two-stage estimation involves a loss of efficiency; however, as Rivers and Vuong (1988) show, this drawback may disappear when the error terms ε_{in} and δ_{in} in (4) and (6) are homoscedastic and not autocorrelated. The other drawback of the two-stage version of the CF method is that standard errors cannot be obtained directly from the information matrix, requiring alternative methods, such as the bootstrap. Nevertheless, as discussed by Guevara (2015), the two-stage version of the CF method is more robust to misspecifications of the distributional assumptions of the model, as well as much easier to apply and requiring fewer computational resources.

It is worth noting that there are other methods to correct for endogeneity in DCM, beyond the CF method. These include, among others, the use of Proxies (Guevara 2015), the Multiple Indicator Solution (Guevara and Polanco 2016; Guevara et al. 2018; Mariel et al. 2018; Fernandez-Antolin et al. 2016), the latent variables approach (Walker 2001) and the BLP method (Berry et al. 1995). Guevara (2015) makes a critical assessment of most of these methods.

Instrumental variables-IV

A fundamental requirement that can turn into a real challenge for applying the CF method is the availability of proper IV. It is achieved if: (i) the IV are correlated with the endogenous variable, and (ii) the IV are independent of the DCM's error. The former is known as *relevance condition* and the second as *exogeneity condition*.

Mumbower et al. (2014) distinguish four possible sources for IV. The first are the *cost-shifting* instruments (Casey 1989), which correspond to variables that impact a product's cost but are uncorrelated with demand shocks. The second are the so-called *Hausman instruments* (Hausman et al. 1994; Hausman 1996), which correspond to prices of the same brand in other geographic contexts. The third are the *Stern instruments* used like measures of the level of market power by multiproduct firms and measures of the level of competition (Stern 1996). Finally, the *BLP instruments* correspond to the average non-price characteristics of other products supplied by the same firm in the same market (Berry et al. 1995).

Tests for the validity of instruments

As mentioned above, a crucial challenge in the correction of endogeneity with the CF resides in finding proper instruments that are sufficiently correlated with the endogenous variable (*strong*) and independent of the error term (*exogenous*).

The strength of an instrument can be assessed by looking at the degree of correlation between the endogenous variable and the instrument, something that has been extensively investigated for linear models, but remains to be fully explored for DCM. Nevertheless, preliminary results suggest that this may be achieved looking at the F test of the first stage regression of the CF method, for which similar thresholds as those reported in linear models seem to be applicable (Guevara and Navarro 2015).

Assessing the exogeneity of instruments is more challenging in some sense, because one needs to test independence with the error term, which is obviously not observed. This requirement may be guessed by the analyst based on his/her understanding of the data generation process but may also be formally tested with overidentification tests that rely on having more instruments than endogenous variables. In the case of linear models, the *Sargan* test (Sargan 1958) is applicable. For DCM, the only test available until recently was the *Amemiya–Lee–Newey test* (Amemiya 1978; Newey 1987; Lee 1992) that requires estimating an auxiliary generalized method of moments (GMM) model, making its application challenging. Guevara (2018) recently proposed two overidentification tests for the exogeneity of the instruments for DCM that are not only easier to apply, but also show better power and less size distortion[1] than the previous tests: the *Refutability Test* (S_{REF}) and its variation, the *Modified Refutability test* (S_{mREF}).

Guevara's (2018) *Refutability Test* (S_{REF}) requires the following two stages:

Stage 1: Estimate the reduced form equation for X_{kin} in (6) by OLS to obtain the residuals $\hat{\delta}$, as shown in (8):

$$X_{kin} = \gamma_{z_1} z_{1in} + \gamma_{z_2} z_{2in} + \delta_{in} \xrightarrow{yields} \hat{\delta}_{in} \tag{8}$$

Stage 2: Estimate the DCM considering $\hat{\delta}_{in}$ and the X_{in} attributes, but also one of the instruments (for example z_{1in}) as an additional variable within the utility function and obtain the log-likelihood $l(\theta)^{CF\_Z1}$, consistent with the utility function shown in (9).

$$U_{in} = ASC_i + \beta'_x X_{in} + \beta_{\hat{\delta}} \hat{\delta}_{in} + \beta_{z_1} z_{1in} + \tilde{\tilde{e}}_{in} \tag{9}$$

Given that in (9) only z_{1in} is used, a log-likelihood $l(\theta)^{CF\_Z1}$ is obtained. The same process must be repeated using z_{2in} as an additional variable within the utility function and obtain a log-likelihood $l(\theta)^{CF\_Z2}$. In this way, two log-likelihood values are computed, by fixing each time all instruments to zero but one (in our case by fixing z_{2in} first, and z_{1in} second).

The second test, S_{mREF}, can also be obtained in two stages. The first is the same as for the S_{REF} test (8); the second stage proceeds as follows:

Stage 2: Estimate the DCM considering the ASC_i, β'_x and $\beta_{\hat{\delta}}$ fixed. Then add all the instruments considered (i.e., z_{1in} and z_{2in}) as additional variables within the utility function and obtain the log-likelihood $l(\theta)^{CF\_Zall}$, consistent with the utility function in (10):

$$U_{in} = ASC_i + \beta'_x X_{in} + \beta_{\hat{\delta}} \hat{\delta}_{in} + \beta_{z_1} z_{1in} + \beta_{z_2} z_{2in} + \tilde{\tilde{\tilde{e}}}_{in} \tag{10}$$

The statistics of the *Refutability Test*—S_{REF} (11) and (12), and its modified version—S_{mREF} (13)—used to test for exogeneity are the following:

$$S_{REF}^{FixingZ1} = -2\left(l(\theta)^{CF} - l(\theta)^{CF\_Z2}\right) \sim \chi_r^2 \tag{11}$$

$$S_{REF}^{FixingZ2} = -2\left(l(\theta)^{CF} - l(\theta)^{CF\_Z1}\right) \sim \chi_r^2 \tag{12}$$

[1] The size distortion corresponds to the difference between the nominal significance of the tests, and the empirical size for the Type I error under the null hypothesis. This type of measure is a standard tool for the assessment of the statistical tests (Guevara 2018).

$$S_{mREF} = -2\left(l(\theta)^{CF} - l(\theta)^{CF\_Zall}\right) \sim \chi_r^2 \tag{13}$$

where $l(\theta)^{CF}$ is the log-likelihood of the corrected model obtained in (7) and χ_r^2 is the value of the chi-squared distribution with degrees of freedom (r) equal to the degrees of overidentification of the model. For the reference tests described in (9) and (10), r is equal to 1 because the model includes one endogenous variable and two instruments (z_{1in} and z_{2in}). The null hypothesis for the S_{REF} and S_{mREF} tests is that both z_{1in} and z_{2in} are valid; the alternative hypothesis is that either z_{1in} and z_{2in}, or both, are endogenous. Thereby, if S_{REF} and S_{mREF} are less than the critical value of χ_r^2 at the required level of significance, the instruments are exogenous and, therefore, they are independent of the DCM's error.

It should be noted that overidentification tests for the exogeneity of the instruments are inconsistent, in the sense that there are null hypotheses for which the tests have no power. This means that there might be cases where the instruments are endogenous and these tests are unable to detect that failure, even if the sample size goes to infinity. Nevertheless, it has been shown that the hypotheses for which overidentification tests of this type are inconsistent, are very peculiar and can be narrowed to cases where both instruments are of the same origin, if they come from the same source. This is something we tried to avoid in the present application. The reader is referred to Guevara (2018, p. 242) for a review and discussion about this topic.

Subjective value of time (SVT) and elasticities

We will estimate the SVT and aggregate elasticities to quantify the impacts of neglecting the problem of endogeneity in the estimation of the parameters of a strategic urban modal choice model. As the representative utility function in most classical models is assumed to be linear and additive in the (fixed) marginal utility parameters, the SVT (Gaudry et al. 1989) usually corresponds to just the ratio between the estimated parameters for travel time β_t and for travel cost β_c, yielding (14):

$$SVT = \frac{\partial V_i / \partial t_i}{\partial V_i / \partial c_i} = \frac{\beta_t}{\beta_c} \tag{14}$$

The aggregate elasticities ($E_{kin}^{\tilde{P}i}$) can be calculated as in (15):

$$E_{in}^{\tilde{P}i} = \frac{\sum_n P_n(i) E_{X_{in}}^{P_n(i)}}{\sum_n P_n(i)} \tag{15}$$

where $E_{X_{in}}^{P_n(i)}$ is the disaggregate direct point elasticity with respect to variable X_{in}, and $P_n(i)$ the probability that individual n chooses alternative i (Ben-Akiva and Lerman 1985).

Application

Great Valparaíso case study

The Great Valparaíso is a conurbation located in the Valparaiso Region of Chile, encompassing the municipalities of Valparaíso, Viña del Mar, Concón, Quilpué and Villa Alemana, an area of some 1130 km^2 (SECTRA 2014a). According to the National Statistics

Institute (INE 2013), it is the third most populated area in the country, after Great Santiago and Great Concepción, but given its strategic location and proximity to the capital, it is the second in importance.

The databank comes from the Great Valparaiso 2014 Origin–Destination Survey and was used by SECTRA[2] (2014a) to estimate—among other things—the DCM embedded in the mode choice stage of ESTRAVAL, the strategic transport model for the Great Valparaiso. ESTRAVAL is a simultaneous supply–demand equilibrium model designed to analyse and evaluate multimodal urban transport systems with multiple user classes (De Cea et al. 2005). This type of approach is also used in packages such as EMME/2 (INRO 1996) or CUBE (Citilabs 2016).

The aim of our research was not to change the model used in ESTRAVAL; we just wanted to examine the consequences of correcting it for endogeneity. The model contemplates seven transport modes: Car driver, Shared car, Bus, Train, Shared taxi, Walking and the combined mode Train/Bus. The survey considered three trip purposes: Work, Study and Other, but in the framework of this research we only considered the correction of the work trips mode choice model for the morning peak period. The sample available for this purpose comprised 2417 observations. A general descriptive statistical analysis of the sample shows that 36% of trips used private modes (Car driver and Shared car), while the most used public transport mode was Bus with 26% of the market shares.

Instrumental variables (IV) used for endogeneity correction

The instruments used to estimate the first stage of the CF method were built resembling what Mumbower et al. (2014) denominate *Hausman type* instruments, that is, values of the endogenous variable in "other markets", that may share marginal costs, but are independent regarding demand shocks. For this mode choice model, we suspect the existence of endogeneity both in travel time and travel cost, and for this reason we propose the following three instruments:

(i) The average travel time of other origin–destination (O–D) pairs with similar length to the O–D pair of the considered trip (*IV_GT*).
(ii) The average travel cost of other O–D pairs with similar length to the O–D pair of the considered trip (*IV_C*).
(iii) The network trip distance between the trip's origin and destination (*IV_D*).

As can be seen, each of these instruments should be correlated with the endogenous variables (cost, time or both) but they do not influence the individuals' choice, being then independent regarding demand shocks. If both properties are fulfilled, the instruments are valid to correct appropriately for endogeneity using the CF approach (Rivers and Vuong 1988; Villas-Boas and Winer 1999).

To verify that the proposed instruments fulfil the *relevance* condition, we considered the results described by Staiger and Stock (1997), which had been preliminary suggested to hold also for DCM by Guevara and Navarro (2015). In this case, if the value of the first stage's F-statistic is less than 10, the instrument is *weak* (i.e., it does not satisfactorily fulfil

[2] SECTRA is the Chilean governmental agency for transport planning and policy formulation.

the condition). However, it should be noted that as this result formally holds only for linear models, this is a limitation of this research, which we intend to explore in the future.

To guarantee the exogeneity of the instruments, we considered using the information of a geographical context different from the Great Valparaiso (i.e., Hausman-type instruments), an approach which has been successfully used in several studies (Mumbower et al. 2014; Guevara and Ben-Akiva 2006; Petrin and Train 2010). In this case, the other geographical context data was the Santiago 2012 Origin–Destination Survey (SECTRA 2014b) also in Chile. The procedure applied to find the instruments considered the zoning system used by SECTRA in their model for Santiago.

In this way, *IV_GT* and *IV_C* were calculated as the average travel time and travel cost, respectively, for every zone included in a band defined by a lower bound of ± 100 m and an upper bound of ± 2.1 km with respect to the distance of the O–D pair under consideration. For example, consider a distance (Euclidean distance, measured between centroids for the given O–D pair) of 5 km; in this case, the lower and upper bounds defined would be as follows: [2.9–4.9 km] and [5.1–7.1 km]. Thereby, any O–D pair, the distance of which is inside any of these two bands, would be part of the average for *IV_GT* or *IV_C*. The argument to sustain the suitability of such instruments is equivalent to that used by Guevara and Ben-Akiva (2012), Hausman (1996) and Nevo (2001) in other modelling contexts.

The lower bound (100 m) guarantees that the O–D pair under consideration is not included because otherwise endogeneity would arise. On the other hand, the upper bound (2.1 km) ensures that every O–D pair has enough data to estimate an average. In this way, we make sure that every O–D pair has a set of O–D pairs inside the bands defined. This fact makes them share marginal travel costs (or travel times) and, therefore, their travel costs (or travel times) are correlated.

Finally, the third instrument used is the *IV_D*, known in the literature as a *cost-shifting* instrument (Casey 1989). *IV_D* was calculated directly from the network defined for ESTRAVAL, thus, from the city of Valparaíso. Instruments of a similar nature (route distance) have also been used successfully for the case of air transport (Hsiao 2008; Granados et al. 2012). We argue that *IV_D* is correlated with the travel time and travel cost, but independent of the error term of the mode choice.

It should be noted that any of the O–D pairs used to build the instruments (*IV_GT* and *IV_C*) could (or not) be overlapping among them. However, this is not an issue because the instruments were constructed as the attributes' average of the O–D pairs that were part of the bands defined above. What may instead be critical in general is that none of the O–D pairs used to build the instruments, overlapped with the O–D pair under analysis (i.e., the incumbent O–D pair for which we needed to address endogeneity). This is not necessarily an issue for *IV_GT* and *IV_C* in this case study, since they come from a different city, but it was nevertheless further enforced by defining the band's lower bound different from zero (100 m) to avoid endogeneity arising due to reflection bias. Regarding the *IV_D* instrument, the overlapping is also possible, but it did not affect the instrument estimation because it only depends on the route determined by the network topology used in ESTRAVAL.

Correction of endogeneity in the strategic urban mode choice models

We assumed that endogeneity affects the travel cost and travel time variables in the Great Valparaíso urban mode choice model because of, as mentioned earlier, the potential erroneous measurement of the relevant variables included in model, the omission of potential relevant variables (such as comfort or reliability) and the fact that the model is embedded in a simultaneous supply–demand equilibrium mechanism. Our hypothesis is that the measurement error due to aggregation may affect both travel time and travel cost. And the omission of attributes and the simultaneity issue may affect travel time further.

Table 1 presents the endogenous and corrected mode choice models estimated with the Great Valparaiso dataset. The left-hand side model (potentially endogenous) is the model currently used by SECTRA in ESTRAVAL. This is the model that we want to correct for endogeneity. It was estimated by SECTRA for two morning peak periods (AM1 and AM2), so 14 (seven modes by two periods) alternative specific constants (ASC) were estimated, fixing one ($ASC_{Walking1}$) to zero, as reference. The parameter $\beta_{Cost/Income}$ corresponds to the marginal utility of the variable Cost divided by Income. The model also includes three different parameters for Generalised Time (i.e. the sum of travel time, access time and waiting time): $\beta_{Generalised\ Time\ Car}$, $\beta_{Generalised\ Time\ Walking}$ and $\beta_{Generalised\ Time\ Public\ transport}$, correspond to the marginal utilities of the private modes (Car driver and Shared car), walk mode

Table 1 Endogenous and corrected mode choice models for Great Valparaíso

Variable	Endogenous model			Corrected model		
	Value	Std. error	t-test	Value	Std. error[a]	t-test
$ASC_{Car\ driver1}$	− 0.425	0.329	− 1.29	− 0.247	0.337	− 0.73
$ASC_{Shared\ car1}$	− 2.980	0.354	− 8.43	− 3.199	0.352	− 9.09
ASC_{Bus1}	− 1.140	0.341	− 3.35	− 1.035	0.331	− 3.13
ASC_{Train1}	− 2.780	0.682	− 4.07	0.839	0.448	1.87
$ASC_{Shared\ taxi1}$	− 2.180	0.356	− 6.12	− 1.705	0.349	− 4.89
$ASC_{Walking1}$	fixed			fixed		
$ASC_{Train+Bus1}$	− 4.930	0.572	− 8.62	− 3.150	1.214	− 2.59
$ASC_{Car\ driver2}$	0.826	0.322	2.56	1.002	0.328	3.05
$ASC_{Shared\ car2}$	− 1.670	0.328	− 5.09	− 1.878	0.334	− 5.62
ASC_{Bus2}	− 0.659	0.343	− 1.92	− 0.539	0.336	− 1.60
ASC_{Train2}	− 1.580	0.47	− 3.37	2.014	0.416	4.84
$ASC_{Shared\ taxi2}$	− 1.590	0.345	− 4.60	− 1.110	0.344	− 3.23
$ASC_{Walking2}$	1.200	0.168	7.15	1.190	0.174	6.84
$ASC_{Train+Bus2}$	− 4.030	0.460	− 8.74	− 2.237	0.489	− 4.57
$\beta_{Cost/Income}$	− 0.015	0.003	− 5.10	− 0.026	0.0067	− 3.94
$\beta_{Generalised\ Time\ Car}$	− 0.031	0.0075	− 4.14	− 0.0308	0.0084	− 3.80
$\beta_{Generalised\ Time\ Walking}$	− 0.113	0.011	− 10.35	− 0.115	0.011	− 10.45
$\beta_{Generalised\ Time\ Public\ transport}$	− 0.0075	0.0017	− 4.44	− 0.0097	0.0017	− 5.71
$\beta_{Distance\ travel\ ST1}$	− 1.560	0.475	− 3.28	− 1.952	0.479	− 4.08
$\beta_{Distance\ travel\ TTB1}$	2.000	0.645	3.10	− 1.568	0.711	− 2.21
$\beta_{Distance\ travel\ ST2}$	− 1.830	0.376	− 4.87	− 2.241	0.299	− 7.49
$\beta_{Distance\ travel\ TTB2}$	1.660	0.382	4.34	− 1.888	0.457	− 4.13
$\beta_{\hat{\delta}CI}$				0.014	0.007	2.00
$\beta_{\hat{\delta}GT}$				0.002	0.0003	6.67
Sample size	2417			2417		
Log-likelihood	− 3266.32			− 3258.79		

[a]Standard errors determined using Bootstrap

(Walking) and public transport modes (Bus, Train, Shared taxi and Train + Bus), respectively. Finally, $\beta_{Distance\ travel\ ST1}$, $\beta_{Distance\ travel\ TTB1}$, $\beta_{Distance\ travel\ ST2}$ and $\beta_{Distance\ travel\ TTB2}$ are parameters associated with dummy variables, which take the value of 1 for Shared taxi ($\beta_{Distance\ travel\ ST1}$ and $\beta_{Distance\ travel\ ST2}$ for the periods AM1 and AM2, respectively), and Train and Train + Bus ($\beta_{Distance\ travel\ TTB1}$ and $\beta_{Distance\ travel\ TTB2}$ for the same periods) if the trip had a distance greater than 10 km. All the level-of-service parameters of the potentially endogenous model in Table 1 have correct signs and are statistically significant at the 95% level. We note that trips with distances greater than 10 km are preferred by Train and Train + Bus users.

The right-hand side of Table 1 shows the model corrected for endogeneity. This includes the parameters $\beta_{\hat{\delta}GT}$ and $\beta_{\hat{\delta}CI}$ (residuals from the first stage of the CF approach) related to the variables Generalised Time and Cost/Income, respectively. The inclusion of these two parameters is required because of our initial hypothesis that the uncorrected model is endogenous in Cost/Income and in the Generalised Times. The verification of this hypothesis is carried out following Rivers and Vuong (1988); so if $\beta_{\hat{\delta}GT}$ and $\beta_{\hat{\delta}CI}$ are significant in the second stage of the CF approach, then there is evidence that the model is endogenous in the variables related with these residuals. As can be seen from the right-hand side model, both $\beta_{\hat{\delta}GT}$ and $\beta_{\hat{\delta}CI}$ are significant.

One practical aspect of the application of the CF method to this case study,[3] worth highlighting, is that the first stage of the CF method was estimated by mode, instead of stacking the information from all available alternatives, as has been done in other cases. This approach was followed because the Shared Car and Walking alternatives have a travel cost of zero in this application, and this would preclude proper estimation of the residuals of a stacked first stage via an OLS if the dependent variable is zero. Nevertheless, the same coefficient for the residual was considered for all modes, as shown in Table 1. An extensive Monte Carlo simulation validated this approach for the practical problem of modes with zero cost.

The validity of the instruments was verified using the overidentification tests for the exogeneity of the instruments in DCM proposed by Guevara (2018). In this case, $l(\theta)^{CF}$ is obtained directly from the model in Table 1 and $l(\theta)^{CF\_Z}$ was obtained by fixing one of the instruments to zero (for example *IV_GT*) in each case and including the other two instruments (i.e. *IV_C* and *IV_D*) as additional variables within the utility function.

The degree of overidentification for this test is equal to one because the model includes two endogenous variables (Travel Cost and Generalised Time) and three instruments (*IV_GT*, *IV_C* and *IV_D*). It is worth noting that, although the model includes three parameters for Generalised Time, differentiated by mode, the variable is the same. Therefore, it is better to consider it as just one parameter when analysing the degrees of freedom for testing instrument validity. Another alternative, technically also valid, would be to consider it as three different variables, but that would be misleadingly much laxer. Indeed, if we considered Generalised Time as three variables, we should do the same with the respective instrument and, therefore, we would have $1+3=4$ endogenous variables and $1+6+1=8$

[3] An additional issue that did not come out in this application, but may be relevant for other cases, is what to do when the endogenous variable interacts with exogenous variables, such as level of income or gender. Bun and Harrison (2018) formally show that, under such circumstances, the endogeneity bias will reduce to zero for the ordinary least squares estimator, as far as the interaction term is concerned. The same holds for the Control Function method in discrete choices, something that has been implicitly used, among others, by Petrin and Train (2010) and Guevara and Ben-Akiva (2006). We thank an anonymous reviewer for having asked this question.

instruments. Such an approach would lead to four degrees of freedom instead of only one, implying a much laxer critical value of 9.49 instead of 3.84 for the tests shown in (16)–(18).

The results of the S_{REF} in (16), (17) and (18) show that in all cases $S_{REF} < \chi_1^2$ (3.84); so, we can conclude that all our instruments are indeed exogenous:

$$S_{REF}^{FixingIV\_GT} = -2(-3258.79 + 3258.14) = 1.31 < 3.84 \tag{16}$$

$$S_{REF}^{FixingIV\_C} = -2(-3258.79 + 3258.51) = 2.57 < 3.84 \tag{17}$$

$$S_{REF}^{FixingIV\_D} = -2(-3258.79 + 3257.99) = 1.59 < 3.84 \tag{18}$$

To apply the S_{mREF}, we considered *IV_GT*, *IV_C*, and *IV_D* as additional variables within the DCM, fixed each of the β parameters of the right-side model of Table 1 and obtained the log-likelihood $l(\theta)^{CF\_Zall}$ (−3258.27). This gives, $S_{mREF} = -2 * (-3258.79 + 3258.27) = 1.05$, less than the critical $\chi_1^2 = 3.84$ value; therefore, we can conclude that all our instruments are valid. To the best of our knowledge, these instruments had not been suggested before to correct for endogeneity in modelling urban mode choice.

The corrected model in Table 1 also has parameters for the level-of-service variables with correct signs and statistically significant (at 95% level). An interesting fact is the change of sign in the parameters $\beta_{Distance\ travel\ TTB1}$ and $\beta_{Distance\ travel\ TTB2}$, which suggests that trips over 10 km are actually not preferred to be made by Train and Train + Bus, contrary to the potentially endogenous model. It is also interesting to note that the parameter $\beta_{Cost/Income}$ in the corrected model is 73% higher than the one estimated in the endogenous model.

On the other hand, although the parameters $\beta_{Generalised\ Time\ Car}$ and $\beta_{Generalised\ Time\ Walking}$ are similar in both models, suggesting low bias (less 1%) in their estimation, the parameter $\beta_{Generalised\ Time\ Public\ transport}$ is 30% higher in the corrected model. Thus, the percentage differences between the generalised time parameters for the endogenous and corrected models are smaller in comparison with those of the cost parameter. This result suggests that the cost parameter appears to be more vulnerable to endogeneity than the time parameters and, thereby, it was more poorly estimated in the original model by SECTRA. This finding is in line with that shown by Varela et al. (2018). It also suggests that problems such as measurement errors, perception errors and omitted variables, affect more the cost parameter than the time parameters. In practice, then, efforts should focus on improving the way the cost variable is collected and measured in our surveys, to achieve more consistent parameters during model estimation. In particular, the bias in the parameter $\beta_{Generalised\ Time\ Public\ transport}$ can also be due to omitted variables that explain mode choice. If attributes like comfort and reliability, often correlated with travel cost and time, are excluded from the mode choice model, this is a potential source for endogeneity and, as a result, the SVT is overestimated (Tirachini et al. 2013). In any case, given the aims of this research we cannot ascertain how much endogeneity is due to some of the sources described previously. It is an interesting research question that is left for future research.

Now, given that the endogenous model is a restricted version of the corrected model, it is possible to apply the likelihood ratio (LR) test (Ortúzar and Willumsen 2011, page 281) to investigate the presence of endogeneity.[4] The null hypothesis, in this case, is that there is no endogeneity. Then both models are equivalent, rejecting it, which implies that the restricted model is erroneous and then endogeneity is present. 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, which in this case corresponds to the number of residuals incorporated into the corrected model.

In our case, the degrees of freedom of the LR test are $r=2$ (because the restrictions are that both $\beta_{\hat{\delta}GT}$ and $\beta_{\hat{\delta}CI}$ are zero). So, $LR = -2(-3266.32 + 3258.79) = 15.07$, and this value must be compared with the critical value for two degrees of freedom at the 95% level ($\chi_2^2 = 5.99$). As LR > χ_2^2 the null hypothesis is confidently rejected, and we can conclude that the corrected model is superior.

It must be noted that CF approach tends to yield variances of the estimators that are often larger than those of the *true* model and usually also larger than those of the endogenous model; therefore, its confidence intervals could be wider. Thus, although the correction may be relatively poorer in this regard (at least in some cases), what is crucial is that the estimators will be consistent with the CF correction. Neglecting it may even result in reversing the effect of the attributes due to a change of sign. In the case study analysed, there was no change of sign. Still, the difference in point estimates were as large as 43% in some cases (see Fig. 1), implying that even if what one cares about is the MSE and not the finite sample bias, the CF results would be preferred. This recommendation is reinforced by the fact that, in strategic transport models the point estimate (i.e., the mean of the estimator distribution) is used for forecasting. For this reason, any bias on the base year values would be exacerbated in future simulations, resulting in poor model forecasting performance.

Quantification of effects due to endogeneity

In this subsection we quantify the impacts of endogeneity in the model. The measures used for this are SVT and aggregate Elasticities. These were calculated for the endogenous and corrected model and later compared. Detailed results are provided in Tables 2, 3 and 4. To estimate each of these measures, we divided the dataset into two samples: 80% for estimation and 20% for validation (holdout sample). Additionally, the process was repeated 100 times (i.e. 100 repetitions), with the aim of guaranteeing randomness in the estimates.

It should be noted that the estimation of the standard errors when applying the two-stage CF method comes with a caveat. Given that the proposed estimator is estimated in two stages, variances cannot be calculated directly from the Fisher-information-matrix. Therefore, to make the inference, the variance–covariance matrix must be determined using non-parametric methods such as bootstrapping (Petrin and Train 2002) or the approach proposed by Karaca-Mandic and Train (2003), or by writing, instead, the full likelihood of both stages together (Train 2003).

In this application, we used the bootstrap approach to calculate the standard errors and the confidence intervals for the estimators reported in Table 1 and for the VST and

[4] Following Rivers and Vuong (1988), note that when using a two-step procedure, the test for the presence of endogeneity does not require correcting the standard errors with bootstrap. This holds because the test is evaluated under the null hypothesis that there is no endogeneity. Therefore, the population coefficient of the residuals is zero. This logic holds for Wald, Lagrange Multiplier and LR tests, when used to evaluate the presence of endogeneity, which is what we use in this section (see, for example, the discussion in Guevara 2010, Ch. 2).

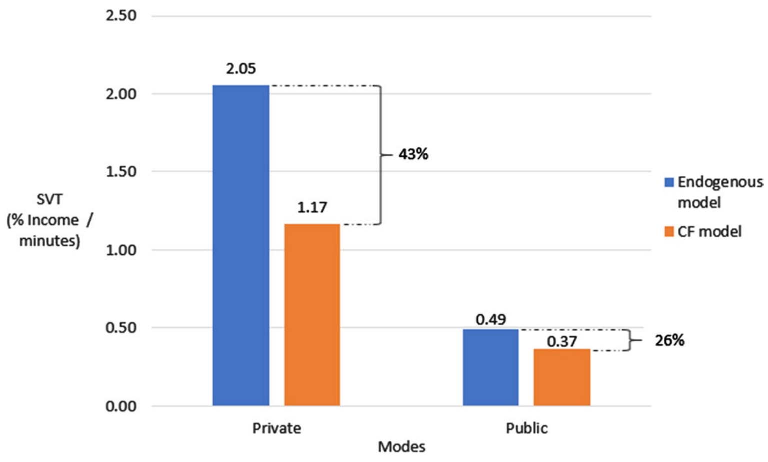

Fig. 1 SVT for private and public modes

elasticities reported in Tables 2, 3 and 4. Confidence intervals for both VST and elasticities were estimated using the percentile bootstrap method (Davison and Hinkley 1997). This approach considers using the percentiles of the bootstrap distribution directly (in our case 2.5% and 97.5%), to represent a confidence interval at the 5% significance level.

Given that the cost variable is really Cost/Income, then SVT is expressed as [% Income/min]. It was possible to estimate it separately for the private and public transport modes, given that the Generalised Time parameters were specific for these modes (see Fig. 1). As can be seen, the SVT of the original model was overestimated in comparison with that obtained for the corrected model. For the private modes, the SVT suffers an overestimation of up to 43%, while in the case of the public transport modes this reaches 26%. These findings are in line with those shown by Varela et al. (2018), who used a case study for Stockholm commuters to assess the magnitude of the measurement errors in travel time and travel cost using latent variables. These differences are important, because measures such as SVT are critical in the social evaluation of transport projects. Given that the bias of the cost parameter is higher than the bias of the time parameters, it makes sense that the SVT estimates are overestimated. If the SVT is biased, the social evaluation of the project will likely be biased too.

Elasticities are frequently used in transport project evaluation (Ortúzar and Willumsen 2011). In the case of the elasticities of the Generalised Times (Fig. 2) and Cost/Income (Fig. 3) for the original model, we can see that these are underestimated in comparison with those obtained using the corrected model. This finding is also consistent with the results of Varela et al. (2018) and Varotto et al. (2017), who observed increases of up to 65% in the time elasticity value and of up to 50% of the price elasticity, when assessing the

Table 2 Mean and confidence intervals for the VST

Model	Private	Public
Endogenous	2.05	0.49
	(1.02–4.19)	(0.31–0.98)
Corrected	1.17	0.37
	(0.18–2.16)	(0.17–0.63)

Confidence interval (in parenthesis)

Table 3 Mean and confidence intervals for the generalised time elasticities

Model\mode	Car driver1	Shared car1	Bus1	Train1	Shared taxi1	Walking1	Train + Bus1
Endogenous	− 0.347	− 0.413	− 0.504	− 0.751	− 0.523	− 1.96	− 0.807
	(− 0.533 to − 0.211)	(− 0.632 to − 0.257)	(− 0.797 to − 0.358)	(− 1.187 to − 0.532)	(− 0.858 to − 0.384)	(− 2.258 to − 1.692)	(− 1.266 to − 0.559)
Corrected	− 0.346	− 0.436	− 0.657	− 0.974	− 0.665	− 1.998	− 1.074
	(− 0.422 to − 0.063)	(− 0.517 to − 0.088)	(− 0.781 to − 0.339)	(− 1.164 to − 0.502)	(− 0.846 to − 0.368)	(− 2.248 to − 1.681)	(− 1.26 to − 0.55)
Model\mode	Car driver2	Shared car2	Bus2	Train2	Shared taxi2	Walking2	Train + Bus2
Endogenous	− 0.268	− 0.468	− 0.451	− 0.555	− 0.48	− 1.336	− 0.771
	(− 0.408 to − 0.162)	(− 0.692 to − 0.295)	(− 0.72 to − 0.317)	(− 0.934 to − 0.414)	(− 0.836 to − 0.371)	(− 1.502 to − 1.134)	(− 1.218 to − 0.537)
Corrected	− 0.269	− 0.494	− 0.588	− 0.722	− 0.604	− 1.365	− 1.029
	(− 0.325 to − 0.053)	(− 0.568 to − 0.101)	(− 0.708 to − 0.302)	(− 0.853 to − 0.368)	(− 0.823 to − 0.354)	(− 1.492 to − 1.128)	(− 1.212 to − 0.528)

Confidence interval (in parenthesis)

Table 4 Mean and confidence intervals for the cost/income elasticities

Model\mode	Car driver1	Bus1	Train1	Shared taxi1	Train + Bus1
Endogenous	− 0.318	− 0.082	− 0.089	− 0.145	− 0.111
	(− 0.422 to − 0.207)	(− 0.114 to − 0.051)	(− 0.118 to − 0.052)	(− 0.213 to − 0.097)	(− 0.153 to − 0.066)
Corrected	− 0.54	− 0.139	− 0.15	− 0.241	− 0.192
	(− 0.821 to − 0.304)	(− 0.216 to − 0.076)	(− 0.232 to − 0.084)	(− 0.411 to − 0.143)	(− 0.293 to − 0.102)
Model\mode	Car driver2	Bus2	Train2	Shared taxi2	Train + Bus2
Endogenous	− 0.196	− 0.066	− 0.071	− 0.133	− 0.112
	(− 0.261 to − 0.125)	(− 0.091 to − 0.042)	(− 0.1 to − 0.045)	(− 0.202 to − 0.092)	(− 0.159 to − 0.068)
Corrected	− 0.333	− 0.112	− 0.12	− 0.22	− 0.197
	(− 0.516 to − 0.189)	(− 0.174 to − 0.063)	(− 0.19 to − 0.066)	(− 0.388 to − 0.136)	(− 0.304 to − 0.105)

Confidence interval (in parenthesis)

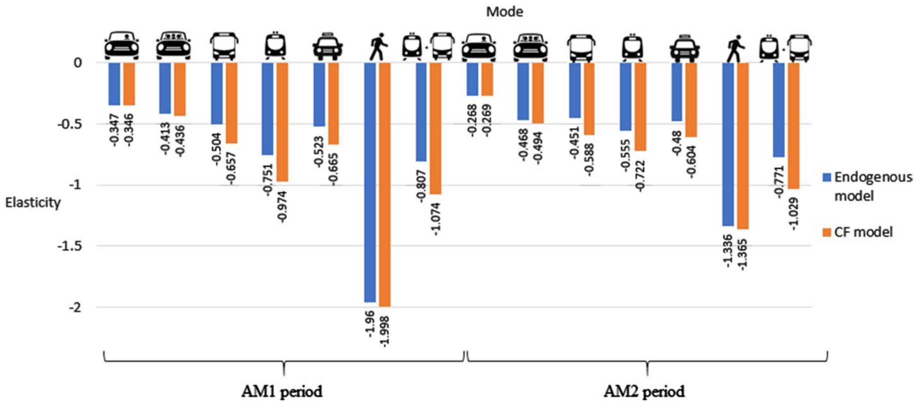

Fig. 2 Generalised time elasticities

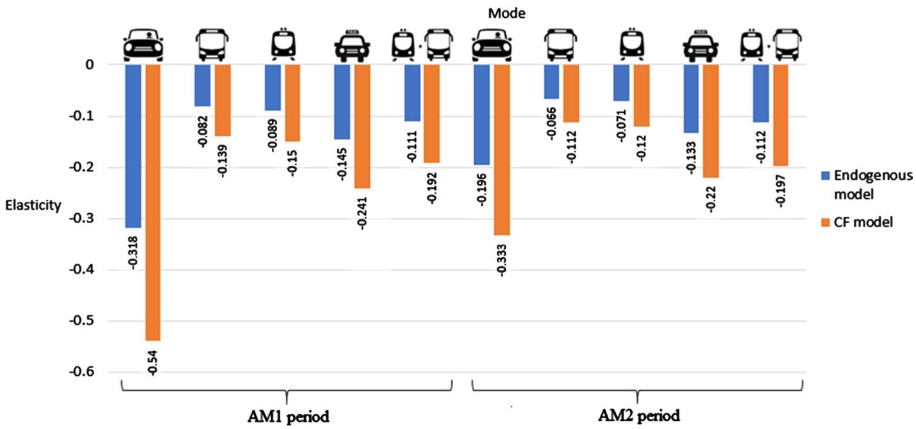

Fig. 3 Cost/income elasticities

magnitude of the measurement errors in these variables (using latent variables), in a large-scale travel demand model.

In our case, the Generalised Time elasticities are underestimated up to 33%, while the Cost/Income elasticities are underestimated up to 75%. The mode with the highest generalised time elasticity is Walking (in both periods). On the other hand, the smallest generalised time elasticity is registered for Car driver (in both periods), but the generalised time elasticities for both private modes (Car driver and Shared car) show no differences between the endogenous and the corrected model.

Cost/Income elasticities for the AM1 period (− 0.54) are higher in the Car driver mode than in the public transport modes (where they vary between −0.139 and − 0.241). These results are also consistent with findings from other studies (Varela et al. 2018). Given that the parameter $\beta_{Cost/Income}$ and $\beta_{Generalised\ Time\ Public\ transport}$ were underestimated in the endogenous model, the underestimation of elasticities was expected. Note that the Generalised time elasticities calculated for both models and for the modes Car driver, Shared car and Walking, are also similar because both parameters $\beta_{Generalised\ Time\ Car}$ and $\beta_{Generalised\ Time\ Walking}$ have a rather low bias.

Conclusions

Endogeneity is an anomaly that also arises in urban mode choice models at the strategic level. It affects the consistency of the model parameters estimated, especially those related to the travel cost and travel time variables. As these are key explanatory variables in strategic mode choice models, not correcting the endogenous models may lead to faulty decision-making.

This paper provides a framework that uses the CF method to correct for the endogeneity of mode choice models at the strategic level using appropriate instrumental variables. The CF method can be considered an adequate methodology in this case. The instruments used were: (i) The average travel time of other origin–destination pairs that have a similar length than the origin and destination of the considered trip; (ii) the average travel cost of other origin–destination pairs that have a similar length than the origin and destination of the considered trip, and (iii) the network trip distance between the origin and the destination for each mode. Defining these instruments is a relevant finding, and they can be considered valid. Government planning agencies (central or local) should begin to consider the CF approach and the instruments used in this research as a guide to correct mode choice models that may present endogeneity.

The confidence in strategic urban mode choice models based on level-of-service variables, such as travel cost and travel time, must be questioned. Our results show that the cost parameters could be more poorly estimated than the time parameters. This may be due to the fact that urban mode choice models at the strategic level may be affected by three sources of endogeneity: measurement errors, omitted variables and simultaneous estimation. We recommended: (i) to use instruments within the framework shown in this paper to improve the estimations, and (ii) to focus the efforts in improving the way the cost variable is collected and measured in surveys, to achieve more consistent parameters during model estimation.

We quantified the effects of endogeneity in strategic urban mode choice models. We found that the SVT was overestimated by 43% and 26% for private and public modes, respectively in our case study. This fact may have a strong influence in the social evaluation of transport projects where the SVT is critical. We also showed the impact on model elasticities, finding that these were underestimated. In particular, the Generalised Time elasticities showed underestimations of up to 33%, while the Cost/Income elasticities reached underestimations of up to 75%.

Three areas for further research can be identified. First, we believe it is important to study how correcting for endogeneity would work in forecasting when the variables that change are endogenous, such as travel times and cost in a strategic transport planning model. We also recommend examining in greater depth how the social evaluation of transport projects may be affected by endogeneity, especially given our findings regarding the changes in SVT. Finally, an exciting topic for further research is the identification of weak and strong instruments for correcting endogeneity, because this has been solved for linear models (Stock and Yogo 2005) but not yet fully extended for DCM.

Acknowledgements The authors would like to thank SECTRA (Secretaría de Planificación de Transporte de Chile) and, in particular Alan Thomas, for the provision of the real data used in the study. Preliminary work for this research was presented at the *6th International Choice Modelling Conference*, Kobe, Japan. This research was partially funded by 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. Finally, we wish to acknowledge the relevant and insightful comments of three unknown referees, which helped us to improve the paper considerably. Of course, any remaining errors are our fault.

Author contributions The authors confirm contributions to the paper as follows: study conception and design: TEG, CÁG and JDO; analysis and interpretation of results: TEG, CÁG, JDO and EC; draft manuscript preparation: TEG, CÁG, JDO and EC. All authors reviewed the results and approved the final version of the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Amemiya, T.: The estimation of a simultaneous equation generalized probit model. Econometrica **46**, 1193–1205 (1978)
Anderson, J., Hernandez, S.: Roadway classifications and the accident injury severities of heavy-vehicle drivers. Anal. Methods Accid. Res. **15**, 17–28 (2017)
Bass, F.: A simultaneous equation regression study of advertising and sales of cigarettes. J. Mark. Res. **6**, 291–300 (1969)
Bass, P., Donoso, P., Munizaga, M.: A model to assess public transport demand stability. Transp. Res. Part A Policy Pract. **45**, 755–764 (2011)
Ben-Akiva, M.E., Lerman, S.R.: Discrete Choice Analysis: Theory and Application to Travel Demand. The MIT Press, Cambridge (1985)
Berry, S., Levinsohn, J., Pakes, A.: Automobile prices in market equilibrium. Econometrica **63**, 841–890 (1995)
Bresnahan, T.F.: The apple–cinnamon cheerios war: valuing new goods, identifying market power, and economic measurement. In: Bresnahan, T.F., Gordon, R.J. (eds.) The Economics of New Goods. NBER Studies in Income and Wealth Number, vol. 58. The University of Chicago Press, Chicago (1997)
Bun, M.J., Harrison, T.D.: OLS and IV estimation of regression models including endogenous interaction terms. Econ. Rev. **38**, 814–827 (2018)
Casey, B.: Early retirement: the problems of "instrument substitution" and "cost shifting" and their implications for restructuring the process of retirement. In: Schmähl, W. (ed.) Redefining the Process of Retirement. Springer, Berlin (1989)
Chou, C., Chen, P.: Preferences and willingness to pay for green hotel attributes in tourist choice behaviour: the case of Taiwan. J. Travel Tour. Mark. **31**, 937–957 (2014)
Citilabs: Discover Cube 6.4 Tutorial. Citilabs, Inc., Sacramento (2016)
Davison, A.C., Hinkley, D.V.: Bootstrap Methods and Their Application, vol. 1. Cambridge University Press, Cambridge (1997)
De Cea, J., Fernández, J.E., Dekock, V., Soto, A.: Solving network equilibrium problems on multimodal urban transportation networks with multiple user classes. Transp. Rev. **35**, 293–317 (2005)
Domencich, T., McFadden, D.: Urban Travel Demand: A Behavioural Analysis. North Holland, Amsterdam (1975)
Fernández-Antolín, A., Guevara, C.A., De Lapparent, M., Bierlaire, M.: Correcting for endogeneity due to omitted attitudes: empirical assessment of a modified MIS method using RP mode choice data. J. Choice Model. **20**, 1–15 (2016)
Gaudry, M.J., Jara-Diaz, S.R., Ortúzar, J. de D.: Value of time sensitivity to model specification. Transp. Res. Part B Methodol. **23**, 151–158 (1989)
Granados, N., Gupta, A., Kauffman, R.J.: Online and offline demand and price elasticities: evidence from the air travel industry. Inf. Syst. Res. **23**, 164–181 (2012)
Guevara, C.A.: Endogeneity and sampling of alternatives in spatial choice models. Ph. D. Thesis, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology (2010)

Guevara, C.A.: Critical assessment of five methods to correct for endogeneity in discrete-choice models. Transp. Res. Part A Policy Pract. **82**, 240–254 (2015)
Guevara, C.A.: Overidentification tests for the exogeneity of instruments in discrete choice models. Transp. Res. Part B Methodol. **114**, 241–253 (2018)
Guevara, C.A., Ben-Akiva, M.E.: Endogeneity in residential location choice models. Transp. Res. Rec. **1977**, 60–66 (2006)
Guevara, C.A., Ben-Akiva, M.E.: Addressing endogeneity in discrete choice models: assessing control-function and latent-variable methods. In: Hess, S., Daly, A. (eds.) Choice Modelling: The State-of-the-Art and The State-of-Practice. Edward Elgar publishers, Cheltenham (2010)
Guevara, C.A., Ben-Akiva, M.E.: Change of scale and forecasting with the control-function method in logit models. Transp. Sci. **46**, 425–437 (2012)
Guevara, C.A., Navarro, P.: Detection of weak instruments when correcting for endogeneity in binary logit models. In: Proceedings 14th International Conference on Travel Behaviour Research (IATBR), Windsor, UK, July 19–23 (2015)
Guevara, C.A., Polanco, D.: Correcting for endogeneity due to omitted attributes in discrete-choice models: the multiple indicator solution. Transp. A Transp. Sci. **12**, 458–478 (2016)
Guevara, C.A., Tirachini, A., Hurtubia, R., Dekker, T.: Correcting for endogeneity due to omitted crowding in public transport choice using the Multiple Indicator Solution (MIS) method. Transp. Res. Part A Policy Pract. (2018). https://doi.org/10.1016/j.tra.2018.10.030
Hausman, J.A.: Specification tests in econometrics. Econometrica **46**, 1251–1271 (1978)
Hausman, J.A.: Valuation of new goods under perfect and imperfect competition. In: Bresnahan, T.F., Gordon, R.J. (eds.) The Economics of New Goods. NBER Studies in Income and Wealth Number, vol. 58. The University of Chicago Press, Chicago (1996)
Hausman, J.A., Leonard, G., Zona, J.D.: Competitive analysis with differentiated products. Ann. Econ. Stat. **34**, 159–180 (1994)
Hess, S., Beharry-Borg, N.: Accounting for latent attitudes in willingness-to-pay studies: the case of coastal water quality improvements in Tobago. Environ. Resour. Econ. **52**, 109–131 (2012)
Hsiao, C.Y.: Passenger demand for air transportation in a hub-and-spoke network. PhD Thesis, Department of Civil and Environmental Engineering, University of California at Berkeley (2008)
Hurtubia, R., Bierlaire, M.: Estimation of bid functions for location choice and price modelling with a latent variable approach. Netw. Spat. Econ. **14**, 47–65 (2014)
INRO: EMME/2 *User's Manual*. INRO Inc., Montreal (1996)
Instituto Nacional de Estadísticas: Chile proyecciones y estimaciones de población (2013). http://www.ine.cl/canales/chile_estadistico/demografia_y_vitales/demografia/demografia.php. Accessed 24 July 2019 (in Spanish)
Jensen, A.F., Cherchi, E., Mabit, S.L.: On the stability of preferences and attitudes before and after experiencing an electric vehicle. Transp. Res. Part D Transp. Environ. **25**, 24–32 (2013)
Karaca-Mandic, P., Train, K.: Standard error correction in two-stage estimation with nested samples. Econom. J. **6**, 401–407 (2003)
Lam, S.K., Ahearne, M., Hu, Y., Schillewaert, N.: Resistance to brand switching when a radically new brand is introduced: a social identity theory perspective. J. Mark. **74**, 128–146 (2010)
Lee, L.: Amemiya's generalized least squares and tests of overidentification in simultaneous equation models with qualitative or limited dependent variables. Econ. Rev. **11**, 319–328 (1992)
Mariel, P., Hoyos, D., Artabe, A., Guevara, C.A.: A multiple indicator solution approach to endogeneity in discrete-choice models for environmental valuation. Sci. Total Environ. **633**, 967–980 (2018)
Mumbower, S., Garrow, L.A., Higgins, M.J.: Estimating flight-level price elasticities using online airline data: a first step toward integrating pricing, demand, and revenue optimization. Transp. Res. Part A Policy Pract. **66**, 196–212 (2014)
Nevo, A.: Measuring market power in the ready-to-eat cereal industry. Econometrica **69**, 307–342 (2001)
Newey, W.: Efficient estimation of limited dependent variable models with endogenous explanatory variables. J. Econ. **36**, 231–250 (1987)
Ortúzar, J. de D., Willumsen, L.G.: Modelling Transport, 4th edn. Wiley, Chichester (2011)
Orozco-Fontalvo, M., Arévalo-Támara, A., Guerrero-Barbosa, T., Gutiérrez-Torres, M.: Bicycle choice modelling: a study of university trips in a small Colombian city. J. Transp. Health **9**, 264–274 (2018)
Petrin, A., Train, K.: Omitted product attributes in discrete choice models. Working Paper, Department of Economics, University of California, Berkeley (2002)
Petrin, A., Train, K.: A control function approach to endogeneity in consumer choice models. J. Mark. Res. **47**, 3–13 (2010)
Rivers, D., Vuong, Q.: Limited information estimators and exogeneity tests for simultaneous probit models. J. Econ. **39**, 347–366 (1988)

Rizzi, L.I., Ortúzar, J. de D.: Road safety valuation under a stated choice framework. J. Transp. Econ. Policy **40**, 69–94 (2006)
Sargan, J.: The estimation of economic relationships using instrumental variables. Econometrica **26**, 393–415 (1958)
SECTRA: Actualización Diagnóstico del S.T.U. del Gran Valparaíso. Etapa I-SECTRA, Valparaiso (2014a) (in Spanish)
SECTRA: Actualización y Recolección de Información del Sistema de Transporte Urbano, IX Etapa: Encuesta Origen Destino Santiago 2012. SECTRA, Santiago (2014b) (in Spanish)
Staiger, D., Stock, J.H.: Instrumental variables regression with weak instruments. Econometrica **65**, 557–586 (1997)
Stern, S.: Market definition and the returns to innovation: substitution patterns in pharmaceutical markets. Working Paper, Sloan School of Management, MIT (1996)
Stock, J.H., Yogo, M.: Testing for weak instruments in linear IV regression. In: Andrews, D.W., Stock, J.H. (eds.) Identification and Inference for Econometric Models, Essays in Honour of Thomas Rothenberg. Cambridge University Press, Cambridge (2005)
Tirachini, A., Hensher, D.A., Rose, J.M.: Crowding in public transport systems: effects on users, operation and implications for the estimation of demand. Transp. Res. Part A Policy Pract. **53**, 36–52 (2013)
Torres, I., Greene, M., Ortúzar, J. de D.: Valuation of housing and neighbourhood attributes for city centre location: a case study in Santiago. Habit. Int. **39**, 62–74 (2013)
Train, K.E.: Discrete Choice Methods with Simulation, 2nd edn. Cambridge University Press, Cambridge (2009)
Varela, J.M., Börjesson, M., Daly, A.: Quantifying errors in travel time and cost by latent variables. Transp. Res. Part B Methodol. **117**, 520–541 (2018)
Varotto, S.F., Glerum, A., Stathopoulos, A., Bierlaire, M., Longo, G.: Mitigating the impact of errors in travel time reporting on mode choice modelling. J. Transp. Geogr. **62**, 236–246 (2017)
Villas-Boas, J.M., Winer, R.S.: Endogeneity in brand choice models. Manag. Sci. **45**, 1324–1338 (1999)
Walker, J.L.: (2001) Extended discrete choice models: integrated framework, flexible error structures, and latent variables. Doctoral dissertation, Massachusetts Institute of Technology
Williams, H.C.W.L.: On the formation of travel demand models and economic evaluation measures of user benefit. Environ. Plan. **9A**, 285–344 (1977)
Wooldridge, J.: Econometric Analysis of Cross-Section and Panel Data, 2nd edn. The MIT Press, Cambridge (2010)
Wooldridge, J.: Control function methods in applied econometrics. J. Hum. Resour. **50**, 420–445 (2015)
Yáñez, M.F., Raveau, S., Ortúzar, J. de D.: Inclusion of latent variables in mixed logit models: modelling and forecasting. Transp. Res. Part A Policy Pract. **44**, 744–753 (2010)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Thomas E. Guerrero holds a BSc in Civil Engineering from the Universidad Francisco de Paula Santander, Cúcuta (Colombia) and an MSc in Civil Engineering from the Universidad del Norte, Barranquilla (Colombia). He is Assistant Professor at the Department of Civil Engineering, Universidad Francisco de Paula Santander in Ocaña (Colombia) and is currently doing a PhD at Pontificia Universidad Católica de Chile and Newcastle University (UK). His research is focused on the assessment and correction of endogeneity in discrete choice models.

C. Angelo Guevara is Associate Professor at the Department of Civil Engineering, Universidad de Chile, research affiliate of the ITS Lab, MIT (USA) and external affiliate of the Choice Modelling Centre, University of Leeds (UK). He holds an MSc in transportation from Universidad de Chile, as well as an MSc and a PhD in the same area from MIT.

Elisabetta Cherchi is Professor of Transport at the School of Engineering, Newcastle University (UK) and Adjunct Professor, School of Economics & Management, Beijing Jiaotong University (China). She currently Chairs the International Association of Travel Behaviour Research (IATBR), is Co-Editor in Chief of Transportation Research Part A: Policy and Practice, and past Associate Editor of Transportation.

Juan de Dios Ortúzar is Emeritus Professor (Transport Engineering and Logistics) at Pontificia Universidad Católica de Chile, Principal Researcher at both the Instituto Sistemas Complejos de Ingeniería (ISCI) and the BRT+Centre of Excellence, Research Fellow at the German Institute for Economic Research (DIW-Berlin) and Adjunct Professor, School of Economics & Management, Beijing Jiatong University (China). He is also Co-Editor in Chief of Transportation Research Part A: Policy and Practice.

Affiliations

Thomas E. Guerrero[1,2,3] · **C. Angelo Guevara**[4] · **Elisabetta Cherchi**[3] · **Juan de Dios Ortúzar**[5]

C. Angelo Guevara
crguevar@ing.uchile.cl

Elisabetta Cherchi
elisabetta.cherchi@ncl.ac.uk

Juan de Dios Ortúzar
jos@ing.puc.cl

1 Department of Civil Engineering, Universidad Francisco de Paula Santander Ocaña, Sede el Algodonal Ocaña, Colombia

2 Department of Transport Engineering and Logistics, Pontificia Universidad Católica de Chile, Avda. Vicuña Mackenna, 4860 Macul, Santiago, Chile

3 School of Engineering, Future Mobility Group, Newcastle University, Newcastle upon Tyne NE1 7RU, United Kingdom

4 Department of Civil Engineering, Instituto Sistemas Complejos de Ingeniería (ISCI), Universidad de Chile, Blanco Encalada, 2002 Santiago, Chile

5 Department of Transport Engineering and Logistics, Instituto Sistemas Complejos de Ingeniería (ISCI), BRT+ Centre of Excellence, Pontificia Universidad Católica de Chile, Avda. Vicuña Mackenna 4860, Macul, Santiago, Chile