Pricing urban congestion: a structural random utility model with traffic anticipation

(Previous Title: Optimal Pricing of Endogenous Congestion: A Disaggregated Approach)

Christelle Viauroux
University of Maryland, Baltimore County*

April 7, 2011

Abstract

We design and estimate a game theoretic congestion pricing mechanism in which the regulator aims at reducing urban traffic congestion by price discriminating travelers according to their Value Of Time (VOT). Travelers' preferences depend on their observable characteristics, on the endogenous amount of congestion anticipated, on their Marginal Utility (MU) of income and on some unobserved factors. Using a French household survey, we estimate the demand models to simulate different pricing mechanisms. We find that unobserved determinants of transportation demand are significant and are used to measure the anticipated time spent in traffic and the comfort of traveling: diverging from these expectations is felt as more discomfort than if no expectations were formed a-priori. However some of this discomfort is derived from travelers' marginal utility of income: the lost time in traffic is clearly "unpleasant" because of its opportunity cost. When the regulator and the transportation provider share common objectives, we show that a great welfare improvement can be achieved when implementing a homogenous pricing that accurately accounts for travelers VOT.

Classification codes: R4, D8, D6.

*Department of Economics, 1000 Hilltop Circle, Baltimore, MD 21250, USA. Telephone: 410-455-3117, Fax: 410-455-1054, Email: ckviauro@umbc.edu. The author is especially grateful to the Editor Z. Eckstein, the Associate Editor H. Sieg and two anonymous referees for their very useful comments. She also thanks V. Komornik, R. Pollard and D. Sappington for their helpful suggestions and comments. An early version of this paper was presented at the Western Econometric Conference, Cincinnati, 2006, and at the European Economic Association Conference, Budapest, 2007.

Keywords: Bayesian game; endogenous congestion; value of time; modes reputation; optimal pricing; welfare; structural estimation.

1 Introduction

The predominant form of urban congestion pricing in Europe is flat rate pricing. For example, car travelers driving in Central London on weekdays between 7:00am and 6:30pm were required to pay 5 pounds, increasing to 8 pounds since 2005 (see also Mannheim, Oslo or Bergen, for similar experiments). This simple pricing scheme was chosen because it was relatively fast and easy to implement. The experiment has been successful in many ways, however, there are still concerns regarding the optimal price and/or price structure. In particular, once travelers pay the fee, they have no incentive to minimize driving. In this context, a variable road use fee that would reflect the heterogeneity of drivers, such as the type of vehicle, time and frequency would seem more appropriate. This approach would most accurately reflect the external social costs imposed by driving and would give travelers an incentive to minimize their negative impact by stimulating them to make choices that maximize both their own utility and society's.

A core issue of concern to urban network users, transportation operators and economists is the constantly changing network conditions arising from aggregate decisions and behavior. The literature on road networks (initiated by Beckmann et al., 1956 and extended by Dafermos, 1973 to heterogenous travelers) and on second best optimal pricing (see Gmez-Ibez and Small, 1994; Arnott et al., 1994, 1998; Emmerinck, 1998; Parry and Small, 2005) is based upon the assumption that congestion is stochastic. In these studies, congestion is usually modeled by a function of time and traffic flow, where congestion externalities are represented by a parameter to be estimated. Using this approach, the regulator is able to control the traffic flow using time varying congestion tolls. We argue that the network behavior depends on the aggregate load of the network - the result of many users' decisions on how to use the network. In this sense, it is important that the regulator is able to control the incentives of the travelers. Since urban congestion is caused by too many travelers competing for a limited road space, our objective is to find an economically efficient way to allocate network resources among travelers. Former studies show that incomplete information about travelers Value Of Time (VOT) or aversion to congestion is significant and that urban travelers are willing to pay a non-negligible amount of money to improve their traffic conditions. However, this goal requires that we know the true value that each traveler places on travelling.

We estimate a game theoretic demand model using a household survey made in the city of Montpellier (France) that can reproduce the Bayesian game that travelers play to decide upon the number of trips they make and the resulting aggregation that forms congestion. Viauroux (2007) uses this approach, derives the average willingness of travelers to pay on the network and derives some measures of welfare. However, this analysis differs by the following two points.

1- Her analysis did not account for travelers' unobserved characteristics or the unobserved attributes of the varying modes of transportation. Reputation, for example, may be perceived as useful information regarding the comfort or accessibility of a mode and this might affect the traveler's willingness to pay as well as reduce the regulator's ability to create the right incentives.

2- Travelers sensitivity to transportation costs is assumed here to vary with income and traffic conditions.

Indeed, Figures 1a-b (Appendix 3) summarize statistics on the number of car and bus trips by income range.[1] They show that car traveling is a normal good (or service) up to an income range between 2439 Euros and 2897 Euros while bus traveling is an inferior good up to an income range of (3,354;3,811] and a normal good thereafter. Our hypothesis is that the bus is cheaper in Montpellier for low income individuals such as the unemployed, students, and retired (see details on bus fares in data section) whose MU of income is high. As their income increases and their marginal utility of income decreases, the comfort of traveling by car outweighs its cost to a greater extent. After a certain income level, however, external factors such as traffic conditions reverse the effect of the MU of income in favor of more bus trips. More specifically, we will consider three scenarios: a- the MU of income is assumed constant; b- the MU of income is a linear function of the observed income range; c- the MU of income is a function of the travelers' VOT (see also Viauroux, 2008).

3- We investigate the link between travelers' heterogeneity and the welfare it generates (see also Small, 1992; Small and Yan, 2001). We derive a pricing mechanism aimed at reducing congestion on the network by promoting the use of public transportation. The derivation distinguishes different weights that the regulator might place on travelers' surplus to account for different market structures: the regulator may, for example, place more weight on the public transportation provider's profit than on travelers' surplus. This approach is close to McKie-Mason and Varian who in 1995 proposed a "smart market" mechanism that suggests an auction-based scheme to price internet congestion. Indeed, the urban European city network is similar, in many ways, to the internet network. The street network in the core areas that concentrates most activities (work, school, leisure and shopping) is rarely expanded.

We show that the regulator's policy will primarily target individuals who are less sensitive to transportation costs and whose probability of generating congestion is large. Intuitively, if the marginal cost of a trip is small, the fare is such that, if most individuals travel more than you (because their VOT is lower than yours), you should be assigned a lower fare as you are not held responsible

[1] Income is defined as the gross monthly income of the household, including the professional income of all members of the household including premiums such as a thirteen month, yield premium as well as other real estate income. For confidentiality reasons, the household head was asked to give only a range of income. Despite this precautionary survey measure, 30% of household heads refused to answer. In this case, we replaced missing values of income by its prediction: observed income has been regressed on variables characterizing the individual, such as age, sex and professional status (executive, farmer, retired, and factory worker).

for as much traffic congestion. Moreover, this fare is lower when the traveler's income is low (or sensitivity to cost is high). Hence, the fee increases with the likelihood of frequent travel that creates congestion and decreases with income. In this case, the regulator's intervention is primarily intended for individuals who are less sensitive to transportation costs and whose probability of generating congestion is large. However, as the marginal cost of a trip increases, more and more travelers will contribute to the finance of the capital/infrastructure by an amount that is lower for lower income travelers.

To the best of our knowledge, this is the first paper to derive and estimate a congestion pricing mechanism that accounts for the endogeneity of congestion using travelers' anticipations of traffic as well as the unobserved determinants of travelers' decisions and different levels of heterogeneity in the MU of income.

Estimation results show that reputation or other unobserved determinants of transportation demand (hereafter referred to as reputation)[2], affect significantly the joint traveling choices on the network. Reputation seems to be perceived as a reference measure of anticipated time spent and comfort of traveling. Hence, diverging from these expectations is felt as more discomfort than if no expectations were formed a-priori. While reputation significantly affects the choice of modes of transportation, it only significantly affects the use of public transportation during peak hours. Hence, for a given infrastructure, we find that the use of public transportation may increase if some improvements in reliability and comfort were made. Interestingly, reputation seems to have significantly more impact during off-peak times when travelers' choices are less constrained. Moreover, travelers are also more sensitive to traffic conditions during off-peak hours: during these times, the opportunity cost of traveling is felt to a greater extent because of the foregone leisure or shopping activities. It is also important to note that some of the nuisance coming from unexpected circumstances of traveling comes from the marginal utility of income: the lost time in traffic is clearly "unpleasant" because of its opportunity cost. In other words, when the heterogeneity in the MU of income is accounted for, tolerance to unforeseen events on the network increases.

Marginal cost pricing accounting for reputation of modes and observed heterogeneity of travelers' sensitivity to price (Model 2) generates a welfare gain relative to the pricing scheme designed in Viauroux (2007) of about 12%. When regulator and transportation provider share common objectives, our results show that welfare can be reduced when too much heterogeneity is accounted for. However, a great welfare improvement can be achieved by implementing a homogenous pricing mechanism that accurately accounts for travelers' VOT.

The paper is organized as follows. Section 2 presents theoretical model of transportation demand. Section 3 presents the data set as well as the estimation procedure, the parametrization of the model and discusses the identification of the model. Section 4 derives the optimal pricing policies in the case of complete and incomplete information about each traveler's aversion to traffic congestion.

[2] We assume that reputation of a mode represents the major unobserved determinant of transportation demand. We do so to ease the exposition of the intuitions and inferences of our results.

Estimation and simulation results are reported and analyzed in Section 5. Section 6 concludes.

2 The Model

Individuals' decisions to travel result from a Bayesian game in which travelers possess private information about their value of time (VOT) or 'type'. [3]

We consider the Bayesian game (Harsányi, 1967–68)

$$\Gamma = (I, Q, \Theta, (p_i)_{i\in I}, (u_{i,t})_{i\in I, t\in T}),$$

where I is the set of individuals i traveling in the city ($i = 1, ...I$); Q is the set of possible number of car and bus trips performed with fare type of j ($j = 1, ...J$) during period t ($t = 1, ...T$) by individual i, denoted by $q_{i,t} = (q^c_{i,t}, q^{bj}_{i,t})$; $\Theta = \Theta_i \times \Theta_{-i}$ is the set of travelers' types $\theta = (\theta_i, \theta_{-i})$ representing their tolerance to traffic congestion or the value that they associate to the time lost in traffic (VOT), where $\Theta_{-i} := \underset{k\in I-i}{\times} \Theta_k$. Hence, a low tolerance for congestion corresponds to a high VOT.

We let θ_i index individual i type (for $i = 1, \ldots, I$) and θ_{-i} index the type of travelers other than i. $\theta_i \in \Theta_i = [\underline{\theta_i}, \overline{\theta_i}]$ where $\underline{\theta_i} > 0$ is a taste parameter indexing the least tolerance for congestion (greatest VOT), while $\overline{\theta_i}$ represents the greatest tolerance for congestion (lowest VOT). The regulator belief about θ_i is reflected in the joint density $f_i(\theta_i)$ and its cumulative distribution function $F_i(\theta_i)$ is assumed to be continuous with support $\Theta \subseteq]0, 1]$ for $i = 1, .., I$. We denote by $p_i(.|\theta_i)$ the probability distribution over Θ_{-i} and by σ the randomized-strategy profile for the game, such that $\sigma = (\sigma_i(q_{i,t}|\theta_i))_{i\in I, q_{i,t}\in Q_{i,t}, \theta_i\in\Theta}$.

Travelers' preferences depend upon the number of trips taken and upon the choice (and perception) of a mode of transportation: the traveler can either walk to his destinations, i.e. use neither the car nor the bus; he can use the car at least once but not the bus; use the bus at least once but not the car. The choice of each of these modes is affected by an unobserved factor, reflecting their reputation, and these factors will also affect transportation usage.

Travelers' preferences depend on the anticipated level of congestion, which in turn is determined by the (past) decisions of all other individuals. Travelers decide on the number of trips to take with a particular mode of transportation. They consider both car and public transportation use, hereafter referred to as "bus", which are both affected by the inconvenience of traffic congestion created, mostly, by cars. We assume that all roads leading to a central business district are congested but that the level of congestion varies; hence the selection is made

[3] Given his type, a traveler determines the optimal probability distribution of use for each relevant mode of transportation and chooses the number of trips he wants to make for a given period of time. For example, if his value of time is low, he may lower his number of car trips and increase his probability of using an alternative mode of transportation, such as the bus. Note, that his best probability distribution depends on the probability of all other travelers (of different possible types) taking trips.

across modes of transportation that could involve more traveling time (such as the bus for example). This assumption is in line with pricing experiments such as Singapore, London or Mannheim, for example, where the optimal regulatory policy was applied to an entire congested district.

Individuals' preferences for traveling are affected by their perception of the associated mode of transportation as well as the choice of consuming other goods in quantity $\nu_{i,t}$. We also assume that there exists unobserved factors relevant to the decision of the individual toward one mode of transportation that will impact the number of trips taken and the quantity $\nu_{i,t}$ of other goods consumed. We assume that these unobserved factors alter preferences in the same way as the quality of the mode, the aversion to traffic congestion and the amount of traffic on the network. We use the random utility approach from Manski (1977) and define the traveler unconditional utility function as a variation $(u_{i,t})_{i\in I}$ of the functions introduced by Blackburn in 1970 and Hanemann in 1984 for $j = 1, ...J$,

$$\begin{aligned} U_{i,t} &= U_{i,t}(q_{i,t}, q^*_{-i,t}, \psi_{i,t}, \nu_{i,t}, \theta_i, \varepsilon_{i,t}) \\ &= \alpha^c q^c_{i,t}[1 + \psi^c_{i,t} + \ln \theta_i - \ln s_{-i,t} - \ln q^c_{i,t} + \varepsilon^c_{i,t}] \\ &\quad +\alpha^b \sum_{j=1}^{J} q^{bj}_{i,t} \left[1 + \psi^b_{i,t} + \ln \theta_i - \ln s_{-i,t} - \ln q^{bj}_{i,t} + \varepsilon^{bj}_{i,t}\right] \\ &\quad +\alpha^w q^w_{i,t}[1 + \ln \theta_i - \ln s_{-i,t} - \ln q^w_{i,t} + \varepsilon^w_{i,t}] \\ &\quad +h_i \nu_{i,t} \end{aligned} \tag{1}$$

where $q_{i,t} \equiv \left(q^c_{i,t}, q^{b1}_{i,t}, ..., q^{bJ}_{i,t}, q^w\right)$, $s_{-i,t} = s_{-i}\left(q^c_{-i,t}\right)$ is a mapping of number of car trips $q^c_{-i,t}$ made by travelers other than i, $\psi_{i,t} \equiv \left(\psi^c_{i,t}, \left(\psi^b_{i,t}\right)_{J\times 1}\right)$ where $\psi^c_{i,t}$ (respectively $\psi^b_{i,t}$) denotes a measure of observed comfort of traveling by car (respectively by bus) for individual i. The letters α^c, α^b and α^w represents the marginal utility of using the car, the bus (or public transport) and the alternative mode (typically walk or two-wheels).

We let $h_i(.)$ denote the marginal utility of income of individual i, which can be seen as a rate of change between money and transportation preferences: a 1 euro spending gives the individual a change in utility equal to h_i. When h_i is high, the individual is very sensitive to an increase in transportation fares. Intuitively, this is the case of lower income individuals. On the contrary, a low h_i means that the individual is hardly sensitive to an increase in transportation fares. Parameter(s) in function $h_i(.)$ as well as α^c, α^b are to be estimated. We denote by $\varepsilon_{i,t} \equiv \left(\varepsilon^c_{i,t}, \varepsilon^{b1}_{i,t}, ..., \varepsilon^{bJ}_{i,t}, \varepsilon^w_{i,t}\right)$ the unobserved factors associated respectively to the choices of car, bus with each fare type $j = 1, ...J$. These unobservables could be characteristics of the consumer and/or attributes of the mode of transportation (and menu of payment options) such as its reputation.

Note that $U_{i,t}(q_{i,t}, q^*_{-i,t}, \psi_{i,t}, \nu_{i,t}, \theta_i, \varepsilon_{i,t})$ is twice differentiable in its arguments and strictly quasi-concave in $q_p, \nu_{i,t}$ and that the model satisfies the assumption of "weak complementarity" (see Maler, 1974) implying that the characteristics of a mode of transportation do not matter unless the mode is actually used.

The letter s_{-i} denotes the level of inconvenience in traffic met by traveler i. We assume that it is represented by the average number of car trips made by travelers different from i, given by[4]

$$s_{-i,t} = \frac{1}{I-1} \sum_{j \in I-i} \int_{\varepsilon^c_{j,p}} q^c_{j,p}\left(\varepsilon^c_{j,p}\right) dF(\varepsilon^r_{j,p})$$

in case of complete information of the regulator on the travelers' VOT, and by

$$s_{-i,t} = \frac{1}{I-1} \sum_{l \in I-i} \int_{\Theta_l} \int_{\varepsilon^c_{l,p}} q^c_{l,p}(\theta_l, \varepsilon^c_{l,p}) dF(\varepsilon^c_{l,p}) dF_\theta(\theta_l)$$

in case of incomplete information. Note that this average number of trips over all travelers of all types does include zeros, i.e. individuals who do not travel by car.

Each traveler i faces the following budget constraint:

$$a^c_{i,t} + p^c_{i,t} q^c_{i,t} + \sum_{j=1}^{J} \left(a^b_{i,t} + p^{bj}_{i,t} q^{bj}_{i,t}\right) + p_\nu \nu_{i,t} \leq W_i \tag{2}$$

where $p^c_{i,t}$(resp. $p^{bj}_{i,t}$) denotes the payment to make one trip by car (respectively the per-unit price of a bus trip for all fare categories $j = 1, ...J$), $a^c_{i,t}$ (resp. $a^{bj}_{i,t}$) denote the fixed charge or subscription fee (if applicable) for car use (respectively for bus use for any $j = 1, ...J$) and p_ν denotes the unit price of the composite good $\nu_{i,t}$ that we normalize to 1 without loss of generality and W_i is the individual's income.

Suppose for the moment that individual i has decided to travel exclusively by car. Conditional on his decision, his utility is $\overline{U}^c_{i,t} = \overline{U}^c_{i,t}(\overline{q}^c_{i,t}, q^*_{-i,t}, \overline{\psi}^c_{i,t}, \nu_{i,t}, \theta_i, \varepsilon_{i,t})$ where $\overline{q}^c_{i,t} \equiv \left(q^c_{i,t}, 0_{J\times 1}\right)$ and by virtue of the assumption of weak complementarity $\overline{\psi}^c_{i,t} \equiv \left(\psi^c_{i,t}, 0_{J\times 1}\right)$. His conditional direct utility function is

$$\begin{aligned} \overline{U}^c_{i,t} &= \overline{U}^c_{i,t}(\overline{q}^c_{i,t}, s_{-i,t}, \overline{\psi}^c_{i,t}, \nu_{i,t}, \theta_i, \varepsilon_{i,t}) \\ &= \alpha^c q^c_{i,t}[1 + \psi^c_{i,t} + \ln\theta_i - \ln s_{-i,t} - \ln q^c_{i,t} + \varepsilon^c_{i,t}] + h_i \nu_{i,t} \end{aligned} \tag{3}$$

and the traveler maximizes $\overline{U}^c_{i,t}$ subject to the constraint

$$a^c_{i,t} + p^c_{i,t} q^c_{i,t} + p_\nu \nu_{i,t} \leq W_i \tag{4}$$

and the non-negativity conditions $q^c_{i,t} \geq 0$ and $\nu_{i,t} \geq 0$.

[4] Congestion initiated by buses is assumed negligible.

In the same way, if the traveler chooses to travel by bus only with fare type j, his conditional utility function is, $\forall j$,

$$\begin{aligned}\overline{U}_{i,t}^{bj} &= \overline{U}_{i,t}^{bj}(\overline{q}_{i,t}^{bj}, s_{-i,t}, \overline{\psi}_{i,t}^{b}, \nu_{i,t}, \theta_i, \varepsilon_{i,t}) \\ &= \alpha^b q_{i,t}^{bj}\left[1+\psi_{i,t}^b+\ln\theta_i-\ln s_{-i,t}-\ln q_{i,t}^{bj}+\varepsilon_{i,t}^{bj}\right]+h_i\nu_{i,t} \end{aligned} \quad (5)$$

where $\overline{q}_{i,t}^{bj} \equiv \left(q_{i,t}^{bj}, 0_{J\times 1}\right)$, $\overline{\psi}_{i,t}^{b} \equiv \left(\psi_{i,t}^{b}, 0_{J\times 1}\right)$ and the traveler maximizes $\overline{U}_{i,t}^{bj}$ subject to the constraint

$$a_{i,t}^b + p_{i,t}^{bj} q_{i,t}^{bj} + p_\nu \nu_{i,t} \leq W_i \quad (6)$$

and the non-negativity conditions $q_{i,t}^{bj} \geq 0$ and $\nu_{i,t} \geq 0$.

Finally, if the traveler chooses to walk, his conditional utility function is

$$\begin{aligned}\overline{U}_{i,t}^{w} &= \overline{U}_{i,t}^{w}(\widetilde{q}_{i,t}^{w}, s_{-i,t}, \nu_{i,t}, \theta_i, \varepsilon_{i,t}) \\ &= (1-\alpha^c-\alpha^b) q_{i,t}^{w}\left[1+\ln\theta_i-\ln s_{-i,t}-\ln q_{i,t}^{w}+\varepsilon_{i,t}^{w}\right]+h_i\nu_{i,t} \end{aligned} \quad (7)$$

and the traveler maximizes $\overline{U}_{i,t}^{w}$ subject to the constraint

$$p_\nu \nu_{i,t} \leq W_i \quad (8)$$

and the non-negativity conditions $q_{i,t}^{w} \geq 0$ and $\nu_{i,t} \geq 0$.

The resulting conditional ordinary demand functions will be denoted $\overline{q}_{i,t}^{c}$ and $\overline{q}_{i,t}^{bj}$. One can derive these functions as follows:

Proposition 1 *Assume that $\varepsilon_{i,t}^c$ follows an Extreme Value distribution with parameters (1,μ_1). Then, conditional on his/her choice $m = c, b$ and fare type j (when $m = b$), traveler i's program of maximization of (3) subject to (4) or (5) subject to (6) gives the following conditional demand functions for any $j = 1, ...J$:*

$$\overline{q}_{i,t}^{c}(\theta_i, \psi_{i,t}^c, s_{-i}^*, \varepsilon_{1i}) = \frac{\theta_i}{s_{-i,t}^*} e^{\psi_{i,t}^c - \frac{h_i p_{i,t}^c}{\alpha^c} + \varepsilon_{i,t}^c}, \quad (9)$$

$$\overline{q}_{i,t}^{bj}(\theta_i, \psi_{i,t}^b, s_{-i}^*, \varepsilon_{2i}^j) = \frac{\theta_i}{s_{-i,t}^*} e^{\psi_{i,t}^b - \frac{h_i p_{i,t}^{bj}}{\alpha^b} + \varepsilon_{i,t}^{bj}}, \quad (10)$$

$$\overline{q}_{i,t}^{w}(\theta_i, \psi_{i,t}^b, s_{-i}^*, \varepsilon_{2i}^j) = \frac{\theta_i}{s_{-i,t}^*} e^{\varepsilon_{i,t}^{w}} \quad (11)$$

with

$$s_{-i,t}^* := \left(\frac{\Gamma\left(2-\frac{1}{\mu_1}\right)}{I}\sum_{j\in I}\int_{\theta_j\in\Theta}\theta_j e^{\psi_j^c - \frac{h_j p_{j,p}^c}{\alpha^c}} dF(\theta_l)\right)^{1/2} \quad (12)$$

in the case of incomplete information and

$$\left(s^*_{-i,t}\right)^2 := \frac{\Gamma\left(2-\frac{1}{\mu_1}\right)}{I}\sum_{j\in I-i}\theta_j e^{\psi^c_j - \frac{h_j p^c_{j,p}}{\alpha^c}} \tag{13}$$

in the case of complete information.

Proof. See Appendix 1. ■

The expression of the conditional indirect utility functions $\overline{V}^c_{i,t}, \overline{V}^{bj}_{i,t}$ $\forall j = 1,...,J$ and $\overline{V}^w_{i,t}$ are

$$\begin{aligned}\overline{V}^c_{i,t} &= \overline{V}^j_{i,t}\left(a^c_{i,t}, p^c_{i,t}, w_i, \psi^c_{i,t}, \theta_i, s^*_{-i,t}, \varepsilon^c_{i,t}\right)\\ &= \alpha^c\frac{\theta_i}{s^*_{-i,t}}e^{\psi^c_{i,t}-\frac{h(\theta_i)p^c_i}{\alpha^c}+\varepsilon^c_{i,t}} + h_i\left[w_i - a^c_{i,t}\right],\end{aligned} \tag{14}$$

for $j = 1, ...J$,

$$\begin{aligned}\overline{V}^{bj}_{i,t} &= \overline{V}^{bj}_{i,t}\left(a^{bj}_{i,t}, p^{bj}_{i,t}, w_i, \psi^{bj}_{i,t}, \theta_i, s^*_{-i,t}, \varepsilon^{bj}_{i,t}\right)\\ &= \alpha^b\frac{\theta_i}{s^*_{-i,t}}e^{\psi^b_{i,t}-\frac{h(\theta_i)p^{bj}_{i,t}}{\alpha^b}+\varepsilon^{bj}_{i,t}} + h_i(w_i - a^{bj}_{i,t}),\end{aligned} \tag{15}$$

and

$$\begin{aligned}\overline{V}^w_{i,t} &= \overline{V}^w_{i,t}\left(w_i, \theta_i, s^*_{-i,t}, \varepsilon^c_{i,t}\right)\\ &= (1-\alpha^c-\alpha^b)\frac{\theta_i}{s^*_{-i,t}}e^{\varepsilon^w_{i,t}} + h_i w_i.\end{aligned} \tag{16}$$

Note that the additive form of the direct utility function is such that maximization of the unconditional problem characterized by (1) and (2) would also lead to following first order conditions

$$1 + \psi^c_{i,t} + \ln\theta_i - \ln s^*_{-i,t} - \ln q^{c*}_{i,t} + \varepsilon^c_{i,t} = 1 + \frac{h_i p^c_{i,t}}{\alpha^c}, \tag{17}$$

$$1 + \psi^b_{i,t} + \ln\theta_i - \ln s^*_{-i,t} - \ln q^{bj*}_{i,t} + \varepsilon^{bj}_{i,t} = 1 + \frac{h_i p^{bj}_{i,t}}{\alpha^b} \text{ for } j = 1, ..., J. \tag{18}$$

and conditional demand functions (9) and (10).

3 Empirical Study

3.1 Data

We use a household survey in the greater Montpellier area (south of France; 229,055 inhabitants) recording the transportation activity of 6341 individuals during a two day period. Hence, we observe each individual's number and characteristics of trips as well as his (and his household's) socioeconomic characteristics. A trip is defined as a more-than-300-meters drive or run between two places on a public road. We focus on trips made for the purposes of work, school, shopping and leisure; return trips home are not accounted for.[5] General statistics on the number of car and bus trips are reported in Tables 4-1 and 4-2 of Appendix 2.

More specifically, each individual possesses $J+2$ mutually exclusive and exhaustive alternative choices for traveling: (1) the individual can travel neither by car nor by bus (the trip could then be by bike or by walk or there could be no trip at all), (2) the individual can travel by car only (with possible combined walk and/or bike trips), (3) the individual can travel by bus only (with possible combined walk or bike trips). Furthermore we assume that the mode of payment chosen is the one that is most used during the period, entailing each traveler chooses only one type of payment. In the case where the individual travels by bus, he can choose among J types of payment: more specifically, these include a unit ticket (valid for one trip) of FF7 (1.07 Euros), a booklet of 3 tickets of FF20 (3.05 Euros) or a booklet of 10 tickets (discounted for handicapped or large families), a 7 day or 30 day lump sum (discounted for students, unemployed, scholars depending on district subventions, retired with and without "Carte Or"(retirement card) subscription) or a yearly pass (discounted for scholars, students and unemployed non student-scholars). Hence, these payment options do not depend on the time of travel but they depend on the observed characteristics of the individual.

3.2 Model Estimation and Identification

3.2.1 Estimation Procedure

In this paragraph, we define by $\mathcal{R}$ the set of all possible choices of modes and fare categories $\mathcal{R} \equiv \{c, (b,1), (b,2), ..., (b,J), w\}$. We denote by r, the typical element, $r = 1, ...J+2$. Using these notations, (1) can be written:

$$U_{i,t} = \sum_{r=1}^{J+2} \alpha^r q_{i,t}^r [1 + \psi_{i,t}^r + \ln \theta_i - \ln s_{-i,t} - \ln q_{i,t}^r + \varepsilon_{i,t}^r] + h(\theta_i) \nu_{i,t}$$

Let us denote the unconditional demand functions for traveler i during period t, $q_{i,t}^r(\theta_i, \psi_{i,t}, s_{-i,t}^*, \varepsilon_{i,t}) \, \forall r$ and by $V_{i,t}(\theta_i) \equiv V_{i,t}\left(a_{i,t}, p_{i,t}, w_i, \psi_{i,t}, \theta_i, s_{-i,t}^*, \varepsilon_{i,t}\right)$, the associated unconditional indirect utility function. Let $a_{i,t} \equiv \left(a_{i,t}^c, a_{i,t}^{b1}, ..., a_{i,t}^{bJ}\right)$

[5] Please refer to Viauroux (2007) for a more detailed description of the data set.

and $p_{i,t} \equiv \left(p^c_{i,t}, p^{b1}_{i,t}, ..., p^{bJ}_{i,t}\right)$ the possible fixed fees and per-trip unit price, of respective typical element $a^r_{i,t}$ and $p^r_{i,t}$. The discrete choice of which mode of transportation to select can be represented by a set of binary valued indices $\delta^r_{i,t} = \delta^r_{i,t}\left(a_{i,t}, p_{i,t}, w_i, \psi_{i,t}, \theta_i, s^*_{-i,t}, \varepsilon_{i,t}\right)$, where

$$\begin{aligned} \delta^r_{i,t} &= 1 \text{ if } \overline{V}^r_{i,t}(\theta_i) > \overline{V}^{r'}_{i,t}(\theta_i) \ \ \forall r' \neq r \\ &= 0 \text{ otherwise} \end{aligned} \tag{19}$$

The relationship between these unconditional functions and the corresponding conditional ones above is that:

$$q^r_{i,t}(\theta_i, \psi_{i,t}, s^*_{-i,t}, \varepsilon_{i,t}) = \delta^r_{i,t}\overline{q}^r_{i,t}(\theta_i, \psi^r_{i,t}, s^*_{-i,t}, \varepsilon_{i,t}) \tag{20}$$

and

$$V_{i,t} = \max_{\overline{V}^r_{i,t}}(\overline{V}^1_{i,t}, \overline{V}^2_{i,t}, ..., \overline{V}^{J+1}_{i,t}). \tag{21}$$

Using the approach of Hanemann (1984), we can write the likelihood function of the model as follows. Let r^* denote the index for the mode of transportation and type of payment selected by the i^{th} individual, let $q^{r*}_{i,t}$ be his observed number of trips. Then,

$$\begin{aligned} L &= \prod_{i=1}^{N} \int_{\underline{\theta}}^{\overline{\theta}} f_{q^{r*}_{i,t}}\left(q^{r*}_{i,t}(\theta_i)\right) dF_i(\theta_i), \\ &= \prod_{i=1}^{N} \int_{\underline{\theta}}^{\overline{\theta}} f_{q^{r*}_{i,t}|\varepsilon \in A^r_{i,t}}\left(q^{r*}_{i,t}(\theta_i)\right) P^r_{i,t}(\theta_i)\, dF_i(\theta_i). \end{aligned}$$

where the expressions of $f_{q^{r*}_{i,t}|\varepsilon \in A^r_{i,t}}$ and $P^r_{i,t}(\theta_i)$ are derived in Appendix 2.

3.2.2 Identification

Identification of the model above raises three main concerns that need to be addressed. The first is the handling of the continuous/discrete choice model structure, the second is the presence of a peer effect affecting the decision of travelers and the third is the presence of incomplete information about the traveler's VOT.

Recall that our model can be summarized by (58), (59) and (60). Individual valuations θ_i are i.i.d random draws from the $Beta(1, \mu)$ distribution, where μ is a parameter to estimated, while the random components $\varepsilon^r_{i,t}$ are assumed independent and identically Gumbel distributed with parameters $(0, \mu_1)$. Hence, $\alpha^r \varepsilon^r_{i,t} \sim G\left(0, \frac{\mu_1}{\alpha^r}\right)$ where μ_1 and α^c, α^{bj} are to be estimated. We specify vectors

$\psi_{i,t}^{b}$ and $\psi_{i,t}^{c}$ by $\psi_{i,t}^{b} \equiv \beta^{b} X_{i,t}^{b}$, and $\psi_{i,t}^{c} \equiv \beta^{c} X_{i,t}^{c}$ where $X_{i,t}^{b}$ and $X_{i,t}^{c}$ are vectors characterizing the trip based upon its origin and its destination as well as socioeconomic characteristics of individual i (see Tables 1 and 2). We assume that the MU of income can take three possible specification forms. The MU can be constant (Model 1),

$$h_i = h_0;$$

the MU of income can vary with the monthly household income (Model 2):

$$\begin{aligned} h_i &= h_0^I \; if \;\; Inc_i \leq \widetilde{Inc}_1 \\ &= h_1^I if \; \widetilde{Inc}_1 \leq Inc_i < \widetilde{Inc}_2 \\ &= h_2^I \; if \;\; Inc_i > \widetilde{Inc}_2 \end{aligned}$$

where $\widetilde{Inc}_1$(1500FF or $\simeq$229 euros) and $\widetilde{Inc}_2$ (2883FF or $\simeq$440 euros) are dividing the monthly income data in three equal shares and h_0^I, h_2^I and h_2^I are parameters to be estimated (Petrin, 2003).

Finally, the MU of income is a function of the individual's VOT (Model 3):

$$h_i = h_0^{\theta} \left(\theta_i\right)^{0.5}$$

where h_0^{θ} is a parameter to be estimated. Note that Model 3 is a special case of the more general functional form $h_i = k_3 \theta_i^{\xi}$, which verifies the assumptions of both propositions below for $0 < \xi < 1$. This specification of the MU of income as a function of the unobserved VOT is based on the assumption that the more tolerant to traffic congestion, the more an individual uses transportation despite traffic conditions and the higher the marginal utility of transportation usage (as we expect h_0^{θ} to be estimated to be positive), however this effect wears out above a threshold of tolerence. Hence, we assume that high income travelers have, on average, a higher VOT than low income travelers. This is consistent with the literature stating that the MU of income decreases with income (see Frisch, 1964 and Clark, 1973 for the construction of empirical measures of MU of income in transportation).

First, let us analyze the complete information case (θ_i known $\forall i$) and let us assume s_{-i} is exogenous.

The most general approach of identification of discrete continuous models has been undertaken by Newey (2007) who shows that identification of both demand funtions and nonparametric indirect utility functions of the Dubin and McFadden (1984) model rely solely on the support of the choice probabilites conditional on the choice index and on the independence between the choice index, the random error term and the exogenous prices and income measures. This is possible because the structure of the model reduces the curse of dimensionality. His approach is borrowed from the identification strategies of sample selection models (Heckman, 1979) and consists in identifying the choice probabilities as a first step and the conditional means given each choice as a second step, where the choice probabilities enter as regressors (see also Ahn and Powell, 1993 and Gayle and Viauroux, 2007 for further adaptations of this approach when developing semiparametric estimators).

When $\varepsilon_{i,t}^{r}$ is close to its mean, our framework enters Newey's (2007) linear index T form where

$$T \approx \alpha^{r} + \alpha^{r}\psi_{i,t}^{r} - h_{i}p_{i,t}^{r} + \alpha^{r}\varepsilon_{i,t}^{r} + h_{i}\left[w_{i} - a_{i,t}^{r}\right]\frac{s_{-i}^{*}}{\theta_{i}}$$

Provided that $\alpha^{r}\varepsilon_{i,t}^{r} \sim G\left(0, \frac{\mu_{1}}{\alpha^{r}}\right)$, $\alpha^{r}\varepsilon_{i,t}^{r} - \alpha^{r'}\varepsilon_{i,t}^{r'}$ is strictly monotonic and logistically distributed. Noting that the derivative of T with respect to $p_{i,t}^{r}$ and $a_{i,t}^{r}$ is monotically decreasing provided that $h_{i} > 0$, the difference $\lambda_{i,t}^{r}(\theta_{i}) - \lambda_{i,t}^{r'}(\theta_{i})$ or the non-random component of the indirect utility function, as well as their derivatives with respect to $p_{i,t}^{r}$ and W_{i} are identified. From this stage, Newey (2007) shows that demand functions as well as indirect utility functions are identified using the quantile structural function in the selected sample (see Imbens and Newey, 2007).

The identification result above, however, relies on the assumption that all variables introduced in the model are exogenous. One possible concern in our model is the presence of the externality term s_{-i}, which is a function of other variables in the model. Brock and Durlauf (2001) derive conditions for global identification of parameters in a binary random utility choice framework where social interactions are accounted for. They show that the percentage change in individual utility (assumed constant) from a change in the mean decision level s_{-i} is identified while the same percentage in the linear-in-means framework of Manski (1993) is not. This is because the relationship between the expected average neighborhood choice in a binary choice model and the regressors which characterize the causal determinants of individual behavior is sufficiently nonlinear to justify variation in the neighborhood characteristics. Our model differs from theirs because of its discrete/continuous nature. In our case, the percentage change in individual utility from a change in the mean decision level s_{-i} would depend on the frequency of use of the transportation mode. We do not estimate this percentage change as s_{-i} does not vary much across individuals due to the large number of travelers in the city. Note however that s_{-i} enters nonlineary both in the utility and demand functions.

In general, there is no global identification proof of discrete/continuous models with both peer effects and incomplete information. Existence however, has been established numerically in computational analysis (see Epple and Sieg, 1999, Calabrese et al., 2001, or Epple et al., 2006). Epple and Sieg (1999) estimate the general equilibrium distribution of households across communities, using a framework that accounts for households' private information on their valuation for the public good. Epple et al. (2006) undertake a general equilibrium approach to study the market for higher education. Each student chooses among a subset of colleges which differ in their quality according to a mean ability and a mean income of the student body while the students' preferences depend on their own ability, income and college quality. For each student, ability and income are observed. Calabrese et al. (2001) study voting decisions in local communities when neighborhood quality depends on a peer effect. Households choose among communities that differ in their amount of public good quality and their prices (including taxes). Public good quality is endogenous and depends on an externality term: the mean income in the community. Household preferences for neighborhoods depend on their own valuation of public good quality as well as the mean income in the neighborhood. Conditional on neighborhood choice, each household chooses an amount of housing and of composite private good that maximize their utility. Thus, our approach is closest to Calabrese et al. where preferences include both an incomplete information parameter and a peer effect. Our identification strategy is similar to theirs as follows.

Let us assume that the $J+2$ choices can be ordered according to their quality and respective prices. Noting that the indirect utility functions satisfy the standard single crossing property provided that $\alpha^r > 0$ i.e. $\frac{\partial \overline{V}^r_{i,t}}{\partial q^r_{i,t} \partial \theta_i} \geq 0$, indifference curves in the (q,p) plane have slopes increasing in θ_i. Then, we can identify $J+1$ points of the distribution of θ_i and identify the mean of the distribution of the incomplete information parameter $\frac{1}{1+\mu}$.

For each individual, we observe the number of trips made during period t, the variables measuring the comfort of each mode of transportation $X^r_{i,t}$, unit prices and subcription fees (Recall that these prices and fees vary on an individual basis). We also observe the household income. Using the known information on the distribution of θ_i above, the variation in the conditional demand functions due to the variation in the observed individual variables above allows us to identify the vector of parameters β^b, β^c,as well as parameters$\frac{h_0}{\alpha^c}$ or $\frac{h_0}{\alpha^b}, \frac{h^I_0}{\alpha^c}, \frac{h^I_1}{\alpha^c}, \frac{h^I_2}{\alpha^c}, \frac{h^I_0}{\alpha^b}, \frac{h^I_1}{\alpha^b}, \frac{h^I_2}{\alpha^b}$. Hence, at this stage, we are only able to identify the difference in marginal utility $\alpha^c - \alpha^b$. Moreover observation of each individual choice of mode of transportation and type of payment allows us to identify $\frac{\mu_1}{\alpha^c}$ and$\frac{\mu_1}{\alpha^b}$ through the nonlinearity of the logit probability, hence identifying α^c and $\alpha^b, h_0, h^I_0, h^I_1$ and h^I_2 using the previous stage results.

4 The mechanism

4.1 Introduction

In this section, we design a pricing mechanism that accounts for the endogeneity of congestion, the reputation of the modes and the incomplete information the regulator has about each traveler's private VOT. Travelers' trips begin at one node as shown in Figure 2 below and progress to a destination by traversing the path between the intervening nodes. Their goal is to find the route from the Origin node (O) to the Destination node (D), with lowest congestion and distance between the intervening nodes. Hence, the least cost path is not necessarily the shortest but rather the summation of the contiguous paths of lowest congestion.

Figure 2: Concentric (city) network

[Insert Figure 2 here]

The problem is that travelers choose a path that is consistent with their best interest without regard for the interest of others traveling the network. The regulator ("principal") works closely with the public transportation provider to implement a mechanism that produces a Pareto optimal distribution of traffic in equilibrium. In the following, we focus exclusively on policies aimed at reducing congestion on the city network by promoting the use of public transportation. Travelers must pay a price (congestion fee) which depend on the amount of congestion anticipated on the network, hence the period (peak/off-peak) they travel; the fee is calculated based on the Vickrey-Clarke-Groves Mechanism (see Vickrey, 1961; Clarke, 1971; Groves, 1973).

Implementing a pricing mechanism entails the regulator to allow each traveler to select a fare (matched to a number of trips) for each mode of transportation regulated. For car trips, they select $(q^c_{i,t}, a^c_{i,t})$, while for bus trips they select $(q^{bj}_{i,t}, a^{bj}_{i,t})$ $\forall j$. The quantity of trips is defined as the number of trips from one origin to one destination within a two day period. The regulator will then set the associated fares $(p^c_{i,t}, a^c_{i,t})$ and $(p^{bj}_{i,t}, a^{bj}_{i,t})$ as a function of the individual's valuation of traveling. The regulatory policy must satisfy the constraint that the traveler should have an incentive to truthfully report his type. Note that the feasibility of this instrument is not obvious but one step in that direction might be for the regulator to require the purchase of an RFID (Radio Frequency IDentification) card as a prerequisite to using toll roads and to link the traveler's identification to his/her observed characteristics. The traveler could then purchase a specific number of trips and choose a maximum amount of spending for the period. For each possible set of observable characteristics, the regulator has some information for θ_i prior to any valuation report from the traveler, which is common knowledge to the firm and all other travelers than i. Recall that the regulator's belief about θ_i is reflected in the density $f_i(\theta_i)$ and a cumulative distribution function $F_i(\theta_i)$. The regulator also has a priori information on the distribution associated to the traveler's perception of each mode of transportation. He knows that $\varepsilon^r_{i,t}$ are identically distributed according to a distribution function F_ε. In the section below, we assume that the marginal utility of income varies with travelers' VOT.

We describe a regulatory policy by the functions $(p_{i,t}, q_{i,t}, a_{i,t})$ which can be interpreted as follows. For any $\widetilde{\theta}_i \in [\underline{\theta_i}, \overline{\theta_i}]$, the regulator proposes a subscription fee of $a_{i,t}$ and a unit price per trip $p_{i,t}\left(\widetilde{\theta}_i\right)$; $q_{i,t}$ is the corresponding vector of quantity of trips satisfying $p_{i,t} = P\left(q_{i,t}\left(\widetilde{\theta}_i, \varepsilon_{i,t}\right)\right)$ where $P\left(q_{i,t}\left(\widetilde{\theta}_i, \varepsilon_{i,t}\right)\right)$ are the individual inverse demand curves resulting from the Bayesian game among travelers defined above. We further assume that each individual chooses one mode of transportation only during the two-day survey period.

Inverse demand functions for transportation usage are obtained from Section 2 above:

$$p_{i,t}^r\left(q_{i,t}^r,\varepsilon_{i,t}^r\right)=\frac{\alpha^r}{h\left(\theta_i\right)}\left[\psi_{i,t}^r+\ln\theta_i-\ln s_{-i,t}^*-\ln q_{i,t}^r+\varepsilon_{i,t}^r\right],$$

We assume that the functions $p_{i,t}^r(q_{i,t}^r,s_{-i,t}^*,\theta_i,\varepsilon_{i,t}^r)\ \forall r$ are common knowledge.

The *net* surplus for trips of each traveler, having chosen the mode of transportation and the type of payment r is,

$$\begin{aligned}S_{i,t}&=S_{i,t}^r(q_{i,t}^r,s_{-i,t}^*,a_{i,t}^r,\theta_i,\varepsilon_{i,t}^r)\\&=\int_0^{q_{i,t}^r}P\left(\widetilde{q}\right)d\widetilde{q}-p_{i,t}^c q_{i,t}^c-a_{i,t}^c\\&=\frac{\alpha^r}{h\left(\theta_i\right)}q_{i,t}^r[1+\psi_{i,t}^r+\ln\theta_i-\ln s_{-i,t}^*-\ln q_{i,t}^r+\varepsilon_{i,t}^r]-p_{i,t}^r q_{i,t}^r-a_{i,t}^r\end{aligned}\tag{22}$$

Note that in this case, it is implicitly assumed that $S_{i,t}^{r'}=0\ \forall r'\neq r$.

Using the abbreviations $a=(a_{i,t})_{i\in I,p\in P}$, $\theta=(\theta_i)_{i\in I}$, $q=(q_{i,t}^1,...,q_{i,t}^r,...q_{i,t}^{J+1})_{i\in I,t\in T,}$, the expected *total surplus* in period t is defined $\forall r$ as

$$S(q,a,\theta):=\sum_{i=1}^{I}\sum_{r=1}^{J+1}\int_{\varepsilon_{i,t}^r}S_i^r(q_{i,t}^r,s_{-i,t}^*,a_{i,t}^r,\theta_i,\varepsilon_{i,t}^r)dF(\varepsilon_{i,t}^r)\tag{23}$$

in the case of complete information and by

$$S(q,a):=\sum_{i=1}^{I}\sum_{r=1}^{J+1}\int_{\theta_i\in\Theta_i}\int_{\varepsilon_{i,t}^r}S_i^r(q_{i,t}^r,s_{-i},a_{i,t}^r,\theta_i,\varepsilon_{i,t}^r)dF_i(\theta_i)dF(\varepsilon_{i,t}^r)\tag{24}$$

in the case of incomplete information. Note that for a given value of θ, the expected total surplus is:

$$S(q,a,\theta):=\sum_{i=1}^{I}\sum_{r=1}^{J+1}\int_{\varepsilon_{i,t}^r}S_i^r(q_{i,t}^r,s_{-i},a_{i,t}^r,\theta_i,\varepsilon_{i,t}^r)dF(\varepsilon_{i,t}^r)$$

Furthermore, the *profit* of the transport authority in period t is given by the formula

$$\begin{aligned}\pi&=\pi(q,a)\\&=\sum_{i=1}^{I}\sum_{r=1}^{J+1}\int_{\theta_i\in\Theta_i}\int_{\varepsilon_{i,t}^r}\left[a_{i,t}^r\left(\theta\right)+\left(p_{i,t}^r-c_p^r\right)q_{i,t}^r\left(\theta,\varepsilon_{i,t}^r\right)\right]dF_i(\theta_i)dF(\varepsilon_{i,t}^r),\end{aligned}$$

where c_p^r denotes the marginal cost associated to choice r during period t. Finally, the *total welfare* in period t is given by the formula

$$W(q,\theta):=\pi(q,a)+S(q,a,\theta).\tag{25}$$

Observe that the total welfare does not depend on a. Given a *welfare weight* $\alpha_1 \in (0,1)$ placed on travelers' surplus, in the case of complete information, the social planner's problem is to maximize the *social value function*

$$U(q,a,\theta) := \alpha_1 S(q,a,\theta) + (1-\alpha_1)\pi(q,a) \tag{26}$$

with respect to

$$q^r_{i,t} = q^r_{i,t}(\theta, \varepsilon^r_{i,t}) \text{ and } \quad a^r_{i,t} = a^r_{i,t}(\theta) \tag{27}$$

under the *participation constraints*

$$S^r_{i,t}(q^r_{i,t}, s^*_{-i,t}, a^r_{i,t}, \theta_i, \varepsilon^r_{i,t}) \geq 0 \tag{28}$$

for every i choosing r. In the case of incomplete information the social planner's problem is to maximize the function

$$\int_\Theta U(q(\theta), a(\theta), \theta) f(\theta)\ d\theta \tag{29}$$

under the participation constraints (28) and the *incentive compatibility constraints*,

$$S^r_{i,t}(q^r_{i,t}(\theta, \varepsilon^r_{i,t}), s_{-i}, a^r_{i,t}(\theta), \theta_i) \geq S^r_{i,t}(q^r_{i,t}(\tilde{\theta}, \varepsilon^r_{i,t}), s_{-i}, a_{i,t}(\tilde{\theta}), \theta_i) \quad \forall \theta, \tilde{\theta} \in \Theta \tag{30}$$

for every i choosing r, where $\tilde{\theta}$ differs from θ only in its i component: $\theta_l = \tilde{\theta}_l$ for every $l \neq i$. By the revelation principle, it suffices to show that the truth telling strategy is a dominant strategy in a direct revelation mechanism. Indeed, the revelation principle states that a dominant strategy equilibrium of any Bayesian game can be represented by an equilibrium in a direct revelation mechanism (see Green and Laffont, 1977; Myerson, 1979).

The timing of the regulation process is as follows: First, each traveler i learns his type θ_i. Second, the regulator announces the regulatory policy. Finally, the number of trips is taken by individuals and a new congestion level results from this policy.

4.2 Complete information

As previously mentioned, we have to maximize the function (26) with respect to the variables (27) under the constraints (28), creating the following result.

Proposition 2 *The regulator implements marginal cost prices. The maximum of the social value function is achieved by a unique solution $(q^r_{i,t}, a^r_{i,t})_{i\in I}$, given by the following formulae:*

$$q^r_{i,t} = \theta_i \left(\frac{\Gamma\left(2 - \frac{1}{\mu_1}\right)}{I} \sum_{j\in I} \theta_j e^{\psi^c_j - \frac{h(\theta_j)}{\alpha^c} p^c_{j,p}} \right)^{-1/2} e^{\psi^r_j - \frac{h(\theta_j)}{\alpha^r} c^r_p + \varepsilon^r_{i,t}}. \tag{31}$$

Let $0 < \alpha_1 < 1/2$. *Then*

$$\begin{aligned} a_{i,t} &= a_{i,t}(q^r_{i,t}) \\ &= \frac{\alpha^r}{h(\theta_i)} q^r_{i,t}, &(32) \\ S(q,a,\theta) &= 0, &(33) \\ W(q,\theta) &= \pi(q,a), &(34) \\ &= \sum_{i=1}^{I} \int_{\varepsilon^r_{i,t}} \frac{\alpha^r}{h(\theta_i)} q^r_{i,t}\left(\theta, \varepsilon^r_{i,t}\right) dF(\varepsilon^r_{i,t}), \\ U(q,a,\theta) &= (1-\alpha_1)\pi(q,a). &(35) \end{aligned}$$

Let $\frac{1}{2} < \alpha_1 < 1$. *Then,*

$$\begin{aligned} a_{i,t} &= 0, &(36) \\ W(q,\theta) &= S(q,a,\theta), &(37) \\ &= \sum_{i=1}^{I} \sum_{r=1}^{J+1} \int_{\varepsilon^r_{i,t}} q^r_{i,t} \left(\frac{\alpha^r}{h(\theta_i)} [1 + \psi^r_{i,t} + \ln \theta_i - \ln s^*_{-i,t} - \ln q^r_{i,t} + \varepsilon^r_{i,t}] - p^r_{i,t} \right) dF(\varepsilon^r_{i,t}), \\ \pi(q,a) &= 0, &(38) \\ U(q,a,\theta) &= \alpha_1 W(q,\theta). &(39) \end{aligned}$$

Proof. See Appendix 1. ■

If the regulator had complete information about each individual's VOT, the optimal policy would be to set a unit price per trip equal to the marginal cost and leave travelers a surplus (if $\alpha_1 > \frac{1}{2}$) or charge them a subscription fee (if $\alpha_1 < \frac{1}{2}$) equal to the exact amount that they are willing to spend on transportation.

In case of $\alpha_1 = 1/2$, the social value function does not depend any more on $a_{i,t}$; otherwise, its maximum is attained for the same values of $q^r_{i,t}$ as in Proposition 1.

Of course, this policy is not feasible for the regulator when θ is unknown because it does not satisfy the incentive compatibility constraints. Each traveler would have positive incentives to misrepresent his/her aversion to congestion by reporting a valuation for transportation $\widetilde{\theta} \neq \theta$. Furthermore, this misrepresentation could take two possible directions. Intuitively, the traveler is likely to report $\widetilde{\theta} > \theta$ if he primarily uses public transportation because it does not generate congestion or if he anticipates that the reduction in fare he may get for traveling more overcomes his anticipated charge for creating congestion. In this case, the traveler tries to take advantage of lower fares for making more trips. On the other hand, one may expect $\widetilde{\theta} < \theta$ if the individual travels mostly by car; that is, the individual knows that the portion of the road he uses or the time of the day during which he travels is highly congested and he anticipates being charged a high amount for the inconvenience that his trips may cause on the network.

4.3 Incomplete information

In this section, we analyze the case where the regulator has incomplete information about the traveler's VOT. The expectation of a regulatory policy on urban congestion is the following: by raising the price above the marginal cost, the mechanism reduces the number of trips made by high valuation travelers. This in turn will decrease the level of congestion on the network, and will increase the incentive to make a trip.

We recall from Section 2 that the regulator maximizes the function (29) with respect to the functions (27) under the participation constraints (28) and the incentive compatibility constraints (30).

We will use the envelope theorem to maximize the social value function. Hence, a traveler' surplus will be maximized when he/she reveals his/her true type.[6]

Recall that the types $\theta_i \in \Theta$ are independently distributed according to the cumulative distribution function F_i of density f_i. We set $\Theta = \prod_{i \in I} \Theta_i$ and $f = \prod_{i \in I} f_i$. Furthermore, we assume that h is continuously differentiable, satisfying

$$1 + \frac{1 - 2\alpha_1}{1 - \alpha_1} \cdot \frac{1 - F_i(\theta_i)}{f_i(\theta_i)} \cdot \frac{h'(\theta_i)}{h(\theta_i)} \neq 0 \tag{40}$$

for all all i and θ_i.

In order to state our result, let us introduce the following notation:

$$\begin{aligned}
s^* &:= \left(\frac{\Gamma\left(2 - \frac{1}{\mu_1}\right)}{I} \sum_{j \in I} \int_{\Theta_j} \theta_j e^{\psi_j^c - \frac{h_j}{\alpha^c} c_{j,p}^c(\theta_j) + \varepsilon_{i,t}^c} f_j(\theta_j)\, d\theta_j \right)^{1/2}, \\
m(\theta_i) &:= \left(1 + \frac{1 - 2\alpha_1}{1 - \alpha_1} \cdot \frac{1 - F_i(\theta_i)}{f_i(\theta_i)} \cdot \frac{h'(\theta_i)}{h(\theta_i)} \right)^{-1}, \\
c_{i1}^r(\theta_i) &:= m(\theta_i) \left(c_p^r + \frac{\alpha^r}{h(\theta_i)} \cdot \frac{1 - 2\alpha_1}{1 - \alpha_1} \cdot \frac{1 - F_i(\theta_i)}{\theta_i f_i(\theta_i)} \right), \\
V_i &:= V_i(q_{i,t}(\theta, \varepsilon_{i,t}^r), \theta_i) \\
&= \frac{1}{\theta_i h(\theta_i)} \alpha^r q_{i,t}^r(\theta, \varepsilon_{i,t}^r) \\
&\qquad - \alpha^r \frac{h'(\theta_i)}{h^2(\theta_i)} q_{i,t}^r(\theta, \varepsilon_{i,t})[1 + \psi_{i,t}^r + \ln \theta_i - \ln s_{-i,t}^* - \ln q_{i,t}^r(\theta, \varepsilon_{i,t}^r) + \varepsilon_{i,t}^r].
\end{aligned} \tag{41}$$

We assume that s^*, as defined above is finite.

The solution of the social planner's problem is different again for $\alpha_1 < 1/2$ and for $\alpha_1 > 1/2$. Let us begin with the first case:

---

[6]From the envelope theorem, the derivative of the social value function with respect to θ takes into account only the direct effect of θ, and not the indirect effect stemming from the adjustment in quantity.

Proposition 3 *Assume that*

$$\alpha^r \left(\frac{1}{\theta_i h(\theta_i)} - \frac{h'(\theta_i)}{h^2(\theta_i)} \right) - \frac{h'(\theta_i)}{h(\theta_i)} c^r_{i1}(\theta_i, \varepsilon^r_{i,t}) \geq 0 \tag{42}$$

for all i choosing r and θ_i. If $0 < \alpha_1 < 1/2$, then the regulator implement prices c^r_{i1} and the maximum of the social value function is achieved by a unique vector $(q^r_{i,t}(\theta, \varepsilon^r_{i,t}), a_{i,t}(\theta, \varepsilon^r_{i,t}))_{i \in I, t \in T}$, given by the following formulae for all $i, r \in I \times \mathcal{R}$ and $\theta \in \Theta$:

$$q^r_{i,t}(\theta, \varepsilon^r_{i,t}) = \frac{\theta_i}{s^*_{-i,t}} e^{\psi^r_{i,t} - \frac{h(\theta_i)}{\alpha^r} c^r_{i1}(\theta_i) + \varepsilon^r_{i,t}}, \tag{43}$$

$$a_{i,t}(\theta, \varepsilon^r_{i,t}) = \frac{\alpha^r}{h(\theta_i)} q^r_{i,t}(\theta, \varepsilon^r_{i,t}) - \int_{\underline{\theta_i}}^{\theta_i} V_i(q_{i,t}(\theta_{-i}, \tilde{\theta}_i), \tilde{\theta}_i) \ d\tilde{\theta}_i. \tag{44}$$

Proof. See Appendix 1. ■

To understand why this regulatory policy may be optimal, observe that the regulator wants to encourage the traveler to admit that he has a high traveling valuation whenever this is true, so that pricing accounts for the congestion costs generated. To prevent a traveler from misrepresenting his true valuation, the regulator punishes him when reporting a low valuation. This punishment takes the form of a per-trip price higher than the marginal cost of a trip.

In this case, the regulator assigns slightly more weight to the transportation provider's surplus that to the travelers, i. e. $\frac{1-2\alpha_1}{1-\alpha_1} > 0$.

First, let us note that $P^\theta = \frac{f_i(\theta_i)}{1-F_i(\theta_i)}$ represents the proportion of individuals of VOT θ_i among those whose VOT is lower than individual i's, while $\frac{h'}{h}$ can be seen as a measure of "acceleration" in the sensitivity to transportation costs.

The pricing is a linear function of the MC of using the mode of transportation where the slope increases with P^θ , the intercept decreases with P^θ, while both slope and intercept decrease with $\frac{h'}{h}$.

Intuitively, if the marginal cost of a trip is small, the fare is such that "if most individuals travel more than you (because their VOT is lower than yours), you should be assigned a lower fare as you are not held responsible for much traffic congestion". Moreover, this fare is lower when the traveler's income is low (or sensitivity to cost is high) and when his sensitivity to prices increases with θ_i; i.e. when traffic conditions "accelerate" the sensitivity to transportation costs. Hence, the fee increases with the likelihood of frequent travel that creates congestion and decreases with income. In this case, the regulator's intervention is primarily intended for individuals who are less sensitive to transportation costs and whose probability of generating congestion is large. Also note, that the intercept shift is decreasing with the true valuation of the individual; in other words, the more obligated the traveler is to his schedule, the less punishment imposed.

However, as the marginal cost of a trip increases, the fee is increased for those travelers in proportion P^θ as a contribution for the use of capital/infrastructure and this fee is lower for lower income travelers whose price sensitivity increases with their use of the network. As a consequence, the less information the regulator has about the individual's MU of income, the lower the distortion to the marginal cost.

Finally, the subscription fee is designed to offset the incentive, of a traveler whose VOT is low (θ_i high), to report a lower type and incur lower punishment. It is equal to the traveler's willingness to pay for total transportation usage minus the surplus that he gets by announcing any type that is lower than his own. Hence, if the traveler's aversion annoucement is high (θ_i is low), then he pays the maximum subscription fee. But the higher θ_i, the smaller the subscription fee that a traveler pays. Thus, the subscription fee is inversely related to the traveler's valuation report.

Turning to the case $\alpha_1 > 1/2$, let us add again the extra condition $a_{i,t}(\theta, \varepsilon_{i,t}^r) \geq 0$ for all i and θ.

Proposition 4 *Let $\frac{1}{2} < \alpha_1 < 1$. Then the maximum of the social value function is achieved by a unique triple $(q_{i,t}^r(\theta, \varepsilon_{i,t}^r), a_{i,t}(\theta, \varepsilon_{i,t}^r))_{i \in I}$, given by the following formulae for all $i, r \in I \times \mathcal{R}$ and $\theta \in \Theta$:*

$$q_{i,t}^r(\theta, \varepsilon_{i,t}^r) = \frac{\theta_i}{s^*} e^{\psi_{i,t}^r - \frac{h(\theta_i)}{\alpha^r} c_{i1}^r(\theta_i) + \varepsilon_{i,t}^r}, \tag{45}$$

$$a_{i,t}(\theta, \varepsilon_{i,t}^r) = 0. \tag{46}$$

Proof. See Appendix 1. ■

Finally, $V_i(q_{i,t}(\theta, \varepsilon_{i,t}), \theta_i)$ simplifies to

$$V_i(q_{i,t}(\theta, \varepsilon_{i,t}^r), \theta_i) := \frac{1}{\theta_i h(\theta_i)} \alpha^r q_{i,t}^r(\theta, \varepsilon_{i,t}^r),$$

so that (42) is redundant.

In this case, the regulator assigns slightly more weight to the travelers' surplus than to the transportation provider's, i. e. $\frac{1-2\alpha_1}{1-\alpha_1} < 0$: the idea is that the regulator wants to reward travelers for announcing a high valuation instead of punishing them for announcing a low one.

Intuitively, if the marginal cost of a trip is small, the fare is such that "if most individuals say they travel more than you (because their VOT is reported lower than yours), you should be less "subsidized" as you are not held responsible for much traffic congestion". Moreover, this subsidy is higher when the traveler's income is high (or the sensitivity to prices is low). Finally, since there is no incentive for high valuation travelers to report a lower type, no fixed grant is necessary. Note that the optimal level of congestion, although endogenous, does not directly enter the regulator's pricing policy.

As in the preceding case, when $\alpha_1 = \frac{1}{2}$, conditions (56) simplifies and gives $p_{i,t}^r = c_p^r$, profit $\pi_p = \sum_{i=1}^{I} \sum_{r=1}^{J+1} a_{i,t}^r = 0$ and by simplification of (44) the individual surplus becomes

$$S_i(q_{i,t}(\theta), s_{-i,t}, a_{i,t}, \theta_i) = \frac{1}{h(\theta_i)} \alpha^r q_{i,t}^r(\theta, \varepsilon_{i,t}^r). \tag{47}$$

Observe that the validity Proposition 4 depends on the specification of travelers' marginal utility of income.

Let us consider two special cases of the proposition. In the case of unobservable heterogeneous marginal utility of income, where, for example, $h(\theta_i) = h_0^\theta \theta_i$, then s^* is finite for all α_1 and Proposition 4 applies. Indeed, the integrals can only blow up for $\theta_i \to 0$. In this case we have

$$m(\theta_i) \approx k_1 \theta_i^2 \quad \text{and} \quad c_p^r + \frac{\alpha^r}{h_0^\theta \theta_i} \frac{1-2\alpha_1}{1-\alpha_1} \frac{1-F_i(\theta_i)}{\theta_i f_i(\theta_i)} \approx \frac{k_2}{\theta_i^2},$$

so that c_{i1}^r remains bounded. Therefore, $\theta_i \exp(\psi_{i,t}^r - (\alpha^r)^{-1} h\theta_i c_{i1}^r(\theta_i))$ is bounded and thus $s_{-i,t}^*$ is finite.

On the contrary, in the case of constant marginal utility of income $h(\theta_i) = k_i$ we have $m(\theta_i) = 1$ and the formulae of $c_{i1}^r(\theta_i)$ simplifies to

$$c_{i1}^r(\theta_i) := c_p^r + \frac{\alpha^r}{h_0^\theta} \cdot \frac{1-2\alpha_1}{1-\alpha_1} \cdot \frac{1-F_i(\theta_i)}{\theta_i f_i(\theta_i)}.$$

It follows from these formulae that

$$c_{i1}^r(\theta_i) \begin{cases} \text{tends to } \infty \text{ if } \alpha_1 < \frac{1}{2}, \\ \text{tends to a finite value if } \alpha_1 = \frac{1}{2}, \\ \text{tends to } -\infty \text{ if } \alpha_1 > \frac{1}{2} \end{cases}$$

and then that s^* is finite for $\alpha_1 \le 1/2$ and infinite for $\alpha_1 > 1/2$. Thus, Proposition 4 does not apply in the latter case.

5 Results

Estimation results of the structural parameters are reported in Tables 1 and 2 of Appendix 3. Below, we compare our results to the results of Viauroux (2007).[7]

As anticipated, we find that reputation significantly affects the joint traveling choices of mode of transportation and categories of payment. Interestingly, the estimated distribution associated to the reputation effects varies significantly from peak to off-peak times. Reputation seems to have significantly more impact during off-peak times when travelers' choices are less constrained.

Once they have established the reputation of a mode of transportation, travelers still value their time in traffic (μ significant); in fact they value it more (μ). We observe a decrease in the average tolerance for traffic congestion going from $\frac{1}{1.646} = 0.608$ (Models 1 and 2) or $\frac{1}{1.684} = 0.594$ (Model 3) during peak hours to $\frac{1}{1.642} = 0.609$, $\frac{1}{1.638} = 0.610$ or $\frac{1}{1.572} = 0.636$ (in Models 1, 2, and 3 respectively) during off-peak hours, and the difference increases with the degree of heterogeneity/uncertainty assumed in the specification of the MU of income.

[7] Note that the random models introduced in this paper are structurally different from the non-random model of Viauroux (2007), which explains the difference in log-likelihood results.

This may, in part, explained by the fact that tolerance to traffic congestion is correlated to the reputation of a mode.

Reputation seems to be perceived as a reference measure of anticipated time spent and comfort of traveling. Hence, diverging from these expectations is felt as more discomfort than no expectation at all. For example, if I know that it takes twenty minutes to go from point A to point B and make all the arrangements (in reference to departure and arrival time for example), the occurrence of an accident on the highway delaying all traffic will impact my preferences for the mode to a greater extent than if I made more flexible arrangements because of the unknown amount of congestion expected. Moreover, travelers are more sensitive to a change in their schedule during off-peak hours; during these times, transportation opportunity cost is felt to a greater extent because of the foregone leisure or shopping activities.

It is also important to note that some of this discomfort is due to the travelers 'marginal utility of income: the lost time in traffic is clearly "unpleasant" because of its opportunity cost; hence, when the heterogeneity in the MU of income is accounted for, tolerance to unforeseen events on the network increases.

Unlike their tolerance to unexpected traffic, travelers' sensitivity to a change in transportation costs of a mode (such as an increase in fare or an increase in gas prices) will not change as much when its reputation is established. The non-random model (Viauroux, 2007) estimated the marginal utility of income at 1.94 during peak period and 2.28 during the off-peak period while the MU of income averages to 0.011, 0.012 and 0.019 (respectively in Models 1,2 and 3) during peak hours and to 0.013, 0.013 and 0.026. Model 1 results are extracted from Tables 1 and 2 Column 7; Model 2 values are obtained by computing the average MU of income as follows

$$\widehat{\widehat{h}} = \widehat{h_0^I} P_0 + \widehat{h_1^I} P_1 + \widehat{h_2^I} P_2$$

with $P_0 \equiv P\left(Inc_i \leq \widetilde{Inc_1}\right), P_1 \equiv P\left(\widetilde{Inc_1} \leq Inc_i < \widetilde{Inc_2}\right)$ and $P_2 \equiv P(Inc_i > \widetilde{Inc_2})$ are the observed proportions of individuals whose income are within the range specified[8]; Model 3 values are obtained using the estimated distribution of θ,

$$\widehat{E(h)} = \widehat{h_0^\theta}\left(\frac{1}{1+\widehat{\mu}}\right)^{0.5}.$$

As for observed factors, we find that the introduction of reputation lowers their effect on traveling decisions, although these effects remain significant.

In particular, we still find that during peak hours, the unemployed, the students and the elderly choose to travel more by bus and less by car than their other socioeconomic status counterparts. This result is consistent with their preferred access to lower bus fares, more specifically free rides for the unemployed and various state subsidies for students and retired travelers.

[8]In our model, these proportions are respectively P_0=89.6%, P_1=7.8% and P_2=2.6% during peak hours and P_0=92.6%, P_1=4.9% and P_2=2.4% during off-peak hours.

As expected, the distance and the anticipated time of travel influence both the frequency of travel and the choice of the mode to a lesser extent during off-peak hours. During these hours the time spent in traffic remains a nuisance to travelers (negative signs); it appears to be more of an issue when traveling by car (with a coefficient of -0.002) than when traveling by bus where it becomes insignificant: "reading a book" can diminish the stress associated to it. Furthermore, the significance of this effect decreases with the introduction of unobserved heterogeneity in the MU of income (Model 3): here again, the inconvenience is not as much the discomfort of "being stuck in traffic" as much as the opportunity cost of the time lost that counts.

Finally, we find that the longer the trip undertaken, the more comfortable the use of a car is perceived, avoiding the multiple bus stops of a bus trip, and this result is stronger when the car used has more horsepower. This last result may be explained by an average strong positive correlation between power and comfort. In contrast, time seems to be taken for granted during peak periods once reputation is established (It is insignificant in the results from EER (07)). The longer the duration of the trip, the more individuals travel during peak times with a preference for the use of public transportation. During this time, the stress generated by difficult driving conditions on a longer trip outweighs the inconvenience of crowded public transportation. Furthermore, a short distance bus trip is preferred in general to a long distance car trip. These results are emphasized by Table 3 where the t-tests show that both the duration of a trip and the distance of a trip impact both the choice and usage of a mode of transportation in a very different way during peak and off-peak times.

Tables 4-5 present different simulated pricing strategies (and the associated welfare) as designed in Section 6, i.e. a pricing scheme that internalizes the social costs that travelers impose on each other. For each period, estimated parameters are used to compute the optimal individual functions presented in Propositions 4 and 5. The average surplus is computed as the sum over all individuals of the expected value (according to the distribution of unobservables $\varepsilon_{i,t}^{r}$ and of θ) of the individual surplus as defined by equations (53) and (57) in Proposition 4 and equation (47) in Proposition 5. For confidentiality reasons, we arbitrarily set a value for the transportation operator's marginal cost, namely 1.07 euros (FF7) for the unit cost of a car trip, 0.76 euros (FF5) for a bus trip; slightly lower than the average cost of a trip made with the respective modes of transportation. The current average price of car usage per trip (on average 2 euros) is equal to the price per kilometer multiplied by a distance from the centroid of the Origin area to the centroid of the destination area. The price varies according to the power of the vehicle and the type of fuel used.[9]

Tables 4-1, 4-2 present the simulated frequencies and welfare when the regulator weighting travelers' surplus and transportation providers' profit equally: $\alpha_1 = 1/2$. In particular, we simulate the frequency of car and bus trips if travel-

[9]When the household possessed more than one car, the survey did not indicate which car was used for the trip considered. Consequently, we assumed that the trips made by the household head were with the most powerful car, trips made by the second household member were made with the second most powerful car, and so on.

ers' preferences were affected by the reputation of all available modes of transportation (namely car, bus and walk) assuming the current set of prices. For presentation purposes, the number of trips based on current pricing will be denoted respectively by $\left(q_{i,t}^{c}\right)^{cp}$ and $\left(q_{i,t}^{bj}\right)^{cp}$ while the optimal number of trips are denoted by $\left(q_{i,t}^{c}\right)^{opt}$ and $\left(q_{i,t}^{bj}\right)^{opt}$ respectively.

Results show that during peak hours, the introduction of "reputation" errors lowers significantly the estimated number of car trips compared to the non-random case or compared to the current number of car trips taken; on the other hand, the number of bus trips estimated becomes closer to the actual number of bus trips taken, while the non-random model overestimates them greatly. In other words, if reputation were to influence the frequency of traveling with all modes, then during peak hours the number of car trips would be significantly lower than the one we observe while the number of bus trips would be well estimated. During off-peak hours, reputation effects would lead to an underestimation of car trips and an overestimation of bus trips compared to the non-random case.

These results lead to an important conclusion: while reputation significantly affects the choice of modes of transportation and category of payments, it only significantly affects the use of public transportation during peak hours. This may be due to several reasons:

First, once individuals weighed the costs and benefits associated to a mode, the use of that mode might be subject to other factors. For example, while positive reputation would tend to increase the use of public transportation during off-peak hours, the infrastructure might be insufficient to cover the variety of leisure and shopping destinations during those times. Moreover, direct access to the car in the garage as opposed to walking toward the next bus stop might explain that travelers use their car slightly more despite traffic conditions.

This conclusion also means that the use of public transportation could be increased if conditions (for a given infrastructure) could be improved during peak hours, for example through additional bus lanes or additional comfort and security.

Should the regulator adopt marginal cost pricing, travelers' surplus would increase by 11.7% and 12.1% compared to the non-random case, with a total (peak and off-peak hours combined) in Models 1 and 2 of 113957 Euros and 114407 Euros respectively instead of 102013 Euros in the non-random case (see Viauroux (2007), columns 5 and 6 of Table 6). Overall, accounting for reputation effect increases significantly our measure of travelers' surplus. This is consistent with the previous conclusion that reputation is felt as useful information, which increases travelers' utility. However, under the assumption that the sensitivity to transportation costs varies with traffic condition, overall travelers' surplus decreases.

Tables 5-1 and 5-2 present statistics of simulated heterogenous pricing and welfare results when the regulator and the transportation provider share common objectives, i.e. $\alpha_1 < 1/2$. As an example, we present the results when

$\alpha_1 = 1/3$.[10]

In this case, the regulator uses the information revealed on travelers' value of time to assess more adequately travelers' MU of income and price in a way to extract as much travelers' surplus as possible and transfer it to the benefit of the transportation provider. This result is especially obvious in Model 3. For comparison purposes with the current pricing strategy, we report car and bus simulated unit prices and subscription amounts for three arbitrary chosen types of travelers: travelers whose VOT is very high ($\theta = 0.25$), travelers whose VOT is average ($\theta = \frac{1}{1+\widehat{\mu}}$) and travelers whose VOT is low ($\theta = 0.75$).

The current regulator underestimates the VOT of travelers and as a consequence does not charge them for the social marginal cost they generate: indeed, the current average bus trip fares of 0.94 Euros and 1.01 Euros for peak and off-peak times respectively are of comparable magnitude with Model 3 simulated prices of 0.8 Euros and 0.82 Euros for a type $\theta = 0.75$, however the two-days subscription fee of 0.15 Euros and 0.05 Euros for peak hours and off-peak hours, respectively, would be greatly underestimated. The models predict that the sample average type traveler ($\theta = \frac{1}{1+\widehat{\mu}} \simeq 0.6$) should optimally be charged the unit prices of 3.38 Euros and 4.23 Euros and a two-days subscription fee of 2.67 Euros and 3.26 Euros respectively during peak and off-peak periods.

Interestingly, Table 5-2 reports welfare results associated to homogenous pricing: in this case, the regulator knows the estimated average VOT of travelers and charges them all the associated optimal fare. Results then show significant improvement in total welfare (peak and off-peak times combined) which respectively increase by 42%, 44% and 101% respectively in Models 1,2 and 3, due to great increase in travelers' surplus. In this case, the regulator does not have the ability to discriminate as much as travelers, leaving them a higher surplus.

In conclusion, marginal cost pricing accounting for reputation of modes and observed heterogeneity of travelers' sensitivity to price (Model 2) shows an overall improvement over the pricing scheme designed in Viauroux (2007) by about 12%. In the case where regulator and transportation provider share common objectives, results show that welfare is reduced when too much heterogeneity is accounted for. However, great welfare improvement can be achieved by implementing a homogenous pricing that accurately accounts for travelers VOT.

6 Conclusion

We undertake a disaggregated approach to estimated transportation demand structural parameters, accounting for the fact that traffic congestion is endogenous and can be represented as a Bayesian Nash game between travelers. In this framework we introduced two types of private information: information on the reputation of each mode of transportation and information on the VOT, e.g. aversion to traffic congestion. Using a cross-sectional two-day period data set,

[10] Note that in Models 1 and 2, assuming $\alpha_1 > 1/2$ implies that s^* is infinite (see above comments on Proposition 4). Simulation of Model 3 welfare results in case $\alpha_1 > 1/2$ is available from the author.

we estimate and compare the results using different assumptions on the MU of income. We find that reputation is an important determinant in the decisions of traveling, particularly in the choice of the mode of transportation and that improving the reputation of public transportation could improve the frequency of trips during peak hours. Contrary to the non-random analysis of Viauroux (2007), results show that welfare can be reduced when too much heterogeneity is accounted for in the transportation pricing. However, great welfare improvement can be achieved by implementing a homogenous pricing that accurately accounts for travelers VOT.

Among the directions for future research, the use of panel data could improve the outcome even further. It would allow the individuals' preferences for traveling to depend not only on the current anticipation of congestion but also on all other travelers' experience of congested areas via their optimal behavior from any point in time onward.

7 Appendix 1. Derivation of the optimal tariff

We denote by σ the randomized-strategy profile for the game

$$\Gamma = \left\{ I, Q, \Theta, (p_{i,t})_{i \in I, t \in T}, \left(u_{i,t}^{j}\right)_{i \in I, t \in T} \right\},$$

such that

$$\sigma = \left(\sigma_i \left(q_{i,t} | \theta_i\right)\right)_{i \in I, q_{i,t} \in Q, \theta_i \in \Theta, t \in T},$$

where $\sigma_i \left(q_{i,t} | \theta_i\right)$ represents the conditional probability that traveler i would do $q_{i,t}$ trips if his type were θ_i (see Myerson, 1991). At a Nash equilibrium σ, one may compute for each $i \in I$ and $\theta_i \in \Theta$ the expected number of trips for traveler i by the formula

$$q_{i,t}^{*}\left(\theta_i\right) = \sum_{q_{i,t} \in Q} q_{i,t} \sigma_i (q_{i,t} | \theta_i).$$

Henceforth, the parameter of aversion to traffic is assumed to be independently and identically distributed in the population according to a Beta $(1, \mu)$ distribution for its simplicity of use and the variety of forms it can represent (exponential, uniform). We recall that, by definition of the equilibrium (see also Viauroux, 2007), for any fixed i, r and $\theta \in \Theta$, putting

$$s_{-i}^{*} = s_{-i}\left(q_{-i,t}^{c*}\right)$$

for brevity, the probabilities $\sigma_i(q_{i,t}|\theta_i)$ have to maximize the expression

$$\sum_{q_{i,t} \in Q} \sigma_i(q_{i,t}|\theta_i) U_{i,t}^{j}(q_{i,t}, q_{-i,t}^{*}, \theta),$$

where the functions $U_{i,t}$ are given by the formula

$$\begin{aligned} U_{i,t} &= U_{i,t}(q_{i,t}, q^*_{-i,t}, \psi_{i,t}, \nu_{i,t}, \theta_i, \varepsilon_{i,t}) \\ &= \alpha^c q^c_{i,t}[1 + \psi^c_{i,t} + \ln\theta_i - \ln s_{-i,t} - \ln q^c_{i,t} + \varepsilon^c_{i,t}] \\ &\quad + \alpha^b \sum_{j=1}^{J} q^{bj}_{i,t}\left[1 + \psi^b_{i,t} + \ln\theta_i - \ln s_{-i,t} - \ln q^{bj}_{i,t} + \varepsilon^{bj}_{i,t}\right] \\ &\quad + \alpha^w q^w_{i,t}[1 + \ln\theta_i - \ln s_{-i,t} - \ln q^w_{i,t} + \varepsilon^w_{i,t}] \\ &\quad + h_i \nu_{i,t} \end{aligned}$$

This definition leads to a pure multistrategy equilibrium corresponding to the value of $q_{i,t}$, which maximizes $U_{i,t}(q_{i,t}, q^*_{-i,t}, \theta)$.

Proof. In order to simplify the computation, let us admit that the variable $q_{i,t}$ can be changed continuously, and let us write down the first-order conditions associated with the maximization of (3) subject to (4). Both partial derivatives with respect to $q^c_{i,t}$ and $q^{bj}_{i,t}$ must vanish at the equilibrium point, namely,

$$\alpha^c\left[\psi^c_{i,t} + \ln\theta_i - \ln s^*_{-i,t} - \ln q^{c*}_{i,t} + \varepsilon^c_{i,t}\right] - h_i p^c_{i,t} = 0, \tag{48}$$

$$\alpha^b\left[\psi^b_{i,t} + \ln\theta_i - \ln s^*_{-i,t} - \ln q^{bj*}_{i,t} + \varepsilon^{bj}_{i,t}\right] - h_i p^{bj}_{i,t} = 0 \tag{49}$$

Solving for $q^{c*}_{i,t}$ and $q^{bj*}_{i,t}$, we obtain the first two equalities of Proposition 1:

$$\overline{q}^{c*}_{i,t}(\theta_i, \varepsilon^c_{i,t}) = \frac{\theta_i}{s^*_{-i,t}} e^{\psi^c_{i,t} - \frac{h_i p^c_{i,t}}{\alpha^c} + \varepsilon^c_{i,t}}, \tag{50}$$

$$\overline{q}^{bj*}_{i,t}(\theta_i, \varepsilon^{bj}_{i,t}) = \frac{\theta_i}{s^*_{-i,t}} e^{\psi^b_{i,t} - \frac{h_i p^{bj}_{i,t}}{\alpha^b} + \varepsilon^{bj}_{i,t}}. \tag{51}$$

In order to determine the value of $s^*_{-i,t}$, integrate both parts of (50) with respect to θ_i of density $f(\theta_i)$; we obtain

$$s^*_{-i,t}(q^{c*}_{-i,t}) = \frac{1}{I-1} \sum_{k \in I-i} \int_{\theta_k \in \Theta} \int_{\varepsilon^r_{k,p}} q^{c*}_k(\theta_k, \varepsilon^c_{k,p}) dF(\varepsilon^c_{k,p}) dF_k(\theta_k)$$

Assuming that one individual is negligible in the continuum of individuals so that $s^*_{-i,t} = s^*_{-j,p}$ and given that $E(\varepsilon^c_{i,t}) = 0 \ \forall i$, we have that

$$\begin{aligned} s^*_{-i,t} &= \frac{1}{I-1} \sum_{k \in I-i} \int_{\theta_k \in \Theta} \int_{\varepsilon^c_{k,p}} \left(\frac{\theta_k}{s^*} e^{\psi^c_k - \frac{h_i p^c_k}{\alpha^c} + \varepsilon^c_{k,p}}\right) dF(\varepsilon^c_{k,p}) dF_k(\theta_k) \\ \left(s^*_{-i,t}\right)^2 &\approx \frac{1}{I} \sum_{k \in I} \int_{\theta_k \in \Theta} \int_{-\infty}^{+\infty} \theta_k e^{\psi^c_k - \frac{h_i p^c_k}{\alpha^c} + \varepsilon^c_{k,p}} dF_k(\theta_k) dF(\varepsilon^c_{k,p}) \end{aligned}$$

Assuming that $\varepsilon_{k,p}^c$ follows an exteme value type I distribution with parameters $(1, \mu_1)$ with density

$$f(\varepsilon_{k,p}^c) = \mu_1 e^{-\mu_1 \varepsilon_{k,p}^c} e^{-e^{-\mu_1 \varepsilon_{k,p}^c}}, \mu_1 > 0,$$

then

$$\int_{-\infty}^{+\infty} e^{\varepsilon_{k,p}^c} dF(\varepsilon_{k,p}^c) = \mu_1 \int_{-\infty}^{+\infty} e^{(1-\mu_1)\varepsilon_{k,p}^c} e^{-e^{-\mu_1 \varepsilon_{k,p}^c}} d\varepsilon_{k,p}^c$$

Let assume $Y = e^{-\mu_1 \varepsilon_{k,p}^c}$ or $e^{\varepsilon_{k,p}^c} = (Y)^{-\frac{1}{\mu_1}}$, then

$$\begin{aligned} \int_{-\infty}^{+\infty} e^{\varepsilon_{k,p}^c} dF(\varepsilon_{k,p}^c) &= \int_{0}^{+\infty} Y^{\left(1-\frac{1}{\mu_1}\right)} e^{-Y} dY \\ &= \left(1 - \frac{1}{\mu_1}\right)! = \Gamma\left(2 - \frac{1}{\mu_1}\right) \end{aligned}$$

where $\Gamma(.)$ is the gamma function.

In the case of complete information, the proof always remains the same, except the determination of s_p^*, where we do not have to integrate over Θ. Then we obtain

$$\begin{aligned} \left(s_{-i,t}^*\right)^2 &= \frac{\Gamma\left(2 - \frac{1}{\mu_1}\right)}{I-1} \sum_{j \in I-i} \theta_j e^{\psi_j^c - \frac{h_j p_{j,p}^c}{\alpha^c}} \\ &\approx \frac{\Gamma\left(2 - \frac{1}{\mu_1}\right)}{I} \sum_{j \in I} \theta_j e^{\psi_j^c - \frac{h_j p_{j,p}^c}{\alpha^c}}. \end{aligned}$$

■

Proof of Proposition 2. Case 1: $0 < \alpha_1 < 1/2$.

This problem enters the convex optimization framework of the Kuhn–Tucker theorem. However, we can also solve it directly as follows.

Let us rewrite the social value function in the form

$$U(q, a, \theta) = (2\alpha_1 - 1)S(q, a, \theta) + (1 - \alpha_1)W(q).$$

For any given $q_{i,t}^r$, since $2\alpha_1 - 1 < 0$, the maximum is achieved if the numbers $a_{i,t}^r$ are chosen so as to make $S(q, a, \theta)$ as small as possible. The best choice is obtained by requiring $S_{i,t}^r = 0$ for all i, proving (33). Then we have to maximize $W(q)$ with respect to $q_{i,t}^r$. Since this function is concave and differentiable, the maximum is achieved if and only if its first partial derivatives all vanish, i.e.,

$$\frac{\alpha^r}{h(\theta_i)} \left[\psi_{i,t}^r + \ln \theta_i - \ln s_{-i,t}^* - \ln q_{i,t}^r + \varepsilon_{i,t}^r\right] - c_p^r = 0.$$

for every i. This implies (**??**) and that $p^r_{i,t}\left(q^r_{i,t}\right) = c^r_p$. It follows that

$$S^r_i = \left[\frac{\alpha^r}{h(\theta_i)} + p^r_{i,t} - c^r_p\right] q^c_{i,t} - a^r_{i,t}$$

for every i choosing option r, and therefore the condition $S^r_i = 0$ implies (32). Equations (34) and (35) then follow by definition.

Case 2: $1/2 < \alpha_1 < 1$.

Writing

$$U(q,a,\theta) = (2\alpha_1 - 1)S(q,a,\theta) + (1-\alpha_1)W(q,\theta)$$

again, we see that for any given q this expression takes its largest value when S is biggest. This leads to the conditions (36) and to the equality

$$U(q,0,\theta) = \alpha_1 \sum_i \left(U^r - p^r_{i,t}q^r_{i,t}\right).$$

where

$$U^r := \int_{\varepsilon^r_{i,t}} \left(\frac{\alpha^r}{h(\theta_i)} q^r_{i,t}[1 + \psi^r_{i,t} + \ln\theta_i - \ln s^*_{-i,t} - \ln q^r_{i,t} + \varepsilon^r_{i,t}]\right) dF(\varepsilon^r_{i,t})$$

Maximizing the last expression with respect to the variables $q^r_{i,t}$ we obtain the following conditions:

$$\frac{\alpha^r}{h(\theta_i)}[\psi^r_{i,t} + \ln\theta_i - \ln s^*_{-i,t} - \ln q^r_{i,t} + \varepsilon^r_{i,t}] = \frac{2\alpha_1 - 1}{\alpha_1} p^r_{i,t} + \frac{1-\alpha_1}{\alpha_1} c^r_p. \qquad (52)$$

Denote by W_{q^r} the partial derivative of W with respect to q^r. Equation (22) implies that $W_{q^r} = p^r_{i,t} - c^r_p$. This implies that the left hand side of these equalities is equal to $p^r_{i,t}$ and that $p^r_{i,t} = c^r_p$ leading to (31). Furthermore, substituting (31) into the definition of S^r_i and using the conditions $a^r_{i,t} = 0$ we obtain (37) and (38). Finally, for any positive function $h(.)$ the participation constraints (28) are satisfied. ■

Proof of Proposition 3. Since the functions S_i are differentiable and concave in $q^r_{i,t}$and $a^r_{i,t}$, condition (42) is equivalent for each i to the following first-order condition:

$$\frac{\partial q^r_{i,t}(\theta,\varepsilon^r_{i,t})}{\partial\theta_i}\left(\frac{\alpha^r}{h(\theta_i)}\left[\psi^r_{i,t} + \ln\theta_i - \ln s^*_{-i,t} - \ln q^r_{i,t}(\theta,\varepsilon^r_{i,t}) + \varepsilon^r_{i,t}\right] - p^r_{i,t}\right) - \frac{\partial a_{i,t}(\theta,\varepsilon^r_{i,t})}{\partial\theta_i} = 0.$$

Hence

$$\frac{\partial S^r_i}{\partial\theta_i}(q_{i,t}(\theta,\varepsilon^r_{i,t}), s^*_{-i,t}, a_{i,t}(\theta,\varepsilon^r_{i,t}),\theta_i) = V_i(q^r_{i,t}(\theta,\varepsilon^r_{i,t}),\theta_i,\varepsilon^r_{i,t})$$

and therefore

$$S_i^r(q_{i,t}^r(\theta,\varepsilon_{i,t}^r), s_{-i,t}^*, a_{i,t}^r, \theta_i, \varepsilon_{i,t}^r) = S_i^r(q_{i,t}(\theta_{-i},\underline{\theta_i}), s_{-i,t}^*, a_{i,t}^r(\theta_{-i},\underline{\theta_i}), \underline{\theta_i}, \varepsilon_{i,t}^r) + \int_{\underline{\theta_i}}^{\theta_i} V_i(q_{i,t}(\theta_{-i},\tilde{\theta}_i,\varepsilon_{i,t}^r),\tilde{\theta}_i) d\tilde{\theta}_i. \tag{53}$$

Substituting this expression into the definition of U we obtain the following equality:

$$\begin{aligned}
&\int_\Theta U(q(\theta), a(\theta), \theta) f(\theta)\ d\theta \\
&= \int_\Theta \left(\begin{array}{c} \alpha_1 \left(\sum_{i=1}^I \int_{\varepsilon_{i,t}^r} S_i^r(q_{i,t}(\theta,\varepsilon_{i,t}^r), s_{-i,t}^*, a_{i,t}^r(\theta), \theta_i) dF(\varepsilon_{i,t}^r) \right) \\ +(1-\alpha_1)\pi(q(\theta), a(\theta)) \end{array} \right) f(\theta)\ d\theta \\
&= \int_\Theta \left(\begin{array}{c} (2\alpha_1 - 1) \left(\sum_{i=1}^I \int_{\varepsilon_{i,t}} S_i^r(q_{i,t}^r(\theta,\varepsilon_{i,t}^r), s_{-i,t}^*, a_{i,t}^r(\theta), \theta_i) dF(\varepsilon_{i,t}^r) \right) \\ +(1-\alpha_1)W(q(\theta), \theta) \end{array} \right) f(\theta)\ d\theta \\
&= \int_\Theta \Big((2\alpha_1 - 1) \Big(\sum_{i=1}^I \int_{\varepsilon_{i,t}^r} S_i^r(q_{i,t}(\theta_{-i},\underline{\theta_i},\varepsilon_{i,t}^r), s_{-i,t}^*, a_{i,t}^r(\theta_{-i},\underline{\theta_i},\varepsilon_{i,t}^r), \underline{\theta_i}) dF(\varepsilon_{i,t}^r) \\
&\qquad + \int_{\underline{\theta_i}}^{\theta_i} V_i(q_{i,t}(\theta_{-i},\tilde{\theta}_i,\varepsilon_{i,t}^r),\tilde{\theta}_i)\ d\tilde{\theta}_i \Big) + (1-\alpha_1)W(q(\theta),\theta) \Big) f(\theta)\ d\theta.
\end{aligned}$$

Integrating by parts we have

$$\int_{\underline{\theta_i}}^{\overline{\theta_i}} \int_{\underline{\theta_i}}^{\theta_i} V_i(q_{i,t}(\theta_{-i},\tilde{\theta}_i),\tilde{\theta}_i,\varepsilon_{i,t}^r)\ d\tilde{\theta}_i f_i(\theta_i)\ d\theta_i = \int_{\underline{\theta_i}}^{\overline{\theta_i}} V_i(q_{i,t}(\theta,\varepsilon_{i,t}^r),\theta_i)(1-F_i(\theta_i))\ d\theta_i.$$

Using this relation we may rewrite the previous equality in the following form:

$$\begin{aligned}
&\int_\Theta U(q(\theta), a(\theta), \theta) f(\theta)\ d\theta \\
&\quad = \int_\Theta \int_{\varepsilon_{i,t}} \Big((2\alpha_1 - 1) \Big(\sum_{i=1}^I S_i(q_{i,t}(\theta_{-i},\underline{\theta_i}), s_{-i,t}^*, a_{i,t}(\theta_{-i},\underline{\theta_i}), \underline{\theta_i}) \\
&\quad + V_i(q_{i,t}(\theta,\varepsilon_{i,t}^r),\theta_i) \frac{1-F_i(\theta_i)}{f_i(\theta_i)} \Big) + (1-\alpha_1)W(q(\theta),\theta) \Big) dF(\varepsilon_{i,t}^r) f(\theta)\ d\theta.
\end{aligned}$$

We are going to maximize pointwise the function under the integral sign. Since $\alpha_1 < 1/2$ by assumption, for any given values of $q(\theta,\varepsilon_{i,t}^r)$ we have to choose $a_{i,t}^r(\theta)$ so as to minimize the nonnegative term

$$S_i^r(q_{i,t}^r(\theta_{-i},\underline{\theta_i}), s_{-i,t}^*, a_{i,t}^r(\theta_{-i},\underline{\theta_i}), \underline{\theta_i}, \varepsilon_{i,t}^r).$$

We choose $a_{i,t}(\theta_{-i}, \underline{\theta_i}, \varepsilon_{i,t})$ such that

$$S_i^r(q_{i,t}^r(\theta_{-i}, \underline{\theta_i}), s_{-i,t}^*, a_{i,t}^r(\theta_{-i}, \underline{\theta_i}), \underline{\theta_i}, \varepsilon_{i,t}^r) = 0. \tag{54}$$

Using the relations (54), these formulae reduces to the following:

$$\int_\Theta U(q(\theta), a(\theta), \theta) f(\theta)\, d\theta = \int_\Theta \int_{\varepsilon_{i,t}^r} \Big((2\alpha_1 - 1)\Big(\sum_{i=1}^{I} V_i(q_{i,t}(\theta, \varepsilon_{i,t}^r), \theta_i) \frac{1 - F_i(\theta_i)}{f_i(\theta_i)}\Big) + (1 - \alpha_1) W(q(\theta), \theta)\Big) dF(\varepsilon_{i,t}^r) f(\theta)\, d\theta. \tag{55}$$

It remains to maximize the function under the integral sign on the right-hand side. A straightforward computation yields

$$\frac{\partial V_i(q_{i,t}^r(\theta, \varepsilon_{i,t}^r), \theta_i)}{\partial q_{i,t}^r(\theta, \varepsilon_{i,t}^r)} = \frac{\alpha^r}{\theta_i h(\theta_i)} - \frac{\alpha^r h'(\theta_i)}{h^2(\theta_i)} [\psi_{i,t}^r + \ln\theta_i - \ln s_{-i,t}^* - \ln q_{i,t}^r(\theta, \varepsilon_{i,t}^r) + \varepsilon_{i,t}^r],$$

$$\frac{\partial W(q_{i,t}^r(\theta, \varepsilon_{i,t}^r), \theta)}{\partial q_{i,t}^r(\theta, \varepsilon_{i,t}^r)} = \frac{\alpha^r}{h(\theta_i)} [\psi_{i,t}^r + \ln\theta_i - \ln s_{-i,t}^* - \ln q_{i,t}^r(\theta, \varepsilon_{i,t}^r) + \varepsilon_{i,t}^r] - c_p^r.$$

and therefore the following first-order conditions (solving the problem point-wise):

$$\frac{\alpha^r}{h(\theta_i)} [\psi_{i,t}^r + \ln\theta_i - \ln s_{-i,t}^* - \ln\left(q_{i,t}^r(\theta, \varepsilon_{i,t}^r)\right) + \varepsilon_{i,t}^r] = c_{i1}^r(\theta_i, \varepsilon_{i,t}^r). \tag{56}$$

Solving this system we obtain (43). Using these results, we deduce from (22) and (54) that

$$a_{i,t}(\theta, \varepsilon_{i,t}^r) = \frac{\alpha^r}{h(\theta_i)} \left(q_{i,t}^r\right) [1 + \psi_{i,t}^r + \ln\theta_i - \ln s_{-i,t}^* - \ln\left(q_{i,t}^r\right) + \varepsilon_{i,t}^r] - p_{i,t}^r q_{i,t}^r - \int_{\underline{\theta_i}}^{\theta_i} V_i(q_{i,t}(\theta_{-i}, \tilde\theta_i), \tilde\theta_i, \varepsilon_{i,t}^r)\, d\tilde\theta_i.$$

It follows from (53) and (56) that

$$V_i(q_{i,t}(\theta, \varepsilon_{i,t}), \theta_i, \varepsilon_{i,t}) = q_{i,t}^r(\theta, \varepsilon_{1i}) \left(\frac{\alpha^r}{\theta_i h(\theta_i)} - \frac{\alpha^r h'(\theta_i)}{h^2(\theta_i)} - \frac{h'(\theta_i)}{h(\theta_i)} c_{i1}^r(\theta_i, \varepsilon_{i,t}^r)\right) \tag{57}$$

Since all factors in this expression are nonnegative by (42), the participation constraints (40) follow from (42). ■

Proof of Proposition 4. By repeating the first part of the proof of the preceding proposition, we are led again to the maximization of the integral

(55). If $\theta \in \Theta$ satisfies $\theta_i > \underline{\theta_i}$ for all i, then our former reasoning still leads to the first-order conditions (56). Since almost all points $\theta \in \Theta$ have this property, we can still solve this system and we obtain (45). Furthermore, since each S_i is a decreasing function of $a_{i,t}$, it remains to show that by choosing $a_{i,t}(\theta, \varepsilon_{i,t}^r) = 0$ the participation constraints are satisfied. Finally, the nonnegativity of S_i follows from its explicit formulae

$$S(q_{i,t}^r\left(\theta, \varepsilon_{i,t}^r\right), s_{-i,t}^*, a_{i,t}(\theta, \varepsilon_{i,t}^r), \theta_i, \varepsilon_{i,t}^r) = \frac{\alpha^r}{h(\theta_i)} q_{i,t}^r.$$

■

8 Appendix 2: Estimation Procedure

Recall that $\mathcal{R}$ is the set of all possible choices of modes and fare categories $\mathcal{R} \equiv \{c, (b,1), (b,2), ..., (b,J), w\}$. We denote by r, the typical element, $r = 1, ...J+2$. Using these notations, (1) can be written:

$$U_{i,t} = \sum_{r=1}^{J+2} \alpha^r q_{i,t}^r [1 + \psi_{i,t}^r + \ln\theta_i - \ln s_{-i,t} - \ln q_{i,t}^r + \varepsilon_{i,t}^r] + h(\theta_i)\nu_{i,t}$$

Recall that the discrete choice of which mode of transportation to select can be represented by a set of binary valued indices $\delta_{i,t}^r = \delta_{i,t}^r\left(a_{i,t}, p_{i,t}, w_i, \psi_{i,t}, \theta_i, s_{-i,t}^*, \varepsilon_{i,t}\right)$, where

$$\begin{aligned} \delta_{i,t}^r &= 1 \text{ if } \overline{V}_{i,t}^r(\theta_i) > \overline{V}_{i,t}^{r'}(\theta_i) \ \ \forall r' \neq r \\ &= 0 \text{ otherwise} \end{aligned} \tag{58}$$

The relationship between these unconditional functions and the corresponding conditional ones above is that:

$$q_{i,t}^r(\theta_i, \psi_{i,t}, s_{-i,t}^*, \varepsilon_{i,t}) = \delta_{i,t}^r \overline{q}_{i,t}^r(\theta_i, \psi_{i,t}^r, s_{-i,t}^*, \varepsilon_{i,t}) \tag{59}$$

and

$$V_{i,t} = \max_{\overline{V}_{i,t}^r}(\overline{V}_{i,t}^1, \overline{V}_{i,t}^2, ..., \overline{V}_{i,t}^{J+1}). \tag{60}$$

Hence, the discrete choice indices are random variables, with mean $E\left(\delta_{i,t}^r\right) = P_{i,t}^r(\theta_i)$, with, $\forall r' \neq r$,

$$\begin{aligned} P_{i,t}^r(\theta_i) &= \Pr\text{ob}\left(\overline{V}_{i,t}^r(\theta_i) > \overline{V}_{i,t}^{r'}(\theta_i) \ \ \forall r' \neq r\right) \\ &= \Pr\text{ob}\begin{pmatrix} \alpha^r \frac{\theta_i}{s_{-i}^*} e^{\psi_{i,t}^r - \frac{h(\theta_i)p_{i,t}^r}{\alpha^r} + \varepsilon_{i,t}^r} + h(\theta_i)\left[w_i - a_{i,t}^r\right] & \\ > \alpha^{r'} \frac{\theta_i}{s_{-i}^*} e^{\psi_{i,t}^{r'} - \frac{h(\theta_i)p_{i,t}^{r'}}{\alpha^{r'}} + \varepsilon_{i,t}^{r'}} + h(\theta_i)[w_i - a_{i,t}^{r'}] & \forall r' \neq r \end{pmatrix} \\ &= \Pr\text{ob}\begin{pmatrix} \alpha^r e^{\psi_{i,t}^r - \frac{h(\theta_i)p_{i,t}^r}{\alpha^r} + \varepsilon_{i,t}^r} - a_{i,t}^r h(\theta_i)\frac{s_{-i}^*}{\theta_i} & \\ > \alpha^{r'} e^{\psi_{i,t}^{r'} - \frac{h(\theta_i)p_{i,t}^{r'}}{\alpha^{r'}} + \varepsilon_{i,t}^{r'}} - a_{i,t}^{r'} h(\theta_i)\frac{s_{-i}^*}{\theta_i} & \forall r' \neq r \end{pmatrix} \\ &\simeq \Pr\text{ob}\begin{pmatrix} \alpha^r\left(1 + \psi_{i,t}^r - \frac{h(\theta_i)p_{i,t}^r}{\alpha^r} + \varepsilon_{i,t}^r\right) - a_{i,t}^r h(\theta_i)\frac{s_{-i}^*}{\theta_i} & \\ > \alpha^{r'}\left(1 + \psi_{i,t}^{r'} - \frac{h(\theta_i)p_{i,t}^{r'}}{\alpha^{r'}} + \varepsilon_{i,t}^{r'}\right) - a_{i,t}^{r'} h(\theta_i)\frac{s_{-i}^*}{\theta_i} & \forall r' \neq r \end{pmatrix} \\ &= \Pr\text{ob}\left(\lambda_{i,t}^r(\theta_i) + \alpha^r\varepsilon_{i,t}^r > \lambda_{i,t}^{r'}(\theta_i) + \alpha^{r'}\varepsilon_{i,t}^{r'} \ \forall r' \neq r\right) \end{aligned} \tag{61}$$

where the fourth line uses a first-order Taylor approximation of the exponentional function, and

$$\lambda_{i,t}^r(\theta_i) \equiv \alpha^r + \alpha^r\psi_{i,t}^r - h(\theta_i)p_{i,t}^r - a_{i,t}^r h(\theta_i)\frac{s_{-i}^*}{\theta_i}.$$

Throughout this section, I assume that $\varepsilon_{i,t}^{r}$ $\forall r$ are independently and identically distributed according to the Extreme Value type I distribution $G\left(0,\mu_1\right)$ so that $\alpha^r\varepsilon_{i,t}^{r}\sim G\left(0,\frac{\mu_1}{\alpha^r}\right)$. Hence, their joint cumulative distribution function is

$$F_{\alpha\varepsilon}=\exp\left[-\sum_{r=1}^{J+1}\exp(-\frac{\mu_1}{\alpha^r}\varepsilon_{i,t}^{r})\right]$$

Let $F_{\alpha\varepsilon}^{r}$ be the derivative of $F_{\alpha\varepsilon}$ with respect to the r^{th} component, then

$$P_{i,t}^{r}\left(\theta_i\right)=\int_{-\infty}^{+\infty}F_{\alpha\varepsilon}^{r}(t+\lambda_{i,t}^{r}\left(\theta_i\right)-\lambda_{i,t}^{1}\left(\theta_i\right),...,t+\lambda_{i,t}^{r}\left(\theta_i\right)-\lambda_{i,t}^{J+1}\left(\theta_i\right))dt$$

with

$$F_{\alpha\varepsilon}^{r}(t+\lambda_{i,t}^{r}\left(\theta_i\right)-\lambda_{i,t}^{1}\left(\theta_i\right),...,t+\lambda_{i,t}^{r}\left(\theta_i\right)-\lambda_{i,t}^{J+1}\left(\theta_i\right))=\frac{\mu_1}{\alpha^r}e^{-\frac{\mu_1}{\alpha^r}t}\exp(-\beta_{i,t}^{r}e^{-\frac{\mu_1}{\alpha^r}t})$$

and

$$\beta_{i,t}^{r}\left(\theta_i\right)=e^{-\frac{\mu_1}{\alpha^r}\lambda_{i,t}^{r}\left(\theta_i\right)}\sum_{r'=1}^{J+2}e^{\frac{\mu_1}{\alpha^r}\lambda_{i,t}^{r'}\left(\theta_i\right)}.$$

Hence,

$$\begin{aligned}
P_{i,t}^{r}\left(\theta_i\right) &= \frac{\mu_1}{\alpha^r}\int_{-\infty}^{+\infty}e^{-\frac{\mu_1}{\alpha^r}t}\exp(-\beta_{i,t}^{r}\left(\theta_i\right)e^{-\frac{\mu_1}{\alpha^r}t})dt \\
&= \left(\beta_{i,t}^{r}\left(\theta_i\right)\right)^{-1}=\frac{e^{\frac{\mu_1}{\alpha^r}\lambda_{i,t}^{r}\left(\theta_i\right)}}{\sum_{r'=1}^{J+1}e^{-\frac{\mu_1}{\alpha^r}\lambda_{i,t}^{r'}\left(\theta_i\right)}} \\
&= \frac{e^{\frac{\mu_1}{\alpha^r}\left(\alpha^r+\alpha^r\psi_{i,t}^{r}-h(\theta_i)p_{i,t}^{r}-a_{i,t}^{r}h(\theta_i)\frac{s_{-i}^{*}}{\theta_i}\right)}}{\sum_{r'=1}^{J+2}e^{\frac{\mu_1}{\alpha^{r'}}\left(\alpha^{r'}+\alpha^{r'}\psi_{i,t}^{r'}-h(\theta_i)p_{i,t}^{r'}-a_{i,t}^{r'}h(\theta_i)\frac{s_{-i}^{*}}{\theta_i}\right)}}.
\end{aligned}$$

Let $A_{i,t}^{r}\left(\theta_i\right)=\left\{\alpha\varepsilon|\lambda_{i,t}^{r}\left(\theta_i\right)+\alpha^r\varepsilon_{i,t}^{r}>\lambda_{i,t}^{r'}\left(\theta_i\right)+\alpha^{r'}\varepsilon_{i,t}^{r'},\forall r'\neq r\right\}$. Then the conditional marginal density of vector $\alpha\varepsilon$ is:

$$\begin{aligned}
f_{\alpha\varepsilon|\alpha\varepsilon\in A_{i,t}^{r}} &= \frac{F_{\varepsilon}^{r}(t+\lambda_{i,t}^{r}\left(\theta_i\right)-\lambda_{i,t}^{1}\left(\theta_i\right),...,t+\lambda_{i,t}^{r}\left(\theta_i\right)-\lambda_{i,t}^{J+1}\left(\theta_i\right))}{P_{i,t}^{r}\left(\theta_i\right)} \\
&= \frac{\mu_1}{\alpha^r}\beta_{i,t}^{r}\left(\theta_i\right)e^{-\frac{\mu_1}{\alpha^r}\varepsilon_{i,t}^{r}}\exp(-\beta_{i,t}^{r}\left(\theta_i\right)e^{-\frac{\mu_1}{\alpha^r}\varepsilon_{i,t}^{r}}).
\end{aligned}$$

Moreover,

$$\overline{q}_{i,t}^{r}=\frac{\theta_i}{s_{-i}^{*}}e^{\psi_{i,t}^{r}-\frac{h(\theta_i)p_{i,t}^{r}}{\alpha^{r}}+\varepsilon_{i,t}^{r}},\ \forall r,$$

which with $\psi_{i,t}^{r}-\frac{h(\theta_i)p_{i,t}^{r}}{\alpha^{r}}=\frac{\lambda_{i,t}^{r}(\theta_i)+a_{i,t}^{r}h(\theta_i)\frac{s_{-i}^{*}}{\theta_i}}{\alpha^{r}}-1$, is equivalent to

$$\varepsilon_{i,t}^{r}=\ln\frac{\overline{q}_{i,t}^{r}s_{-i}^{*}}{\theta_i}-\frac{\lambda_{i,t}^{r}(\theta_i)+a_{i,t}^{r}h(\theta_i)\frac{s_{-i}^{*}}{\theta_i}}{\alpha^{r}}+1.$$

Finally, using a change in variable from ε_j to q_j, we have:

$$\begin{aligned}
f_{q_{i,t}^{r}|\alpha\varepsilon\in A_{i,t}^{r}}\left(q\left(\theta_i\right)\right) &= \frac{\mu_1}{\alpha^{r}}\beta_{i,t}^{r}\left(\theta_i\right)e^{-\frac{\mu_1}{\alpha^{r}}\left(\ln\frac{qs_{-i}^{*}}{\theta_i}-\frac{\lambda_{i,t}^{r}(\theta_i)+a_{i,t}^{r}h(\theta_i)\frac{s_{-i}^{*}}{\theta_i}}{\alpha^{r}}+1\right)}\\
&\quad\times\exp\left[-\beta_{i,t}^{r}e^{-\frac{\mu_1}{\alpha^{r}}\left(\ln\frac{qs_{-i}^{*}}{\theta_i}-\frac{\lambda_{i,t}^{r}(\theta_i)+a_{i,t}^{r}h(\theta_i)\frac{s_{-i}^{*}}{\theta_i}}{\alpha^{r}}+1\right)}\right]\\
&= \frac{\mu_1}{\alpha^{r}}\frac{\sum_{r'=1}^{J+2}e^{\frac{\mu_1}{\alpha^{r'}}\lambda_{i,t}^{r'}(\theta_i)}}{e^{\frac{\mu_1}{\alpha^{r}}\lambda_{i,t}^{r}(\theta_i)}}\left(\frac{qs_{-i}^{*}}{\theta_i}\right)^{-\frac{\mu_1}{\alpha^{r}}}e^{-\frac{\mu_1}{\alpha^{r}}\left(1-\frac{\lambda_{i,t}^{r}(\theta_i)+a_{i,t}^{r}h(\theta_i)\frac{s_{-i}^{*}}{\theta_i}}{\alpha^{r}}\right)}\\
&\quad\times\exp\left[-\frac{\sum_{r'=1}^{J+1}e^{\frac{\mu_1}{\alpha^{r'}}\lambda_{i,t}^{r'}(\theta_i)}}{e^{\frac{\mu_1}{\alpha^{r}}\lambda_{i,t}^{r}(\theta_i)}}\left(\frac{qs_{-i}^{*}}{\theta_i}\right)^{-\frac{\mu_1}{\alpha^{r}}}e^{-\frac{\mu_1}{\alpha^{r}}\left(1-\frac{\lambda_{i,t}^{r}(\theta_i)+a_{i,t}^{r}h(\theta_i)\frac{s_{-i}^{*}}{\theta_i}}{\alpha^{r}}\right)}\right].
\end{aligned}$$

Note that

$$\begin{aligned}
E\left(q_{i,t}^{r}|\varepsilon\in A_{i,t}^{r}\right) &= \frac{\theta_i}{s_{-i}^{*}}e^{\psi_{i,t}^{r}-\frac{h(\theta_i)p_{i,t}^{r}}{\alpha^{r}}}E\left[e^{\varepsilon_{i,t}^{r}}|\varepsilon\in A_{i,t}^{r}\right]\\
&= \frac{\theta_i}{s_{-i}^{*}}e^{\psi_{i,t}^{r}-\frac{h(\theta_i)p_{i,t}^{r}}{\alpha^{r}}}\beta_{i,t}^{r}\left(\theta_i\right)\Gamma\left(1-\frac{1}{\mu}t\right),
\end{aligned}$$

where the second equality is derived from the moment generating function of a univariate Extreme Value distribution with scale parameter $\frac{\mu_1}{\alpha^{r}}$ and location parameter $\frac{\mu_1}{\alpha^{r}}\ln\beta_{i,t}^{r}$:

$$E\left(e^{t\varepsilon_{i,t}^{r}}|\varepsilon\in A_{i,t}^{r}\right)=\left(\beta_{i,t}^{r}\left(\theta_i\right)\right)^{\frac{\mu_1}{\alpha^{r}}t}\Gamma\left(1-\frac{\alpha^{r}}{\mu_1}t\right)$$

Let r^{*} denote the index for the mode of transportation and type of payment selected by the i^{th} individual, let $q_{i,t}^{r*}$ be his observed number of trips. Then,

$$
\begin{aligned}
L &= \prod_{i=1}^{N} \int_{\underline{\theta}}^{\overline{\theta}} f_{q^{r*}_{i,t}} \left(q^{r*}_{i,t} \left(\theta_i \right) \right) dF_i(\theta_i), \\
&= \prod_{i=1}^{N} \int_{\underline{\theta}}^{\overline{\theta}} f_{q^{r*}_{i,t} | \varepsilon \in A^{r}_{i,t}} \left(q^{r*}_{i,t} \left(\theta_i \right) \right) P^{r}_{i,t} \left(\theta_i \right) dF_i(\theta_i).
\end{aligned}
$$

9 Appendix 3. Tables and Figures

Figure 1a: Number of car trips by income range

[Insert Figure 1a here]

Figure 1b: Number of bus trips by income range

[Insert Figure 2a here]

Table 1: Estimation results, Peak Period (standard errors in parentheses).

$Variable$	EER(07)	Mod. 1	Mod. 2	Mod. 3
$MU.car$	0.895^*	1.213^*	1.221^*	1.153^*
(α_1)	(0.012)	(0.141)	(0.140)	(0.122)
$MU.\ Bus$	0.105^*	0.746^*	0.774^*	0.728^*
(α_2)	(0.012)	(0.082)	(0.081)	(0.087)
$Distr.\ \theta$	0.300^*	0.646^*	0.646^*	0.684^*
(μ)	(0.002)	(0.034)	(0.031)	(0.032)
$MU\ Inc.$	1.941^*	0.011^*	0.011^*	0.021^*
$(h_0^I), (h_0^\theta)$	(0.259)	(0.001	(0.001)	(0.002)
$MU\ Inc.$	—	—	0.020^*	—
$(h_1^I),(h_1^\theta)$			(0.004)	
$MU\ Inc.$	—	—	0.006	—
(h_2^I)			(0.005)	
$Dist.\ \varepsilon_i$	—	4.630^*	4.652^*	4.364^*
(μ_1)		(0.380)	(0.409)	(0.427)
$B.Const.$	1.618^*	0.532^*	0.533^*	0.513^*
(β_1^b)	(0.177)	(0.154)	(0.152)	(0.194)
$B.Freq.$	0.106^*	0.067^*	0.066^*	0.069^*
(β_2^b)	(0.017)	(0.009)	(0.009)	(0.009)
$B.\ Dist.$	-0.011	-0.014^*	-0.014^*	-0.011
(β_3^b)	(0.015)	(0.008)	(0.008)	(0.008)
$B.Time$	0.005^*	0.003^*	0.003^*	0.003^*
(β_4^b)	(9^{-4})	(5^{-4})	(5^{-4})	(5^{-4})
$B.Stud.$	0.444^*	.019	0.004	-0.008
(β_5^b)	(0.104)	(0.096)	(0.091)	(0.116)
$B.Unemp.$	—	0.050	0.034	0.038
(β_6^b)		(0.088)	(0.083)	(0.099)
$B.Age$	—	-0.004^*	-0.003^*	-0.004^*
(β_7^b)		(0.002)	(0.002)	(0.002)
$C.Const.$	1.518^*	0.916^*	0.910^*	0.842^*
(β_1^c)	(0.179)	(0.141)	(0.143)	(0.163)
$C.Power$	0.034^*	0.034^*	0.034^*	0.032^*
(β_2^c)	(0.006)	(0.005)	(0.005)	(0.005)
$C.Dist.$	0.082^*	0.002	0.003	0.021^*
(β_3^c)	(0.009)	(0.007)	(0.007)	(0.009)
$C.Time$	5^{-4}	7^{-4*}	7^{-4*}	7^{-4*}
(β_4^c)	(7^{-4})	(4^{-4})	(4^{-4})	(4^{-4})
$C.Stud$	-0.490^*	-0.265^*	-0.272^*	-0.297^*
(β_5^c)	(0.099)	(0.080)	(0.078)	(0.086)
$C.Unemp.$	-0.200^*	-0.114^*	-0.123^*	-0.124^*
(β_6^c)	(0.093)	(0.079)	(0.076)	(0.079)
$C.Age$	0.048^*	0.028^*	0.028^*	0.032^*
(β_7^c)	(0.009)	(0.006)	(0.028)	(0.006)
$(C.Age)^2$	-6^{-4*}	-4^{-4*}	-4^{-4*}	-4^{-4*}
(β_8^c)	(1^{-4})	(1^{-4})	(1^{-4})	(1^{-4})
Log L.	-7.7	-157.04	-157.04	-157.03

Table 2: Estimation results, Off-Peak Period (standard errors in parentheses).

Variable	EER(07)	Mod. 1	Mod. 2	Mod. 3
MU.car	0.894^*	1.739^*	1.678^*	2.133
(α_1)	(0.010)	(0.144)	(0.136)	(1.870)
MU.Bus	0.106^*	1.107^*	1.119^*	1.191^*
(α_2)	(0.010)	(0.070)	(0.069)	(0.698)
Distr.θ	0.386^*	0.642^*	0.638^*	0.572^*
(μ)	(0.003)	(0.050)	(0.050)	(0.244)
MU Inc.	2.286^*	0.013^*	0.013^*	0.025
$(\mathrm{H}_0^I),(\mathrm{H}_0^\theta)$	(0.283)	(9^{-4})	(0.001)	(0.017)
MU Inc.	—	—	0.020^*	—-
$(\mathrm{H}_1^I),(\mathrm{H}_1^\theta)$			(0.004)	
MU Inc.	—	—	0.009^*	—
(H_2^I)			(0.012)	
Dist. ε_i	—	7.561^*	7.382^*	7.2269^*
(μ_1)		(0.545)	(0.521)	(1.842)
B.Const.	0.058	0.686^*	0.666^*	0.534
(β_1^b)	(0.192)	(0.081)	(0.082)	(0.482)
B.Freq.	0.206^*	0.054^*	0.056^*	0.058^*
(β_2^b)	(0.024)	(0.007)	(0.007)	(0.018)
B.Dist.	0.078^*	0.002	5^{-4}	0.012
(β_3^b)	(0.019)	(0.006)	(0.006)	(0.032)
B.Time	-0.001	-4^{-4}	-4^{-4}	2^{-4}
(β_4^b)	(0.002)	(3^{-4})	(3^{-4})	(8^{-4})
B.Student	0.348^*	-0.020	-0.010	-0.082
(β_5^b)	(0.132)	(0.037)	(0.037)	(0.331)
B.Unemp.	—	-0.110^*	-0.079^*	-0.130
(β_6^b)		(0.042)	(0.041)	(0.092)
B.Age	—	-0.007^*	-0.007^*	-0.007
(β_7^b)		(6^{-4})	(6^{-4})	(0.004)
C.Const.	0.620^*	0.965^*	0.958^*	0.868^*
(β_1^c)	(0.204)	(0.067)	(0.068)	(0.377)
C.Power	0.048^*	0.021^*	0.021^*	0.019
(β_2^c)	(0.009)	(0.003)	(0.003)	(0.015)
C.Dist.	0.125^*	0.038^*	0.038^*	0.055^*
(β_3^c)	(0.011)	(0.005)	(0.005)	(0.027)
C.Time	-0.012^*	-0.002^*	-0.002^*	-0.001
(β_4^c)	$(9\mathrm{E}^{-4})$	(2^{-4})	(2^{-4})	(0.001)
C.Student	-0.416^*	-0.235^*	-0.233	-0.329
(β_5^c)	(0.118)	(0.036)	(0.037)	(0.293)
C.Unemp.	-0.220^*	-0.273^*	-0.250	-0.310^*
(β_6^c)	(0.078)	(0.037)	(0.037)	(0.084)
C.Age	0.026^*	0.008^*	0.008	0.006
(β_7^c)	(0.009)	(0.002)	(0.002)	(0.010)
$(C.Age)^2$	-4^{-4*}	-2^{-4*}	-2^{-4}	-2^{-4}
(β_8^c)	(1^{-4})	(3^{-5})	(3^{-5})	(1^{-4})
Log L.	-6.6	-376.38	-376.38	-376.36

Table 3: T-tests Peak/Non-Peak Period.

$Variable$	EER(07)	Mod. 1	Mod. 2	Mod. 3
$MU.car$ (α_1)	0.057	-2.612	-2.349	-0.523
$MU.$ Bus (α_2)	0.057	-3.350	-3.261	-0.657
$Distr.$ θ (μ)	-22.183	0.073	0.122	0.455
MU $Inc.(h_0)$	-0.901	-1.243	-1.284	-0.255
MU $Inc.$ $(h_1^I) or (h_1^\theta)$	——	——	-0.003	——
MU $Inc.$ (h_2^I)	——	——	-0.239	——
$Dist.$ ε_i (μ_1)	——	-4.409	-4.125	-1.514
$B.Const.$ (β_1^b)	5.975	-0.904	-0.772	-0.041
$B.Freq.$ (β_2^b)	-3.466	1.145	0.905	0.542
$B.Dist.$ (β_3^b)	-3.622	-1.573	-1.533	-0.705
$B.Time$ (β_4^b)	3.323	6.276	6.015	3.167
$B.Stud.$ (β_5^b)	0.571	0.375	0.150	0.209
$B.Unemp.$ (β_6^b)	——	1.637	1.217	1.240
$B.Age$ (β_7^b)	——	1.773	1.832	0.646
$C.Const.$ (β_1^c)	3.309	-0.314	-0.304	-0.063
$C.Power$ (β_2^c)	-1.352	2.405	2.355	0.780
$C.Dist.$ (β_3^c)	-3.023	-4.167	-3.981	-1.204
$C.Time$ (β_4^c)	11.753	5.945	5.852	1.326
$C.Student$ (β_5^c)	-0.484	-0.336	-0.459	0.105
$C.Unemp.$ (β_6^c)	0.162	1.826	1.498	1.613
$C.Age$ (β_7^c)	1.760	3.155	3.125	2.078
$(C.Age)^2$ (β_8^c)	-1.414	-2.651	-2.578	-1.432

Table 4-1: Welfare, peak period,$\alpha_1 = \frac{1}{2}$.

	Current	EER(07)	Mod. 1	Mod. 2	Mod. 3
$\left(q^c_{i,t}\right)^{cp}$	3.26 (3.23)	3.07 (1.25)	1.27 (0.29)	1.26 (0.30)	1.27 (0.33)
$\left(q^{bj}_{i,t}\right)^{cp}$	0.58 (1.49)	1.01 (0.93)	0.64 (0.13)	0.64 (0.13)	0.60 (0.13)
$\left(q^c_{i,t}\right)^{opt.}$	—	2.93 (1.37)	1.30 (0.28)	1.30 (0.29)	1.33 (0.33)
$\left(q^{bj}_{i,t}\right)^{opt}$	—	0.84 (0.32)	0.63 (0.13)	0.63 (0.13)	0.59 (0.12)
$a_{i,t}$	0.15 (0.47)	16.97 (7.49)	0	0	0
$p^c_{i,t}$	2.01 (1.79)	1.07	1.07	1.07	1.07
$p^{bj}_{i,t}$	0.94 (0.33)	0.76	0.76	0.76	1.07
π_{total}	4847 (2.77)	63207 (8.82)	0	0	0
$S^{total}_{i,t}$	—	13709 (2.11)	64022 (5.08)	65358 (7.03)	40663 (3.68)
$Welfare$	—	76917 (10.83)	64022 (5.08)	65358 (7.03)	40663 (3.68)

Prices and welfare are in euros. Subscription covers a two days period only.

Table 4-2: Welfare, off- peak period, $\alpha_1 = \frac{1}{2}$.

	Current	EER(07)	Mod. 1	Mod. 2	Mod. 3
$\left(q^c_{i,t}\right)^{cp}$	1.15 (1.76)	1.09 (1.18)	0.80 (0.27)	0.81 (0.28)	0.76 (0.27)
$\left(q^{bj}_{i,t}\right)^{cp}$	0.26 (0.91)	0.44 (0.58)	0.59 (0.13)	0.59 (0.13)	0.56 (0.11)
$\left(q^c_{i,t}\right)^{opt.}$	—	1.08 (1.17)	0.82 (0.28)	0.83 (0.29)	0.79 (0.28)
$\left(q^{bj}_{i,t}\right)^{opt}$	—	0.40 (0.29)	0.59 (0.12)	0.59 (0.12)	0.56 (0.10)
$a_{i,t}$	0.05 (0.30)	5.18 (5.44)	0	0	0
$p^c_{i,t}$	2.18 (1.50)	1.07	1.07	1.07	1.07
$p^{bj}_{i,t}$	1.01 (0.21)	0.76	0.76	0.76	0.76
π_{total}	5980 (2.65)	20163 (6.66)	0	0	0
$S^{total}_{i,t}$	—	4932 (1.67)	49935 (6.46)	49090 (6.59)	35429 (4.98)
$Welfare$	—	25096 (8.30)	49935 (6.46)	49049 (6.59)	35429 (4.98)

Table 5-1 : Welfare, heterogenous pricing, $\alpha_1 = \frac{1}{3}^{(1)}$.

	Mod. 1	Mod. 2	Mod. 3	Mod. 1	Mod. 2	Mod. 3
	Peak	Peak	Peak	Off-P.	Off-P.	Off-P.
$\left(q_{i,t}^{c}\right)^{opt.}$	1.12 (0.24)	1.12 (0.25)	1.18 (0.29)	0.71 (0.24)	0.71 (0.25)	0.71 (0.25)
$\left(q_{i,t}^{bj}\right)^{opt.}$	0.54 (0.11)	0.55 (0.11)	0.53 (0.11)	0.51 (0.11)	0.50 (0.11)	0.50 (0.09)
$\left(a_{i,t}^{c}\right)\Big\|_{\theta=\frac{1}{1+\widehat{\mu}}}$	$0^{(2)}$	$0^{(2)}$	9.06	$0^{(2)}$	$0^{(2)}$	8.08
$\left(a_{i,t}^{c}\right)\Big\|_{\theta=0.25}$	$0^{(2)}$	$0^{(2)}$	2.69	$0^{(2)}$	$0^{(2)}$	1.28
$\left(a_{i,t}^{c}\right)\Big\|_{\theta=0.75}$	$0^{(2)}$	$0^{(2)}$	12.58	$0^{(2)}$	$0^{(2)}$	10.54
$\left(a_{i,t}^{bj}\right)\Big\|_{\theta=\frac{1}{1+\widehat{\mu}}}$	$0^{(2)}$	$0^{(2)}$	2.67	$0^{(2)}$	$0^{(2)}$	3.26
$\left(a_{i,t}^{bj}\right)\Big\|_{\theta=0.25}$	$0^{(2)}$	$0^{(2)}$	0.90	$0^{(2)}$	$0^{(2)}$	0.69
$\left(a_{i,t}^{bj}\right)\Big\|_{\theta=0.75}$	$0^{(2)}$	$0^{(2)}$	3.61	$0^{(2)}$	$0^{(2)}$	4.21
$\left(p_{i,t}^{c}\right)\Big\|_{\theta=\frac{1}{1+\widehat{\mu}}}$	9.38	9.56	5.24	11.14	10.88	7.34
$\left(p_{i,t}^{c}\right)\Big\|_{\theta=0.25}$	39.64	40.55	22.73	48.11	47.17	36.92
$\left(p_{i,t}^{c}\right)\Big\|_{\theta=0.75}$	5.35	5.45	1.13	6.29	6.19	1.18
$\left(p_{i,t}^{bj}\right)\Big\|_{\theta=\frac{1}{1+\widehat{\mu}}}$	5.87	6.15	3.38	7.17	7.30	4.23
$\left(p_{i,t}^{bj}\right)\Big\|_{\theta=0.25}$	24.47	25.80	14.39	30.70	31.49	20.67
$\left(p_{i,t}^{bj}\right)\Big\|_{\theta=0.75}$	3.40	3.54	0.80	4.09	4.18	0.82
π_{total}	11593 (0.92)	11829 (1.27)	41626 (3.77)	8982 (1.20)	17623 (2.28)	36789 (5.43)
$S_{i,t}^{total}$	22844 (1.82)	23311 (2.51)	580 (0.05)	17936 (2.25)	8792 (1.22)	850 (0.66)
$Welfare$	34437 (2.74)	35140 (3.78)	42206 (3.82)	26918 (3.42)	26415 (3.47)	37639 (5.28)

$^{(1)}$Recall that Model 1 and Model 2 are not defined for $\alpha_1 > \frac{1}{2}$; $^{(2)}$ Expression of the subscription simplifies to 0 in the special case of constant marginal utility of income (with respect to θ)

Table 5-2 : Welfare, Homogenous pricing, $\alpha_1 = 0.33$

	Mod. 1	Mod. 2	Mod. 3	Mod. 1	Mod. 2	Mod. 3
	Peak	Peak	Peak	Off-P.	Off-P.	Off-P.
$\left(q^c_{i,t}\right)^{opt.}$	1.45	1.45	1.56	0.66	0.70	0.57
	(0.32)	(0.32)	(0.38)	(0.22)	(0.24)	(0.20)
$\left(q^{bj}_{i,t}\right)^{opt}$	0.68	0.68	0.67	0.65	0.65	0.65
	(0.14)	(0.13)	(0.14)	(0.13)	(0.14)	(0.12)
$a^c_{i,t}$	0	0	9.66	0	0	8.59
$a^{bj}_{i,t}$	0	0	2.85	0	0	3.46
$p^c_{i,t}$	9.38	9.56	5.24	11.14	10.88	7.34
	(0)	(1.55)	(0)	(0)	(0.98)	(0)
$p^{bj}_{i,t}$	5.87	6.15	3.38	7.17	7.30	4.23
	(0)	(0.98)	(0)	(0)	(0.66)	(0)
π_{total}	35533	36311	47493	20057	20862	35877
	(2.84)	(3.92)	(4.29)	(2.53)	(2.78)	(5.18)
$S^{total}_{i,t}$	17469	17824	49645	13789	13564	27745
	(1.39)	(1.92)	(4.51)	(1.73)	(1.75)	(3.49)
$Welfare$	53003	54135	97138	33846	34426	63622
	(4.23)	(5.84)	(8.80)	(4.17)	(4.43)	(8.62)

References

[1] Ahn, H. and J. L. Powell, 1993. Semiparametric Estimation of Censored Selection Models with a Nonparametric Selection Mechanism. Journal of Econometrics, 58, 3-29.

[2] Arnott, R., de Palma A., and R. Lindsey, 1994. The Welfare Effects of Congestion Tolls with Heterogenous Commuters. Journal of Transportation Economics and Policy, 28, 139-161.

[3] Arnott, R., A. de Palma and R. Lindsey 1998. Recent developments in the bottleneck model. In: Road Pricing, Traffic Congestion and the Environment: Issues of Efficiency and Social Feasibility, K.J. Button and E.T. Verhoef (eds.). Edward Elgar, Cheltenham, UK.

[4] Beckmann, M., C.B. McGuire, and C.B.Winsten, 1956. Studies in the Economics of Transportation. Yale University Press, NewHaven, CT.

[5] Blackburn, A. J., 1970. An Alternative Approach to Aggregation and Estimation in the Non-Linear Model in the Demand for Travel: Theory and Measurement. D. C. Heath, Lexington, KY.

[6] Brock, W.A. and S. N. Durlauf, 2001. Discrete Choice with Social Interactions. Review of Economic Studies, 68, 2, 235-260.

[7] Calabrese, S., Epple, D., Romer, T. and H. Sieg, 2006. Local Public good provision: Voting, Peer Effects and Mobility. Journal of Public Economics, 90, 959-981.

[8] Clarke, E. H., 1971. Multipart Pricing of Public Goods. Public Choice, 2, 19–33.

[9] Dafermos, S.C., 1973. Toll Patterns for Multiclass-User Transportation Networks. Transportation Science, 211—223.

[10] Dubin, J. and D. McFadden, 1984. An Econometric Analysis of Residential Electric Appliance Holdings and Consumption. Econometrica, 50, 345-362.

[11] Epple, D. and H. Sieg, 1999. Estimating Equilibrium Models of Local Jurisdictions. Journal of Political Economy,107, 645-681.

[12] Epple, D., Romano, R. and H. Sieg, 2006. Admission, Tuition, and Financial Aid Policies in the Market for Higher Education. Econometrica, 74, 4, 885-928.

[13] Emmerink, R. H. M., 1998. Information and Pricing in Road Transportation. Springer-Verlag, New York, NY.

[14] Fudenberg, D. and J. Tirole, 1991. Game Theory. MIT Press, Cambridge, MA.

[15] Gayle, G. and C. Viauroux, 2007. Root N Consistent Semiparametric Estimators of a Dynamic Panel-sample-selection Model, 141, 179-212.

[16] Green, J. R. and J.-J. Laffont, 1977. Characterization of Satisfactory Mechanism for the Revelation of Preferences for Public Goods. Econometrica, 45, 427–438.

[17] Heckman, J. J., 1979. Sample Selection Bias as a Specification Error. Econometrica, 47, 153-161.

[18] Groves, T., 1973. Incentives in Teams, Econometrica, 41, 617–631.

[19] Hanemann, W. M., 1984. Discrete/continuous models of consumer demand. Econometrica 52, 541–561.

[20] Harsányi, J.C., 1967–68. Games with incomplete information played by "Bayesian" players. Management Science 14,159-182, 320-334, 486-529.

[21] Gmez-Ibez, J.A. and K.A. Small, 1994. Road pricing for congestion management: A survey of international practice. National Cooperative Highway Research Program, Synthesis of Highway Practice 210, TRB, National Academy Press, Washington, D.C.

[22] Imbens, G. and W. Newey, Identification and Estimation of Triangular Simultaneous Equation models without Additivity, Econometrica, 77, 5, 1481-1512.

[23] Maler, K., 1974. Environmental Economics. Baltimore, Johns Hopkins.

[24] McKie-Mason, J. and H. Varian, 1995. Pricing the Internet In: B. Kahin and J. Keller (Eds.), Public Access to the Internet. Englewood Cliffs, Prentice Hall, NJ.

[25] Myerson, R. B., 1979. Incentive Compatibility and the Bargaining Problem. Econometrica, 47, 61–73.

[26] Newey, W. K., 2007. Nonparametric Continuous/Discrete Choice Models. International Economic Review, 48, 4, 1429–1439.

[27] Small, K. A., 1992. Urban Transportation Economics. In: J. Lesourne and H. Sonnenshein (eds.), Fundamentals of Pure and Applied Economics, vol. 51. Harwood Academic Publishers, Chur, Switzerland.

[28] Parry, I.W.H. and Small, K. A., 2005. Does Britain or the United States Have the Right Gasoline Tax? American Economic Review, 95:1276–1289.

[29] Petrin, Amil, 2002. Quantifying the Benefits of New Products: The Case of the Minivan. Journal of Political Economy, 110, 4, 705-729.

[30] Small, K. A. and Yan, J., 2001. The value of "value pricing" of roads: second-best pricing and product differentiation. Journal of Urban Economics 49, 310–336.

[31] Viauroux, C., 2007. Structural estimation of congestion costs. European Economic Review, 51, 1–25.

[32] Viauroux, C., 2008. Marginal Utility of Income and Value of Time in Urban Transport. Economics Bulletin, 4, No 3, 1-7.

[33] Vickrey, W., 1961. Counterspeculation, Auctions and Competitive Sealed Tenders. Journal of Finance, 16, 8–37.