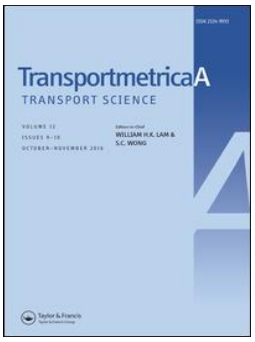

Transportmetrica A: Transport Science

ISSN: 2324-9935 (Print) 2324-9943 (Online) Journal homepage: https://www.tandfonline.com/loi/ttra21

Intercity truck route choices incorporating toll road alternatives using enhanced GPS data

Tomer Toledo, Bilge Atasoy, Peiyu Jing, Jing Ding-Mastera, Jorge Oliveira Santos & Moshe Ben-Akiva

To cite this article: Tomer Toledo, Bilge Atasoy, Peiyu Jing, Jing Ding-Mastera, Jorge Oliveira Santos & Moshe Ben-Akiva (2020) Intercity truck route choices incorporating toll road alternatives using enhanced GPS data, Transportmetrica A: Transport Science, 16:3, 654-675, DOI: 10.1080/23249935.2020.1722284

To link to this article: https://doi.org/10.1080/23249935.2020.1722284

Accepted author version posted online: 27 Jan 2020.
Published online: 10 Feb 2020.

TRANSPORTMETRICA A: TRANSPORT SCIENCE
2020, VOL. 16, NO. 3, 654–675
https://doi.org/10.1080/23249935.2020.1722284

Taylor & Francis
Taylor & Francis Group

Check for updates

Intercity truck route choices incorporating toll road alternatives using enhanced GPS data

Tomer Toledo [a], Bilge Atasoy[b], Peiyu Jing[c], Jing Ding-Mastera[c], Jorge Oliveira Santos[d] and Moshe Ben-Akiva[c]

[a]Faculty of Civil and Environmental Engineering, Technion – Israel Institute of Technology, Haifa, Israel; [b]Department of Maritime and Transport Technology, Delft University of Technology, Delft, The Netherlands; [c]Intelligent Transportation Systems Lab, Massachusetts Institute of Technology, Cambridge, MA, USA; [d]Mobile Market Monitor, Brookline, MA, USA

ABSTRACT
This research presents the data collection, specification and estimation of a route choice model for intercity truck trips, with a focus on toll road usage. The data was obtained from driver-validated and enhanced GPS records. A mixed logit model with a path-size factor is specified. It accounts for heterogeneity among drivers using distributed coefficients for travel time and its variability. The estimation results show wide heterogeneity among drivers based on employment type and availability of electronic toll collection tags. Toll value of time and toll value of reliability distributions are derived. The model application is demonstrated on several trip corridors.

ARTICLE HISTORY
Received 23 February 2019
Accepted 21 December 2019

KEYWORDS
Intercity freight; truck route choice; GPS tracking

1. Introduction

Truck traffic accounts for about 9% of the distance driven on highways in the USA. The total truck flows have been increasing steadily. They are expected to increase a further 43% between 2015 and 2045 (BTS 2017). Truck traffic has a substantial effect on traffic flow. Therefore, understanding trucks' route choices is important in order to forecast truck traffic and model traffic and freight systems. Specifically, trucks, through higher annual distance traveled and higher toll rates, often contribute a significant share of revenues in toll roads (Bain and Polakovic 2005). However, forecasts of trucks' use of toll roads have been shown to overestimate actual use (Bain 2009). This may result in loss of revenue for the developers.

Substantial literature on truck route choices focuses on value of time (VOT), namely the trade-off between the cost and travel time (e.g. Wynter 1995; Jong 2000; Kawamura 2000; Bergkvist 2001; Smalkoski and Levinson 2005; Zamparini and Reggiani 2007; Ismail, Sayed, and Lim 2009; Miao, Wang, and Adams 2011). While these are consistently important determinants of trucks' route choices, they are not the only important ones. Other factors that have been found to affect truck route choices include measures of the travel time reliability (e.g. Jovicic 1998; Small et al. 1999; Kurri, Sirkia, and Mikola 2000; Austroads 2003; Jong et al. 2004; Danielis, Marcucci, and Rotaris 2005; Fowkes and Whiteing 2006;

CONTACT Tomer Toledo toledo@technion.ac.il Technion City, 3200003 Haifa, Israel

Toledo et al. 2013), travel distance (Knorring, He, and Kornhauser 2005; Quattrone and Vitetta 2011; Wood 2011; Toledo et al. 2013; Hess et al. 2015), road types and characteristics (e.g. Hunt and Abraham 2004; Hyodo and Hagino 2010; Arentze et al. 2012; Rowell, Gagliano, and Goodchild 2014; Hess et al. 2015; Tahlyan et al. 2017), facilities along the road (Arentze et al. 2012; Feng, Arentze, and Timmermans 2013; Rowell, Gagliano, and Goodchild 2014), type of freight service (Austroads 2003).

Beyond the explanatory variables, several authors addressed the effect of similarities among routes due to overlap. Quattrone and Vitetta (2011), Hess et al. (2015) and Tahlyan et al. (2017) used Path-size logit (PSL) and C-logit models to capture the effect of similarities among overlapping route alternatives, showing improvement in model fit over the model that does not include these terms. Hess et al. (2015) developed an error components model, in which correlations among routes stem from similarities in road types being driven. In their estimation results, this approach yielded better fit to the data compared to PSL.

The need to address the existence of a heterogeneity in route choice preferences among truckers has also been noted. To capture this, Kurri, Sirkia, and Mikola (2000), Danielis, Marcucci, and Rotaris (2005) and Fowkes and Whiteing (2006) estimated different route choice parameters for different sub-populations of drivers based on their industry. Feng, Arentze, and Timmermans (2013) and Rowell, Gagliano, and Goodchild (2014) estimated latent class models, in which class membership was mostly explained by truck size or travel distance. Quattrone and Vitetta (2011) and Kim, Pasco, and Kothapalli (2017) used fuzzy logic structures, and Kawamura (2000) and Toledo et al. (2013) used random coefficient models to estimate distributed VOTs. All found wide heterogeneity in preferences. Marcucci and Gatta (2012) compared different methods to capture heterogeneity in the context of airport choices. They found that a mixed logit model with random coefficients in the utility function outperformed other approaches.

In terms of data sources, in recent years there has been a shift from using stated preferences (SP) data, which has been the prevailing source of data for estimation of truck route choice models, to revealed preferences (RP), mostly from large-scale GPS records (e.g. Knorring, He, and Kornhauser 2005; Quattrone and Vitetta 2011; Wang and Goodchild 2014; Hess et al. 2015; Tahlyan et al. 2017). Beyond the better response realism, these data offer advantages in terms of accuracy and sample size. However, data derived from GPS records often does not include information about the driver and trip circumstances, or about alternatives to the chosen route.

This paper reports on research to collect data and develop a route choice model for truck drivers that make intercity trips, with a focus on toll road usage. The developed model accounts both for correlations among alternatives due to overlap and for heterogeneity in preferences in VOT and towards toll road use among drivers. The data used in the study was collected using GPS loggers that were installed in trucks that traveled throughout the USA and Canada. These data were verified by the drivers using a web interface and combined with drivers' socio-economic and shipments characteristics that were also collected through the interface.

This work extends the literature on truck route choices in several directions. The data collection combines route tracking using GPS with the use of a web interface to solicit additional information on the drivers and their trips. Most previous models only used variables related to route attributes, levels of service and costs that can be derived from the GPS traces and map databases. The data collection method used in this study, supports model specifications that are able to capture systematic differences in preferences among driver groups based on their characteristics, such as their years of experience and employment type, and additional terms, such as availability of toll tags that are derived from the responses on the web interface. The model also accounts for unobserved heterogeneity through the specification of distributed individual-specific travel time and travel time reliability coefficients. In the context of truck route choices, this has been previously done with SP data, but not with GPS traces.

The use of GPS records collected from trucks undertaking intercity travel over a longer period of time provides data on a wide geographic scale. This is in contrast with most previous studies that focused on specific regions or corridors. The result is a better representation of the non-recurring travel patterns of trucks, unlike those of commuters. Furthermore, it increases the variability in the values of the explanatory variables and improves the efficiency of the parameter estimates. However, the wide geographic coverage makes it difficult to both generate alternative routes and to estimate their attributes, which are required for the modeling task. This problem has not received adequate attention. Some previous truck routing research using GPS traces focused on specific corridors (e.g. Knorring, He, and Kornhauser 2005; Wang and Goodchild 2014) and so used judgement to pre-define a small number of reasonable routes and extracted their attributes from the map database. Quattrone and Vitetta (2011) and Hess et al. (2015) generated network-wide routes using the labeling and link-elimination approaches, respectively. However, neither used navigation map databases to estimate the values of attributes, including those used to generate routes. The former used a coarse national traffic assignment model. The latter estimated travel times based on assumptions on the travel speed on various road types. This paper proposes using openly available capabilities of commercial navigation map databases for the tasks of generation of routes and estimation of their characteristics.

The remainder of this paper is organized as follows. The next section presents the overall route choice modeling methodology. Section 3 presents the data collection and processing methodologies and reports descriptive statistics of the resulting dataset. Section 4 presents the specification details and estimation results of the truck driver route choice model. Next, the application of the model is demonstrated with specific truck trips. Finally, the main findings are summarized and discussed together with future research directions.

2. Route choice model

A mixed logit model with a path-size factor is proposed to predict the route choices of truck drivers. In order to make the route choice model applicable to trips that differ from the ones that were used to estimate the model, the utility specifications are generic (i.e. do not include any parameters that are specific to an alternative). Furthermore, routes in the choice set may partially overlap. The similarity in the common segments causes the error terms of overlapping routes to be correlated. Following Ben-Akiva and Bierlaire (1999) and Ramming (2002), a path-size (PS) variable is used to capture the effects of the similarity among routes in the model:

$$PS_{int} = \sum_{a \in \Gamma_{int}} \left(\frac{l_a}{L_{int}} \right) \frac{1}{\sum_{j \in C_{nt}} \delta_{ajnt}} \tag{1}$$

where, PS_{int} is the path-size value for route i and trip t of driver n. L_{int} is the length of the route. l_a is the length of link a that belongs to the set of links Γ_{int} that comprises route i and trip t of driver n. δ_{ajnt} is an indicator variable, which takes value 1 if link a is part of route j, and 0 otherwise.

It should be noted that several nonlinear PS formulations that were proposed in the literature (using a power parameter γ, see e.g. Prato 2009) were also tried. The formulation in Equation 1 yielded the best fit. The estimated coefficient of the PS term did not change substantially with other formulations. The utility functions are therefore given by:

$$U_{int} = V_{int}(X_{int}, \beta_n) + \beta_{PS} \ln PS_{int} + \varepsilon_{int} \tag{2}$$

where U_{int} is the utility of alternative route i of driver n in trip t. V_{int} is the systematic part of the utility function. X_{int} and β_n are the explanatory variables and the corresponding coefficients, respectively. β_{PS} is the coefficient of the path-size term. ε_{int} is an error term, which is assumed to be independently and identically drawn from the Gumbel distribution.

The probability that driver n chooses route i in trip t is given by:

$$P_n(i_t|\beta_n) = \frac{\exp(V_{int}(X_{int}, \beta_n) + \beta_{PS} \ln PS_{int})}{\sum_{j \in C_{nt}} \exp(V_{jnt}(X_{jnt}, \beta_n) + \beta_{PS} \ln PS_{jnt})} \tag{3}$$

Previous research has shown large heterogeneity in route tastes among truck drivers (e.g. Kawamura 2000; Toledo et al. 2013). Ignoring this taste heterogeneity can lead to inconsistent estimates of the model coefficients and deteriorate prediction power (Ben-Akiva, Bolduc, and Park 2008). In order to capture taste heterogeneity, the coefficients of two variables in the model are assumed to be distributed in the drivers' population: the log of travel time and the square of the travel time range. Both are assumed to follow log-normal distributions (with a negative sign in the utility function) in order to ensure that they are always negative, which indicates that drivers prefer shorter travel times and lower travel time variability. The distributions are assumed to be correlated with each other. Therefore, their joint distribution is given by:

$$\ln\left(\begin{bmatrix} \beta_{LnTT,n} \\ \beta_{ttRangeSq,n} \end{bmatrix}\right) \sim N\left(\begin{bmatrix} \mu_{\beta_{LnTT}} \\ \mu_{\beta_{ttRangeSq}} \end{bmatrix}, \begin{bmatrix} \sigma^2_{\beta_{LnTT}} & \sigma_{\beta_{LnTT}\beta_{ttRangeSq}} \\ \sigma_{\beta_{LnTT}\beta_{ttRangeSq}} & \sigma^2_{\beta_{ttRangeSq}} \end{bmatrix}\right) \tag{4}$$

where $\beta_{LnTT,n}$ and $\beta_{ttRangeSq,n}$ are the coefficients of log of travel time and the square of the travel time range for individual n, respectively. β_{LnTT} and $\beta_{ttRangeSq}$ are the corresponding mean parameters of the lognormal distributions. $\sigma^2_{\beta_{LnTT}}$ and $\sigma^2_{\beta_{ttRangeSq}}$ are the variances of the distributions, and $\sigma_{\beta_{LnTT}\beta_{ttRangeSq}}$ is their covariance.

The random coefficients are assumed to vary among drivers but are constant in all the observations from the same driver. Thus, they capture inter-participant and assume no intra-participant heterogeneity. Although this is the common approach to modeling heterogeneity (Hess and Rose 2009), it may also be possible to relax this assumption (e.g. Hess and Rose 2009; Becker et al. 2018) using a more complex model structure. Estimation of this model may be more difficult and require a simpler utility specification.

Under these assumptions, the conditional probability of the chosen alternative in trip t is given by:

$$P_n(Y_t|\beta_n) = \prod_i [P_n(i_t|\beta_n)]^{Y_{\text{int}}} \tag{5}$$

where Y_{int} is equal 1 for the chosen alternative, and 0 otherwise.

The conditional joint probability of all the observations from the same driver is given by:

$$P_n(Y_1 \ldots Y_T|\beta_n) = \prod_{t=1}^{T} P_n(Y_t|\beta_n) \tag{6}$$

The unconditional joint probability is given by:

$$P_n(Y_1 \ldots Y_T) = \iint_{\beta_n} P_n(Y_1 \ldots Y_T|\beta_n) f(\beta_n) d\beta_n \tag{7}$$

where, $f(\beta_n)$ is the joint distribution of the individual-specific parameters, given in Equation (4).

The intergral in equation(7) may be evaluated using simulation:

$$P_n(Y_1 \ldots Y_T) = \frac{1}{R} \sum_{r=1}^{R} P_n(Y_1 \ldots Y_T|\beta_{nr}) \tag{8}$$

where, β_{nr} are coefficients drawn in replication r from the distribution in equation (4). R is the number of replications.

Finally, the loglikelihood function to be maximized is given by:

$$LL = \sum_{n} \ln(P_n(Y_1 \ldots Y_T)) \tag{9}$$

3. Data collection and processing

This section presents the data collection and processing that led to the final dataset that was later used in model estimation.

3.1. Data collection

The study collected trucks' GPS data and elicited additional information from the drivers through a web interface. Truck drivers were recruited to participate in the experiment in roadside intercepts or by phone calls to lists of drivers in areas of Texas, Indiana, Ontario, New Jersey and Massachusetts. The trucks of recruited drivers were equipped with GPS loggers that continuously collected data on the location and movement of the trucks and transmitted this information through wireless networks to an application server. At the server, the GPS traces were matched to road segments on a Geographic Information System (GIS) map and stop locations were identified. The resulting routes and stops were presented to the drivers on a personal web interface. The drivers were then asked to provide additional information about their trips (e.g. schedule for delivery or pickup, characteristics of the freight being transferred) and the stops they made (e.g. pick-up, delivery and other activities). Figure 1 shows an example of the web interface. A truck's route is shown. One of the stops along the route is highlighted and a question about this stop is displayed. At the end of the tracking period, which was up to a month, the drivers completed an exit survey soliciting their socio-economic characteristics. The drivers were compensated up to \$100 for their participation. The compensation rate depended on the period of participation (one to four weeks) and on providing the additional information requested. Additional details on the data collection methodology and tools are presented in Ben-Akiva et al. (2016).

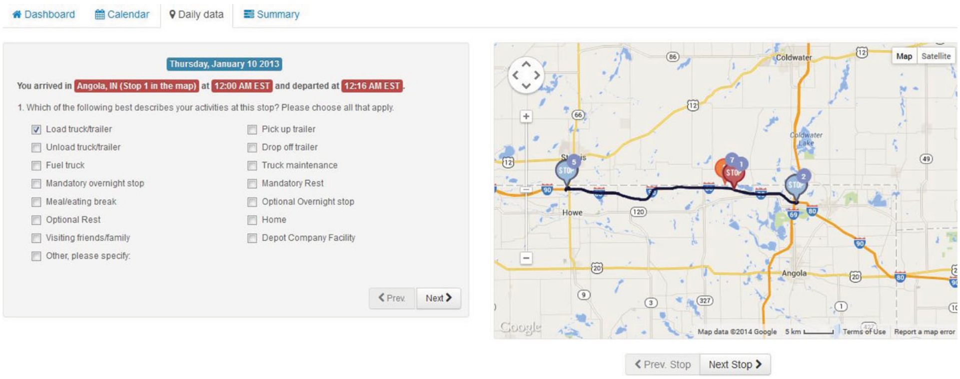

Figure 1. Trip, stops and the related questions displayed on the personal webpage.

3.2. Data processing

In total data from 107 drivers was collected. It covers 2,255 driving days. 12,617 stops were detected. These stops were not only for loading and drop-off, but also for rest, service, fuel, depots, visit home and so on. These data were processed in preparation for the modeling task to generate a database with a choice set of alternative routes for each trip, their attributes and the characteristics of the driver and shipment for the trip and identify the chosen route. To that end, he following steps were taken:

(1) **Identification of trips**: for the analysis, trips were defined as travel between loading points as origins and the following drop-off points as destination stops. In cases when multiple loading or drop-off stops were identified in a sequence, the trip was defined from the last loading stop to the first drop-off stop in the sequence. Trips with substantial portions of missing GPS data gap or missing information from the drivers about stops were removed.
(2) **Choice set generation**: A labeling approach was used to generate a choice set of route alternatives. Four available navigation applications were used: Google maps, Bing, MapQuest, and INRIX. Labeled routes were generated by running route recommendation queries in these applications with different options that they support: routes with or without tolls, preferring or avoiding highways, shortest distance or travel times based on free flow or time-dependent conditions. This approach allowed generating alternative routes with the level of detail of navigation map databases at a wide geographic scale. Using transportation planning models would not yield the same level of detail. The use of other generation approaches, such as simulation or link elimination, would require accessing and manipulating a navigation map database with travel times and other relevant attributes. However, these would be prohibitively expensive to obtain. For example, Hess et al. (2015) advocate using navigation maps for route generation. They used a link elimination approach. But, they did not have access to travel time information and so compromised on assuming average speeds for different road classes to calculate travel times. The routes that were accumulated from the various queries were evaluated to remove duplicate routes. Alternative routes with overlap of over 80% of their length were considered duplicates. A similar overlap threshold

was recommended by Ramming (2002). It reflects the interest in inter-urban travel, for which the network is relatively sparse. Trips for which a single route was generated or when none of the generated routes had at least 80% overlap with the GPS observed route were discarded.

(3) **Route attributes**: For the resulting sets of routes, their attributes were collected from various sources:
   (a) Time-dependent travel times and travel time variability for trucks were queried from the INRIX database.
   (b) Tolls were calculated from tables of point-to-point tolls by vehicle type on major toll roads in North America. These tables were extracted from the websites of various road operators.
   (c) Network attributes such as distances and road classifications were extracted from the OpenStreetMap (OSM) GIS database.

(4) **Chosen route**: As noted above, the observed GPS routes were matched to a navigation map database. The matching accuracy is expected to be high because of the high frequency of the GPS records and the inter-urban nature of trips, which means that the road network is relatively sparse. Among the routes that were generated, the route that had the highest overlap with the matched routes (at least 80% of its length) was determined to be the chosen route.

3.3. Descriptive statistics

The resulting dataset, which was used for model estimation included a total of 1,021 trips made by 99 drivers. 9,902 alternative routes were generated for these trips. On average there are 9.7 routes per trip. This is a relatively large number for an inter-urban network. Due to their experience, it is plausible that professional truck drivers are knowledgeable about a larger number of alternative routes. For shorter truck trips, which are more likely to use dense urban networks, Hess et al. (2015), generated an average of 15 routes per trip. Frejinger (2007) shows that optimal route choice estimation results are obtained when the full universal choice set is included, which is expected to be large. Bovy (2009) shows that addition of irrelevant routes to the choice set should not affect the estimation results. However, this argument is not supported by empirical results (Prato and Bekhor 2007; Bliemer and Bovy 2008).

Descriptive statistics of these routes and drivers are shown in Table 1. The table separately shows statistics for the chosen routes and for the full route choice set. A wide variability exists in all variables: travel times, distances, types of roads being driven and tolls. The reported travel times are the time-dependent expected values from the INRIX database. The travel time range is defined by the difference between the longest and shortest expected travel times reported over the day. Thus, it captures the within-day variability in travel time and can be viewed as a proxy to the congestion levels and risk of delays. As can be expected, the statistics show that on average, the chosen routes are shorter than alternative routes in terms of travel times, travel time ranges and travel distances. In contrast, chosen routes are more likely to involve tolls compared to other routes. While 31% of chosen routes involve tolls, only 24% of all generated routes do. As a result, the mean toll for chosen routes is higher than that of the other routes. Another difference between chosen and other routes is in their road class composition. Road segments that comprise the

Table 1. Descriptive statistics of the routes and drivers in the estimation dataset.

		Route attributes				
		Minimum	Median	Mean	Maximum	Std. dev.
Number of routes		2	10	9.70	22	14.05
Travel time (hours)	Chosen	0.02	2.23	4.05	45.48	5.22
	All	0.02	3.23	5.46	70.27	6.55
Travel time range (hours)	Chosen	0	0.08	0.14	0.93	0.15
	All	0	0.13	0.17	1.13	0.14
Distance (km)	Chosen	0.5	192.3	381.8	4651.6	521.5
	All	0.5	243.1	450.8	6547.0	606.4
Toll cost (USD)	Chosen	0.0	0.0	1.9	244.1	12.0
	All	0.0	0.0	1.2	244.1	7.6
Fraction on major interstate roads	Chosen	0.00	0.00	0.24	1.00	0.43
	All	0.00	0.00	0.12	1.00	0.32
Fraction on highways	Chosen	0.00	0.84	0.69	1.00	0.34
	All	0.00	0.22	0.35	1.00	0.36
Fraction on trunk roads	Chosen	0.00	0.00	0.08	1.00	0.18
	All	0.00	0.01	0.09	1.00	0.15
Fraction on primary or secondary roads	Chosen	0.00	0.02	0.19	1.00	0.20
	All	0.00	0.27	0.47	1.00	0.25
Fraction on tertiary or unclassified roads	Chosen	0.00	0.01	0.04	0.89	0.10
	All	0.00	0.03	0.09	1.00	0.15
		Driver characteristics				
Employment type fractions	Hired: 0.65;		Owner-operator: 0.33;		Unknown: 0.02	
Years of experience fractions	0–1: 0.01;		1–2: 0.02;		3–5: 0.01;	
	6–9: 0.16;		10+: 0.79;		Unknown: 0.01	
Fraction of trucks with toll tags	Yes: 0.76;		No: 0.23;		Unknown: 0.01	

route are classified in four categories based on the US administrative system implemented in OSM. Chosen routes tend to heavily use highways (e.g. Interstate, freeways and other divided and grade separated roads), whereas other generated routes also use lower class roads, especially primary and secondary ones (e.g. US and state roads).

The sample characteristics are consistent with industry statistics. About two-third of the drivers are hired drivers, and the rest are owner-operators. This is comparable with estimates by Global Insight (2005) that 30% of heavy truck drivers are owner-operators. The vast majority are experienced drivers: 95% have at least 6 years of experience and 80% have been driving trucks for ten years or more. The median drivers' age in the sample is 50 years. Costello and Suarez (2015) report that the industry's median is 49 years. Finally, 76% of the trucks are equipped with electronic toll collection tags, which make using toll roads simpler. The sample may be subject to self-selection. However, the effect of self-selection bias is expected to be minor as there is no evident connection between volunteering to participate in the survey and route choice behaviors. For a detailed presentation of self-selection bias and techniques to mitigate it see the review in Mokhtarian and Cao (2008).

4. Model specification and estimation

Table 2 lists the variables used in the final specification of the model. The independent variables that are of interest capture the trade-offs between travel times, costs, distances and variability of travel times. The travel cost considered is the direct toll cost. The model also captures the effect of the use of a toll road, regardless of the toll cost. The travel time variability is captured by the square of the difference between the minimum and the maximum

Table 2. Definitions of variables used in the model specification.

Variable	Definition
SPtt	Dummy variable: 1 if route is shortest travel time route, 0 otherwise
LSnhwy	Dummy variable: 1 if route has the least distance driven on non-highway roads, 0 otherwise
TollRoute	Dummy variable: 1 if route involves tolls, 0 otherwise
tt110%	Dummy variable: 1 if route is up to 10% longer than shortest travel time route, 0 otherwise
nhwy+30	Dummy variable: 1 if the non-highway distance on the route is up to 30 kilometers longer than route with least non-highway distance, 0 otherwise
HiredTag	Dummy variable: 1 if driver is a hired driver for a company and truck is equipped with an electronic toll collection tag, 0 otherwise
HiredNoTag	Dummy variable: 1 if driver is a hired driver for a company and truck is not equipped with an electronic toll collection tag, 0 otherwise
lnTT	Log of travel time (hours)
ttRangeSq	Square of the difference between lowest and highest travel times within the day (hours)
Toll	Toll amount (2014 USD)
Dist	Length of route (100-kilometers)
Hired	Dummy variable: 1 if driver is a hired driver for a company, 0 otherwise
Experience10	Dummy variable: 1 if driver has over 10 years of experience, 0 otherwise
Road1	Fraction of route distance that used class 1 roads (highways)
Road2	Fraction of route distance that used class 2 roads (Trunk roads)
Road3	Fraction of route distance that used class 3 roads (primary and secondary roads)
Road4	Fraction of route distance that used class 4 roads (Tertiary and Unclassified roads)
MajorIS	Dummy variable: 1 if a third or more of the route is on major Interstate roads, 0 otherwise.
Google	Dummy variable: 1 if route was recommended by Google, 0 otherwise.
Bing	Dummy variable: 1 if route was recommended by Bing, 0 otherwise.
Mapquest	Dummy variable: 1 if route was recommended by Mapquest, 0 otherwise.
Inrix	Dummy variable: 1 if route was recommended by Inrix, 0 otherwise.

travel times that were measured over the day in the time-dependent travel time data. This measures the variability of travel times over the day. A measure of the variability of travel time at a time of day period may be more appropriate for the route choice model. However, it would require a much richer source of data, which may not be readily available in the context of a large geographic area.

The first four variables listed in the table are labels, identifying specific routes as having the best value in some attribute: shortest travel time, the shortest distance, the least number of non-highway kilometers and as being a route involving tolls. The next group of variables are related to routes that are near-best, that is within a certain distance metric from the best routes, with respect to an attribute: routes that are up to 10% and 5% longer than the shortest in terms of travel time or distance, respectively, and routes that involve up to 30 more non-highway kilometers compared to the route with the least non-highway distance. Routes that are included in these categories may be perceived as better based on their superior properties in the specific attributes. The tolerance allowed in these near-best values reflects people's imperfect knowledge of the values of these attributes, and their inherent variability. The thresholds used with these variables were selected, after some trial and error, such that they provide the best fit to observed choices.

Attributes of the routes were also interacted with characteristics of the driver in order to capture the different sensitivities of the various groups of drivers. In the final model, these characteristics are the type of driver (whether the driver is an owner-operator or a hired driver), existence of an electronic toll tag in the truck, and the level of experience the driver has (less or over 10 years).

The routes used in the model were generated using four navigation systems that incorporate different route planning capabilities (with and without tolls, avoiding or preferring highways and so on). For each route, dummy variables for the system(s) that generated it are introduced. These are meant to implicitly capture the underlying criteria that the navigation systems use when making their route recommendations.

The utility functions for the resulting model are given by:

$$\begin{aligned} U_{int} = {} & \beta_{SPtt}\delta_{int}^{SPtt} + \beta_{LSnhwy}\delta_{int}^{LSnhwy} + \beta_{tt110\%}\delta_{int}^{tt110\%} + \beta_{nhwy+20}\delta_{int}^{nhwy+20} \\ & + \beta_{tollHiredNoTag}\delta_{nt}^{hiredNoTag}\delta_{int}^{tollRoute} - \beta_{LnTT,n}\ln TT_{int} - \beta_{ttRangeSq,n}ttRangeSq_{int} \\ & + (Toll/Dist)_{int}(\beta_{tollPerMile} + \beta_{tollPerMileHiredTag}\delta_{n}^{HiredTag}) \\ & + Dist_{int}(\beta_{dist} + \beta_{distHired}\delta_{n}^{Hired} + \beta_{distExperiance10}\delta_{n}^{Experiance10}) + \beta_{road2}fr_{int}^{road2} \\ & + \beta_{road3}fr_{int}^{road3} + \beta_{road4}fr_{int}^{road4} + \beta_{majorIS}\delta_{int}^{majorIS} + \beta_{Google}\delta_{int}^{Google} \\ & + \beta_{Bing}\delta_{int}^{Bing} + \beta_{Mapquest}\delta_{int}^{Mapquest} + \beta_{Inrix}\delta_{int}^{Inrix} + \beta_{PS}\ln PS_{int} + \varepsilon_{int} \end{aligned} \tag{10}$$

Proposed models were estimated using the MLOGIT package (Croissant 2012) in the R statistical software. The method of simulated maximum likelihood was used with 10,000 Halton draws. The model estimation results are presented in Table 3. To further evaluate the effect of the various groups of variables, Table 4 presents estimation results of models that omit specific groups of variables from the full model:

(1) A model that excludes all route dummy variables (labels and maps sources). It only retains the variables related to the characteristics of the route alternatives.
(2) A model that excludes all systematic heterogeneity variables that capture differences among driver types and availability of ET toll tags
(3) A model that excludes random heterogeneity in travel time and travel time heterogeneity

Overall, the estimates are consistent with prior expectations. The signs for the coefficients of travel time, toll cost and squared travel time range are all negative. These imply that increases in the values of these variables for a specific route alternative reduce the utility of that route and the probability that it is chosen.

The label variables in the model capture the preference of drivers to routes that are best or near best with respect to some attribute. The coefficients of these variables are mostly positive. Drivers prefer routes that are shorter in terms of travel time and that involve as little driving on non-highway roads as possible. For these variables, routes that are near-best are also preferred to those that are not. This may be capturing screening criteria that drivers use in order to reduce the set of alternative routes that they consider in making their final selection. These effects are stronger than the preferences for the best routes with the corresponding criteria. They may also reflect the imperfect information drivers have on the exact attributes for the various routes and measurement errors. The model fit is reduced substantially when these variables are excluded from the model. The parameters of several of the other variables become insignificant.

A toll route dummy variable is significant only when interacted with hired drivers with no tags. The coefficient is negative, suggesting avoiding toll routes. This may reflect policies or decisions of the companies that do not provide drivers with tag, strongly discouraging the use of toll roads. For other driver types, no significant result was found, which indicates

Table 3. Estimation results.

Parameter	Estimate	*t*-statistic	*p*-value
Shortest travel time (β_{SPtt})	0.316	3.16	0.002
Least non highway distance (β_{LSnhwy})	0.281	2.46	0.014
Travel time within 10% ($\beta_{tt110\%}$)	0.332	2.26	0.024
Non highway distance plus 30 ($\beta_{nhwy+30}$)	0.853	5.03	$<$ 0.001
Toll route – Hired No tag ($\beta_{tollHiredNoTag}$)	−1.276	−2.78	0.005
Log Travel time (β_{lnTT})	−1.860	39.8	$<$ 0.001
Std – log travel time	0.525	14.2	$<$ 0.001
Travel time range squared ($\beta_{ttRangeSq}$)	−0.703	1.69	0.090
Std – travel time range squared	2.045	5.99	$<$ 0.001
Covariance – travel time and range	0.814	2.27	0.023
Toll per km ($\beta_{tollPerKm}$)	−0.127	−2.09	0.037
Toll per km – Hired with tag ($\beta_{tollPerKmHiredTag}$)	0.106	1.71	0.087
Distance (β_{dist})	−1.160	−35.9	$<$ 0.001
Distance - Hired ($\beta_{distHired}$)	−0.654	−26.2	$<$ 0.001
Distance – Experience 10 ($\beta_{distExperience10}$)	1.021	35.7	$<$ 0.001
Road class 2 (β_{road2})	−0.945	−1.80	0.072
Road class 3 (β_{road3})	−2.179	−5.68	$<$ 0.001
Road class 4 (β_{road4})	−5.091	−6.46	$<$ 0.001
Major interstate ($\beta_{majorIS}$)	−0.556	−2.67	0.008
Google (β_{Google})	0.564	4.16	$<$ 0.001
Bing (β_{Bing})	1.230	9.56	$<$ 0.001
Mapquest ($\beta_{Mapquest}$)	0.602	5.84	$<$ 0.001
Inrix (β_{Inrix})	0.449	4.18	$<$ 0.001
Ln (Path Size) (β_{PS})	−0.474	−2.77	0.006
|Observations: 1,021	Adjusted rho square: 0.594		
Initial log-likelihood: −2,222.10	Final log-likelihood: −877.69		

that, everything else being equal (including the cost), they do not have a strong preference for or against toll roads.

The toll costs are normalized in the model by the distance. As expected, it has a negative effect on the utility of the route. The variable was also interacted with indicators for hired drivers and for owner-operators, both with and without ETC toll tags. A difference was found only for hired drivers who drive trucks that are equipped with ETC tags. Compared to the other driver types, they have a significantly lower coefficient for the toll cost (in absolute value), which indicates that they are almost indifferent to the toll costs.

The coefficient of the travel distance is negative, as expected. This variable also captures indirect costs that are strongly correlated with it, such as fuel and wear. Interactions of this variable with hired driver status and with trucking experience of over 10 years are also included in the model. The coefficient for hired drivers is negative, which means that they are more sensitive to increasing travel distances compared to owner-operators. For experienced drivers, the coefficient of this interaction is positive, indicating that these drivers are less sensitive to travel distances compared to less experienced drivers.

Variables of the interaction of both toll and distance with driver characteristics re used in the model. These capture systematic heterogeneity in preferences. When these are eliminated from the model, the model fit, expressed by the adjusted rho square measure slightly decreases. The marginal effect of the toll cost variable is substantially smaller. The loss of the systematic heterogeneity also increased the random heterogeneity in the model, expressed by an increase in the variances of the random heterogeneity parameters for travel time and its variability.

Table 4. Estimation results for reduced models.

Parameter	Excluding labels	Excluding systematic heterogeneity	Excluding random heterogeneity
Shortest travel time (β_{SPtt})	-	**0.307**	**0.328**
Least non highway distance (β_{LSnhwy})	-	**0.259**	**0.287**
Travel time within 10% ($\beta_{tt110\%}$)	-	**0.362**	0.358
Non highway distance plus 30 ($\beta_{nhwy+30}$)	-	**0.789**	**0.851**
Toll route – Hired No tag ($\beta_{tollHiredNoTag}$)	-	-	**−1.317**
Log Travel time (β_{lnTT})	**−2.776**	**−1.797**	**−1.808**
Std – log travel time	**0.301**	**0.879**	-
Travel time range squared ($\beta_{ttRangeSq}$)	−0.551	**−0.433**	**−0.818**
Std – travel time range squared	0.143	**3.373**	-
Covariance – travel time and range	**1.971**	**0.844**	-
Toll per km ($\beta_{tollPerKm}$)	**−0.078**	**−0.017**	**−0.076**
Toll per km – Hired with tag ($\beta_{tollPerKmHiredTag}$)	0.061	-	0.063
Distance (β_{dist})	**−1.728**	**−1.094**	**−1.726**
Distance – Hired ($\beta_{distHired}$)	−0.745	-	−0.995
Distance – Experience 10 ($\beta_{distExperience10}$)	**1.202**	-	1.475
Road class 2 (β_{road2})	**−2.941**	**−1.067**	−0.945
Road class 3 (β_{road3})	**−4.582**	**−2.258**	**−2.178**
Road class 4 (β_{road4})	**−8.181**	**−5.230**	**−5.072**
Major interstate ($\beta_{majorIS}$)	-	**−0.596**	**−0.571**
Google (β_{Google})	-	**0.580**	**0.564**
Bing (β_{Bing})	-	**1.219**	**1.224**
Mapquest ($\beta_{Mapquest}$)	-	**0.609**	**0.604**
Inrix (β_{Inrix})	-	**0.441**	**0.449**
Ln (Path Size) (β_{PS})	**−1.294**	**−0.492**	**−0.476**
Adjusted rho square	0.534	0.590	0.592
Final log-likelihood	−1021.60	−890.51	−884.44

Parameters significant with p-value < 0.05 are in bold.

Travel times and the variability of travel times (captured by the travel time range, as discussed above) are included in the model with logarithmic and square transformations, respectively. These provided the best fit compared to other functional forms. The coefficients of these two variables were estimated as random parameters that follow a bi-variate log-normal distribution with estimated values:

$$\ln\left(\begin{bmatrix}\beta_{LnTT,n}\\ \beta_{ttRangeSq,n}\end{bmatrix}\right) \sim N\left(\begin{bmatrix}1.860\\ 0.703\end{bmatrix},\begin{bmatrix}0.276 & 0.814\\ 0.814 & 4.184\end{bmatrix}\right) \tag{11}$$

With this distribution the mean, median and standard deviation of the log of travel time coefficient are 7.371, 6.422 and 5.133. The corresponding values for the coefficient of travel time range square are 16.355, 2.019 and 80.363, respectively. The signs to the coefficients corresponding to log of travel time and travel time range squared are negative, because the two attributes are given negative signs in the utility function. The coefficient of travel time range squared has a very large variance. This indicates a high degree of heterogeneity among drivers with respect to the preference to avoid routes with high time variability. Replacement of the random parameters with fixed ones has a marginal effect on the model fit. But, several of the remaining parameters are insignificant, including those that capture systematic heterogeneity.

The estimates suggest significant trade-offs among travel time, travel time variability and toll costs. The estimation of random travel time and travel time variability coefficients leads to distributions of toll VOT and value of reliability (VOR). The use of nonlinear forms

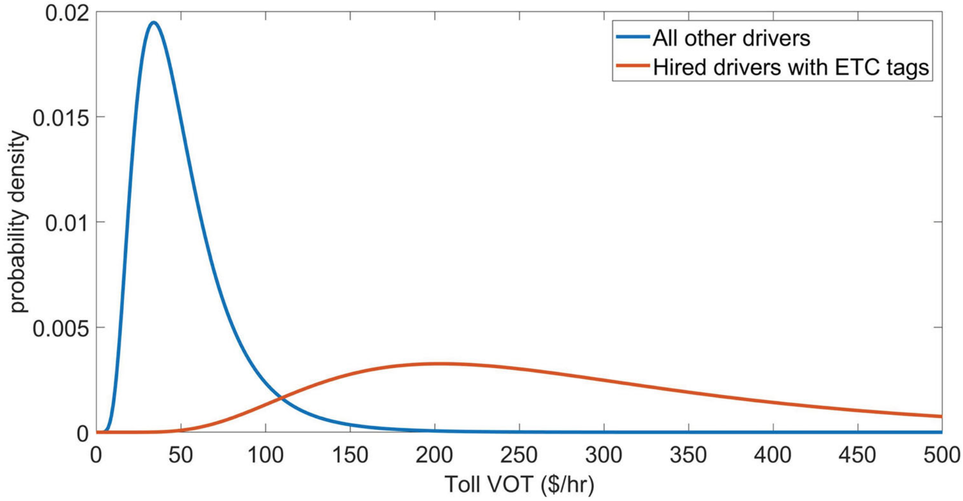

Figure 2. Toll VOT distributions for different driver types.

also makes their trade-offs dependent on the route travel time and travel time variability values. Furthermore, since the toll cost is interacted with driver characteristics, toll VOT also depends on the type of driver and ownership of ETC tag. The toll VOT reflected in this model is given by (variables and their units are as defined in Table 2):

$$VOT_n = \frac{\beta_{LnTT,n} Dist_n}{tt_n(\beta_{tollPerMile} + \beta_{tollPerMileHiredTag}\delta_n^{HiredTag})} \tag{12}$$

The toll VOR is the trade-off between travel time variability and the toll cost. It is given by:

$$VOR_n = \frac{2\beta_{ttRangeSq,n} ttRange_n Dist}{\beta_{tollPerMile} + \beta_{tollPerMileHiredTag}\delta_n^{HiredTag}} \tag{13}$$

To illustrate the variability of VOT and VOR, a 5-hour trip at an average speed of 90 kilometers per hour and a with variability of 0.2 h is assumed. For this case, the estimated VOT and VOR distributions, for hired drivers with ETC tag and all other drivers (hired drivers without ETC tag and owner operators) are plotted in Figures 2 and 3, respectively.

For the scenario described above, the median VOTs and VORs and their interquartile ranges (IQR) were calculated. Bliemer and Rose (2013) recommend calculation of medians over means when the range of parameters in the denominator of VOT includes zero. The calculations are based on the K&R method (Krinsky and Robb 1986, 1990) that uses simulation of the parameter values. This method was applied by Bliemer and Rose (2013) for discrete choice models with random parameters. See Gatta, Marcucci, and Scaccia (2015) for a review and evaluation of VOT confidence intervals estimation methods.

For hired drivers with ETC tags, the median toll VOTs is 272 $/Hr. The IQR is 164–469 $/hr. For all other drivers, the median is 46 $/Hr and the IQR is 29–78 $/hr. Using the same assumptions, the median toll VOR for hired drivers with ETC tags is 174 $/Hr. The corresponding IQR is 38–766 $/hr. For all other drivers, the median toll VOR is 30 $/Hr. The IQR is 7–129$/hr. The VOTs and VORs for hired drivers with ETC tags are high because

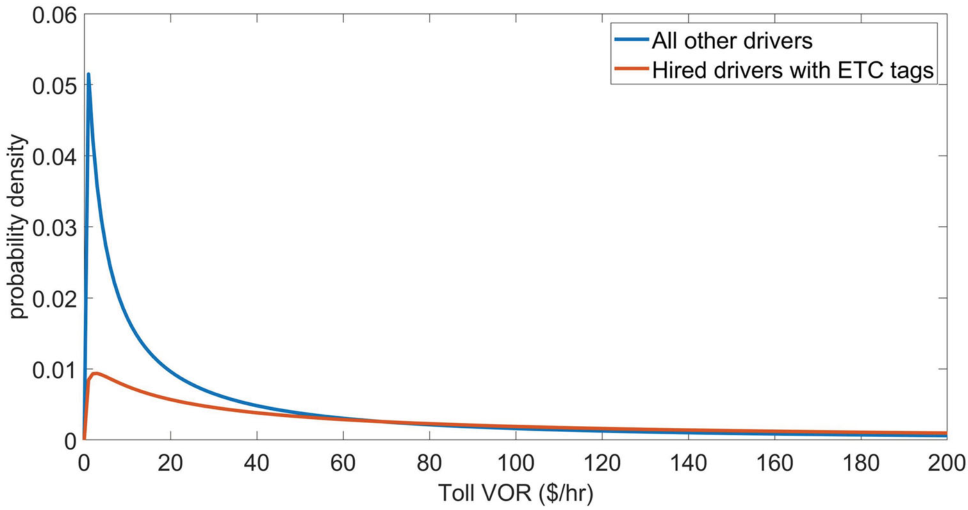

Figure 3. Toll VOR distributions for the different driver types.

the coefficients suggest that they are practically indifferent to the toll cost. Therefore, the denominators in the VOT and VOR expressions may be very small, which explains their wide IQRs. The standard deviation of the travel time variability coefficient is very large. This reflects a high level of heterogeneity among drivers in their toll VOR, which suggests against the use of deterministic VOR.

The path-size variable is designed to capture the effect of overlap among routes. Its coefficient is expected to be positive and theoretically equal 1. But, with the relatively sparse inter-urban network, attractive routes tend to overlap with other routes in the choice set. This is amplified by the use of the navigation applications to generate routes. Therefore, preferred routes tend to have high levels of overlap and lower path-size values, which explains the negative estimate of the value of the coefficient of this variable. A similar result, with a similar interpretation, was reported in Frejinger and Bierlaire (2007). Several other methods to capture the effects of overlap among routes have been proposed in the literature, each with several alternative correction factors. Most notably, the C-logit model (Cascetta et al. 1996), which uses commonality factors instead of path-size variables to capture similarities among routes. Starting from a different set of behavioral assumptions about the choice process, the quantum utility model (Vitetta 2016) obtains an interference term, which is another type of correction factor. A C-logit model was estimated. It yielded a slightly worse maximum likelihood of −878.29. The estimated commonality factor (CF) coefficient was 1.109 (p-value < 0.001). Other parameter values did not substantially differ from the path-size logit model results reported in Table 3. This result is consistent with earlier studies (e.g. Ramming 2002; Prato 2014) that did not find substantial differences in the performance of models with different correction factors.

5. Model application

In order to demonstrate the use of the model, it is applied to evaluate the fraction of various driver groups that are expected to use a toll road alternative in several specific trips. In each

Table 5. Attributes of the alternative routes.

	San Antonio – Dallas		Toledo – Chicago		Hamilton – Peterborough	
Attribute	Route A	Route B	Route A	Route B	Route A	Route B
Distance (Kms)	434	457	391	396	203	204
Peak travel time (Hr)	4:26	4:20	4:48	3:56	2:42	2:25
Off-peak travel time (Hr)	4:05	3:58	4:19	3:40	2:03	2:11
Travel time variability (Hr)	0:21	0:22	0:29	0:16	0:39	0:14
Toll – Tag ($)	0	22.92	0	70.48	0	99.07
Toll – no tag ($)	0	25.31	0	72.95	0	149.07
Road class 1	0.9958	0.9849	0.4167	0.9954	0.8622	0.6995
Road class 2	0	0	0.0555	0	0.0915	0
Road class 3	0.0040	0.0149	0.5220	0.0023	0.0457	0.3000
Road class 4	0.0002	0.0002	0.0058	0.0023	0.0006	0.0005

case two alternative routes are considered. It should be noted that the application with only two alternatives is useful to illustrate the effects of the various variables in the model. However, in an application, routes should be generated using the same method that was used in the estimation. The route choice fractions are predicted for different driver groups and conditions along several dimensions: Peak and off-peak period travel, truckers that are hired drivers or owner-operators and truckers with and without ETC tags. The situations evaluated are:

(1) A hired driver, with a toll tag, driving during the peak period. This is the base case.
(2) Same as the base case, but with travel during the off-peak period.
(3) Same as the base case, but the truck is not equipped with an ETC tag.
(4) Same as the base case, but the driver is an owner-operator.
(5) Same as the base case, but the driver is an owner-operator and the truck is not equipped with an ETC tag.

Three corridors are used in the analysis. The relevant attributes of the routes in each corridor are shown in Table 5:

(1) **San Antonio TX – Dallas TX (Texas corridor)**: Route A is a non-toll route that uses I-35. Route B is a tolled alternative that uses SH-130 toll road to bypasses Austin TX. The two alternatives are shown in Figure 4. Route B is longer by distance but offers travel times and travel time reliability savings. Both routes almost exclusively use highways.
(2) **Toledo OH – Chicago IL**: Route A is a non-toll route that uses lower class roads (US-20 and state road IN-2) for large parts of the trip. Route B uses the tolled highways (I-80/90 and Chicago Skyway). The two alternatives are shown in Figure 5. The two routes are similar in distance, but the toll road offers lower travel times and better travel time reliability.
(3) **Hamilton ON – Peterborough ON (Ontario corridor)**: This route crosses the Toronto metropolitan area. Route A is a non-toll route that uses the highways system that crosses the city (ON-403, QEW, ON-427, ON-401 and ON-115). Route B uses a toll road to cross the city (ON-407ETR) and connects via regional roads to the highway system. The two alternatives, which are roughly equal in distance, are shown in Figure 6. Route B uses class 3 roads to connect between the toll road and highway system. The travel time on the free route is shorter in the off-peak period, but longer in the peak period.

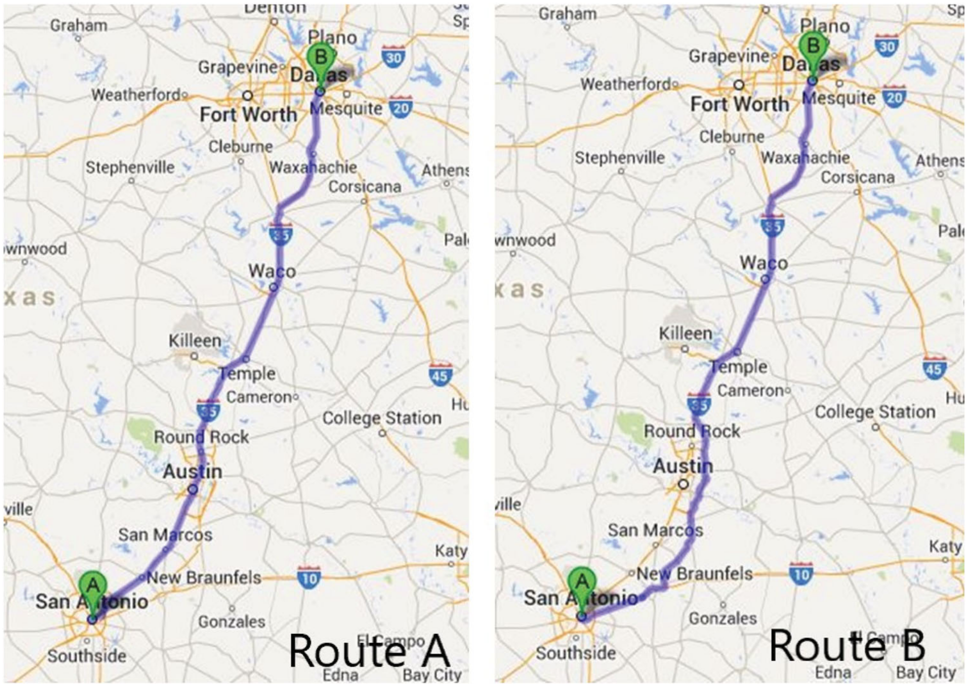

Figure 4. Alternative routes for the San-Antonio TX – Dallas TX trip (Google Maps 2017).

Figure 7 presents box plots of the probabilities of choosing route B (the toll route) for each drivers' segments in each of the three corridors. The different values were generated by drawing values from the distributions of the random coefficients. The expected values of the probabilities are marked with an 'x' sign. The plots demonstrate the high variability in preferences that is manifested in the random coefficients. With respect to the expected values of the probabilities, only the Ontario corridor shows differences in the route choice fractions between the peak and the off-peak periods. This is largely because it is the only corridor that exhibits large differences in travel time savings between the two periods. In all cases, hired drivers with trucks that are equipped with ETC tags are the most likely to use the toll routes. Those without ETC tags are the least likely to use toll roads. For owner-operators, the presence of ETC tags does not substantially affect their probability of choosing the tolled route. Among the corridors, a large majority of drivers are predicted to use the tolled route in the Toledo – Chicago corridor, which offers large travel times savings, improved reliability and higher-class roads. In the other corridors, the tolled route choice probabilities are much lower. In the Texas trip this is due to the minimal travel time savings and longer distance. In Ontario, the tolled route also offers smaller travel time savings, and only in the peak period. It also makes more use of lower-class roads.

The toll elasticity of demand captures the effect of changes in the toll cost on the probability of choice and market share of an alternative. Figure 8 presents the (negative of) toll elasticity of demand for the toll route alternatives in the various segments and corridors. The elasticity is in general low, demonstrating low sensitivity of the demand to the toll cost. In particular, the price elasticity of the demand for toll roads for hired drivers with ETC tags is very low. It is highest for hired drivers in trucks without ETC tags. The Ontario trip exhibits

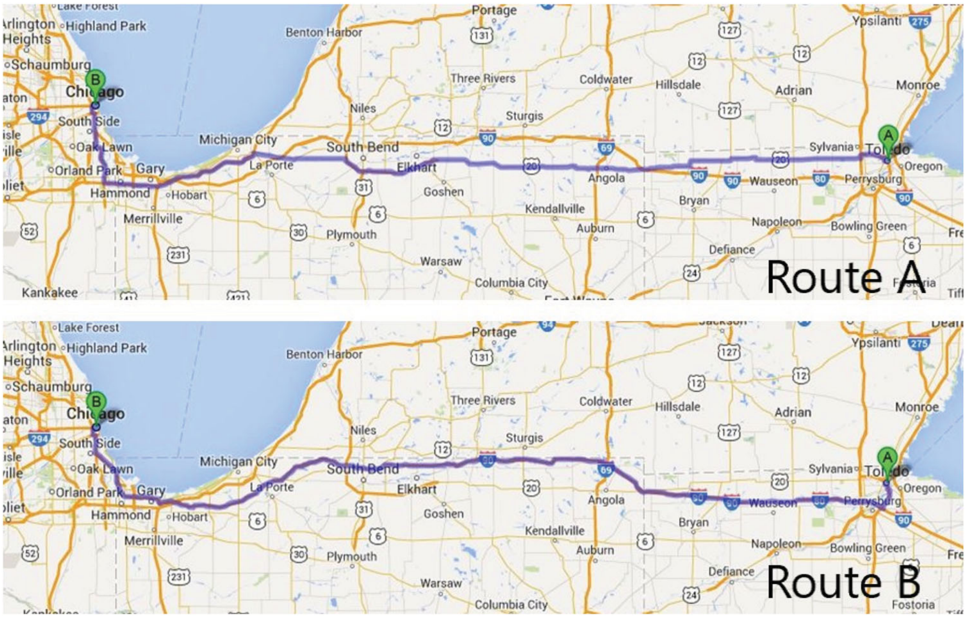

Figure 5. Alternative routes for the Toledo OH – Chicago IL trip (Google Maps 2017).

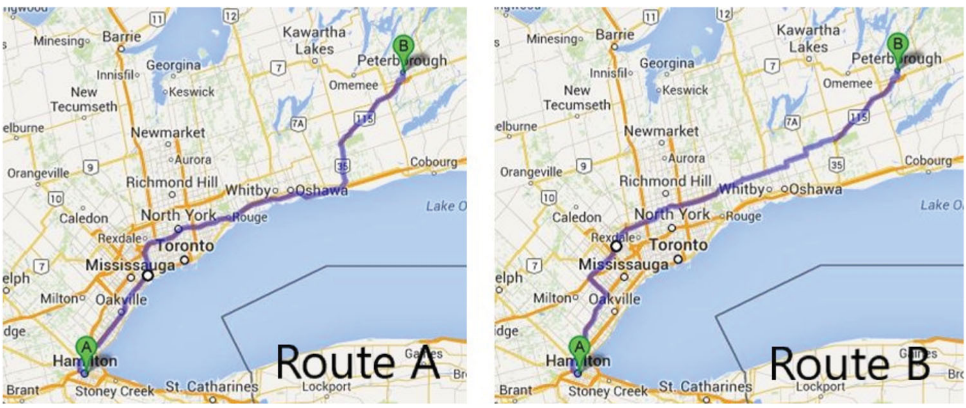

Figure 6. Alternative routes for the Hamilton ON – Peterborough ON trip (Google Maps 2017).

a different behavior. There, the demand by owner-operators is elastic. This is a result of the high toll cost and relatively short trip in this corridor. This combination makes the value of the cost per distance variable high and influential in the model.

6. Summary and conclusions

The research represents an attempt to better understand the route choices that intercity truck drivers make. A data collection methodology based on GPS data and user verification through a web interface was developed and implemented. The RP data collected was used to develop a route choice model that accounts for the attributes of the trip (e.g. travel time,

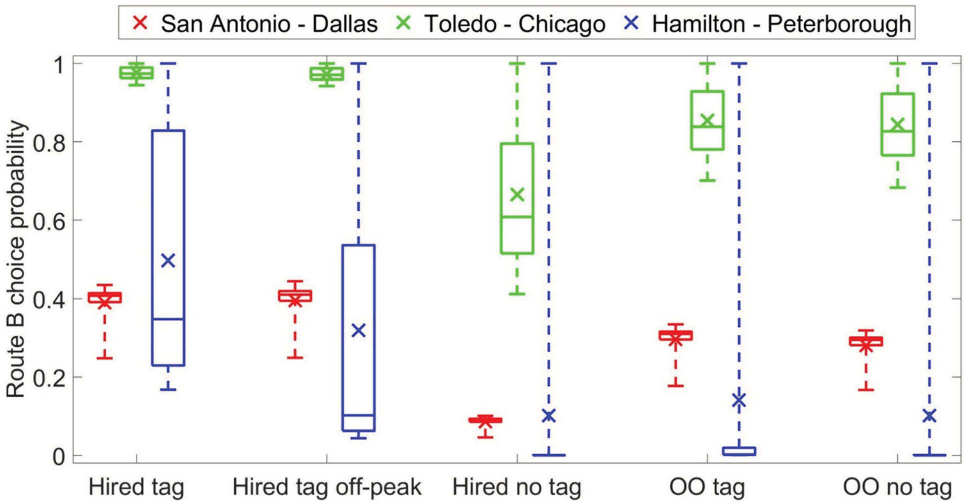

Figure 7. Probabilities of choosing route B (the toll route) for various driver segments.

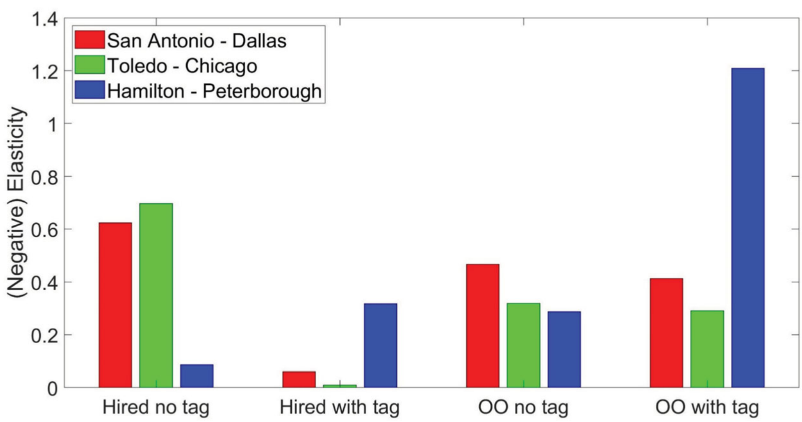

Figure 8. Toll (negative) elasticity of demand.

travel time range, distance, and classes of the roads used) and the characteristics of the truck drivers (e.g. owner operators and hired drivers, trucks with and without ETC tags, and trucking experience).

For various driver types, random coefficients for both travel time and travel time variability were used in order to capture the heterogeneity of preferences. Both are significant in the model and show large standard deviations of the coefficients. These suggest large differences in preferences among drivers. The model captures inter-participant heterogeneity. The large values estimated suggest that it may be useful to also model intra-respondent heterogeneity. Appropriate models and estimation techniques were proposed, for example, by Hess and Rose (2009) and Becker et al. (2018).

Large differences in toll VOT and toll VOR were also found between owner operators and hired drivers with or without ETC tags. Based on the estimation results and model application, the willingness to pay for tolls is substantially higher for hired drivers with ETC tag compared to the other drivers (i.e. owner operators and hired drivers without ETC tag). The underlying reason may relate to that hired drivers with ETC tag do not need to pay for tolls out of their pockets. In addition, the equipment of ETC tags improves the convenience of using toll facilities and may also imply that the trucking companies encourage toll road usage. In contrast, Toll VOT and toll VOR for hired drivers without ETC tag and for owner operators are similar. It may be that companies that do not provide ETC tags to hired drivers may also not cover the toll cost, so the hired drivers show similar behavior as owner-operators that bear the toll costs themselves. These results imply that differentiated toll pricing based on driver type and ETC equipment can increase the toll road use.

Another set of variables that were found to affect route choices are the road classes. Drivers showed strong preference towards higher class roads, and in particular highways (class 1). Routes that extensively use lower-class roads are considered substantially inferior. In terms of travel distance, hired drivers were found to be more sensitive to the travel distance compared to owner-operators. Experienced drivers put much lower weight on the travel distance in choosing their routes compared to less experienced drivers.

The model was applied to predict individuals' choices of routes in real-world corridors under different scenarios. Toll road operators may find the results instructive to predict use of tolled and toll-free roads by different population groups. It may also help design incentives and personalized tolling policies and marketing efforts to affect these choices based on the driver's characteristics.

The model suffers several weaknesses. There may exist a sampling bias as participation was encouraged by providing the drivers with monetary incentives, which may contribute to a lower estimate of VOT and VOR than the population. A choice set generation technique was used to ensure that a variety of route information can be obtained. But, the alternatives recommended by navigation engines largely overlap with one another.

Several directions for future research may be suggested. Other model structures such as latent class models and error component models may be considered as suggested in Marcucci and Gatta (2012). Other approaches for capture the effects of route overlap could also be tested. In addition, intra-driver heterogeneity in route choices may be investigated. In the current study, this was not possible due to insufficient number of data points for individual drivers.

Most drivers in the sample do not make fixed trips. As a result, there were not many opportunities in the data to observe repeated route choices on similar origins and destinations. If participants would be tracked over longer periods, these could be observed, and support modeling of dynamic behaviors, such as tendency to keep to the same route. Furthermore, the estimated model is based on the random utility theory with a mixed error structure. Other models that are not based on random utility may also be useful in this context. Examples include random regret (Chorus 2010; Prato 2014), fuzzy logic (Lotan and Koutsopoulos 1993; Henn 2000; Ridwan 2004) and quantum utility model (Vitetta 2016).

Acknowledgement

The authors thank the anonymous referees for constructive comments and suggestions. Any remaining errors are our own.

Disclosure statement

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

ORCID

Tomer Toledo http://orcid.org/0000-0003-3611-6024

References

Arentze, T., T. Feng, H. Timmermans, and J. Robroeks. 2012. "Context-Dependent Influence of Road Attributes and Pricing Policies on Route Choice Behavior of Truck Drivers: Results of a Conjoint Choice Experiment." *Transportation* 39 (6): 1173–1188.

Austroads. 2003. Economic Evaluation of Road Investment Proposals: Valuing Travel Time Savings for Freight, Austroads Publication No. AP–R230/03, Sydney, Australia.

Bain, R. 2009. "Error and Optimism Bias in Toll Road Traffic Forecasts." *Transportation* 36 (5): 469–482.

Bain, R., and L. Polakovic. 2005. Traffic Forecasting Risk Study 2005: Through Ramp-Up and Beyond, Standard & Poor's, London UK.

Becker, F., M. Danaf, X. Song, B. Atasoy, and M. Ben-Akiva. 2018. "Bayesian Estimator for Logit Mixtures with Inter- and Intra-Consumer Heterogeneity." *Transportation Research Part B: Methodological* 117: 1–17.

Ben-Akiva, M., and M. Bierlaire. 1999. "Discrete Choice Methods and Their Applications to Short-Term Travel Decisions." In *Handbook of Transportation Science*, edited by R. Hall, 5–34. Boston, MA: Kluwer Academic Publishers.

Ben-Akiva, M., D. Bolduc, and J. Q. Park. 2008. "Discrete Choice Analysis of Shippers' Preferences." In *Recent Developments in Transport Modeling: Lessons for the Freight Sector*, edited by by M. Ben-Akiva, H. Meersman, and E. Van de Voorde, 135–155. Bingley UK: Emerald Group.

Ben-Akiva, M., T. Toledo, J. Santos, N. Cox, F. Zhao, Y. J. Lee, and V. Marzano. 2016. "Freight Data Collection Using GPS and Web-Based Surveys: Insights from US Truck Drivers' Survey and Perspectives for Urban Freight." *Case Studies on Transport Policy* 4 (1): 38–44.

Bergkvist, E. 2001. Freight Transportation: Valuation of Time and Forecasting of Flows, Umea Economic Studies 549, Umea University, Sweden.

Bliemer, M. C. J., and P. H. L. Bovy. 2008. "Impact of Route Choice set on Route Choice Probabilities." *Transportation Research Record* 2076: 10–19.

Bliemer, M. C. J., and J. M. Rose. 2013. "Confidence Intervals of Willingness-to-Pay for Random Coefficient Logit Models." *Transportation Research Part B: Methodological* 58: 199–214.

Bovy, P. H. L. 2009. "On Modelling Route Choice Sets in Transportation Networks: A Synthesis." *Transport Reviews* 29 (1): 43–68.

BTS. 2017. Freight Facts and Figures 2017, Bureau of Transportation Statistics, US Department of Transportation. http://www.bts. gov/programs/freight_transportation/html/trucking.html.

Cascetta, E., A. Nuzzolo, F. Russo, and A. Vitetta. 1996. "A Modified Logit Route Choice Model Overcoming Path Overlapping Problems: Specification and Some Calibration Results for Interurban Networks." In Transportation and Traffic Theory. Proceedings from the 13th International Symposium on Transportation and Traffic Theory, Lyon, France. J.B. Lesort, ed. Pergamon.

Chorus, C. G. 2010. "A New Model of Random Regret Minimization." *European Journal of Transportation and Infrastructure Research* 10 (2): 181–196.

Costello, B., and R. Suarez. 2015. Truck Driver Shortage Analysis 2015, American Trucking Associations.

Croissant, Y. 2012. Estimation of Multinomial Logit Models in R: The mlogit Packages, R package version 0.2–2. http://cran.r-project.org/web/packages/mlogit/vignettes/mlogit.pdf.

Danielis, R., E. Marcucci, and L. Rotaris. 2005. "Logistics Managers' Stated Preferences for Freight Service Attributes." *Transportation Research Part E: Logistics and Transportation Review* 41: 201–215.

Feng, T., T. Arentze, and H. Timmermans. 2013. "Capturing Preference Heterogeneity of Truck Drivers' Route Choice Behavior with Context Effects Using a Latent Class Model." *European Journal of Transport and Infrastructure Research* 13 (4): 259–273.

Fowkes, A. S., and A. Whiteing. 2006. "The Value of Freight Travel Time Savings and Reliability Improvements – Recent Evidence from Great Britain." Proceedings of the European Transport Conference 2006.

Frejinger, E. 2007. "Random Sampling of Alternatives in a Route Choice Context." Proceedings of the European Transport Conference, Leeuwenhorst, The Netherlands.

Frejinger, E., and M. Bierlaire. 2007. "Capturing Correlation with Subnetworks in Route Choice Models." *Transportation Research Part B: Methodological* 41 (3): 363–378.

Gatta, V., E. Marcucci, and L. Scaccia. 2015. "On Finite Sample Performance of Confidence Intervals Methods for Willingness to Pay Measures." *Transportation Research Part A* 82: 169–192.

Global Insight. 2005. "The U.S. Truck Driver Shortage: Analysis and Forecasts." Report Submitted to American Trucking Association, Arlington VA.

Google Maps. 2017. San-Antonio TX – Dallas TX, Toledo OH – Chicago IL, Hamilton ON – Peterborough ON corridors.

Henn, V. 2000. "Fuzzy Route Choice Model for Traffic Assignment." *Fuzzy Sets and Systems* 116 (11): 77–101.

Hess, S., M. Quddus, N. Rieser-Schüssler, and A. Daly. 2015. "Developing Advanced Route Choice Models for Heavy Goods Vehicles Using GPS Data." *Transportation Research Part E: Logistics and Transportation Review* 77: 29–44.

Hess, S., and J. M. Rose. 2009. "Allowing for Intra-Respondent Variations in Coefficients Estimated on Repeated Choice Data." *Transportation Research Part B: Methodological* 43 (6): 708–719.

Hunt, J. D., and J. E. Abraham. 2004. "Stated Response Examination of Factors Influencing Commercial Movement Route Choice Behavior.", Proceedings of the 8th International Conference on Applications of Advanced Technologies in Transportation Engineering, Beijing, China, 72–77.

Hyodo, T., and Y. Hagino. 2010. "Modeling Truck Route Choice Behavior by Traffic Electronic Application Data." Paper presented at the Transportation Research Board SHRP 2 Symposium: Innovations in Freight Demand Modeling and Data, Washington DC.

Ismail, K., T. Sayed, and C. Lim. 2009. "A Study of the Commercial Value of Time for Operation at Border Crossings." Proceedings of the 2009 annual Conference of the Transportation Association of Canada, Vancouver BC.

Jong, G. de. 2000. "Value of Freight Travel Time Savings." In *Handbook of Transport Modelling*. 2nd ed, edited by D. A. Hensher, and K. J. Button, 553–564. New York: Elsevier.

Jong, G. de, S. Bakker, M. Pieters, and P. W. Donselaar. 2004. "New Values of Time and Reliability in Freight Transport in the Netherlands." Proceedings of the European Transport Conference 2004.

Jovicic, G. 1998. "Application of Models Based on Stated and Revealed Preference Data for Forecasting Danish International Freight Transport." Proceedings of the Annual Transport Conference at Aalborg University.

Kawamura, K. 2000. "Perceived Value of Time for Truck Operators." *Transportation Research Record* 1725: 31–36.

Kim, J. J., W. Pasco, and S. K. Kothapalli. 2017. "Route Choice Characteristics of Owner-Operated Trucks on Southern California Freeways." Report METRANS 15–10, California State University Long Beach CA.

Knorring, J. H., R. He, and A. L. Kornhauser. 2005. "Analysis of Route Choice Decisions by Long-Haul Truck Drivers." *Transportation Research Record* 1923: 46–60.

Krinsky, I., and A. L. Robb. 1986. "On Approximating the Statistical Properties of Elasticities." *Review of Economics and Statistics* 68 (4): 715–719.

Krinsky, I., and A. L. Robb. 1990. "On Approximating the Statistical Properties of Elasticities: A Correction." *Review of Economics and Statistics* 72 (1): 189–190.

Kurri, J., A. Sirkia, and J. Mikola. 2000. "Value of Time in Freight Transport in Finland." *Transportation Research Record* 1725: 26–30.

Lotan, T., and H. Koutsopoulos. 1993. "Models for Route Choice Behavior in the Presence of Information Using Concepts From Fuzzy set Theory and Approximate Reasoning." *Transportation* 20 (2): 129–155.

Marcucci, E., and V. Gatta. 2012. "Dissecting Preference Heterogeneity in Consumer Stated Choices." *Transportation Research Part E: Logistics and Transportation Review* 48 (1): 331–339.

Miao, Q., B. X. Wang, and T. M. Adams. 2011. "Assessing the Value of Delay to Truckers and Carriers." Final Report CFIRE 03–15 and UTCM 09-00-45, University Transportation Research Program.

Mokhtarian, P. L., and X. Cao. 2008. "Examining the Impacts of Residential Self-Selection on Travel Behavior: A Focus on Methodologies." *Transportation Research Part B: Methodological* 42 (3): 204–228.

Prato, C. G. 2009. "Route Choice Modeling: Past, Present and Future Research Directions." *Journal of Choice Modelling* 2 (1): 65–100.

Prato, C. G. 2014. "Expending the Applicability of Random Regret Minimization for Route Choice Analysis." *Transportation* 41 (2): 351–375.

Prato, C. G., and S. Bekhor. 2007. "Modeling Route Choice Behavior: How Relevant is the Choice Set Composition?" *Transportation Research Record* 2003: 64–73.

Quattrone, A., and A. Vitetta. 2011. "Random and Fuzzy Utility Models for Road Route Choice." *Transportation Research Part E: Logistics and Transportation Review* 47 (6): 1126–1139.

Ramming, M. S. 2002. "Network Knowledge and Route Choice." PhD dissertation, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology.

Ridwan, M. 2004. "Fuzzy Preference Based Traffic Assignment Problem." *Transportation Research Part C: Emerging Technologies* 12 (3-4): 209–233.

Rowell, M., A. Gagliano, and A. Goodchild. 2014. "Identifying Truck Route Choice Priorities: The Implications for Travel Models." *Transportation Letters* 6 (2): 98–106.

Smalkoski, B., and D. Levinson. 2005. "Value of Time for Commercial Vehicle Operators." *Journal of the Transportation Research Forum* 44: 89–102.

Small, K. A., R. Noland, X. Chu, and D. Lewis. 1999. "Valuation of Travel Time Savings and Predictability in Congested Conditions for Highway User Cost Estimation." NCHRP Report 431, US Department of Transportation, Washington DC.

Tahlyan, D., A. R. Pinjari, T. D. Luong, and S. Ozkul. 2017. "Truck Route Choice Modeling using Large Streams of GPS Data." Report CAIT-UTC-NC32, University of Florida, Tampa FL.

Toledo, T., Y. Sun, K. Rosa, M. Ben-Akiva, K. Flanagan, R. Sanchez, and E. Spissu. 2013. Decision-Making Process and Factors Affecting Truck Routing, in Freight Transport Modelling, 233–249.

Vitetta, A. 2016. "A Quantum Utility Model for Route Choice in Transport Systems." *Travel Behavior and Society* 3: 29–37.

Wang, Z., and A. Goodchild. 2014. "GPS Data Analysis of the Impact of Tolling on Truck Speed and Routing: Case Study in Seattle, Washington." *Transportation Research Record* 2411: 112–119.

Wood, H. P. 2011. "Truck Tolling: Understanding Industry Tradeoffs When Using or Avoiding Toll Facilities." Final Report NCFRP Project 19 and NCHRP Project 19–09, US Department of Transportation, Washington DC.

Wynter, L. M. 1995. "Stated Preference Survey for Calculating Values of Time of Road Freight Transport in France." *Transportation Research Record* 1477: 1–6.

Zamparini, L., and A. Reggiani. 2007. "Freight Transport and the Value of Travel Time Savings: A Meta-Analysis of Empirical Studies." *Transport Reviews* 27: 621–636.