Transportation Research Part A 212 (2026) 105169



Contents lists available at ScienceDirect

Transportation Research Part A

journal homepage: www.elsevier.com/locate/tra





A New Approach to collinearity and elasticity problems in value of travel time savings estimation with multinomial logit models

Louis de Grange [a,*], Matthieu Marechal [b,c], Raúl Pezoa [a,c], Felipe González [a], Rodrigo Troncoso [d]

[a] Escuela de Ingeniería Industrial, Universidad Diego Portales, Santiago, Chile
[b] Instituto de Ciencias Básicas, Universidad Diego Portales, Santiago, Chile
[c] Centro Interdisciplinario de Data Science, Universidad Diego Portales, Santiago, Chile
[d] Facultad de Gobierno, Universidad del Desarrollo, Santiago, Chile

ARTICLE INFO

Keywords:
Value of travel time savings
Collinearity
Elasticity
Multinomial logit models
Constrained estimation
Hausman test

ABSTRACT

The value of travel time savings (VTTS) is a central parameter in transport-project appraisal. When estimated from a multinomial logit (MNL) model of mode choice, the VTTS can become imprecise under two structural conditions of the data: strong collinearity between travel time and travel cost, common in revealed-preference samples where cost is imputed at a fixed rate per kilometer; and low absolute price-elasticity of demand, which makes the estimated cost coefficient close to zero and inflates the variance of the VTTS estimator, formed as a ratio with that coefficient in the denominator. Both conditions share a common cause: the cost coefficient is not well identified by the data. This paper diagnoses both conditions analytically and through simulation, and proposes an elasticity-constrained estimator that addresses this common cause by tying the cost coefficient to a value consistent with external price-elasticity evidence. The estimator is compared through simulations to maximum likelihood and ridge estimation. The results show that under adverse collinearity and demand elasticity conditions, maximum likelihood estimates exhibit high variance and errors, even with large samples, while ridge estimates display less variance and improved stability but also significant levels of bias. By contrast, the proposed method, although it introduces bias if the actual demand elasticity is distant from that of the applied restriction, achieves lower values for mean squared error. The three methods are also compared in an application to the Swissmetro mode choice dataset. The results in this case show that although multinomial logit can, in theory, be consistent, the elasticity-constrained method again obtains the lowest mean squared error by reducing the variance of the VTTS estimates. The ridge estimator, meanwhile, proved to be less competitive. Complementing these comparative findings are the results of the Hausman test. The proposed estimator is recommended for VTTS estimation in revealed-preference contexts where strong collinearity and inelastic demand jointly compromise identification, and we provide guidance on the contexts in which it is most useful and on its relationship with alternatives such as willingness-to-pay-space estimation.

1. Introduction

Travel time savings typically account for the largest share of monetized benefits in transport-project appraisals. In the United Kingdom and other European jurisdictions, roughly 80% of the forecast benefits of a new transport project derive from valuations of

* Corresponding author.
E-mail addresses: louis.degrange@udp.cl (L. de Grange), matthieu.marechal@udp.cl (M. Marechal), raul.pezoa@udp.cl (R. Pezoa), felipe.gonzalezr@udp.cl (F. González), rtroncoso@udd.cl (R. Troncoso).

https://doi.org/10.1016/j.tra.2026.105169
Received 10 December 2025; Received in revised form 8 July 2026; Accepted 17 July 2026
Available online 27 July 2026
0965-8564/© 2026 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

time savings (Mackie et al., 2001; Tsoleridis et al., 2022). The parameter that converts those savings into money equivalents, called the value of travel time savings (VTTS) in the modern literature and the subjective value of time (SVT) in earlier work (DeSerpa, 1971; Jara-Díaz and Farah, 1987), is therefore decisive for which transport investments are recommended for funding and which are deferred. National value-of-time studies, which inform official appraisal guidance for periods of a decade or more (Wardman et al., 2016), depend on this parameter being estimated reliably. Yet despite a mature literature on discrete choice modelling, VTTS estimates from standard maximum likelihood remain unstable under empirical conditions that arise routinely in transport data, particularly in revealed-preference samples where the monetary cost of a trip is imputed rather than directly observed.

When the VTTS is estimated from a multinomial logit (MNL) model of mode choice, it can become highly imprecise under two structural conditions of the data: strong collinearity between the travel time and travel cost variables, and a low absolute level of price-elasticity of demand. These conditions share a common cause. The VTTS is formed as a ratio with the estimated cost coefficient $\widehat{\beta}_c$ in the denominator, and both conditions push $\widehat{\beta}_c$ toward zero: collinearity by inflating its sampling variance, inelasticity by making the true magnitude $|\beta_c|$ itself small. Both conditions are common in revealed-preference (RP) samples where travel cost is imputed at a fixed rate per kilometer (Tsoleridis et al., 2022; Ortúzar and Willumsen, 2024). In revealed-preference data a related mechanism reinforces the same effect, because measurement error in the imputed cost attenuates the estimated cost coefficient toward zero (Walker et al., 2010; Lorenzo Varela et al., 2018). This paper addresses the question of how to estimate the VTTS precisely in MNL mode-choice models when the data alone do not identify it well.

Existing approaches to VTTS estimation in MNL models fall into three families. Standard maximum likelihood combined with the delta method provides consistent estimates and asymptotic confidence intervals (Daly et al., 2012a), but its variance estimates diverge whenever $\widehat{\beta}_c$ is close to zero, regardless of sample size. The ridge estimator, applied to discrete choice by de Grange et al. (2015), stabilizes estimates by penalizing coefficient magnitudes, but the penalty treats all coefficients alike and shrinks β_c toward zero, which can introduce substantial bias when the true $|\beta_c|$ is small. Willingness-to-pay-space estimation reparameterizes the utility function so that the VTTS distribution itself becomes the primary object of estimation (Train and Weeks, 2005; Fosgerau, 2007; Scarpa et al., 2008), avoiding the ratio-of-coefficients problem at the cost of imposing parametric structure on the VTTS distribution. No estimator can recover identifying information the data do not contain. Each of the methods reviewed above brings in a different form of external structure to compensate. The present paper proposes a complementary route: bringing external evidence on the level of price-elasticity of demand directly into the estimation of β_c, drawing on decades of empirical evidence (Holmgren, 2007; Brons et al., 2008; Berry and Börjesson, 2024).

Among the many quantities that could be brought in as external evidence, the price-elasticity of demand is particularly well-suited for three reasons. First, the connection to β_c is mechanical in the MNL: the estimated elasticity is a direct function of β_c and observable sample quantities, so a target on the elasticity translates exactly into a target on the cost coefficient. Second, the literature on transport-demand price-elasticity has accumulated decades of empirical evidence with relatively narrow consensus ranges across contexts (Holmgren, 2007; Brons et al., 2008; Gillingham and Munk-Nielsen, 2019; Berry and Börjesson, 2024), making the external target both available and credible. Third, price-elasticity is the quantity that downstream applications rely on anyway: cost-benefit analyses of price changes, congestion charges, toll roads, and fare reforms all depend on it, so an estimator that respects external evidence on this quantity produces VTTS estimates that are coherent with the broader empirical record.

This paper makes two contributions. First, we propose an elasticity-constrained estimator for the VTTS in MNL models that addresses the joint imprecision caused by collinearity and inelastic demand. The estimator ties β_c to a value consistent with external empirical evidence on price-elasticity of demand, and we evaluate it against maximum likelihood and ridge estimation in simulation and in an application to the Swissmetro mode choice dataset; Hausman tests provide evidence on the bias introduced by the constraint. The estimator admits an equivalent maximum a posteriori interpretation under a Gaussian prior on the cost coefficient, providing a rigorous Bayesian and penalized-likelihood framing. The resulting estimate is interpretable as the value of travel time savings conditional on the price-elasticity being consistent with external evidence, not as an unconditional point identification. Second, we provide guidance on the contexts in which the proposed estimator is most useful, namely revealed-preference samples with moderate N where structural collinearity and inelastic demand are common, and discuss its relationship with alternative approaches such as willingness-to-pay-space estimation.

The paper proceeds as follows. Section 2 reviews the literature on VTTS estimation in discrete choice models and on the use of restrictions in econometric estimation. Section 3 derives the analytic expression for VTTS variance and develops the proposed estimator alongside the ridge alternative. Sections 4 and 5 report the simulation results and the Swissmetro application, respectively. Section 6 concludes with guidance on when the proposed estimator is most useful and its relationship with WTP-space alternatives.

2. Literature review

2.1. Foundations and the estimation problem

VTTS is defined in microeconomic terms as the marginal rate of substitution between travel time and cost, that is, the willingness to pay for saving a unit of time (Becker, 1965; DeSerpa, 1971; Jara-Díaz and Farah, 1987; Gaudry et al., 1989; Hensher, 2001; Mackie et al., 2001). Good reviews of value of time theory include González (1997) and Small (2012); the empirical evidence is discussed in Noland and Polak (2002), Li et al. (2010), Abrantes and Wardman (2011), Festjens and Janiszewski (2015), Gardes (2018) and Agness et al. (2025). The Swedish national value-of-time study (Börjesson and Eliasson, 2014) provides a recent example of how these methods inform appraisal practice.

For transport models with linear explanatory variables (e.g., for travel time and cost), the marginal rate of substitution is obtained as the ratio of the parameter for travel time to the parameter for travel cost. In formal terms, VTTS's economic definition and corresponding estimate are specified as

$$\text{VTTS} = \phi = \frac{\beta_t}{\beta_c}, \qquad \widehat{\text{VTTS}} = \hat{\phi} = \frac{\hat{\beta}_t}{\hat{\beta}_c} \tag{1}$$

where β_t is the time parameter and β_c the cost parameter.

If VTTS is modelled and estimated using, for example, a multinomial-type discrete choice model (logit, probit, mixed logit, nested logit, among others), a specific VTTS estimate can be obtained for each alternative present (e.g., mode choice):

$$\text{VTTS}^m = \phi^m = \frac{\beta_t^m}{\beta_c^m}, \qquad \widehat{\text{VTTS}}^m = \hat{\phi}^m = \frac{\hat{\beta}_t^m}{\hat{\beta}_c^m} \tag{2}$$

where the superscript m specifies the mth alternative or transport mode.

The microeconomic interpretation of the VTTS as a marginal rate of substitution requires the marginal utility of income to be positive, so β_c^m is a strictly negative parameter (DeSerpa, 1971; Jara-Díaz and Farah, 1987; Jara-Diaz, 2024).

Maximum likelihood estimation delivers consistent estimators $\hat{\beta} = (\hat{\beta}_t^m, \hat{\beta}_c^m)$ that are asymptotically normal around the population values (β_t^m, β_c^m). The VTTS estimator $\hat{\phi}^m = \hat{\beta}_t^m / \hat{\beta}_c^m$ is a smooth function of the MLE estimators, and the delta method gives its exact asymptotic standard error (Daly et al., 2023). The asymptotic variance of $\hat{\phi}^m$ scales with $1/\beta_c^{m\,2}$ and is amplified by the covariance between $\hat{\beta}_t^m$ and $\hat{\beta}_c^m$, which becomes strongly negative under collinearity (Greene, 2003). The reliability of asymptotic intervals depends on the t-ratio of $\hat{\beta}_c^m$ being high; lower precision calls for alternative procedures such as likelihood-ratio or bootstrap intervals.

Two structural features of the data jointly produce a low t-ratio for $\hat{\beta}_c^m$: strong collinearity between travel time and cost inflates its sampling variance, and low absolute price-elasticity of demand makes the population $|\beta_c^m|$ small. Both routes share a common cause, the under-identification of β_c by the data, and in this regime the point estimate itself is imprecise, not only its standard error.

A further mechanism operates in revealed-preference data and pushes the cost coefficient in the same direction, though for a different reason. When the monetary cost of a trip is imputed rather than directly observed, the cost variable carries measurement error, and classical measurement error attenuates the coefficient of the affected variable toward zero (Walker et al., 2010; Lorenzo Varela et al., 2018). This attenuation is a bias rather than a loss of precision, so it is distinct from collinearity and from low elasticity, but the observable consequence is the same, a cost coefficient that is small or unreliable and a value of travel time savings that is correspondingly unstable. Walker et al. (2010) and Lorenzo Varela et al. (2018) address it directly, using latent-variable and hybrid-choice models to correct the measurement error and recover coefficients less biased toward zero.

2.2. Existing methodological approaches

2.2.1. Variance estimation methods

A first family of methods takes the point estimate $\hat{\beta}$ as given and seeks reliable confidence intervals for $\hat{\phi}$.

Among delta-method applications, Yadlin and Yadlin (1999) constructs confidence intervals for transport-user VTTS based on the asymptotic normality of the maximum likelihood estimator, comparing it with the alternative approach developed by Ortúzar and Armstrong (1995) and finding the former superior in accuracy and computational simplicity. Daly et al. (2012a) argue based on simulations for the delta method's reliability under standard regularity. Bliemer and Rose (2013) apply the same method in mixed logit by taking into account the variance-covariance matrix of the estimates.

Simulation-based alternatives include the bootstrap approach of Krinsky and Robb (1986), which produces smaller confidence intervals at higher computational cost; Espino et al. (2006) apply this method and find the interval size heavily affected by atypical values; and Hole (2007) compares delta, Krinsky-Robb and bootstrap, finding that the delta method is most accurate when the data are well-conditioned while the bootstrap is most robust to noisy data and model misspecification.

Armstrong et al. (2001) compare alternative methods based on the t-test and likelihood-ratio test, noting that the interval's midpoint does not coincide with the VTTS point estimate. Bolduc et al. (2010) compare the delta method with the generalized Fieller method, concluding that the latter is more reliable when the denominator of the parameter ratio is close to zero. Scaccia et al. (2023) propose an improved delta method (DSMG) for mixed logit, offering an efficient balance between accuracy and computational cost.

Carson and Czajkowski (2019) reparametrize the cost coefficient using an exponential transformation so that it remains negative throughout estimation. Daly et al. (2023) show, however, that this reparameterization is mathematically equivalent to the standard preference-space specification under maximum likelihood and offers no small-sample advantage, while introducing numerical difficulties in optimization.

Hess et al. (2005) studies VTTS estimation in discrete choice models with random heterogeneity of preferences, noting that bounded distributions such as a scaled Johnson's S_B (Train and Sonnier, 2005) avoid the interpretation problems and extreme values that arise with the typical normal distribution. McFadden (2022) argues that some econometric models such as mixed logit may be unstable and produce statistically unreliable conclusions even if market equilibria exist, while Biswas et al. (2024) concludes that incorporating stochastic explanatory variables with random coefficients improves the accuracy of discrete choice models.

These methods address the computation of standard errors and confidence intervals given the underlying estimate of $\hat{\beta}_c$. When the t-ratio of $\hat{\beta}_c$ is low, the source of imprecision is the point estimate itself, and a different kind of intervention is required.

2.2.2. Reparameterization in willingness-to-pay space

An alternative family of methods reparameterizes the utility function so that the willingness-to-pay for travel time savings is estimated as a single coefficient (Train and Weeks, 2005). In this approach, called estimation in willingness-to-pay space (or VTTS space), the utility is written as $U = \beta_c(\phi t + c)$ with ϕ the VTTS itself, and the delta method applies directly to the standard error of ϕ. Daly et al. (2023) show that preference-space and WTP-space specifications are mathematically equivalent under maximum likelihood, with identical fit and identical asymptotic standard errors. A related strand starting with Fosgerau (2007) estimates the willingness-to-pay distribution itself, modelling choice as a comparison between an individual's WTP and the bid implied by each scenario; the distribution can be specified parametrically or recovered non-parametrically from binary choice data with sufficient bid variation.

In mixed logit specifications, the WTP-space approach allows preference heterogeneity by treating ϕ as a random parameter across individuals. The distribution chosen for ϕ determines whether moments of the WTP distribution exist. Daly et al. (2012b) provide a theorem stating necessary and sufficient conditions for the existence of inverse moments. The lognormal, Johnson S_B, and distributions bounded strictly away from zero produce finite moments under stated conditions, while common alternatives that place positive mass near zero (such as the normal or the triangular bounded at zero) produce a WTP distribution whose moments diverge. Hess and Train (2017) reinforce this conclusion in the context of correlated random coefficients.

Empirical comparisons of preference-space and WTP-space specifications report mixed conclusions on fit but consistent advantages on the realism of the implied WTP distribution. Scarpa et al. (2008) find WTP-space fits better and yields more realistic estimates on revealed-preference data for destination choice in the Alps. Sonnier et al. (2007) report similar conclusions on conjoint data. Hole and Kolstad (2012) compare both regimes on a health-related choice experiment and find that preference-space fits slightly better in their case, while WTP-space produces less skewed WTP distributions whose means are more plausible.

Recent applications of mixed logit in travel behaviour continue to engage with these issues. The Mabit (2026) review of mixed logit applications documents that the handling of WTP-related issues remains uneven across the published literature, with roughly two-thirds of recent applications addressing the moment-existence problem substantively. Bouscasse and de Lapparent (2019) provide an example that combines random coefficients with latent variables to capture perceived comfort effects on VTTS.

WTP-space methods and the elasticity-constrained estimator proposed in this paper both bring external information into VTTS estimation, but the information is different. WTP-space methods constrain the parametric form of preference heterogeneity across individuals; this is most useful in moderate-to-large samples. Our proposal constrains the level of β_c using empirical evidence on price-elasticity. The two approaches are complementary.

2.2.3. Penalty and constraint methods

A third family imposes constraints on the parameters of the model itself, either through penalty norms on coefficient magnitudes or through structural restrictions on the choice probabilities or the parameter space.

The most traditional formulations addressing data collinearity use ridge or lasso estimators (Mauerer et al., 2015; Pavlou et al., 2016). In the optimization process the ridge estimator constrains the sum of squares of the parameters to be estimated whereas the lasso estimator puts restrictions on the sum of the parameters' absolute values.

In a multinomial logit context, Castro et al. (2013) introduce a constrained multinomial logit model (CMNL) that incorporates thresholds on certain attributes as an endogenous process in the formation of the choice set. Restrictions are imposed on the attributes of the alternatives in order to better represent the limits facing individuals on their decisions. de Grange et al. (2015) apply the ridge regression technique to VTTS estimation, finding that the ridge estimator produces estimates with lower variance and mean squared error than ordinary least squares. Delle Site and Parmar (2024) use constrained maximum likelihood estimators (CMLE) within the linear probability model, ensuring valid probabilities while studying welfare implications.

The methods just reviewed impose constraints on coefficients directly. The estimator proposed in this paper belongs to this same family but imposes its constraint on a derived quantity (the price-elasticity of demand) rather than on a coefficient.

2.3. Synthesis

VTTS estimation in MNL becomes imprecise when the t-ratio of the cost coefficient is low, a regime reached when travel time and cost are strongly collinear or when the absolute price-elasticity of demand is small. Three families of methods respond to this problem: better confidence intervals around the point estimate, reparameterization of the utility function to estimate the WTP directly, and penalty norms or structural constraints on coefficient magnitudes.

The two structural conditions discussed above arise differently across data sources. In revealed-preference (RP) samples from household travel surveys, the monetary cost of a trip is rarely observed directly and is typically imputed at a fixed rate per kilometer, which introduces collinearity between cost and the underlying network distance, and through it with travel time (Tsoleridis et al., 2022; Ortúzar and Willumsen, 2024). Stated-preference (SP) experiments can break this collinearity by design through orthogonal variation of cost and time, although they remain subject to the inelastic-demand component of the problem and to the experimental-design choices that shape the distribution of cost values shown to respondents (Wardman et al., 2016). Pooled SP-RP approaches combine both data types to jointly examine VTTS and price-elasticity (Ilahi et al., 2021). Large passive-tracking datasets such as mobile-network call detail records, platform trip records, or toll-corridor transponder data reduce the small-sample component by enlarging N by orders of magnitude (Calabrese et al., 2013; Bwambale et al., 2019; Lin et al., 2026), but the structural collinearity does not vanish and these data sources typically lack direct cost information. The methods developed in this paper are most relevant

in RP contexts with moderate sample sizes, where the joint effect of structural collinearity and inelastic demand on VTTS precision is most acute.

The methodological innovation of the present paper is to bring external empirical evidence on the level of price-elasticity of demand into the estimation of β_c (Holmgren, 2007; Brons et al., 2008; Wardman et al., 2018; Bull et al., 2021; Gillingham and Munk-Nielsen, 2019; Berry and Börjesson, 2024). In MNL the price-elasticity is a function of the cost coefficient, so constraining one constrains the other.

3. Methodology

We begin the presentation of our methodology by developing an analytic expression with which we can directly visualize the relation between VTTS variance and both the degree of collinearity between the travel time and cost variable and the level of price-elasticity of demand. This relation highlights the efficiency problems in the form of large variances that may arise when estimating VTTS in the presence collinearity and inelastic demand.

We then discuss two alternative estimation methods involving constraints on the likelihood function. The first method is based on the classic ridge estimator, which is typically used to improve estimation when there is strong collinearity between explanatory variables. The second method is a novel one we propose that incorporates exogenous elasticity constraints into the maximum likelihood problem. In Section 4 the two approaches are compared with the classic maximum likelihood method, which does not apply restrictions.

3.1. Analytic expression for VTTS variance

VTTS as defined in (1) is amenable to the application of the delta method, which approximates the variance of a function of random variables. In the present case, the estimators $(\widehat{\beta}_t, \widehat{\beta}_c)$ represent random variables with a given mean and variance. The delta approximation of the VTTS variance is given by the following expression:

$$\mathrm{Var}(\widehat{\phi}) \;=\; \mathrm{Var}\left(\frac{\widehat{\beta}_t}{\widehat{\beta}_c}\right) \;\approx\; \nabla\phi^{\top}\; \mathrm{Var}(\widehat{\beta}_t, \widehat{\beta}_c)\, \nabla\phi \tag{3}$$

where ϕ = VTTS, $\widehat{\phi}$ is the VTTS estimator, $\mathrm{Var}(\widehat{\phi})$ is the estimator's variance and $\mathrm{Var}(\widehat{\beta}_t, \widehat{\beta}_c)$ is the 2×2 variance–covariance matrix.

Taking the gradient of the ratio of the β parameters defining VTTS, we obtain

$$\nabla\phi \;=\; \nabla\left(\frac{\beta_t}{\beta_c}\right) \;=\; \begin{bmatrix} \dfrac{1}{\beta_c} \\ -\dfrac{\beta_t}{\beta_c^2} \end{bmatrix} \tag{4}$$

Then, substituting (4) into (3), we get

$$\mathrm{Var}(\widehat{\phi}) \;\approx\; \begin{bmatrix} \dfrac{1}{\beta_c} & -\dfrac{\beta_t}{\beta_c^2} \end{bmatrix} \begin{bmatrix} \mathrm{Var}(\widehat{\beta}_t) & \mathrm{Cov}(\widehat{\beta}_t, \widehat{\beta}_c) \\ \mathrm{Cov}(\widehat{\beta}_t, \widehat{\beta}_c) & \mathrm{Var}(\widehat{\beta}_c) \end{bmatrix} \begin{bmatrix} \dfrac{1}{\beta_c} \\ -\dfrac{\beta_t}{\beta_c^2} \end{bmatrix} \tag{5}$$

$$\mathrm{Var}(\widehat{\phi}) \;\approx\; \frac{1}{\beta_c^2}\left[\mathrm{Var}(\widehat{\beta}_t) - 2\;\mathrm{Cov}(\widehat{\beta}_t, \widehat{\beta}_c)\,\frac{\beta_t}{\beta_c} + \mathrm{Var}(\widehat{\beta}_c)\left(\frac{\beta_t}{\beta_c}\right)^2\right] \tag{6}$$

$$\mathrm{Var}(\widehat{\phi}) \;\approx\; \frac{1}{\beta_c^2}\left[\mathrm{Var}(\widehat{\beta}_t) - 2\;\mathrm{Cov}(\widehat{\beta}_t, \widehat{\beta}_c)\,\phi + \mathrm{Var}(\widehat{\beta}_c)\,\phi^2\right] \tag{7}$$

Finally, the sample variance estimator of (7) is

$$\widehat{\mathrm{Var}}(\widehat{\phi}) \;=\; \frac{1}{\widehat{\beta}_c^{\,2}}\left[\widehat{\mathrm{Var}}(\widehat{\beta}_t) - 2\,\widehat{\mathrm{Cov}}(\widehat{\beta}_t, \widehat{\beta}_c)\,\widehat{\phi} + \widehat{\mathrm{Var}}(\widehat{\beta}_c)\,\widehat{\phi}^{\,2}\right] \tag{8}$$

If instead of using the VTTS definition in (1) we had used the more general case in (2), the same steps would be followed, giving as the VTTS sample variance estimator the following expression (different for each transport mode m):

$$\widehat{\mathrm{Var}}(\widehat{\phi}^{\,m}) \;=\; \frac{1}{\widehat{\beta}_c^{\,m\,2}}\left[\widehat{\mathrm{Var}}(\widehat{\beta}_t^{\,m}) - 2\,\widehat{\mathrm{Cov}}(\widehat{\beta}_t^{\,m}, \widehat{\beta}_c^{\,m})\,\widehat{\phi}^{\,m} + \widehat{\mathrm{Var}}(\widehat{\beta}_c^{\,m})\,\widehat{\phi}^{\,m\,2}\right] \tag{9}$$

Now recall that for an MNL model that is linear in its variables, the elasticity of attribute or variable k of alternative m is expressed as (McFadden, 1972):

$$\eta_{ki}^{m} \;=\; \frac{\partial P_i^m}{\partial x_{ki}^m} \cdot \frac{x_{ki}^m}{P_i^m} \;=\; \beta_k^m\, P_i^m\,(1 - P_i^m)\,\frac{x_{ki}^m}{P_i^m} \;=\; \beta_k^m\, x_{ki}^m\,(1 - P_i^m) \tag{10}$$

where P_i^m is the probability that individual i chooses alternative m and β_k^m are the parameters to be estimated representing the relative weights on each attribute k and alternative m. These individual attributes or variables x_{ki}^m typically include travel time and travel cost for each individual i for each of the available m alternatives.

Recall further that for an MNL, P_i^m is given by the following expression (McFadden, 1972):

$$P_i^m = \frac{\exp\left(\sum_k \beta_k^m x_{ki}^m\right)}{\sum_{m'} \exp\left(\sum_k \beta_k^{m'} x_{ki}^{m'}\right)} \tag{11}$$

From (10) it can be seen that the average elasticity η_k^m of the kth attribute for the mth alternative is given by

$$\eta_k^m = \frac{\sum_{i=1}^N \eta_{ki}^m}{N} = \frac{\sum_{i=1}^N \beta_k^m x_{ki}^m (1 - P_i^m)}{N} \tag{12}$$

where N is the number of individuals i in the sample.

It is also known that the maximization problem for the log-likelihood function that estimates the β_k^m parameters and their respective variances and covariances is

$$\max_{\beta_k^m} \ln L(\beta) = \sum_i \sum_m \delta_i^m \ln P_i^m(\beta) \tag{13}$$

where δ_i^m is equal to 1 if alternative m was chosen by individual i, otherwise it is 0.

The optimality conditions for (13) are

$$\frac{\partial \ln L(\beta)}{\partial \beta_k^m} = \sum_i \left(\delta_i^m - P_i^m\right) x_{ki}^m = 0 \tag{14}$$

From (14) we directly obtain that

$$\sum_i \delta_i^m x_{ki}^m = \sum_i P_i^m x_{ki}^m \tag{15}$$

This expression ensures that if there are model constants (or intercepts), the modelled and observed portions of the market are the same. It further guarantees that the modelled and observed average attributes (e.g., travel time, fares, etc.) are also the same.

If we substitute (15) into (12), we obtain directly

$$\eta_k^m = \frac{\sum_{i=1}^N \beta_k^m x_{ki}^m (1 - \delta_i^m)}{N} \tag{16}$$

Now defining as η_c^m the price-elasticity of demand associated with travel cost for mode m, (16) reduces to

$$\eta_c^m = \frac{\sum_{i=1}^N \beta_c^m x_{ci}^m (1 - \delta_i^m)}{N} \tag{17}$$

where β_c^m is the parameter associated with variable x_{ci}^m, the travel cost of mode m for individual i.

Finally, from (17) we obtain directly the following expression for said parameter:

$$\beta_c^m = \frac{N \eta_c^m}{\sum_{i=1}^N x_{ci}^m (1 - \delta_i^m)} = \gamma_c^m \eta_c^m, \qquad \text{where} \quad \gamma_c^m = \frac{N}{\sum_{i=1}^N x_{ci}^m (1 - \delta_i^m)} \tag{18}$$

Note that in (18), the β_c^m parameter is directly proportional to the price-elasticity of demand η_c^m. Through the delta-method variance (9), which contains the factor $1/\beta_c^{m\,2}$, this proportionality directly links the absolute elasticity to the precision of $\widehat{\phi}^m$. In other words, the greater is the elasticity, the larger is β_c^m, and therefore the more accurate will be the VTTS variance estimate. Analogously, if η_c^m is small, demand is more inelastic and the variance estimate is less accurate. Importantly, the substitution above uses the first-order conditions of the unrestricted likelihood (15), so the identity $\beta_c^m = \gamma_c^m \eta_c^m$ in (18) holds exactly at the maximum likelihood estimate and only approximately under estimators that do not satisfy these conditions exactly, such as ridge or the elasticity-constrained estimator we propose.

If we substitute the expression for β_c^m in (18) into (9), which gives the sample variance estimator for each mode m, we get

$$\widehat{\mathrm{Var}}(\widehat{\phi}^m) = \frac{1}{\left(\hat{\gamma}_c^m \hat{\eta}_c^m\right)^2}\left[\widehat{\mathrm{Var}}(\widehat{\beta}_t^m) - 2\,\widehat{\mathrm{Cov}}(\widehat{\beta}_t^m, \widehat{\beta}_c^m)\,\widehat{\phi}^m + \widehat{\mathrm{Var}}(\widehat{\beta}_c^m)\,\widehat{\phi}^{m\,2}\right] \tag{19}$$

From (19) we may state the following:

i. Since the travel time and cost variables tend to be highly correlated, the $\mathrm{Cov}(\widehat{\beta}_t^m, \widehat{\beta}_c^m)$ term will be negative. The presence of strong collinearity generates negative covariance in the adjacent coefficients, a phenomenon occurring both in linear regression and discrete choice models (Greene, 2003). This implies that the greater is the correlation between the two variables, the greater will be the variance of the VTTS estimator.
ii. If price-elasticity of demand is low, that is, if the absolute value (modulus) of η_c^m is small, the variance of the VTTS estimator will be large. In the hypothetical case where demand is perfectly inelastic ($\eta_c^m = 0$), the variance will be infinite. This is consistent with $\beta_c^m = 0$, meaning demand is not dependent on travel cost.

Table 1
Parameter Values Obtained in Preliminary Simulations.

η_c^{car}	ϕ^{car}	β_0^{car}	β_0^{train}	β_c^{car}	β_t^{car}	β_t^{train}
−1.0	5.0	0	−0.50	−1.30	−6.50	−4.550
−0.7	5.0	0	0.00	−0.90	−4.50	−3.150
−0.5	5.0	0	0.05	−0.70	−3.50	−2.450
−0.3	5.0	0	0.50	−0.40	−2.00	−1.400
−0.1	5.0	0	1.00	−0.15	−0.75	−0.525

The foregoing two points are very important when estimating VTTS using MNL-type transport demand models given that the travel time and cost variables typically exhibit high correlation (strong collinearity) and demand levels tend to be inelastic. Although such models have been extensively used to estimate the value of time, causal interpretation of the results depends on the validity of the assumptions. As noted by Graham (2025), if the conditions for the observational design of the data collection are not ideal, careful consideration must be given to the techniques of causal inference in order to avoid spurious conclusions.

3.2. Preliminary analysis based on simulations

Preliminary simulations were carried out in order to confirm the impact of strong collinearity and price-inelasticity of demand on VTTS variance and determine their respective estimates according to the formula in (19). The results obtained, presented below, identify the different ranges of collinearity and elasticity for which problems in estimating VTTS were detected. As will be seen, the problematic ranges were associated with high levels of variance.

The simulations were performed for two transport modes, private car and rail, each with its own utility function specified as follows:

$$U_i^{car} = \beta_0^{car} + \beta_t^{car}\, t_i^{car} + \beta_c^{car}\, c_i^{car} + \varepsilon_i^{car}, \qquad \varepsilon_i^{car} \sim \text{Gumbel}(0,1) \tag{20}$$

$$U_i^{train} = \beta_0^{train} + \beta_t^{train}\, t_i^{train} + \varepsilon_i^{train}, \qquad \varepsilon_i^{train} \sim \text{Gumbel}(0,1) \tag{21}$$

where t_i^{car} is car travel time and c_i^{car} is the monetary cost of car travel. For the rail mode, only travel time t_i^{train} is included as the fare was assumed to be fixed and independent of travel time (i.e., flat fare).

The car mode VTTS was set at 5 ($\phi^{car} = 5.0$) and results were generated for five different price-elasticity of demand levels: $\eta_c^{car} = -0.1; -0.3; -0.5; -0.7; -1.0$. The values obtained for the parameters $\beta_0^{car}, \beta_0^{train}, \beta_t^{car}, \beta_t^{train}$ and β_c^{car} are shown in Table 1.

The variables for rail travel time t_i^{train}, car travel time t_i^{car} and car travel cost c_i^{car} follow gamma distributions with the following parameters:

$$t_i^{train} \sim \Gamma(7;\, 4.25) \tag{22}$$

$$t_i^{car} \sim \Gamma(5;\, 5) \tag{23}$$

$$c_i^{car} \sim \Gamma(4;\, 0.25) \tag{24}$$

Also included in the simulations were different degrees of strong collinearity between the variables representing car travel time and car travel cost as well as between those representing car travel time and rail travel time. Denoting the correlation between the car travel time and car travel cost variables as $\rho_1 = \text{Corr}(t_i^{car}, c_i^{car})$ and the correlation between the car travel time and rail travel as $\rho_2 = \text{Corr}(t_i^{car}, t_i^{train})$, we considered for each correlation type in each of the five elasticity scenarios defined above ($\eta_c^{car} = -0.1; -0.3; -0.5; -0.7; -1.0$) the following three correlation values: $\rho_1 \in \{0.5, 0.7, 0.9\}$ and $\rho_2 \in \{0.5, 0.7, 0.9\}$. Thus, we obtained a total of $5 \times 3 \times 3 = 45$ different elasticity scenarios.

Each of the 45 scenarios was replicated 1000 times for each of three sample sizes of 1,000, 3000 and 10,000 observations, the number of different simulations thus totalling 135,000. For each one, maximum likelihood was employed to obtain the various parameters and the respective VTTS estimate. The results are displayed in Table 2,[1] which also shows the standard deviation (the square root of the variance) for each scenario. The cells in grey are the cases where the standard deviation is greater than the mean (S.E. > 5.0). The results in bold type are those where $\rho_1 = 0.9$.

From these results we may draw the following conclusions:

- For the sample with 1000 observations, the VTTS standard deviation behaved rather erratically, especially in cases of strong collinearity ($\rho_1 = 0.9$) and medium or low elasticity ($|\eta_c^{car}| \leq 0.7$).
- Strong collinearity between explanatory variables of the same mode (i.e., correlation of 0.9 between car travel time and car travel cost) generated a very significant rise in the VTTS variance estimates.

[1] For each scenario, replications whose VTTS estimate lies more than three standard deviations from the replication mean are excluded before computing the reported standard deviations in Tables 2, 4, 5 and 6. When $\hat{\beta}_c^m$ is poorly identified, individual replications can fall close to zero and produce very large VTTS values that dominate the empirical standard deviation; the filter limits the influence of these extreme draws, and the reported standard deviations are conservative lower bounds on the unfiltered values. The effect is stronger for maximum likelihood than for the constrained estimators, since the latter keep $\hat{\beta}_c^m$ away from zero by construction; the relative advantage of the ridge and elasticity-constrained methods reported in Section 4 is therefore at least as large as the tables show.

Table 2
VTTS Estimates and their Standard Deviations in Preliminary Simulations.

$N = 1{,}000$				$N = 3{,}000$				$N = 10{,}000$			
η_c^{car}	ρ_1	ρ_2	$\text{SE}(\widehat{\phi^{\text{car}}})$	η_c^{car}	ρ_1	ρ_2	$\text{SE}(\widehat{\phi^{\text{car}}})$	η_c^{car}	ρ_1	ρ_2	$\text{SE}(\widehat{\phi^{\text{car}}})$
−1.0	0.5	0.5	1.08	−1.0	0.5	0.5	0.62	−1.0	0.5	0.5	0.33
−1.0	0.5	0.7	1.12	−1.0	0.5	0.7	0.61	−1.0	0.5	0.7	0.34
−1.0	0.5	0.9	1.17	−1.0	0.5	0.9	0.68	−1.0	0.5	0.9	0.38
−1.0	0.7	0.5	1.39	−1.0	0.7	0.5	0.81	−1.0	0.7	0.5	0.42
−1.0	0.7	0.7	1.42	−1.0	0.7	0.7	0.79	−1.0	0.7	0.7	0.43
−1.0	0.7	0.9	1.45	−1.0	0.7	0.9	0.88	−1.0	0.7	0.9	0.46
−1.0	**0.9**	0.5	**3.10**	−1.0	**0.9**	0.5	**1.34**	−1.0	**0.9**	0.5	**0.74**
−1.0	**0.9**	0.7	**3.20**	−1.0	**0.9**	0.7	**1.35**	−1.0	**0.9**	0.7	**0.71**
−1.0	**0.9**	0.9	**2.70**	−1.0	**0.9**	0.9	**1.32**	−1.0	**0.9**	0.9	**0.72**
−0.7	0.5	0.5	1.58	−0.7	0.5	0.5	0.82	−0.7	0.5	0.5	0.44
−0.7	0.5	0.7	1.66	−0.7	0.5	0.7	0.88	−0.7	0.5	0.7	0.47
−0.7	0.5	0.9	1.73	−0.7	0.5	0.9	0.95	−0.7	0.5	0.9	0.54
−0.7	0.7	0.5	2.20	−0.7	0.7	0.5	1.07	−0.7	0.7	0.5	0.60
−0.7	0.7	0.7	2.11	−0.7	0.7	0.7	1.13	−0.7	0.7	0.7	0.61
−0.7	0.7	0.9	2.25	−0.7	0.7	0.9	1.15	−0.7	0.7	0.9	0.66
−0.7	**0.9**	0.5	**23.90**	−0.7	**0.9**	0.5	**1.77**	−0.7	**0.9**	0.5	**0.97**
−0.7	**0.9**	0.7	**19.67**	−0.7	**0.9**	0.7	**1.87**	−0.7	**0.9**	0.7	**0.97**
−0.7	**0.9**	0.9	**15.41**	−0.7	**0.9**	0.9	**1.87**	−0.7	**0.9**	0.9	**1.02**
−0.5	0.5	0.5	2.02	−0.5	0.5	0.5	1.01	−0.5	0.5	0.5	0.54
−0.5	0.5	0.7	2.49	−0.5	0.5	0.7	1.01	−0.5	0.5	0.7	0.56
−0.5	0.5	0.9	2.63	−0.5	0.5	0.9	1.11	−0.5	0.5	0.9	0.61
−0.5	0.7	0.5	3.07	−0.5	0.7	0.5	1.32	−0.5	0.7	0.5	0.72
−0.5	0.7	0.7	2.93	−0.5	0.7	0.7	1.35	−0.5	0.7	0.7	0.70
−0.5	0.7	0.9	3.15	−0.5	0.7	0.9	1.37	−0.5	0.7	0.9	0.79
−0.5	**0.9**	0.5	**16.92**	−0.5	**0.9**	0.5	**2.45**	−0.5	**0.9**	0.5	**1.20**
−0.5	**0.9**	0.7	**20.60**	−0.5	**0.9**	0.7	**2.41**	−0.5	**0.9**	0.7	**1.18**
−0.5	**0.9**	0.9	**18.54**	−0.5	**0.9**	0.9	**2.50**	−0.5	**0.9**	0.9	**1.27**
−0.3	0.5	0.5	24.14	−0.3	0.5	0.5	1.78	−0.3	0.5	0.5	0.97
−0.3	0.5	0.7	20.64	−0.3	0.5	0.7	1.79	−0.3	0.5	0.7	0.98
−0.3	0.5	0.9	19.42	−0.3	0.5	0.9	2.00	−0.3	0.5	0.9	1.12
−0.3	0.7	0.5	15.77	−0.3	0.7	0.5	2.45	−0.3	0.7	0.5	1.24
−0.3	0.7	0.7	22.71	−0.3	0.7	0.7	2.47	−0.3	0.7	0.7	1.21
−0.3	0.7	0.9	25.63	−0.3	0.7	0.9	2.52	−0.3	0.7	0.9	1.31
−0.3	**0.9**	0.5	**74.54**	−0.3	**0.9**	0.5	**16.87**	−0.3	**0.9**	0.5	**2.76**
−0.3	**0.9**	0.7	**160.67**	−0.3	**0.9**	0.7	**16.81**	−0.3	**0.9**	0.7	**2.46**
−0.3	**0.9**	0.9	**56.35**	−0.3	**0.9**	0.9	**17.54**	−0.3	**0.9**	0.9	**2.72**
−0.1	0.5	0.5	74.13	−0.1	0.5	0.5	16.97	−0.1	0.5	0.5	6.46
−0.1	0.5	0.7	66.07	−0.1	0.5	0.7	14.90	−0.1	0.5	0.7	5.94
−0.1	0.5	0.9	65.18	−0.1	0.5	0.9	14.31	−0.1	0.5	0.9	6.89
−0.1	0.7	0.5	143.98	−0.1	0.7	0.5	16.31	−0.1	0.7	0.5	12.92
−0.1	0.7	0.7	78.54	−0.1	0.7	0.7	18.28	−0.1	0.7	0.7	11.60
−0.1	0.7	0.9	309.00	−0.1	0.7	0.9	16.12	−0.1	0.7	0.9	12.59
−0.1	**0.9**	0.5	**56.16**	−0.1	**0.9**	0.5	**26.75**	−0.1	**0.9**	0.5	**54.43**
−0.1	**0.9**	0.7	**89.04**	−0.1	**0.9**	0.7	**26.06**	−0.1	**0.9**	0.7	**56.03**
−0.1	**0.9**	0.9	**30.25**	−0.1	**0.9**	0.9	**26.58**	−0.1	**0.9**	0.9	**79.19**

Table 3
Parameter Estimates, Multiple Regression Model.

Parameter	Estimate	*t*-test	*p*-value
α_0	34.0283	1.88*	0.060
α_1	31.8800	1.93*	0.053
α_2	7.4983	0.45	0.649
α_3	−48.7343	5.66***	< 0.001
α_4	−0.3395	3.58***	< 0.001
R^2	0.27		
df	130		

* Significant at 90%; ** Significant at 95%; *** Significant at 99%.

- Strong collinearity between explanatory variables of different modes (i.e., car travel time and rail travel time) did not have a significant impact on the VTTS variance estimates.
- In the more inelastic demand scenarios (i.e., elasticity values of −0.5, −0.3 and −0.1), the VTTS variance estimates were very high.
- As was expected, the greater was the sample size, the lower was the VTTS variance estimate. However, even with a very large sample (N = 10,000), when collinearity between variables of the same mode was strong (0.9) and demand was very inelastic (−0.1), the VTTS variance estimates were very high.

To complement the above analysis, we formulated and estimated a multiple linear regression model relating VTTS standard deviation (the explained variable) to the explanatory variables in Table 2. The model was specified as follows:

$$y_i = \alpha_0 + \alpha_1 \rho_{1i} + \alpha_2 \rho_{2i} + \alpha_3 |\eta_{ci}^{\text{car}}| + \alpha_4 \sqrt{N_i} + \varepsilon_i \tag{25}$$

where, for simulation scenario i,

- y_i is the standard deviation of the VTTS estimate,
- ρ_{1i} is the correlation between car travel time and car travel cost,
- ρ_{2i} is the correlation between car travel time and rail travel time,
- $|\eta_{ci}^{\text{car}}|$ is the absolute value of the price-elasticity of demand for the car mode,
- and N_i is the number of observations (using the values 1,000, 3000 and 10,000).

As can be seen in Table 2, the number of observations (i.e., the number of different scenarios) for (25) was $45 \times 3 = 135$.

The parameter estimates obtained for (25) are set out in Table 3. They indicate that correlation between car travel time and car travel cost, inelastic demand, and sample size (the control variable) all have statistically significant effects on the VTTS standard deviation estimate. These results are consistent with the theoretical results of (19) discussed in Section 3.1.

The impact of these effects may be strong enough to invalidate the VTTS estimates obtained using classic estimation methods such as maximum likelihood. Thus, when the variables for a given transport mode are strongly collinear (e.g., correlations of about 0.9 or more) and absolute price-elasticity of demand is low (e.g., less than |0.3|), sample sizes should by very large (in the tens of thousands). Often, however, in practical terms this will not be possible. In such a case, resort should be had to alternative estimation approaches based on the incorporation of constraints that can produce a robust analysis and generate reliable VTTS estimates.

The next section presents two such approaches, whose objective is precisely to mitigate the impact on VTTS estimates of strong collinearity and inelastic demand. Both methods are based on the classic technique of maximum likelihood but impose restrictions on the parameters to be estimated. The first method incorporates the ridge estimator, widely used when collinearity between explanatory variables is strong. The second method is a novel approach developed by the present authors that imposes exogenous constraints on price-elasticity of demand using external information.

3.3. Constrained estimation: Maximum likelihood with restrictions

The imposition of restrictions on parameters to be estimated is common practice in econometrics. It has proved to be useful in various situations depending on the theoretical or practical context of the modelling process in question. In particular, it has been employed for the purposes of strengthening consistency with economic theory (for example, the homogeneity constraint in demand models or the non-negativity restriction on production function coefficients), reducing model complexity (fewer parameters), mitigating collinearity and convergence problems, improving model stability and robustness (greater efficiency) and incorporating practical and regulatory considerations, among other possibilities.

In the case of transport demand models, since price-elasticity of demand values tend to be relatively homogeneous and robust (Graham and Glaister, 2004; Hörcher and Tirachini, 2021), they constitute a reasonable value range that can be applied as exogenous restrictions to assist in alleviating some of the above-discussed estimation problems.

The next two subsections present two constrained estimation methods: the classic ridge estimator and the proposed method involving direct exogenous restrictions on elasticity.

3.3.1. Ridge estimator

The ridge estimator is obtained by maximizing the likelihood function (13) but with the addition of quadratic restrictions on the parameters. The problem is specified as follows:

$$\begin{aligned} \max_{\beta_k^m} \quad & \ln L(\beta) \;=\; \sum_i \sum_m \delta_i^m \, \ln P_i^m(\beta) \\ \text{s.t.:} \quad & \sum_{m,k} \left(\beta_k^m\right)^2 \;\le\; C(\lambda) \end{aligned} \tag{26}$$

The penalized form of problem (26) that is then maximized is

$$\max_{\beta_k^m, \lambda} \quad \sum_i \sum_m \delta_i^m \, \ln P_i^m(\beta) \;+\; \lambda \sum_{m,k} \left(\beta_k^m\right)^2 \tag{27}$$

The value of λ is obtained using the cross-validation method (Martín-Baos et al., 2021).

3.3.2. Elasticity-Constrained estimator

For this proposed approach we bring the empirical evidence on the level of price-elasticity of demand into the maximum likelihood problem as prior information on the cost coefficient. Following the Bayesian approach to using prior information (Rossi et al., 2024), we place a Gaussian prior on β_c^m, namely $\beta_c^m \sim \mathcal{N}(\gamma_c^m \, \bar{\eta}_c^m, \, \tau^2)$, where γ_c^m is the sample-determined factor of eq. 18 (a known function of the data, not a parameter), and $\bar{\eta}_c^m$ is the external prior elasticity, different from the in-sample elasticity η_c^m of eq. (17). The prior mean is the value of β_c^m implied by eq. (18) when the elasticity is set at $\bar{\eta}_c^m$, and the prior variance τ^2 controls how strongly the prior pulls the estimate. The remaining coefficients are given flat priors, so the data alone identify them.

The choice of price-elasticity as the constrained quantity reflects its behavioural content. Price-elasticity measures the response of trip demand to price, grounded in the same consumer-theoretic requirement that $\beta_c^m < 0$, namely the positive marginal utility of income (Jara-Diaz, 2024). The prior $\bar{\eta}_c^m$ is chosen from this body of evidence, so any specific value can be justified by reference to published work, and a value outside the established range can be flagged as such. This prior expresses pre-data uncertainty about the fixed population value of β_c^m and is conceptually distinct from the distribution of β_c^m across individuals discussed in Section 2. Because $\gamma_c^m \, \bar{\eta}_c^m < 0$ and the prior is informative (τ^2 small), the mass placed on positive values is negligible; equivalently, the prior can be written as a normal truncated to $\beta_c^m < 0$, which leaves eq. (28) unchanged up to an additive constant.

The maximum a posteriori (MAP) estimator of β maximizes the log-posterior, which up to an additive constant equals the log-likelihood minus a quadratic penalty in $\beta_c^m - \gamma_c^m \, \bar{\eta}_c^m$,

$$\max_{\beta_k^m} \quad \sum_i \sum_m \delta_i^m \, \ln P_i^m(\beta) \;-\; \frac{1}{2\tau^2} \left(\beta_c^m - \gamma_c^m \, \bar{\eta}_c^m\right)^2 \tag{28}$$

which is the standard form of penalized maximum likelihood with a quadratic penalty (Hastie et al., 2009), with penalty weight $\mu \equiv 1/(2\tau^2) \ge 0$. The estimator differs from the classical ridge of eq. (27) in two respects. First, the penalty is centered at the external target $\gamma_c^m \, \bar{\eta}_c^m$ rather than at zero, so the estimator shrinks β_c^m toward a value supported by the literature rather than toward zero. Second, only the cost coefficient is penalized; the remaining coefficients are estimated without restrictions, keeping their usual interpretation.

The MAP problem (28) can be viewed as the quadratic-penalty relaxation of the equality-constrained maximum likelihood problem

$$\begin{aligned} \max_{\beta_k^m} \quad & \ln L(\beta) \;=\; \sum_i \sum_m \delta_i^m \, \ln P_i^m(\beta) \\ \text{s.t.:} \quad & \beta_c^m \;=\; \gamma_c^m \, \bar{\eta}_c^m \qquad (\mu) \end{aligned} \tag{29}$$

with μ playing the role of the Lagrange multiplier. The relaxation is exact in the limit $\tau^2 \to 0$ ($\mu \to \infty$), where the prior becomes infinitely strong and the equality is enforced; for finite $\tau^2 > 0$ the prior is informative but not absolute, and the MAP estimate is pulled toward the unconstrained maximum likelihood estimate as the data become more informative.

The Bayesian and penalized-likelihood interpretations make explicit the trade-off introduced by the constraint. The MAP combines two sources of information about β_c^m: the data, summarized in the unconstrained maximum likelihood estimate $\widehat{\beta}_c^{m,\,\mathrm{ML}}$, and the prior, the external target $\gamma_c^m \, \bar{\eta}_c^m$. Their relative weights are determined by the Fisher information from the sample and the prior precision $1/\tau^2$. When the data are informative on β_c^m (high t-ratio of $\widehat{\beta}_c^m$), the data dominate and the prior has little influence. When the data are weakly informative (the regime of strong collinearity and inelastic demand identified in Section 2 and Section 3), the prior dominates and substantially reduces the variance of the estimate. Bias is introduced whenever the prior mean differs from the true population value; its size is proportional to $\gamma_c^m \, (\bar{\eta}_c^m - \eta_c^{m,\,\mathrm{true}})$, weighted by the prior's relative influence over the data. In the asymptotic limit with fixed prior, the MAP converges to the MLE and the bias vanishes; in the finite-sample setting we focus on (weak identification of β_c^m), the prior has real weight by design, and so does the bias if the prior is wrong. This bias is controlled by the choice of $\bar{\eta}_c^m$, which should be drawn from the empirical literature on price-elasticity of demand, so the choice can be checked against published evidence rather than chosen arbitrarily. The simulations in Section 4 quantify the trade-off across scenarios; the Hausman test in Section 5 checks whether the bias is statistically detectable in any given application.

The interpretation of the resulting estimate follows from this construction. The elasticity-constrained value of travel time savings is the value of time conditional on the price-elasticity of demand being consistent with the external evidence, rather than an unconditional point identification of a population preference parameter. This conditional reading is appropriate for policy applications that require an internally coherent value-of-time and price-elasticity pair, such as cost-benefit analysis of toll roads, congestion-charging

schemes and fare reforms. The sensitivity analysis of Section 4.2 describes when the estimator is preferable: it varies $\bar{\eta}_c^m$ across the modal range and the true elasticity across a wider grid, quantifies the bias under misspecification, and identifies the regimes in which the constraint improves on maximum likelihood.

The value of $\bar{\eta}_c^m$ will depend on the particular context, but for urban transport systems the price-elasticity of demand in the short run typically varies between -0.2 and -0.4 (Goodwin, 1992; Holmgren, 2007; Brons et al., 2008; De Grange et al., 2013; Davis, 2021; Gillingham and Munk-Nielsen, 2019; Berry and Börjesson, 2024). Moreover, in a recent document by Bain and Cimon (2026), which considered a meta-analysis of several hundred scientific articles and technical studies, they conclude that the median of price-demand elasticity is -0.21, and that 92% of the reviewed values fall between -0.5 and -0.03. For purposes of estimating the VTTS, this elasticity restriction is considered to be exogenous.

In what follows we focus on strong priors: the penalty is set large enough for the implied elasticity to reach the target $\bar{\eta}_c^m$ up to a small tolerance. Intermediate τ^2 is permitted by the framework, with cross-validation providing a data-driven choice.

4. Comparative analysis of estimation methods by simulation

Simulations were carried out to compare the performances of the classic (unconstrained) maximum likelihood method and the two alternative methods described in Section 3.3, that is, the ridge estimator and our proposed elasticity-constrained approach. An MNL model was used with two mode alternatives, private car and rail, and two explanatory variables, travel time and travel cost (i.e., price).

In light of the preliminary results reported above in Section 3.2, a total of 24 different model scenarios were simulated covering all those for which the combinations of sample size, degree of collinearity and level of elasticity as reported in Table 2 resulted in standard deviations for VTTS that were greater than its population mean (VTTS = 5). Values were derived for variance, bias and mean squared error. The results obtained with the three methods are reported and compared in Section 4.1 and a number of recommendations based on the comparisons are discussed in Section 4.2.

4.1. Simulations results for the three estimation methods

The simulation results on bias, standard error and mean squared error for the 24 scenarios are shown in Table 4 for sample sizes of 1000 observations. As can be seen, in the case of bias, classic maximum likelihood (hereafter simply “maximum likelihood” or ML) tended to generate lower values, the results in 10 of the 24 scenarios being the lowest of all three methods. The ridge estimator showed the lowest bias in 9 scenarios while the proposed elasticity-constrained method (Elast-0.4) exhibited the lowest bias in only 5. This last result was expected given that they were the scenarios with elasticities between -0.5 and -0.3, that is, the values increasingly approaching the cutoff of -0.4 imposed by the restriction. Thus, on bias the maximum likelihood and ridge methods displayed better relative performance.

As regards standard deviation, the Table 4 results reveal that maximum likelihood clearly generated the highest values. The proposed method displayed the lowest values in 18 scenarios and the ridge method in 6. For mean squared error, the table again shows that maximum likelihood performed worst whereas the proposed method obtained the lowest value in 21 of the 24 scenarios.

Table 5 reports the comparative simulation results for the 12 scenarios in Table 2 with a sample size of 3000 observations for which the standard deviation of VTTS was greater than its mean. Given that the value set for the latter was 5, the results for bias were in general relatively high. The ridge method tended to produce the lowest bias values, outperforming the other two methods in 8 of the scenarios, while maximum likelihood had the best performance in the remaining 4. The values for the proposed method were systematically higher than those with the other two approaches. Thus, on the bias criterion the results point mainly to the ridge method as the best performer. Regarding standard deviation, the Table 5 results show plainly that in all 12 scenarios the proposed method obtained smaller standard deviations. The same was true for mean squared error, which is the sum of the squares of bias and standard deviation.

Table 6 sets out the comparative simulation results for the 9 scenarios in Table 2 where the standard deviation of VTTS was greater than its mean with a sample size of 10,000 observations. As is evident, these results were similar to those for 3000 observations in Table 5. Thus, the lowest bias levels were produced by ridge in all cases, followed by maximum likelihood (3 of 9 cases). Furthermore, in the most critical scenarios with greatest demand inelasticity, ridge was always the least biased. As regards standard deviation and mean squared error, maximum likelihood again performed the worst while our proposed method obtained the best outcomes.

4.2. Sensitivity to a misspecified prior elasticity

A reasonable objection to the elasticity-constrained estimator is that the prior $\bar{\eta}$, drawn from the empirical literature, may differ from the true elasticity that generated the data at hand. Because the constraint pins $\hat{\beta}_c^{\text{car}}$ to the value implied by $\bar{\eta}$, a wrong prior introduces a bias that the estimator itself does not undo. The questions of interest are then how large that bias can plausibly be, and whether the resulting estimator still outperforms maximum likelihood in the regimes where the constraint is useful.

The Elast-0.4 column of Tables 4–6 provides part of the answer by construction: the prior is held fixed at $\bar{\eta} = -0.4$ while the true population elasticity η_c^{true} varies across rows, so each entry of that column already quantifies a misspecification by a known amount. To complete the picture we also vary the prior. We sweep $\bar{\eta} \in \{-0.5, -0.4, -0.3, -0.2\}$, a grid that covers the usual short-run values reported by the empirical literature on price-elasticity of urban transport demand (Holmgren, 2007; Brons et al., 2008; Bain and Cimon, 2026), and let the true elasticity take the levels of Table 1, $\eta_c^{\text{true}} \in \{-0.7, -0.5, -0.3, -0.1\}$, for three levels of within-mode

Table 4
Bias, SE and MSE by VTTS Estimation Method (Simulations with N = 1,000).

			Bias			Standard Error			MSE		
η_c^{car}	ρ_1	ρ_2	ML	Ridge	Elast-0.4	ML	Ridge	Elast-0.4	ML	Ridge	Elast-0.4
−0.7	0.9	0.5	1.81	−3.09	5.00	6.95	0.62	1.30	51.57	9.93	26.70
−0.7	0.9	0.7	1.79	−3.11	5.07	6.39	0.65	1.54	43.98	10.08	28.07
−0.7	0.9	0.9	1.73	−3.32	5.11	6.76	0.68	2.01	48.65	11.50	30.21
−0.5	0.9	0.5	3.14	−2.82	2.10	14.82	0.92	1.09	229.63	8.80	5.58
−0.5	0.9	0.7	2.63	−2.92	2.10	13.72	0.92	1.25	195.22	9.39	5.97
−0.5	0.9	0.9	2.42	−3.06	2.15	12.72	1.00	1.84	167.53	10.37	8.04
−0.3	0.5	0.5	1.79	−0.48	−1.23	7.21	3.61	1.01	55.21	13.29	2.53
−0.3	0.5	0.7	1.92	−0.45	−1.17	63.49	7.12	1.20	4,034.59	50.85	2.83
−0.3	0.5	0.9	1.58	−1.53	−1.22	6.25	2.64	1.94	41.59	9.30	5.25
−0.3	0.7	0.5	3.59	−0.72	−1.43	53.17	4.41	1.03	2,839.72	19.96	3.10
−0.3	0.7	0.7	1.06	−0.67	−1.40	21.79	6.39	1.15	475.73	41.34	3.31
−0.3	0.7	0.9	1.22	−2.17	−1.34	21.44	9.35	1.75	461.01	92.17	4.86
−0.3	0.9	0.5	0.96	−2.31	−1.63	34.22	3.47	1.05	1,171.75	17.37	3.77
−0.3	0.9	0.7	−0.59	−2.26	−1.68	33.95	3.58	1.12	1,153.12	17.90	4.08
−0.3	0.9	0.9	2.07	−2.25	−1.69	88.22	12.62	1.84	7,787.32	164.28	6.26
−0.1	0.5	0.5	−1.64	−1.37	−4.08	52.94	32.68	1.14	2,805.30	1,069.79	17.95
−0.1	0.5	0.7	−0.04	−1.71	−4.04	40.20	18.00	1.31	1,615.99	326.89	18.01
−0.1	0.5	0.9	9.13	−3.20	−4.04	204.08	13.18	2.10	41,732.76	184.05	20.72
−0.1	0.7	0.5	−1.24	−2.31	−4.55	31.82	37.72	1.08	1,014.25	1,427.76	21.84
−0.1	0.7	0.7	−7.44	−3.45	−4.54	112.72	33.65	1.26	12,762.21	1,143.94	22.25
−0.1	0.7	0.9	−2.96	−3.76	−4.58	65.77	23.36	1.98	4,334.44	559.71	24.92
−0.1	0.9	0.5	−4.39	−4.64	−5.09	26.94	17.67	1.01	744.76	333.75	26.96
−0.1	0.9	0.7	−5.42	−4.77	−5.10	14.38	17.18	1.22	236.35	317.81	27.51
−0.1	0.9	0.9	−5.45	−4.95	−5.21	23.16	10.34	1.88	566.26	131.30	30.64

time-cost collinearity, $\rho_1 \in \{0.5, 0.7, 0.9\}$, with the cross-mode time correlation fixed at $\rho_2 = 0.7$, at the most adverse sample size studied, N = 1,000. Each cell is computed over 1,000 Monte Carlo replications with the same $\pm 3\sigma$ outlier filter applied to Tables 4–6. Fig. 1 reports the resulting mean squared error of $\widehat{\phi}^{car}$.

The bias of $\widehat{\phi}^{car}$ is approximately monotone in the misspecification distance $\bar{\eta} - \eta_c^{true}$, with signs that match the analytic prediction of Section 3: when $|\bar{\eta}| > |\eta_c^{true}|$, the prior pulls $|\widehat{\beta}_c^{car}|$ above its true value and $\widehat{\phi}^{car} = \widehat{\beta}_t^{car} / \widehat{\beta}_c^{car}$ is biased downward in magnitude; the pattern reverses on the opposite side. The minimum MSE within each panel lies on the diagonal of approximately correct priors and is below 6 across the three collinearity regimes.

Comparison with the reference columns separates two regimes. In the adverse regime ($\rho_1 = 0.9$), maximum likelihood MSE ranges from 234 to 828 and the elasticity-constrained estimator's MSE is one to three orders of magnitude lower in nearly every cell of the grid, even when the prior is moderately misspecified; the constraint underperforms ML only at the most extreme misspecification on the grid ($\bar{\eta} = -0.2$ against $\eta_c^{true} = -0.7$, distance 0.5). In the non-adverse regime ($\rho_1 = 0.5$), the picture changes: maximum likelihood is competitive when $|\eta_c^{true}| \geq 0.5$ (MSE between 3 and 4), and the constrained estimator with a misspecified prior can underperform it. The operational implication is consistent with the design intent of the proposal: the elasticity-constrained method is the appropriate tool in the adverse regime, where the data alone do not identify β_c^{car} precisely; in well-behaved data it is not the default choice.

Note that at $\eta_c^{true} = -0.1$ the maximum-likelihood MSE decreases as ρ_1 rises (from 3,795 at $\rho_1 = 0.5$ to 459 at $\rho_1 = 0.9$), which seems to contradict the claim that more collinearity worsens identification. However, the second panel of Fig. 1 shows that the proportion of replications with $\widehat{\beta}_c^{car} > 0$ rises from 21% to 33% for maximum likelihood (and stays around 18%–22% for ridge), all of them behaviourally invalid. Thus, this pattern reveals two failure modes, both undesirable. At low ρ_1, values of $\widehat{\beta}_c^{car}$ cluster around the true value of β_c^{car}, which is small in magnitude when $\eta_c^{true} = -0.1$. This produces a low share of wrong-signed $\widehat{\beta}_c^{car}$, but, due to the small

Table 5
Bias, SE and MSE by VTTS Estimation Method (Simulations with $N = 3{,}000$).

			Bias			Standard Error			MSE		
η_c^{car}	ρ_1	ρ_2	ML	Ridge	Elast-0.4	ML	Ridge	Elast-0.4	ML	Ridge	Elast-0.4
−0.3	0.9	0.5	4.44	−0.41	−1.52	76.07	10.88	0.69	5,805.73	118.46	2.79
−0.3	0.9	0.7	1.23	−1.07	−1.57	9.19	5.86	0.78	85.91	35.45	3.07
−0.3	0.9	0.9	1.62	−1.42	−1.57	14.48	2.62	1.10	212.21	8.88	3.67
−0.1	0.5	0.5	1.84	−0.23	−3.56	13.24	12.55	0.85	178.62	157.54	13.40
−0.1	0.5	0.7	−0.03	−1.73	−3.57	21.69	98.09	0.99	470.55	9,624.77	13.73
−0.1	0.5	0.9	0.93	−0.78	−3.59	30.57	18.55	1.58	935.15	344.65	15.38
−0.1	0.7	0.5	−1.65	−4.33	−4.12	33.17	49.89	0.75	1,103.27	2,507.91	17.54
−0.1	0.7	0.7	0.85	−0.40	−4.11	46.02	23.22	0.92	2,118.99	539.25	17.72
−0.1	0.7	0.9	2.54	−1.95	−4.11	41.20	18.39	1.45	1,703.87	341.89	19.02
−0.1	0.9	0.5	−3.83	−5.00	−4.90	29.81	68.98	0.66	903.31	4,783.43	24.42
−0.1	0.9	0.7	−1.98	−3.13	−4.94	80.74	26.81	0.79	6,523.09	728.58	25.08
−0.1	0.9	0.9	−5.48	−2.82	−4.94	31.60	24.35	1.15	1,028.45	601.09	25.72

Table 6
Bias, SE and MSE by VTTS Estimation Method (Simulations with $N = 10{,}000$).

			Bias			Standard Error			MSE		
η_c^{car}	ρ_1	ρ_2	ML	Ridge	Elast-0.4	ML	Ridge	Elast-0.4	ML	Ridge	Elast-0.4
−0.1	0.5	0.5	0.81	0.58	−2.45	3.21	3.00	0.84	10.96	9.33	6.68
−0.1	0.5	0.7	0.69	0.43	−2.47	3.28	3.05	0.91	11.25	9.47	6.94
−0.1	0.5	0.9	0.68	0.21	−2.44	3.64	3.26	1.26	13.74	10.68	7.56
−0.1	0.7	0.5	2.83	1.52	−3.08	16.96	8.79	0.78	295.79	79.52	10.09
−0.1	0.7	0.7	3.19	1.37	−3.09	24.23	11.04	0.88	597.31	123.70	10.35
−0.1	0.7	0.9	1.41	0.62	−3.11	5.60	4.55	1.13	33.37	21.14	10.93
−0.1	0.9	0.5	1.24	0.11	−4.30	49.33	42.32	0.55	2,434.85	1,791.05	18.76
−0.1	0.9	0.7	1.73	−0.40	−4.35	21.71	12.50	0.59	474.24	156.53	19.26
−0.1	0.9	0.9	0.74	−0.03	−4.35	32.93	18.45	0.84	1,084.87	340.34	19.65

magnitude, the VTTS estimates $\widehat{\beta}_t^{car}/\widehat{\beta}_c^{car}$ explode and produce extreme values of MSE. Conversely, at high ρ_1, values of $\widehat{\beta}_c^{car}$ have a wider distribution due to poor identification. This produces a higher share of wrong-signed $\widehat{\beta}_c^{car}$, but, since these values are farther from zero, the VTTS estimates are lower in magnitude (although carrying the wrong sign). In summary, a lower MSE is not the only relevant metric in this regime, which motivates the microeconomic sign test on $\widehat{\beta}_c^{car}$ applied in the empirical applications of Section 5.

4.3. Summary of results and recommendations for choice of method

The results presented above point to a number of specific recommendations regarding the appropriate choice of method among the three applied in our simulations for estimating VTTS in different transport demand modelling contexts.

To begin with, maximum likelihood is particularly useful when the data are well behaved, that is, when the travel time and cost variables are not strongly collinear (correlation below 0.7) and price-elasticity is high in absolute value (i.e., above |0.7|), that is, when demand is relatively sensitive to cost. Under such conditions, maximum likelihood tends to produce estimates low in bias. However, since it is highly sensitive to collinearity and inelastic demand, when such problems are present the variance of the estimates explodes and the results are not reliable, even if bias is low. Furthermore, in difficult conditions maximum likelihood performance does not improve substantially even when sample size is large. Thus, in contexts where collinearity is strong and demand is inelastic, maximum likelihood is not recommended.

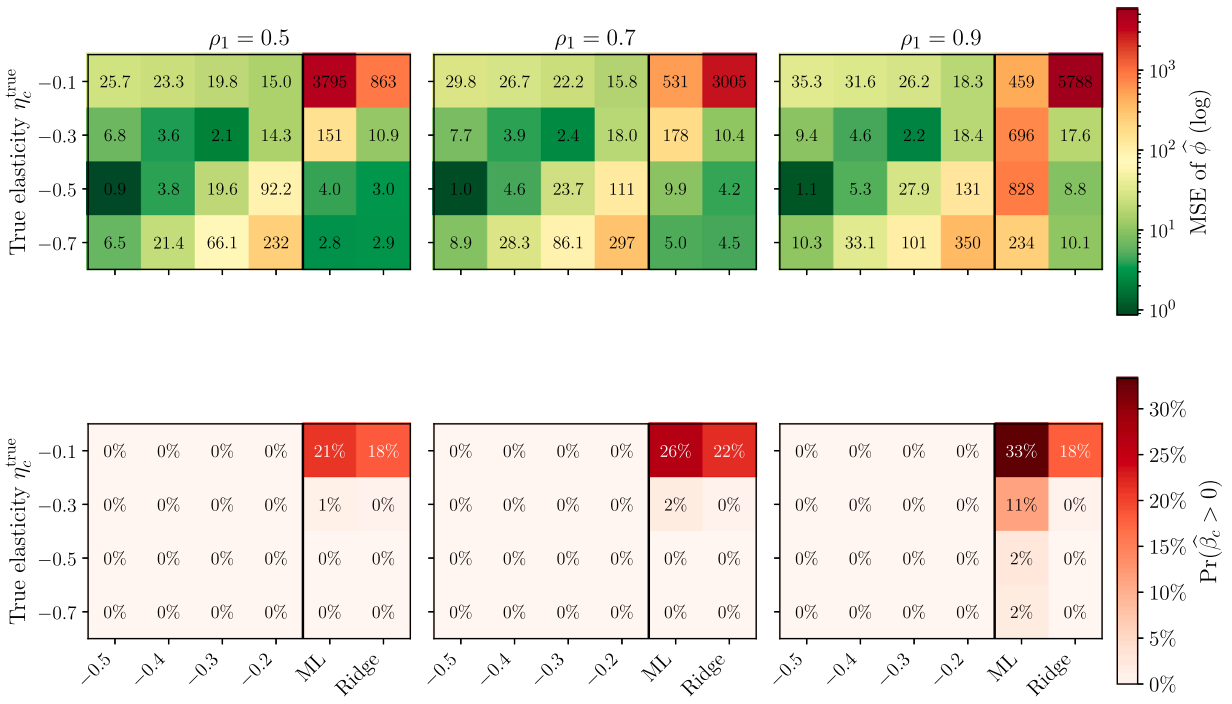

Fig. 1. MSE of $\hat{\phi}^{\text{car}}$ (top) and probability of sign-flip $\Pr(\hat{\beta}_c^{\text{car}} > 0)$ (bottom), for combinations of true elasticity η_c^{true} (rows) and prior $\bar{\eta}$ (Prior columns), at three levels of time-cost collinearity ρ_1 (panels). Maximum likelihood and ridge shown as reference. MSE on log colour scale; sign-flip on linear scale.

As for ridge, it is an intermediate alternative in the sense that it is suitable for contexts where collinearity is moderate or even strong but no precise information on elasticities is available. Its main strength is that it significantly reduces variance without imposing external restrictions, thereby improving the estimates' mean squared error. Its results do exhibit a certain bias, but the amount is generally acceptable and less than that of the proposed elasticity-constrained method, particularly when the latter's imposed restriction is distant from the actual elasticity level. Ridge is especially useful when sample size is small and it is desired to stabilize the parameter estimates.

Finally, the proposed method imposing an exogenous restriction on a specific elasticity level (for example, −0.4) is the most efficient in adverse conditions where collinearity is strong (correlation ≈ 0.9) and demand is inelastic (absolute values of less than |0.3|). Although as noted it can introduce bias if the restriction is distant from the actual value, it obtains notable reductions in the estimates' variance and mean squared error. It is thus a valuable tool when information on elasticities is available and achieving a robust VTTS estimate is required, especially with small samples. Care should nevertheless be taken in its application, as its use where not properly justified could lead to distorted results.

A summary of the results for the three discussed methods on the different statistical criteria is set out in Table 7. Briefly put, they indicate that the choice of method must balance the need for accuracy and robustness with the availability of external information. Maximum likelihood is often appropriate for ideal contexts, ridge is recommended when improvement over maximum likelihood is desired through the imposition of hard restrictions, and the proposed elasticity-constrained method is the best option for complex settings where control of variance is important.

5. Applications with real data

This section evaluates the proposed estimator on two real-world datasets chosen to cover different collinearity regimes. The Swiss-metro stated-preference dataset represents the moderate-collinearity regime ($\rho = 0.72$), in which the unrestricted maximum likelihood estimator is well behaved and the constrained estimator competes on efficiency grounds. The London Passenger Mode Choice (LPMC) revealed-preference dataset covers the high-collinearity regime ($\rho = 0.94$), in which the unrestricted maximum likelihood estimator produces a wrong-signed cost coefficient and the question shifts from efficiency to behavioural validity.

Three diagnostics accompany the parameter estimates in each application. The first is a microeconomic sign test on the cost coefficient: $H_0 : \beta_c \geq 0$ against $H_1 : \beta_c < 0$, with rejection at the 5% level when the t-statistic $\hat{\beta}_c/\text{SE}(\hat{\beta}_c)$ falls below −1.645. The test examines the requirement that the marginal utility of income be positive (DeSerpa, 1971; Jara-Díaz and Farah, 1987; Jara-Diaz, 2024), and depends only on the estimator under examination, not on any benchmark. The second is the mean squared error of each estimator, computed as the sum of variance and squared bias relative to the ML estimate taken as a proxy for the true parameter. This proxy works when ML is consistent and reasonably precise, but loses its meaning when the ML estimate is itself behaviourally invalid.

Table 7
Comparative Summary of VTTS Estimation Method Results.

	Method		
Criterion	Classic ML	Ridge	Elast-0.4
Bias	Generally the least biased. Best results when actual elasticity is between |0.7| and |1.0|.	Performs well, especially with strong collinearity or low elasticity.	Generally the most biased. Less biased if elasticity restriction is close to the actual level.
Standard error	Highest in almost all scenarios.	Better than ML, worse than elasticity-restricted.	Least in most scenarios.
MSE	The highest (worst) in most scenarios.	Performs well, competitive in moderate scenarios.	The lowest (best) in most scenarios.
Robustness to collinearity	Very sensitive to strong collinearity.	Robust.	Very robust.
Robustness to low elasticity	Performs poorly with elasticities of less than |0.3|.	Performs considerably better than ML.	Performs well; specifically designed for such scenarios.
Effect of increasing sample size	Improves, but remains the least efficient.	Improves notably.	Is already efficient even with small samples.

The third is the Hausman test against the ML baseline on the joint parameter (β_t, β_c), which we interpret as a detectability diagnostic rather than a bias test: a small H indicates that the constrained and unrestricted estimators are statistically indistinguishable at the current sample size, while a large H indicates a detectable departure whose direction depends on which estimator is held to be more reliable. The reinterpretation is needed because the classical Hausman framework assumes ML consistency under the null, an assumption that does not always hold in the regimes the paper targets. When identification of β_c is weak, the ML estimator, although consistent in principle, may show substantial finite-sample bias and produce behaviourally invalid estimates, weakening its role as a reliable null benchmark.

The informativeness of each diagnostic depends on whether ML is behaviourally valid in the regime at hand. When the ML cost coefficient satisfies the microeconomic sign requirement, ML is taken to be approximately unbiased: MSE relative to ML measures the efficiency gain of the constrained estimator, the Hausman test recovers its conventional interpretation as a bias detector, and the sign test serves as a minimal check that all estimators are expected to pass. When the ML cost coefficient violates the sign requirement, this assumption is dropped: the sign test becomes the discriminating diagnostic, MSE loses its reference point because distance from a behaviourally invalid baseline is uninformative, and the Hausman statistic must be read as detectability rather than bias. Reporting in each subsection follows this logic.

5.1. *The swissmetro application*

We begin with the Swissmetro dataset, developed by Bierlaire et al. (2001) from stated preference data collected on trains running between the Swiss cities of St. Gallen and Geneva during March 1998. The responses relate to user reactions to the Swissmetro system's innovative mode of transport compared to the established alternatives of private car and rail.

Our focus is on the mode choice between private car and Swissmetro. We select for inclusion in the application those individuals who do not have an annual fare pass, as having one considerably lowers the user's perceived travel cost. We then filter the data to keep only individuals who have practical access to both modes, ensuring the one they opt for reflects a genuine choice between two real options.

The utility functions specified for the two modes are:

$$U_i^{\text{car}} = \beta_0^{\text{car}} + \beta_t^{\text{car}}\, t_i^{\text{car}} + \beta_c^{\text{car}}\, c_i^{\text{car}} + \varepsilon_i^{\text{car}}, \qquad \varepsilon_i^{\text{car}} \sim \text{Gumbel} \tag{30}$$

$$U_i^{\text{sm}} = \beta_0^{\text{sm}} + \beta_t^{\text{sm}}\, t_i^{\text{sm}} + \beta_c^{\text{sm}}\, c_i^{\text{sm}} + \varepsilon_i^{\text{sm}}, \qquad \varepsilon_i^{\text{sm}} \sim \text{Gumbel} \tag{31}$$

where t_i^{car} and c_i^{car} are individual i's travel time (in minutes) and cost (in CHF) variables for the car mode while t_i^{sm} and c_i^{sm} are i's travel time and cost variables for the Swissmetro mode. The correlation between the car mode variables is 0.72.

In addition to maximum likelihood and ridge, we estimate the VTTS parameters with the proposed method for elasticity restrictions of −0.4 and −0.3. The results for the car mode are shown in Table 8. The car travel time parameter β_t^{car} is significantly negative with all methods, consistent with economic theory in that an increase in travel time reduces utility. The car travel cost parameter β_c^{car} is also negative and statistically significant in every case.

The VTTS parameter ϕ^{car} ranges from 140.80 (Elast-0.4) to 162.28 (ML and Elast-0.3) CHF per hour, and is statistically significant in all cases. The mean squared error of the ML estimate is 1,098, ridge 1,084, and Elast-0.4 obtains 896, an 18% reduction over ML. Elast-0.3 yields estimates very close to ML because the ML elasticity is already near the target of −0.3; the constraint still provides a slight variance reduction. The MSE gain for Elast-0.4 comes mainly from variance reduction: the standard error of $\widehat{\phi}^{\text{car}}$ falls from 33.11 (ML) to 20.86 (Elast-0.4) CHF per hour, a 37% reduction.

The price-elasticity values $\widehat{\eta}_c^{\text{car}}$ are negative across all methods, consistent with economic theory. Ridge shifts the ML value of −0.296 slightly to −0.292, and Elast-0.4 reaches −0.336. All four estimators pass the sign test (Table 8). Table 9 reports the Hausman test against the ML baseline on the joint parameters β_t^{car} and β_c^{car}.

Table 8
Parameter Estimates for the Swissmetro Application.

Parameter	Maximum Likelihood			Ridge			Elast-0.4			Elast-0.3		
	Estimate	Std Err	t-test	Estimate	Std Err	t-test	Estimate	Std Err	t-test	Estimate	Std Err	t-test
β_t^{car}	−0.01301	0.00063	−20.54	−0.01279	0.00062	−20.51	−0.01278	0.00060	−21.20	−0.01301	0.00063	−20.55
β_c^{car}	−0.00481	0.00084	− 5.71	−0.00474	0.00083	− 5.72	−0.00545	0.00067	− 8.19	−0.00481	0.00084	− 5.72
ϕ^{car} (*)	162.28	33.11	4.90	161.76	32.92	4.91	140.80	20.86	6.75	162.27	33.04	4.91
MSE	1,098			1,084			896			1,091		
η_c^{car}	−0.296			−0.292			−0.336			−0.296		
Sign test	$\beta_c < 0$			$\beta_c < 0$			$\beta_c < 0$			$\beta_c < 0$		

(*) VTTS is expressed in CHF per hour. VTTS variance was obtained using the delta method (Huang and Qian, 2025).

Table 9
Hausman Test Results for the Swissmetro Application.

Hypothesis	Estimate (H)	Critical value (95%)	Result
$\widehat{\beta}^{\text{ML}} = \widehat{\beta}^{\text{Ridge}}$	7.767	5.991	Rejected
$\widehat{\beta}^{\text{ML}} = \widehat{\beta}^{\text{Elast-0.4}}$	1.420	5.991	Not rejected
$\widehat{\beta}^{\text{ML}} = \widehat{\beta}^{\text{Elast-0.3}}$	0.000	5.991	Not rejected

Table 10
Parameter Estimates for the LPMC Application.

Parameter	Maximum Likelihood			Ridge			Elast-0.4			Elast-0.3		
	Estimate	Std Err	t-test	Estimate	Std Err	t-test	Estimate	Std Err	t-test	Estimate	Std Err	t-test
β_t^{drive}	−9.9945	0.1584	−63.08	−0.0302	0.0048	− 6.25	−5.1735	0.1031	−50.16	−5.6356	0.1026	−54.92
β_c^{drive}	1.1087	0.0429	25.85	−0.0622	0.0045	−13.92	−0.7858	0.0109	−72.04	−0.5731	0.0108	−53.07
ϕ^{drive} (*)	−9.01	0.25	−35.62	0.49	0.09	5.61	6.58	0.18	36.43	9.83	0.29	33.63
η_c^{drive}	+0.450			−0.022			−0.341			−0.246		
Sign test	$\beta_c > 0$			$\beta_c < 0$			$\beta_c < 0$			$\beta_c < 0$		

(*) VTTS is expressed in GBP per hour. VTTS variance was obtained using the delta method (Huang and Qian, 2025).

The ridge departure from ML is detectable at the 5% level, reflecting the shrinkage of β_t^{car} toward zero. The Elast-0.4 departure from ML is not detectable ($H = 1.42$), so imposing the elasticity restriction produces no statistically distinguishable departure from the ML estimate at this sample size. The Elast-0.3 departure from ML is also not detectable, consistent with the proximity of the ML elasticity to the target discussed above. The empirical picture for Swissmetro is variance reduction at target −0.4 without a detectable departure from ML.

5.2. The london passenger mode choice application

The LPMC dataset (Hillel et al., 2018) contains 81,086 trips made by Londoners between 2012 and 2015 across four modes (walking, cycling, public transit, driving). We restrict the analysis to a binary choice between driving and public transit, dropping trips with a congestion-charge component to isolate the standard revealed-preference regime in which driving cost is imputed as a fixed rate per kilometre of fuel. The resulting sample has $N = 57{,}626$ trips, with within-driving correlation between cost and time of $\rho = 0.94$, well above the level the simulations in Section 4 identify as adverse. The utility specification follows Eqs. (30)–(31), with public transit in place of Swissmetro; cost is in GBP and time in hours.

Table 10 reports estimates for the driving alternative. The unrestricted maximum likelihood yields $\widehat{\beta}_c^{\text{drive}} = +1.109$ with a t-statistic of +25.85: a cost coefficient that is positive, statistically distinguishable from zero, and inconsistent with the microeconomic sign requirement. The implied own-price elasticity is +0.450, and the implied VTTS is −9.01 GBP per hour. The wrong sign is not a finite-sample artefact at this N but a structural consequence of the within-mode collinearity ($\rho = 0.94$) between fuel cost and driving time, the regime that Section 3.2 identifies analytically.

Ridge shrinks both coefficients aggressively. The cost coefficient becomes negative (−0.062) but the time coefficient is pushed close to zero, producing a VTTS estimate of 0.49 GBP per hour that is far below plausible commute-time valuations. The elasticity-constrained estimator approaches each literature target. Elast-0.4 yields $\widehat{\phi}^{\text{drive}} = 6.58$ GBP per hour with $\widehat{\eta}_c = -0.341$, and Elast-0.3 yields 9.83 GBP per hour with $\widehat{\eta}_c = -0.246$. Both VTTS values are consistent with meta-analytic evidence on UK car commute travel (Wardman et al., 2016), and both estimators pass the microeconomic sign test.

The Hausman test (Table 11) rejects every comparison against ML by a wide margin. Under the detectability framing of Section 5 the test correctly registers that all alternatives differ statistically from a maximum likelihood baseline that is itself wrong-signed. The empirically meaningful diagnostics in this regime are the sign test (ML fails, the alternatives pass) and the consistency of the implied VTTS with the meta-analytic literature.

Table 11
Hausman Test Results for the LPMC Application.

Hypothesis	Estimate (H)	Critical value (95%)	Result
$\widehat{\beta}^{\mathrm{ML}} = \widehat{\beta}^{\mathrm{Ridge}}$	5,242	5.991	Rejected
$\widehat{\beta}^{\mathrm{ML}} = \widehat{\beta}^{\mathrm{Elast\text{-}0.4}}$	41,306	5.991	Rejected
$\widehat{\beta}^{\mathrm{ML}} = \widehat{\beta}^{\mathrm{Elast\text{-}0.3}}$	6,390	5.991	Rejected

Table 12
Willingness-to-Pay-Space Estimates.

	Swissmetro (car)				LPMC (drive)			
	σ_ϕ^m=0		σ_ϕ^m free		σ_ϕ^m=0		σ_ϕ^m free	
Parameter	Est.	Std Err	Est.	Std Err	Est.	Std Err	Est.	Std Err
β_c^m	−0.00481	0.00080	−0.00521	0.00091	-8.28×10^{-4}	3.20×10^{-5}	−0.01796	0.00655
μ_ϕ^m	0.9950	0.1932	1.0122	0.2023	9.0333	0.0395	7.9760	0.3700
σ_ϕ^m	0	—	0.0761	0.0527	0	—	0.2170	0.0223
$\widehat{\mathbb{E}}[\phi^m]$ (*)	162.28	—	165.58	33.50	8,378	—	2,980	1,101
$\widehat{\eta}_c^m$	−0.296		−0.320		−0.0003		−0.008	
Sign test	$\beta_c < 0$		$\beta_c < 0$		$\beta_c < 0$		$\beta_c < 0$	

(*) $\widehat{\mathbb{E}}[\phi^m]$ in CHF per hour for Swissmetro, GBP per hour for LPMC. Estimation by simulated maximum likelihood with $R = 250$ Halton draws; standard errors from a numerical Hessian at the converged point.

5.3. Comparison with the willingness-to-Pay-Space approach

A well-established route in the literature is willingness-to-pay-space estimation (Train and Weeks, 2005). Instead of estimating β_t and β_c separately and deriving the VTTS as their ratio, the model is reparameterized so that the VTTS appears directly as a parameter, typically with a lognormal or similar distribution capturing preference heterogeneity across individuals. A related route, starting with Fosgerau (2007), models choice as $y = \mathbf{1}[\mathrm{WTP} > v]$ with WTP itself as the random object. This subsection compares both estimators with the elasticity-constrained estimator on the same two datasets used in Sections 5.1 and 5.2, focusing on the same car (Swissmetro) and drive (LPMC) alternatives.

The specification uses mode-specific cost coefficients and mode-specific lognormal random willingness-to-pay,

$$U_i^{\mathrm{car}} = \beta_0^{\mathrm{car}} + \beta_c^{\mathrm{car}}\left(\phi_i^{\mathrm{car}}\, t_i^{\mathrm{car}} + c_i^{\mathrm{car}}\right) + \varepsilon_i^{\mathrm{car}}, \tag{32}$$

$$\phi_i^m \sim \mathrm{Lognormal}\left(\mu_\phi^m,\, (\sigma_\phi^m)^2\right), \quad \phi_i^{\mathrm{car}} \perp \phi_i^{\mathrm{alt}}, \tag{33}$$

with an analogous expression for the alternative mode. The location parameter μ_ϕ^m sets the level of the willingness-to-pay (the median equals $\exp(\mu_\phi^m)$), and the scale parameter σ_ϕ^m governs the dispersion of preferences across individuals. The lognormal parameterization imposes $\phi^m > 0$ at every value of σ_ϕ^m, including zero. We report two configurations: $\sigma_\phi^m = 0$, in which the willingness-to-pay is deterministic and the specification is a one-to-one reparameterization of the preference-space ML up to the sign constraint on ϕ^m; and σ_ϕ^m free, the standard mixed-logit willingness-to-pay-space estimator. Both configurations are estimated by simulated maximum likelihood with $R = 250$ two-dimensional Halton draws.

Table 12 reports the estimates. In Swissmetro the deterministic limit $\sigma_\phi^m = 0$ reproduces the preference-space ML of Table 8: same point estimates of β_c^{car} and same implied $\widehat{\phi}^{\mathrm{car}}$ (162.28 CHF per hour), as expected from the reparameterization. The mixed-logit specification with σ_ϕ^m free captures modest heterogeneity in the car alternative ($\widehat{\sigma}_\phi^{\mathrm{car}} = 0.08$) and a higher implied mean willingness-to-pay of 165.58 CHF per hour. The implied own-price elasticity moves from −0.296 to −0.320, within the range used to motivate the elasticity-constrained estimator. The willingness-to-pay-space approach and the elasticity-constrained approach reach similar substantive conclusions in this dataset.

In LPMC the willingness-to-pay-space specification produces a sign-consistent β_c^{drive} at $\sigma_\phi^m = 0$, in contrast to the wrong-signed preference-space ML. The mechanism is the lognormal parameterization: the constraint $\phi^{\mathrm{drive}} > 0$ combined with the reparameterization $\beta_t^{\mathrm{drive}} = \beta_c^{\mathrm{drive}}\,\phi^{\mathrm{drive}}$ forces β_c^{drive} to inherit the sign of the well-identified β_t^{drive}. The sign is rescued, but the level of ϕ^{drive} is not: the data inform the product $\beta_c^{\mathrm{drive}}\,\phi^{\mathrm{drive}}$ but not its separation, so the optimizer compensates a small $\widehat{\beta}_c^{\mathrm{drive}} = -8.28\times10^{-4}$ with an implausibly large $\widehat{\phi}^{\mathrm{drive}} = 8{,}378$ GBP per hour, two orders of magnitude above any plausible commute valuation. The mixed-logit specification does not change this pattern: $\widehat{\mathbb{E}}[\phi^{\mathrm{drive}}]$ remains far above plausible levels and the implied elasticity $\widehat{\eta}_c^{\mathrm{drive}} = -0.008$ is two orders of magnitude smaller than the literature suggests.

We also report a parametric random-WTP specification inspired by Model B of Fosgerau (2007), which we call Fosgerau B. Instead of writing utility in terms of β_t and β_c, the choice is modelled directly as a comparison between an individual's (random) WTP and the bid $v_i = \Delta C_i / \Delta T_i$:

$$y_{\mathrm{fe},i} = \mathbf{1}[\mathrm{WTP}_i > v_i], \tag{34}$$

$$\log \mathrm{WTP}_i = \alpha + \gamma \log \Delta T_i + \delta\, D_i + u_i, \quad u_i \sim \mathrm{Logistic}, \tag{35}$$

Table 13
Fosgerau B Estimates.

	Swissmetro		LPMC	
Parameter	Est.	Std Err	Est.	Std Err
α	4.5597	0.1592	−1.2241	0.0349
γ	−1.5744	0.0381	−1.0599	0.0178
δ	0.9443	0.1242	−0.5607	0.0732
Median WTP	27.61	1.02	1.80	0.03

Median WTP is evaluated at the sample median of $\log \Delta T$ and the modal value of D; it is reported in CHF per hour for Swissmetro and GBP per hour for LPMC.

where (34) is the random-WTP choice rule of Fosgerau (2007), (35) is a parametric version of his log-WTP index from Section 4.5 with a logistic residual, $\log \Delta T_i$ is the absolute time difference between the two alternatives, and $D_i = 1$ if the labelled alternative (Swissmetro or public transit) is the fast option in scenario i and 0 otherwise. The coefficient γ on $\log \Delta T$ corresponds to the empirical dependence of WTP on the absolute time difference noted by Fosgerau (2007); δ absorbs asymmetric mode preference in our labeled-alternatives data, an adaptation of Fosgerau (2007), which uses unlabelled alternatives. Substituting (35) into (34) gives a binary logit on $\log v$, $\log \Delta T$ and D; under the unit-scale logistic normalisation for u_i, the coefficient on $\log v$ is fixed at -1. Median WTP at reference covariates is $\exp(\alpha + \gamma \log \Delta T_{\text{ref}} + \delta D_{\text{ref}})$ in the same time unit used to construct v_i (the median of the log-logistic distribution implied by (35)); for Swissmetro we multiply this value by 60 to report CHF per hour Table 13 reports the estimates for both datasets.

The Fosgerau B median is 27.61 CHF per hour in Swissmetro and 1.80 GBP per hour in LPMC, with the estimated $\hat{\gamma}$ negative in both datasets (-1.57 Swissmetro, -1.06 LPMC). Fosgerau (2007) reports $\hat{\gamma} \approx +0.3$ on Danish bus and train data; our sign-reversed estimates depart from his empirical pattern, with the steeply negative $\log \Delta T$ slope pulling the implied median WTP well below plausible commute valuations (Wardman et al., 2016, report a UK car commute range of roughly 5–15 GBP per hour). Note that the mixed-logit and Fosgerau B parametric forms give LPMC medians that differ by three orders of magnitude (8,378 versus 1.80 GBP per hour). These results suggest that, in these data, the median VTTS is highly sensitive to the imposed parametric form and to external information about cost sensitivity. In this context, the elasticity-constrained estimate provides a more plausible estimate.

In consequence, our approach is complementary to the WTP-space approach. The lognormal bound on ϕ rescues the sign of β_c at any σ_ϕ^m, including zero, because the sign constraint is structural to the parameterization. The elasticity-constrained estimator adds a different piece of external information, the level of the price-elasticity drawn from the empirical literature, which ties β_c to a value consistent with that elasticity. In Swissmetro both sources of information are mutually consistent. In LPMC the willingness-to-pay-space specification rescues the sign but leaves the level of ϕ^{drive} effectively unidentified; the elasticity-constrained estimator at target -0.4 delivers $\hat{\phi}^{\text{drive}} = 6.58$ GBP per hour (Section 5.2), consistent with meta-analytic evidence on UK car commute travel. Combining the two routes, by placing an elasticity-constrained prior on the location of the willingness-to-pay distribution, is left for future work.

6. Conclusions

This paper has addressed the imprecision of value of travel time savings (VTTS) estimates in multinomial logit models when two structural conditions of the data hold jointly: strong collinearity between travel time and travel cost, and a low absolute level of price-elasticity of demand. Both conditions push the cost coefficient $\hat{\beta}_c$ toward zero, the first by inflating its sampling variance and the second by making the true magnitude small. Because the VTTS is formed as a ratio with $\hat{\beta}_c$ in the denominator, both routes degrade the precision of the VTTS estimate. No estimator can recover identifying information the data do not contain; what an estimator can do is channel external information into the estimation in an explicit form. We propose an elasticity-constrained maximum-likelihood estimator that brings external evidence on the price-elasticity of demand into the estimation of β_c, supported by an analytic characterization of the VTTS variance through the delta method and by Monte Carlo simulations across combinations of correlation, elasticity and sample size. Two empirical applications, on the Swissmetro stated-preference dataset and the London Passenger Mode Choice revealed-preference dataset, evaluate the proposed estimator against maximum likelihood, ridge, and willingness-to-pay-space mixed logit.

The proposed estimator has limitations the analysis makes explicit. The constraint pins the cost coefficient to a value implied by an externally chosen elasticity target $\bar{\eta}_c$, and a target that differs from the true population elasticity introduces bias. The sensitivity analysis in Section 4.2 shows that in adverse data (strong collinearity, low elasticity) the constrained estimator dominates maximum likelihood across the modal range of priors reported in the empirical literature, but in well-behaved data a misspecified prior can make the constrained estimator underperform maximum likelihood. The estimate produced by the method is therefore the value of travel time savings conditional on the price-elasticity being consistent with external evidence, not an unconditional identification of a population preference parameter. The Hausman test, interpreted as a detectability diagnostic rather than a classical bias test, plays a different role across regimes: in Swissmetro it confirms that the constraint introduces no statistically distinguishable departure from maximum likelihood at the sample size at hand; in the London application it registers that the constrained estimator deliberately diverges from a wrong-signed baseline, and the appropriate diagnostics in that regime become the microeconomic sign test on β_c and the consistency of the implied VTTS with published evidence.

The implications for transport-project appraisal are concrete. Revealed-preference samples with imputed cost typically exhibit the structural collinearity that motivates the method, and inelastic demand is the empirical regularity in urban transport that compounds the problem. In these settings the elasticity-constrained estimator delivers a VTTS that is consistent with the price-elasticity of demand used elsewhere in the appraisal, for example in demand forecasts for toll roads, congestion-charging schemes and fare reforms (Mackie et al., 2014). The London application illustrates the operational difference: unrestricted maximum likelihood there produces a positively-signed cost coefficient and a VTTS of −9.01 GBP per hour, which would be unusable for appraisal; the elasticity-constrained estimator at target −0.4 recovers a VTTS of 6.58 GBP per hour, inside the range that UK appraisal practice supports. For research questions about pure preference parameters of a specific population, in datasets where the cost coefficient is well identified, unconstrained maximum likelihood remains the natural choice and the constrained estimator is not the default.

Two directions for future work follow naturally. The first is methodological. Section 5.3 compares the elasticity-constrained estimator with the willingness-to-pay-space mixed logit of Fosgerau (2007) and Train and Weeks (2005) and finds that the two routes bring external information about different objects: the lognormal parameterization of WTP-space constrains the form of preference heterogeneity, and the elasticity constraint introduced here constrains the level of β_c. The two pieces of external information are complementary, and a willingness-to-pay-space specification with an elasticity-constrained prior on the location of the WTP distribution is a natural next step. Extensions to nested logit and more flexible mixed logit specifications are likewise straightforward in principle. A further methodological question concerns the measurement error in imputed cost that Section 2 identifies as a third source of a downward-biased cost coefficient. The elasticity constraint addresses a cost coefficient that is small or uncertain whatever the reason, so it does not separate measurement-error attenuation from weak identification due to collinearity, and it does not reveal how large the measurement-error component is. The latent-variable and hybrid-choice methods of Walker et al. (2010) and Lorenzo Varela et al. (2018) are designed for that diagnosis, and combining them with an elasticity-constrained target is a natural next step. The second direction is empirical. Large passive-tracking datasets such as mobile-network call detail records, platform trip records and toll-corridor transponder data (Calabrese et al., 2013; Bwambale et al., 2019; Lin et al., 2026) reduce the small-sample component by enlarging N by orders of magnitude, but the structural collinearity between time and cost does not vanish and these sources typically lack direct cost information; the relevance of the proposed estimator in those settings, and its interaction with imputed-cost schemes, is open. The broader point is that collinearity and low price-elasticity should not be treated as merely technical complications but as structural conditions that determine the validity of the value of travel time savings as a welfare measure and as an input to project appraisal, and that the explicit incorporation of external evidence into estimation, in a form that can be checked against the published literature, is the principled response.

CRediT authorship contribution statement

Louis de Grange: Writing – review & editing, Writing – original draft, Methodology, Conceptualization; **Matthieu Marechal:** Writing – review & editing, Validation, Software, Formal analysis, Data curation; **Raúl Pezoa:** Writing – review & editing, Visualization, Validation, Software, Methodology, Formal analysis, Data curation; **Felipe González:** Writing – review & editing, Visualization, Validation, Data curation; **Rodrigo Troncoso:** Writing – review & editing, Writing – original draft, Formal analysis.

Data availability

Data will be made available on request.

References

Abrantes, P. A.L., Wardman, M.R., 2011. Meta-analysis of uk values of travel time: an update. Transp. Res. A: Policy Pract. 45 (1), 1–17.
Agness, D., Baseler, T., Chassang, S., Dupas, P., Snowberg, E., 2025. Valuing the time of the self-employed. Rev. Econ. Stud. 92 (6), 3471–3503.
Armstrong, P., Garrido, R., de Dios Ortúzar, J., 2001. Confidence intervals to bound the value of time. Transp. Res. E: Logist. Transp. Rev. 37 (2–3), 143–161.
Bain, R., Cimon, B., 2026. The toll price elasticity of demand: a guide for practitioners. Technical Report. Route One Publishing Ltd (UK), CSRB Group.
Becker, G.S., 1965. A theory of the allocation of time. Econ. J. 75 (299), 493–517.
Berry, C., Börjesson, M., 2024. Income and fuel price elasticities of car use on micro panel data. Energy Econ. 135, 107661.
Bierlaire, M., Axhausen, K., Abay, G., 2001. The acceptance of modal innovation: the case of swissmetro. In: Swiss Transport Research Conference.
Biswas, M., Bhat, C.R., Ghosh, S., Pinjari, A.R., 2024. Choice models with stochastic variables and random coefficients. J. Choice Model. 51, 100488.
Bliemer, M. C.J., Rose, J.M., 2013. Confidence intervals of willingness-to-pay for random coefficient logit models. Transp. Res. B: Methodol. 58, 199–214.
Bolduc, D., Khalaf, L., Yélou, C., 2010. Identification robust confidence set methods for inference on parameter ratios with application to discrete choice models. J. Econom. 157 (2), 317–327.
Börjesson, M., Eliasson, J., 2014. Experiences from the swedish value of time study. Transp. Res. a: Policy Pract. 59, 144–158.
Bouscasse, H., de Lapparent, M., 2019. Perceived comfort and values of travel time savings in the rhône-alpes region. Transp. Res. a: Policy Pract. 124, 370–387.
Brons, M., Nijkamp, P., Pels, E., Rietveld, P., 2008. A meta-analysis of the price elasticity of gasoline demand. a SUR approach. Energy Econ. 30 (5), 2105–2122.
Bull, O., Muñoz, J.C., Silva, H.E., 2021. The impact of fare-free public transport on travel behavior: evidence from a randomized controlled trial. Reg. Sci. Urban Econ. 86, 103616.
Bwambale, A., Choudhury, C.F., Hess, S., 2019. Modelling long-distance route choice using mobile phone call detail record data: a case study of senegal. Transp. a: Transp. Sci. 15 (2), 1543–1569.
Calabrese, F., Diao, M., Di Lorenzo, G., Ferreira, J., Ratti, C., 2013. Understanding individual mobility patterns from urban sensing data: a mobile phone trace example. Transp. Res. C: Emerg. Technol. 26, 301–313.
Carson, R.T., Czajkowski, M., 2019. A new baseline model for estimating willingness to pay from discrete choice models. J. Environ. Econ. Manag. 95, 57–61.
Castro, M., Martínez, F., Munizaga, M.A., 2013. Estimation of a constrained multinomial logit model. Transportation 40 (3), 563–581.
Daly, A., Hess, S., De Jong, G., 2012a. Calculating errors for measures derived from choice modelling estimates. Transp. Res. B: Methodol. 46 (2), 333–341.
Daly, A., Hess, S., de Dios Ortúzar, J., 2023. Estimating willingness-to-pay from discrete choice models: setting the record straight. Transp. Res. a: Policy Pract. 176, 103828.

Daly, A., Hess, S., Train, K., 2012b. Assuring finite moments for willingness to pay in random coefficient models. Transportation 39 (1), 19–31.
Davis, L.W., 2021. Estimating the price elasticity of demand for subways: evidence from mexico. Reg. Sci. Urban Econ. 87, 103651.
De Grange, L., González, F., Muñoz, J.C., Troncoso, R., 2013. Aggregate estimation of the price elasticity of demand for public transport in integrated fare systems: the case of transantiago. Transp. Policy 29, 178–185.
Delle Site, P., Parmar, J., 2024. On the linear probability model as binary choice random utility model. J. Choice Model. 52, 100505.
DeSerpa, A.C., 1971. A theory of the economics of time. Econ. J. 81 (324), 828–846.
de Grange, L., Fariña, P., de Dios Ortúzar, J., 2015. Dealing with collinearity in travel time valuation. Transp. a: Transp. Sci. 11 (4), 317–332.
Espino, R., de Dios Ortúzar, J., Román, C., 2006. Confidence interval for willingness to pay measures in mode choice models. Netw. Spat. Econ. 6 (2), 81–96.
Festjens, A., Janiszewski, C., 2015. The value of time. J. Consum. Res. 42 (2), 178–195.
Fosgerau, M., 2007. Using nonparametrics to specify a model to measure the value of travel time. Transp. Res. a: Policy Pract. 41 (9), 842–856.
Gardes, F., 2018. On the value of time and human life. Documents de travail. Centre d'Économie de la Sorbonne.
Gaudry, M. J.I., Jara-Diaz, S.R., de Dios Ortúzar, J., 1989. Value of time sensitivity to model specification. Transp. Res. B: Methodol. 23 (2), 151–158.
Gillingham, K., Munk-Nielsen, A., 2019. A tale of two tails: commuting and the fuel price response in driving. J. Urban Econ. 109, 27–40.
González, R.M., 1997. The value of time: a theoretical review. Transp. Rev. 17 (3), 245–266.
Goodwin, P.B., 1992. A review of new demand elasticities with special reference to short and long run effects of price changes. J. Transp. Econ. Policy 26 (2), 155–169.
Graham, D.J., 2025. Causal inference for transport research. Transp. Res. A: Policy Pract. 192, 104324.
Graham, D.J., Glaister, S., 2004. Road traffic demand elasticity estimates: a review. Transp. Rev. 24 (3), 261–274.
Greene, W.H., 2003. Econometric Analysis. Prentice Hall.
Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, New York. 2nd edition.
Hensher, D.A., 2001. The valuation of commuter travel time savings for car drivers: evaluating alternative model specifications. Transportation 28 (2), 101–118.
Hess, S., Bierlaire, M., Polak, J.W., 2005. Estimation of value of travel-time savings using mixed logit models. Transp. Res. a: Policy Pract. 39 (2–3), 221–236.
Hess, S., Train, K., 2017. Correlation and scale in mixed logit models. J. Choice Model. 23, 1–8.
Hillel, T., Elshafie, M.Z., Jin, Y., 2018. Recreating passenger mode choice-sets for transport simulation: a case study of london, uk. Proc. Inst. Civ. Eng. - Smart Infrastruct. Constr. 171 (1), 29–42.
Hole, A.R., 2007. A comparison of approaches to estimating confidence intervals for willingness to pay measures. Health Econ. 16 (8), 827–840.
Hole, A.R., Kolstad, J.R., 2012. Mixed logit estimation of willingness to pay distributions: a comparison of models in preference and wtp space using data from a health-related choice experiment. Empir. Econ. 42 (2), 445–469.
Holmgren, J., 2007. Meta-analysis of public transport demand. Transp. Res. A: Policy Pract. 41 (10), 1021–1035.
Hörcher, D., Tirachini, A., 2021. A review of public transport economics. Econ. Transp. 25, 100196.
Huang, Y., Qian, L., 2025. Consumer preferences and willingness to pay for data privacy in automated vehicles. Transp. Res. A: Policy Pract. 199, 104585.
Ilahi, A., Belgiawan, P.F., Balac, M., Axhausen, K.W., 2021. Understanding travel and mode choice with emerging modes: a pooled sp and rp model in greater jakarta, indonesia. Transp. Res. A: Policy Pract. 150, 398–422.
Jara-Diaz, S., 2024. The value(s) of travel time savings considering in-vehicle activities. Transp. Res. A: Policy Pract. 184, 104092.
Jara-Díaz, S.R., Farah, M., 1987. Transport demand and users' benefits with fixed income: the goods/leisure trade off revisited. Transp. Res. B: Methodol. 21 (2), 165–170.
Krinsky, I., Robb, A.L., 1986. On approximating the statistical properties of elasticities. Rev. Econ. Stat. , 715–719.
Li, Z., Hensher, D.A., Rose, J.M., 2010. Willingness to pay for travel time reliability in passenger transport: a review and some new empirical evidence. Transp. Res. E: Logist. Transp. Rev. 46 (3), 384–403.
Lin, S., Spissu, E., Cirillo, C., 2026. New metrics and estimates of travel time reliability using large passive datasets. Transp. Res. A: Policy Pract. 209, 105021.
Lorenzo Varela, J.M., Börjesson, M., Daly, A., 2018. Quantifying errors in travel time and cost by latent variables. Transp. Res. B: Methodol. 117, 520–541.
Mabit, S.E., 2026. Mixed logit applications in travel behaviour research. Transp. Res. A: Policy Pract. 203, 104728.
Mackie, P., Worsley, T., Eliasson, J., 2014. Transport appraisal revisited. Res. Transp. Econ. 47, 3–18.
Mackie, P.J., Jara-Díaz, S., Fowkes, A.S., 2001. The value of travel time savings in evaluation. Transp. Res. E: Logist. Transp. Rev. 37 (2–3), 91–106.
Martín-Baos, J.Á., García-Ródenas, R., Rodriguez-Benitez, L., 2021. A python package for performing penalized maximum likelihood estimation of conditional logit models using kernel logistic regression. Transp. Res. Procedia 58, 61–68.
Mauerer, I., Pöß necker, W., Thurner, P.W., Tutz, G., 2015. Modeling electoral choices in multiparty systems with high-dimensional data: a regularized selection of parameters using the lasso approach. J. Choice Model. 16, 23–42.
McFadden, D., 1972. Conditional logit analysis of qualitative choice behavior .
McFadden, D., 2022. Instability in mixed logit demand models. J. Choice Model. 43, 100353.
Noland, R.B., Polak, J.W., 2002. Travel time variability: a review of theoretical and empirical issues. Transp. Rev. 22 (1), 39–54.
Ortúzar, J. d.D., Armstrong, P.M., 1995. Confidence intervals and the social value of travel time savings. In: Models and Applications. Proceedings of Seminar F Held at the 23Rd European Transport Forum, University of Warwick, England, September 11–15, 1995. Volume P393.
Ortúzar, J. d.D., Willumsen, L.G., 2024. Modelling Transport. Wiley. 5th edition.
Pavlou, M., Ambler, G., Seaman, S., De Iorio, M., Omar, R.Z., 2016. Review and evaluation of penalised regression methods for risk prediction in low-dimensional data with few events. Stat. Med. 35 (7), 1159–1177.
Rossi, P.E., Allenby, G.M., McCulloch, R., 2024. Bayesian Statistics and Marketing. Wiley, Chichester. 2nd edition.
Scaccia, L., Marcucci, E., Gatta, V., 2023. Prediction and confidence intervals of willingness-to-pay for mixed logit models. Transp. Res. B: Methodol. 167, 54–78.
Scarpa, R., Thiene, M., Train, K., 2008. Utility in willingness-to-pay space: a tool to address confounding random scale effects in destination choice to the alps. Am. J. Agric. Econ. 90, 994–1010.
Small, K.A., 2012. Valuation of travel time. Econ. Transp. 1 (1), 2–14.
Sonnier, G., Ainslie, A., Otter, T., 2007. Heterogeneity distributions of willingness-to-pay in choice models. Quant. Mark. Econ. 5 (3), 313–331.
Train, K., Sonnier, G., 2005. Mixed logit with bounded distributions of correlated partworths. In: Applications of Simulation Methods in Environmental and Resource Economics. Springer, pp. 117–134.
Train, K., Weeks, M., 2005. Discrete choice models in preference space and willingness-to-pay space. In: Scarpa, R., Alberini, A., (Eds.), Applications of Simulation Methods in Environmental and Resource Economics. Springer, Dordrecht, pp. 1–16.
Tsoleridis, P., Choudhury, C.F., Hess, S., 2022. Deriving transport appraisal values from emerging revealed preference data. Transp. Res. a: Policy Pract. 165, 225–245.
Walker, J.L., Li, J., Srinivasan, S., Bolduc, D., 2010. Travel demand models in the developing world: correcting for measurement errors. Transp. Lett. 2 (4), 231–243.
Wardman, M., Chintakayala, V. P.K., de Jong, G., 2016. Values of travel time in europe: review and meta-analysis. Transp. Res. A: Policy Pract. 94, 93–111.
Wardman, M., Toner, J., Fearnley, N., Flögel, S., Killi, M., 2018. Review and meta-analysis of inter-modal cross-elasticity evidence. Transp. Res. A: Policy Pract. 118, 662–681.
Yadlin, A., Yadlin, M., 1999. Confidence intervals for the subjective value of time of transportation users. Braz. J. Probab. Stat. 13 (2), 193–200.