

Estimating potential demand for Autolib' – a new transport system for Paris

AUTHORS

Marco Kouwenhoven¹⁾, Eric Kroes^{1,2)}, Cyrille Gazave³⁾, Eric Tardivel⁴⁾

¹⁾ Significance, The Hague (Netherlands)

²⁾ VU University, Amsterdam (Netherlands)

³⁾ CDG Conseil, Paris (France)

⁴⁾ RATP, Paris (France)

Corresponding author: kroes@significance.nl

ABSTRACT

The City of Paris, together with surrounding “communes”, has created a public authority to investigate the possibility to launch by the end of 2011 a new transport system: Autolib'. The project is related to the highly successful Velib' project that was installed in Paris a few years ago. Autolib' is essentially a system of 4,000 “shared” electric cars that can be used for one-way trips of limited distance between 1,400 parking points within central Paris and the surrounding regions. Details of the fare system are still being studied, but it is envisioned that the user would pay a subscription fee and a variable cost depending on the duration of use. And importantly: there would be a guaranteed parking space at the destination of the trip.

Avis, RATP, SNCF and VINCI Park formed a consortium to bid for the operation of the Autolib' system. They have commissioned research to estimate the potential demand and revenue for the new Autolib' service with the highest possible accuracy. This to help them to shape the service in the best possible way, to determine the financial conditions and the economic basis of the project.

In the paper we briefly introduce the proposed new system, and report the stated choice research that was carried out to estimate the potential demand. The following three experiments were conducted:

- A stated intentions exercise to measure the absolute willingness to subscribe to the new Autolib' concept for different specifications of the system and its pricing;
- A stated choice experiment to measure preferences for different combinations of characteristics in the specification of the Autolib' system and its pricing;
- A stated choice experiment investigating mode choice among three alternatives: chosen mode, best alternative mode, newly proposed mode in varying specifications.

In the paper we describe the chosen methodology, the way in which the results of the three experiments have been integrated, the implementation of the population simulator and what we feel are five important elements in estimating potential demand for a new transport mode using stated choice experiments.

1. Introduction

In order to reduce air pollution and traffic congestion Bertrand Delanoë, the mayor of Paris since 2001, has pursued a policy limiting the use of the passenger car and promoting other transport modes that are more apt for use in city centres. The idea is to offer to the Parisians a large choice of modes of transport, and to facilitate those who want to respect the urban environmental capacity.

In order to do that, the Mairie de Paris has issued the following policy measures:

- Expanding the bus and taxi lanes, while reducing the amount of space available for the private car;
- Re-introducing the tramway, and extending its peripheral network;
- Increasing the parking cost;
- Enhancing the night bus service, and extending the operating hours of the underground on Friday and Saturday evenings;
- Developing a network of bicycle paths, and installing the Velib' bicycle system¹ in 2007;
- Increasing the space available for parking bicycles and motor cycles;
- Launching Autolib' in 2010, together with the communes surrounding the city of Paris.

Autolib' is a sort of electric car sharing system, which can be used for trips between 700 stations inside Paris and another 700 stations in the communes surrounding Paris. Unlike the existing car sharing systems Autolib' does not require its user to return the vehicle to the point where it was picked up. In order to use Autolib' one must be pre-registered. It is a transport system that has to be paid for, either with or without a subscription.

The operator for the installation and exploitation of Autolib' has been selected in a competitive tender process. A consortium consisting of four complementary companies, Avis, RATP, SNCF et Vinci Park, has been formed to respond to the tender. The consortium wanted to estimate the potential demand and revenue of Autolib' as accurately as possible. In order to do this it was necessary to:

- Define the tariff structure and levels for the service ;
- Research the maximum distance that would be accepted by the potential customers to collect (and drop) the vehicles ;
- Define the safety requirements for the vehicle stations as expected by the target group ;
- Define and quantify the types of clients that were interested most in the service.

Below we shall describe the methodology that we have used for this study, starting with the sorts of problems that need to be addressed when one wants to estimate the potential demand for an entirely new transport service. Then we shall explain the method that we have used, and the way in which we have used Stated Preference experiments to estimate demand. Finally we shall describe how we have developed a practical demand and revenue simulation tool for Paris, and what lessons we have learnt from this exercise.

¹ Velib' is a subscription based system of public bicycles, that can be used free of charge for up to 30 minutes, and rented for longer periods of use. Originally 7,500 bicycles were available from 750 stations, nowadays 18,000 bicycles are provided at 1,200 stations. In total some 22 million bicycle trips are made each year, and the number of subscriptions varies between 160,000 and 200,000.

2. The problem: estimating potential demand for a new transport mode

When estimating potential demand for a new mode, it is common practice to transfer as much as possible the knowledge that is already available for the existing modes. For the region of Paris, Ile-de-France, several transport data bases and demand models were available that could be used as a starting point. In this study, RATP's IMPACT 4 model has been used as such. IMPACT 4 uses discrete choice models for mode and destination choice, which are applied in a large sample enumeration system in connection with the large EGT² transport survey data base for the Ile-de-France region.

Given that Autolib' is essentially a car based mode of transport, one could simply extend the utility function structure that is available for the existing passenger cars, with the exception of the mode-specific constant: it is likely that the new transport mode will have a very different mode specific constant, given the very different nature of the Autolib' system, its characteristics and its availability. Therefore the key challenge is to estimate an updated alternative specific constant (ASC) for Autolib' for application in the logit demand model.

In principle this sounds like a fairly straightforward task, in practice there are many reasons why it is not so easily done. We mention three important complications.

A first complication is the fact that Autolib' is a transport mode that does not yet exist in reality. That forces us to use Stated Preference data to estimate the ASC, with all the usual caveats surrounding the use of that type of data for estimating real world market shares. In particular, how does one correct for possible differences in error and hence the scale of the model between the hypothetical choice context and the expected real world choice situation. And also the classical SP validity questions arises: how will we know that everybody who indicates that he/she will choose Autolib' in the survey will in fact also choose Autolib' in reality? It is well known that respondents in West European countries, and certainly in France, tend to overstate their willingness to switch in SP studies. And the scope to correct this effect by using for instance combined SP/RP estimation is very limited here.

A related issue is the fact that the Parisians have a very large number of alternative modes of transport from which they can choose. At the level of model application this means that there are many alternatives for which utilities have to be computed, at the level of data collection this means that the Stated Preference survey should ideally present (and vary) a large number of alternatives, variables and levels. In practice the limited size of the computer screen and more importantly the limited ability of respondents to process information forces us to artificially reduce the choice set. But that creates new problems that must be addressed in turn.

A second and more practical complication is the fact that the Autolib' system will require all users to be pre-registered, and possibly to pay a monthly subscription rate in addition to pay for its actual use. So the question is not one of just choosing Autolib' or not, but one of first registering and buying a subscription, and conditional on that the choice whether to use Autolib' or not for a specific journey. This poses a complication for the SP survey (how many experiments? How structured?) but also for the analysis (separate or joint analysis? Integration for application?).

A third and obvious complication is the spatial coverage of the system: Autolib' is only a relevant choice alternative when both origin and destination of the journey are near Autolib' stations, where a vehicle is available (origin) for pick up and where a space is available for parking (destination) the vehicle. The Ile-de-France region consists of three concentric components:

- Paris *intra-muros*, the central area within the Boulevard Périphérique;

² Enquete Globale des Transports, global transport survey

- the *Petit Couronne*, a series of about 80 communes consisting of the suburbs close to Paris;
- the *Grand Couronne*, a large number of communes further away from Paris.

At the time of the study about half of the communes of the Petit Couronne had signed the co-operation agreement to participate in the Autolib' project, but this number is expected to increase in the future. That means that the SP experiment needs to address the uncertainty in spatial coverage of the system.

In addition to these specific complications, we had to take into account the usual elements in choice modelling, such as the taste variation (heterogeneity) of potential users. This is a classical problem, certainly not unique for Autolib', but that does not make it less important or less difficult to deal with. The fact that this is a proposed new transport system adds to the uncertainty and possibly to the heterogeneity. Also the fact that the market share that can be achieved by the system has to remain low, due to reasons of limited availability of vehicles and stations, does not make it easy to deal with this issue.

3. Our approach: a combination of SP with EGT based simulation

In order to estimate potential demand for Autolib', despite the problems and complications mentioned in the previous section, we have developed a research approach consisting of the following key elements:

1. A large-scale SP and RP survey among 3,924 persons with driving license, living within the Ile-de-France region³;
2. The development of interrelated subscription and mode choice models;
3. A market potential simulation procedure using the Ile-de-France *Enquete Global de Transport* data base.

3.1 Choice experiments

We conducted a computerised survey among members of a large (>200,000) French internet panel. First we asked three screening questions: (1) age of 18 years or more, (2) possession of driving license, and (3) living in Ile de France) and questions to check the completion of the quota: (i) zone of residence, (ii) age class and (iii) gender. Target numbers were set for each segment. If a respondent belonged to a segment for which the target was not yet reached, he was forwarded to the main questionnaire. The structure of this questionnaire will be presented in section 4 in more detail.

3.2 Subscription and mode choice models

After completion of the survey and extensive data quality checking a data base was created for analysis (see section 5). This data was used to estimate a series of models, initially three separate models, later these different models have been integrated for application in the market potential simulator. In section 6 the analysis including the different component models and the combined models will be described further.

3.3 Market potential simulation

After estimation of the models a population simulator has been created to enable us to estimate the future market potential for Autolib' for the entire Ile-de-France region. We did this by applying the estimated models to a sample of about 80,000 travellers and trips, which was available from the EGT. The development and implementation of this market simulator will be described in section 7.

³ We have decided to exclude potential demand from visitors living outside Ile-de-France

4. Design of the Survey

4.1 Questionnaire

The key issue here was how to present the new Autolib' mode, together with the need to buy a subscription, and then to ask questions about its potential use for journeys where it could potentially be used. We did this in the following steps:

1. ask in-scope respondents about their vehicle possession and travel behaviour in general;
2. present the concept of Autolib'; first the overall concept was presented, then more detailed elements were described; after each characteristic the respondents were asked to rate the importance of that characteristic;
3. ask for the respondents' stated intentions to buy a subscription for Autolib' (choice experiment 1; will be elaborated in section 4.2.1)
4. ask respondents to choose between two different specifications of Autolib' (choice experiment 2; will be elaborated in section 4.2.2)
5. select and describe a journey that could have been made using Autolib' (RP); details asked included origin, destination, distance, group size, presence of luggage, animals, type of day, time period, weather type, transport modes used, main mode, access mode(s), egress mode(s), journey duration by components, cost of journey, part of journey cost paid by others, frequency of journey, respondents were asked for their chosen mode and also for their second best mode, as an approximation of their choice set;
6. ask respondents to indicate their mode choice including the Autolib' option (choice experiment 3; will be elaborated in section 4.2.3)
7. ask respondent for their expected frequency of use of Autolib' (for a given Autolib' specification and for a series of journey purposes) and the maximum monthly budget that they would be prepared to spend on use of Autolib' (over and above the specified Autolib' subscription cost).
8. ask respondent to rate a list of typical statements for the typological classification (will be elaborated in section 4.3)
9. Finally some socio-demographic questions were asked.

4.2 Detailed description of three choice experiments

4.2.1 Choice experiment 1: Subscription to Autolib'

The first experiment was a series of stated intention questions. The main objective of this experiment was to obtain an indication of the respondents' absolute likelihood to buy a subscription for Autolib' for different specifications of the system (subscription and usage). This part of the survey was also intended as an introduction to the choice questions that would follow later in the questionnaire. In total four choice questions were asked to each respondent, using the following attributes and levels:

- Subscription cost (4 different levels)
- Usage cost (4 different levels)
- Maximum distance to an Autolib' station (4 different levels)
- Level of security at stations (4 different levels)

Respondents were asked to express their strength of intention using the following answer categories:

- Yes, certainly
- Yes, probably
- No, probably not
- No, certainly not

Figure 1 presents an example of a choice of the first experiment.

choix 1/4 19%

	Offre Autolib'
Coût d'abonnement	20€ par mois
Coût d'utilisation	0,10€ /minute (= 3€ / demi-heure)
Distance maximale jusqu'à une station	3 minutes de marche (200 mètres environ)
Présence en stations	Les stations sont surveillées en permanence par des caméras et un employé peut être sur place en moins de 15 minutes

Pensez-vous que vous souscririez à Autolib' ?

Oui, certainement
 Oui, probablement
 Non, probablement pas
 Non, certainement pas

Figure 1: Example of first choice experiment

For the scenarios offered about one-sixth of all respondents indicated “certainly yes”, and about one third “probably yes”. Each respondent gave four answers, and it turned out that a large majority of the respondents were actually trading (86% varied their responses with different scenarios).

4.2.2 Choice experiment 2: Subscription to Autolib'

The second experiment was a series of stated choice questions. The main objective of this experiment was to estimate the relative importance of each of the characteristics and levels in the choice process of Autolib' subscription. For this the respondent was asked to choose between pairs of different specifications of Autolib' (subscription and usage). using the same attributes and levels as used in choice experiment 1. In total five stated choice questions were asked to each respondent, of which on question had a dominating alternative (i.e. all attributes of one alternative were better than or equal to the other alternative).

We used a classical orthogonal design, as we had no prior information about the utility values to expect and we wanted to be free to use non-linear specifications if we felt we needed to.

Respondents were asked to indicate their choice. Figure 2 presents an example of a choice of the second experiment.

Choix entre formules Autolib' - Choix 1 / 5 22%

Laquelle de ces offres préférez-vous ?

	Offre Autolib' A	Offre Autolib' B
Coût d'abonnement	20€ par mois	10€ par mois
Coût d'utilisation	0,15€ /minute (= 4.50€ / demi-heure)	0,20€ /minute (= 6€ / demi-heure)
Distance maximale jusqu'à une station	3 minutes de marche (200 mètres environ)	12 minutes de marche (700 mètres environ)
Présence en stations	Pas de présence en station (pas de caméra, pas de possibilité d'intervention)	Pas de présence en station (pas de caméra, pas de possibilité d'intervention)

Offre Autolib' A Offre Autolib' B

Figure 2: Example of second choice experiment

From the responses it was clear that the subscription cost was very important here, and in fact dominated the choice for over half the respondents. But there was also substantial trading, for all four variables included in the experiment.

4.2.3 Choice experiment 3: Mode choice

In order to specify the context for a realistic mode choice experiment the survey contained questions about a recent journey that could have been made using Autolib'. The questions proceeded in the following steps:

- Can you think of a recent journey that could have been made using Autolib'? 81% answered yes, and this trip was used for SP3
- Those who answered no were asked: can you think of a possible future trip for which you could use Autolib'? Another 13% answered yes, and this trip was used for SP3
- The remaining respondents were asked to describe a trip at night (no public transport), or with heavy luggage (car needed) or a trip for which their usual mode was not available. 5% could not think of such a trip, and no SP questions were asked.

In this stated choice experiment all information collected in the previous SP experiments and in the RP description of the existing journey is used to describe how the Autolib' alternative would compare with the existing chosen mode and the second best mode. By varying the conditions for all three modes nine choices were presented to each respondent. All three mode alternatives were presented using the following attributes:

- Total travel time door-to-door, which was split into
 - o Acces time from origin to each transport mode
 - o In vehicle travel time
 - o Egress time from each transport mode to destination
- Total cost of travel, which was split into
 - o Out of pocket travel cost
 - o Parking cost (if relevant)

Respondents were asked to indicate which mode they would choose.

Choix 1/9 66%

Supposons que vous soyez inscrit à Autolib'
(avec une inscription gratuite),
quel mode choisiriez-vous pour effectuer le déplacement que vous venez de me décrire ?

	Voiture particulière - conducteur	Vélib	Autolib'
Durée pour vous rendre jusqu'à votre mode de déplacement	5 minutes	2 minutes	3 minutes de marche (200 mètres)
Durée de déplacement	25 minutes	40 minutes	18 minutes
Durée pour aller de votre mode de déplacement jusqu'à votre destination finale	10 minutes	1 minutes	3 minutes de marche (200 mètres)
Durée totale du trajet	40 minutes	43 minutes	24 minutes
Coût d'utilisation	7.50€	0.00€	9.60€
Coût de stationnement	3.00€	-	-
Coût total du déplacement	10.50€	0.00€	9.60€

Voiture particulière - conducteur
 Vélib
 Autolib'

Figure 3: Example of third choice experiment

To stimulate the trading between the choice options, we usually made the characteristics (travel time and cost) of the used alternative worse than their current values. As a result, in 44% of the cases the used mode was chosen, in 14% the second best mode and in 42% Autolib'. Most respondents (65%) were actually trading (i.e. they did not always choose the same answer for all mode choice questions).

4.3 Attitudes based segmentation

In order to inform the marketing and communication strategies for Autolib' the study was required to produce a typology of potential customers based upon their attitudes underlying their modal choices. On the one hand, this typology was needed to understand the motivations and possible reservations of each market segment with respect to Autolib', and to identify the levers that would trigger their participation. On the other hand the typology would make it possible to characterise them in order to assess their market potential.

In order to do this, we have used the results of qualitative research that has taken place before our study, which had produced a first typology for the same objectives. Based upon this research, and other previous qualitative research we have established a list of 30 attitude items, typically statements to which the respondent had to express his or her agreement or disagreement. These had been selected based upon their ability to explain mode choice behaviour and the underlying factors. Subsequently a correlation analysis allowed us to reduce the list slightly, to 27 items.

Our list of items included questions concerning the attachment to the passenger car, the interest in and ease of using public transport, the time constraints, the orientation towards innovation, the orientation towards public services in general, the importance of ecology and the reduction of the use of the passenger car in cities, travel behaviour in general, the anticipation in transport mode choice, and attachment or not of private ownership. The respondent was asked to rate each item along a scale from 0 to 10.

For example :

- "Without a car my daily life would be hard to organise"
- "As soon as something goes not as I planned it I easily get upset"

- “I agree with everything that enables people to reduce their use of the car in Ile de France”
- “I refuse to step into a taxi, or in the car of somebody else if the vehicle is not really clean”
- “I love to be able to use something without actually being the owner (for instance like a rented car, or Velib’)”
- “If I use something that does not belong to myself I am always worried that it does not function properly”
- “Whenever there is something new I like to try it”
- “I know public transport inside out”
- “When I drive by car in Paris or in the Paris region I can easily find my way”
- “Where I live it is hard to travel by public transport”

5. Survey fieldwork and data base

In total 3,926 respondents completed the questionnaire before end June, the deadline for the fieldwork. Before proceeding with further analysis, we first conducted five tests to verify the quality of the obtained data:

1. In the SP survey we included one dominant question, where Autolib’ option B was cheaper, and at shorter distance than Autolib’ option A, while the safety situation was equal. Respondents who chose the non-intuitive (dominated) alternative (i.e. 154 respondents, or 3.9%) were removed from the data base used for analysis.
2. At the end of the questionnaire we asked the respondents whether they had thought that the questionnaire had been clear. In total 152 respondents indicated that they thought it was not clear. These respondents have been removed from the data base used for analysis.
3. We also asked the respondents whether they thought the questionnaire had been too long. 85 respondents thought it had been too long, and indicated that their answers had been less precise towards the end of the survey. These respondents have been eliminated.
4. 76 of the remaining respondents completed the survey in 8 minutes or less. We believe they did not really pay the necessary attention to all questions, and therefore we removed them from the sample used for analysis.
5. We tested the variation in answers to the attitude questions. When respondents gave an identical answer to each of the 27 statements they were asked to review, we think they did not pay enough attention to the survey and consequently we have eliminated them from our sample for analysis.

After these five checks we had removed 478 respondents (about 12% of the original sample), and our cleaned up sample for analysis consisted of 3,448 respondents. Table 1 shows how these respondents were distributed over the gender/age/residence location segments that were used in this study.

		Paris	Petite Couronne		Grande Couronne	Total
			<i>zone Autolib</i>	<i>non-Autolib</i>		
18 - 25 years	Homme	68	54	13	23	158
	Femme	97	137	29	18	281
26 - 39 years	Homme	180	253	55	114	602
	Femme	232	363	71	67	733
40 - 59 years	Homme	189	199	59	111	558
	Femme	320	340	71	58	789
60+ years	Homme	53	60	8	41	162
	Femme	50	60	13	42	165
Total		1189	1466	319	474	3448

Table 1: distribution of final sample over the gender/age/residence location segments

6. Analysis

6.1 Attitude questions

The typology analysis subdivides the survey sample into a predefined number of segments (“types”) in such a way that the differences *between* the types are as large as possible, but that the differences *within* each type are as small a possible. The typology analysis starts from the individual observations, and classifies them according to their similarity in responses to the so-called “active” variables. For this we have used *baysean network* software from BayesiaLab⁴. Unlike the usual cluster methods this approach allowed us to obtain the best compromise between the “purity” of the clusters and the number of clusters.

The software has retained a segmentation of 4 types, with a reported mean “purity” of 93%. This can be interpreted as a very robust classification, based upon strong links between the variables. Two types of outputs are obtained from the typology:

- The contribution of the active variables to the creation of the types. This is an expression of the importance of each item for the typology;
- The description of the types, based upon the differences between their dominant tendencies. For the further description and interpretation of the types we use also the other questions (the “passive” variables).

The typology analysis has enabled us to identify four strategic target groups for Autolib’, two because of their high interest in Autolib’ and a fourth one less interested in Autolib’ but interesting in terms of market potential for its sheer volume:

1. The “nomads” with an orientation towards personal efficiency (13%); they are not dependent on the car, they use public transport with ease and pleasure, they are interested in reducing the use of their car and increasing the use of slow modes in the public interest, and they use all sorts of different modes themselves depending on what is most efficient for them; these are often women living just outside Paris, 40-60 years old, with adapted working hours, living with partner and children;
2. The “nomads” with an orientation towards collective efficiency (23%); they do not (want to) own a car, the notion of ownership is not important for them, they easily use public transport, they are interested in reducing the use of the car and increasing the use of slow modes in general, and they use all sorts of different modes themselves depending on which mode is most efficient for them; these are often women living inside Paris, studying or unemployed, living alone without children, and having a low income;

⁴ <http://www.bayesia.com/en/products/bayesialab.php>

3. The “flaneurs” (26%); they prefer the car or taxi over public transport, they see the public transport as not adapted to their needs and social status, the public interest is relevant for them only if it suits their personal needs, and their mobility is limited; these are often men living just outside Paris, often working in the private sector, with some control over their working hours, and living with a partner and children;
4. The commuters (38%); they view the car as an instrument that brings them freedom and flexibility, they do not enjoy using public transport, they are not very sensitive to the common interest, and they tend to use always the same mode; these are often people living further away from Paris, working in the public sector, with no control over their working hours, who start their work early in the morning.

For each segment the typology has enabled us to understand the motivations and reservations with regard to Autolib’, and to identify tariff structures and policies that enhance these potential users’ interest. But also this has enabled us to advice on the desired characteristics of the Autolib’ product. During the model estimation stages of the project we have estimated separate models for each of these four attitude segments, to take account of their differences in preferences and likelihood to use Autolib’.

6.2 Model estimates

First, we developed models for each experiment separately. In order to use the typology information in the application tool as well, we estimated a model that explained the typology class of a respondent from other socio-economic characteristics.

6.2.1 Choice experiment 1: Subscription to Autolib’

We estimated base utility models that could be used to describes the likeliness that a respondent would choose any of the four possible answers (“certainly yes, probably yes, probably no, certainly no”). The resulting coefficients were intuitive: the lower the price for the subscription or for the usage, the more likely a person was to buy a subscription. If the density of Autolib’ stations were higher and/or if the security was better, the respondent was more likely to buy a subscription. We only were not able to distinguish a preference for either the security level “The sites are under permanent surveillance by cameras and an employee can be present within 5 minutes” or for the level “The sites are under permanent surveillance by cameras and an employee can be present within 15 minutes”.

In addition, the models showed a number of significant socio-economic interaction terms:

- the older the person is, the less likely he is to buy a subscription
- living further away from Paris (Grand Couronne, and Petit-Couronne non-Autolib’ zones) makes it less likely that a person would buy a subscription
- if the respondent owned (at least) 1 car, it was less likely that he would buy an Autolib’ subscription. owning a motor, scooter, or bicycle makes it more likely that he would buy a subscription. Also PT subscription holders are more likely to buy an Autolib’ subscription.
- Females are less likely to buy a subscription. Persons in households with children of age 10 or below are more likely to buy a subscription.
- Respondents with an income of 11,000 euro and more are more likely to buy a subscription

6.2.2 Choice experiment 2: Subscription to Autolib’

Standard utility models were developed to describe the choices made by respondents. These models contained linear terms for subscription and usage costs. and a linear term for the

maximum distance to the nearest Autolib' station. Tests have been done that confirmed that the costs and distance attributes can indeed best be described with a linear function (i.e. a constant disutility per euro / per minute).

Tests for personal characteristics (income, age, gender and a combination of age and gender), have been made for each of the above described coefficient, in order to determine differences in respondents' behaviour. Particularly the following effects have been found:

- Segment of population with an income equal or higher than 16,500 euros/year have a dislike for higher subscription costs;
- female older than 40 years are less sensitive to a higher subscription costs, compare to the base level;
- For cost of utilization, different behaviour patterns have been found for combinations of age and gender. The older the person, the less sensitive to higher cost of utilization is.
- Females are more sensitive to a higher degree of surveillance of the station.

6.2.3 Game 3 mode choice (conditional on subscription)

Sixteen different modes were included in the model (Autolib, car driver, motor driver, car/motor passenger, rental car, shared car, taxi, metro, RER (i.e. the local suburban train), train, tram, bus, night bus, bike, Velib' and walking.

A standard mode-choice utility model was estimated. This used linear terms for both the Autolib-subscription cost and the (general) usage cost. A separate parking cost coefficient was estimated which turned out to be about twice the normal usage cost coefficient. All other coefficients were purpose-specific. Five purpose groups were distinguished:

- home-work & home-education;
- home-shopping & home-personal business
- home-leisure
- business (both home-business & non-home-business)

For each purpose, five mode-specific time coefficients were included in the model (car driver (incl. Autolib'), passenger, bus, metro, slow mode). Separate coefficients on access-time (for public transport modes) and on maximum-distance-to-the-nearest-station (Autolib') were estimated. Finally, purpose/mode-specific constants were estimated. For Autolib', this mode-specific constant was also specific for time-of-day, area, age, luggage-situation and typology class. Given the large number of respondents that each answered nine questions, we were able to estimate significant coefficients and constants in almost every case. Insignificant coefficients were combined with others to get significant estimates.

6.2.5. Typology model

Four different typology classes were identified:

- the “nomads” with an orientation towards personal efficiency
- the “nomads” with an orientation towards collective efficiency
- the “flaneurs”
- the commuters

Because of the definitions of the typology classes, a person has a 100% likeliness of belonging to typology class 2 “Nomades tendance efficacité collective” if the household to which he belongs does not own a car. For the other three classes, a utility function describes the probability of belonging to each class. These utility functions are a simple sum of constants dependent on socio-economic status that apply to the person. The model estimates show that

- people that do not live in Paris are more likely to belong to the “flaneurs” and especially to the commuters
- females are more likely to belong any of the two “nomads” classes
- people between 31 and 50 are more likely to belong to the “flaneurs” or to the commuters
- if there is more than one adult in the household, the probability that all adults are commuters is reduced
- if there is any child below 10, the probability of being a commuter is increased; if any child is between 11 and 17 this probability is reduced (this is probably an interaction with other constants that apply for those respondents)
- people with a public transport subscription are more likely to belong to a “nomads” class
- people in the lowest income class are less likely to be a “flaneur” and are more likely to be a commuter

7. Development and implementation of the demand forecasting tool

7.1 Structure of the demand forecasting tool

In order to forecast demand, a market simulator was developed. This simulator is a Visual Basic for Applications (VBA) tool that runs under Microsoft Excel so that it can easily be run by the client as well. The basis structure of the tool is showed in Figure 4.

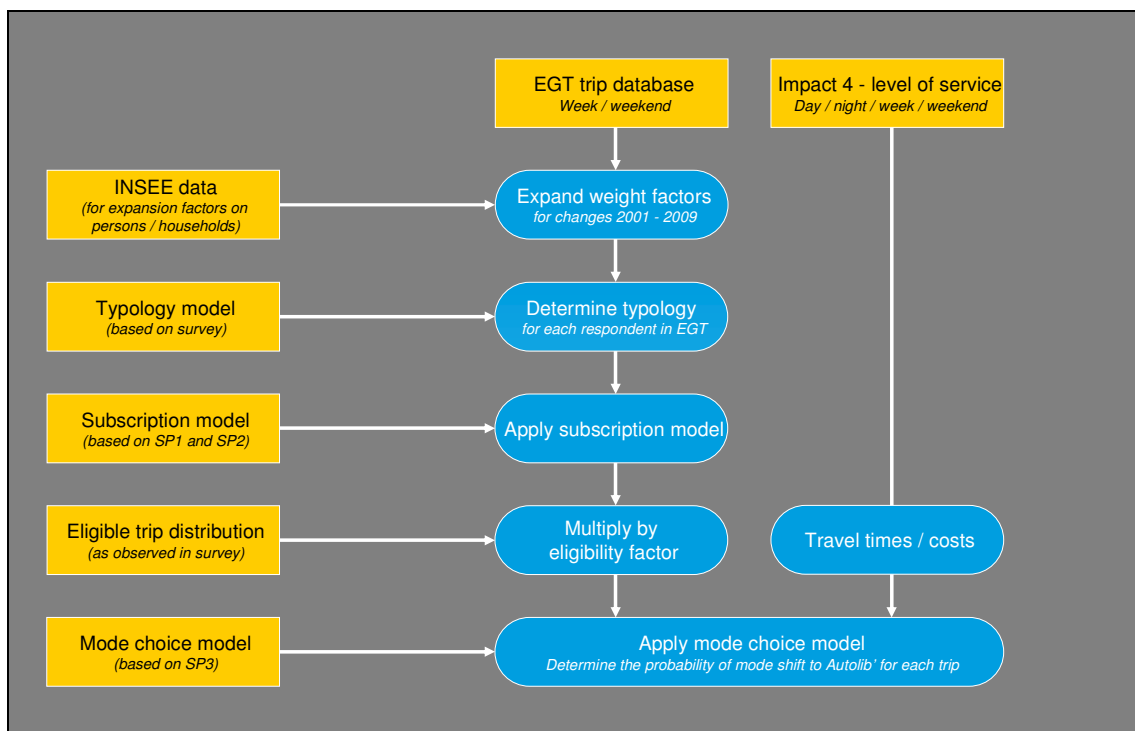


Figure 4: Structure of the demand simulation tool

Essentially, it is a sample enumeration tool based on the EGT database. For each trip that is recorded in this database, the tool calculates:

1. an expansion factor to correct the record weight factor as present in the EGT database for any development in the period between 2001 and 2009. The EGT was compiled in 2001, so the weight factors of the records need to be changed for any change in population between 2001 and 2009. For this, we used data from INSEE. For each gender (male/female), each age category (0-19, 20-39, 40-59, 60-74, 75+), each department in the Ile-de-France, we determined the growth of the population between 2001 and 2007 and extrapolated this until 2009. The EGT weight factors were multiplied with these extrapolated growth factors;
2. the probability that the person that makes the trip belongs to any of the four typology classes;
3. for each typology class, the probability that the person would buy an Autolib' subscription, based on the results from choice experiments 1 and 2 (this is elaborated in section 7.2);
4. the probability that a trip is an "eligible trip" and Autolib' might be considered as a plausible mode for it. This probability is related to the fact that choice experiment 3 is not about a random or a most recent trip, but about a trip "that could have been made using Autolib". Corrections for other biases are included in this probability as well. This is elaborated in section 7.3);
5. the probability that Autolib' is indeed chosen for this trip, given the travel times and costs for Autolib' and other modes, as taken from the Impact 4 travel demand model.

7.2 Calculation of the probability of buying a subscription

Both choice experiments 1 and 2 were about the choice of buying an Autolib' subscription, but neither could be used in the model directly. A new joint model had to be estimated based on the choices from both experiments. To get a prudent estimation and to correct for optimism biases, we only considered the answer "Yes, certainly" to be a clear choice to by an Autolib' subscription. All other answers were interpreted as "not buying an Autolib' subscription" As a result, we used the tree structure as displayed in figure 5 for our joint model.

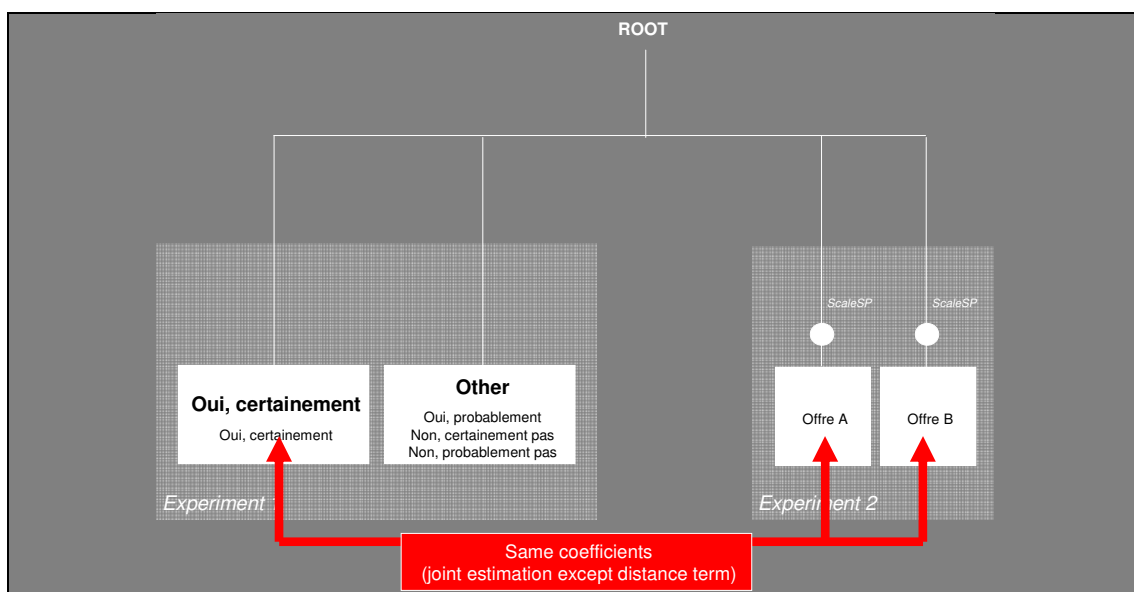


Figure 5: tree structure for joint model of data from choice experiments 1 and 2

We estimated general models, but also separate models for each of the four typology classes (for compatibility reasons, we kept any insignificant coefficients in these typology class models if that coefficient was significant in the main model as described in sections 6.2.1 and 6.2.2).

7.3 Eligibility factor

This factors consists of three parts.

1. We reduce the total probability with 4%. This is the drop-out rate of respondents during the survey after the Autolib' concept has been introduced. It is likely that these respondents are not average respondents, but they are probably not interested in this concept.
2. We reduce the total probability with 56%. This is based on the percentage of respondents that
 - in the choice experiment 1 regularly indicated that they would by a subscription, but when asked to describe a trip that they would possibly make with Autolib, could not think of such a trip;
 - could think of a trip that they would possibly make with Autolib', but in choice experiment 3 never choose the Autolib' option, even not when the time and costs for the current alternative were worsened;
 - would choose Autolib' at least once per month, but indicated that they expected to spend a monthly budget of less than 5 euro, in which case it is not very likely to buy an subscription.
3. We multiply the probability with an eligibility factor, which corresponds to the probability that the trip was indicated when the respondent was asked to describe a trip that he would possibly make with Autolib'. In order to determine this eligibility factor, we segmented all trips in the EGT database and in the RP database from our survey by the following factors:
 - Purpose
 - Short/long trips
 - Regular / infrequent trips
 - Current mode
 - Week / weekend day
 - Morning/afternoon/evening/night
 - Group size
 - Carrying a heavy object / animal

We noted that taxi trips, public transport trips for business purposes during day-time and public transport trips during night times occurred relatively frequent in our survey. Since these are typical trips that could be made with Autolib', we assumed that these trips are 100% eligible. We derived eligibility factors for all other trips by comparing the relative frequency of these trips in our own survey compared to these reference trips. Given the low number of trips compared to the large number of combinations of the trip characteristics mentioned above, we used an IPF (iterative proportional fitting) technique for this.

From this analysis, it became clear that for instance weekend public transport trips are far more likely to be replaced by Autolib' than trips that are currently made by walking, or as a car/motor passenger. This is all intuitively very acceptable.

7.3 Tool characteristics

The final product is a user-friendly tool where he can select a price level for the subscription, usage, for the density of the Autolib' stations and the security level. He is also able to select which out-of-Paris communes are included in the Autolib' scheme. Trips are only eligible for Autolib' if the respondent lives in a commune in which Autolib' has been introduced and both the origin and destination are in areas where Autolib' stations are present. Figure 6 shows the user interaction screen. The typical run time for the model is about two minutes.

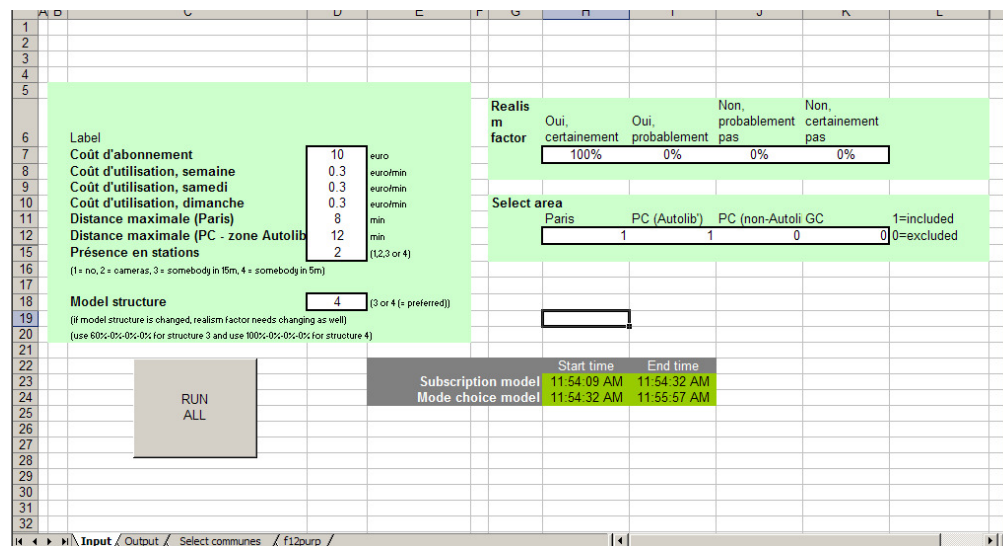


Figure 6: user interaction screen of the demand forecasting tool

The tool reports the number in percentage of subscriptions that are bought by persons and households, the number of trips made by Autolib' (split by current mode, trip purpose, period of travel, area, typology class) and some important indicators such as the average number of monthly trips made with Autolib' by subscription holders, monthly budget spent by Autolib' holders, average number of trips made per Autolib' vehicle per day, total revenue (from subscriptions and from usage. Figure 7 shows the revenue as a function of subscription costs (for reasons of confidentiality, no numbers have been indicated along the axes). For this situation, maximum revenue is obtained for very low subscription costs. Other scenarios produce different results.

These last indicators produced plausible numbers:

- the daily number of trips made per vehicle were such that enough time is allowed for battery charging;
- the monthly number of trips made per subscription holder is such that it is indeed worth buying a subscription for him;
- the monthly budget spent per subscription holder is in agreement with the average budget that respondents indicated that they were likely to spent on this;

and these indicators were very important for the acceptance of the results.

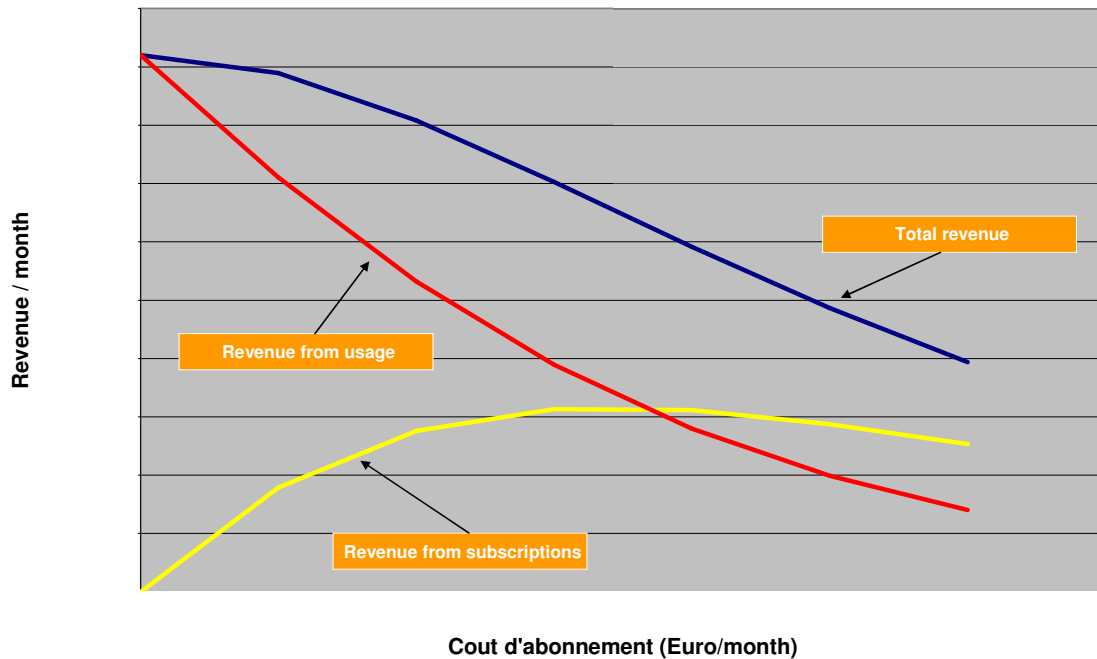


Figure 7: results from the forecasting tool: revenue as a function of subscription cost for certain assumptions of the other parameters

For the interpretation of these results, it has to be taken into account that all market potential forecasts for any non-existing product are uncertain. Especially in this project, we have used the prudent assumption that only respondents that indicated that they are certainly buying a subscription will in reality indeed buy such a description. Furthermore, a number of factors affecting demand have not been taken into account, such as

- image effects (marketing)
- technical functioning (and timing)
- availability of vehicles (the tool assumes perfect supply both in time and space)

and these will have a large impact on demand.

8. Discussion and conclusions

In this paper we have described how we developed a simulation model to predict the market potential for a new transport mode called Autolib'. We used Stated Preference experiments in order to obtain the necessary insights for this newly proposed transport mode. Estimating market share forecasts from SP is never an easy task, but in this case there were some extra complications: (1) the Autolib' mode was entirely new, and added to an already very large choice set of transport modes available to the Parisians, (2) the Autolib' mode requires pre-registration and subscription, so there are two interrelated choices: the choice to subscribe and mode choice, and (3) the spatial coverage of the Autolib' system was unknown at the time of the survey, and subject to variation over time (it was expected that more communities would join the initiative in the future).

We have tried to address and solve the many theoretical and practical problems that arose during this project, and the different steps have been summarised in the previous sections. We want to conclude this paper by presenting some concluding remarks and discussion on five

points that we feel were important elements in this study and in studies with a similar nature. We hope these are of some value for other researchers.

First we want to highlight the importance to build in a large number of *checks and balances* in a heavily SP based project like this. We summarise here some of the important tests that we have applied:

- five tests to verify the basic data quality, and rejection of suspect data where we felt this was necessary (correct answer to dominant choice question, assessment of clarity of the questions, assessment of length of the questionnaire, time taken to complete the questionnaire, and sufficient variation in answers to attitude questions);
- the usual evaluation of the reasonableness of the estimation results, including the obtained values-of-travel time (essentially comparison to other published values for similar population, and possible explanation for observed systematic differences);
- three tests for coherence in the results of the simulation (check that predicted number of trips per Autolib' vehicle per day falls within acceptable interval, comparison of the predicted monthly cost of Autolib not too different from stated budget, comparison between predicted monthly frequency of use and subscription choice).

We highlight that it is very important in a study like this to be very critical and aware of the many possible problems that may arise during the stages of data collection, survey analysis, model estimation and application, and to build in all checks and balances deemed necessary.

Second we point at the importance of a carefully designed, custom-made series of *interrelated SP experiments* to obtain the necessary information. In our case we needed to estimate the probability to buy a subscription and (conditional on that) the probability to use Autolib' for a given journey. The mutual dependence of subscription and mode choice was clearly a complication, particularly as our client wanted to test price sensitivity for both subscription and usage together with the sensitivity to various more subtle service specification elements. By simultaneously analysing the SP data in a suitable tree-structure a satisfactory model has been obtained. However, there is one element where we feel we are not entirely certain about the results: we have had to make a judgement on the interpretation of the SP experiment dealing with the level of absolute demand: how likely are you to buy a subscription for Autolib'? After very careful analysis we have decided to be conservative in our interpretation, and only include the "would certainly buy subscription" answers as valid statements of choice. In the end this was based more on judgement rather than on hard empirical evidence.

Third we think that the use of *attitudinal information* to help explain the heterogeneity in preferences of different potential travellers has been of benefit in this study. And this benefit can possibly be further increased in the future. In our project we have used the attitudinal information in a fairly traditional way, by inferring attitude-based market segments and by estimating separate models for these segments. In fact we have also estimated attitude-segment membership models, in order to be able to apply the models to the EGT data base for which no attitudinal information was available.

Recently examples of more advanced approaches by several authors have become available, where attitude-based latent classes and the choice models themselves are estimated simultaneously. We feel that this is very promising direction that deserves to be further explored in the future. After a period where often mixed logit models were fitted to *describe* unobserved heterogeneity we welcome the return of attempts to *explain* as much possible of the heterogeneity that appears to exist in reality.

Fourth we mention the *identification and where necessary the correction of biases* that (may have) arisen during various stages of the work. In this project we have

- six tests for biases during the survey procedure, and correction where necessary (drop out at opening of the survey, bias due to license holding, bias due to area and age/gender target quota, bias due to drop-outs during two different stages of the

questionnaire, bias due to data quality exclusions; where possible we have made rational assumptions, and where possible we tested model estimations with and without the excluded respondents; in the end we found remarkably small differences, and applied only one simple correction for the EGT application sample: an overall 3.9% of all people are not interested in Autolib' and will not buy a subscription)

- the necessary weighing in the expansion of the “eligible trips” to all trips made within the EGT survey (the eligible trips where Autolib' could be used are not a random sample of all reported trips in the EGT data base; so we have derived weighing factors to correct for this in the model application)
- correction for the size of the choice set in model application (rather than simulating a binary choice between chosen mode and Autolib', we used three alternatives: chosen mode, second best mode, and Autolib' both in the SP mode choice questions and in the application model; ideally the application model should have included *all* available modes, but we have experienced difficulties with the technique recommended by Daly and Rohr (1998), and we feel that our three modes choice was a workable approximation.

Fifth we have to acknowledge that we *cannot give any guarantees* that the market potential forecasts produced by our simulator will in fact become reality in the future. There are many factors which influence the realisation of real demand, including:

- The interpretation of the SP responses: our general impression is that there was some optimism bias in the results of the survey, and therefore we chose a rather prudent hypothesis concerning the actual likelihood to subscribe;
- The marketing of the new product, the publicity and image effects, the technical functioning (particularly in the early stages), the economic climate during the introduction, and perhaps most importantly: the availability of enough Autolib' vehicles where needed, in space and time.

References

Daly, A.J. and C. Rohr (1998) Forecasting Demand for New Travel Alternatives. In: T Gärling, T Laitila, K Westin (ed.) Theoretical Foundation for Travel Choice Modelling, Pergamon.