ANTONIN: UPDATING AND COMPARING A TRANSPORT MODEL FOR THE PARIS REGION

Jan Gerrit Tuinenga and Marits Pieters
RAND Europe
Laurence Debrincat
Syndicat des Transports d'Ile de France

1. THE CONTEXT

1.1 Public transport network in Ile de France: key features

The Ile de France region holds approximately 11 million inhabitants (20 % of France's population in a territory which represents only 2 % of the total country) and 5 million jobs. Paris has always concentrated an important part of the inhabitants and jobs in the region but new developments in urbanisation have been concentrated in the suburbs since 30 years.

Public transport networks play a major role in Ile de France. They offer a wide variety of lines and modes (buses, tramway, metro and railway lines). State-owned and private companies operate the public transport services in the region. The public companies, RATP and SNCF, carry more than 90 % of public transport travellers (RATP operates subway lines, regional express network lines and busses; SNCF Ile-de-France operates suburban rail and regional express network lines). 80 private companies operate bus services in the suburbs.

STIF (Syndicat des Transports d'Ile de France) is the public transport authority in the Ile-de-France region (Paris region). It co-ordinates the activities of all operators which means particularly that it defines fares, allows the creation of new infrastructures, and set the service level for the different public transport lines.

1.2 Mobility in Ile de France: main results from the Enquêtes Globales de Transport

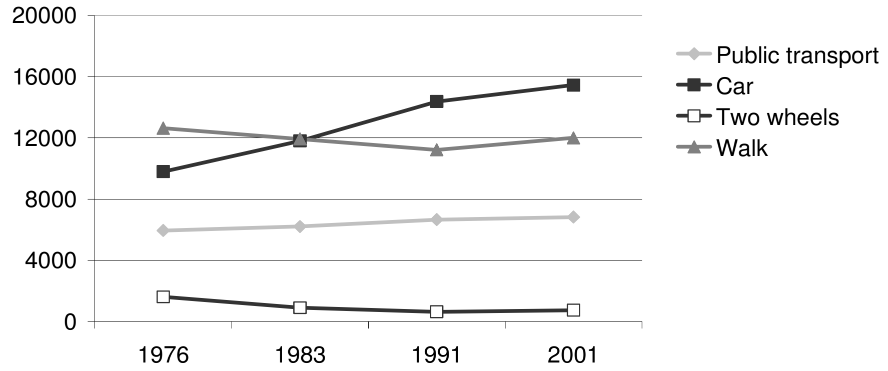
The *Enquêtes Globales de Transport* (EGT) are household transport surveys in the Ile de France region. EGT surveys were conducted in 1976, 1983, 1991, 1997 and 2001. In 2001, some 10,500 households were interviewed, which means that some 25,000 persons (above six years old) described their travel behaviour. Some of the main results of these surveys are summarised below.

The individual mobility of Ile de France inhabitants has remained unchanged since 1976, equal to 3.5 trips per person and per day. However, this stability is only true on average and results are the sum of important evolutions in modal choice, trips purposes and origins and destinations within the region linked to the changes in urbanisation and way of life.

Multiplied by the number of inhabitants in Ile de France, the individual mobility leads in 2001 to 35 million trips per day (cf. Figure 1), one third of which are walk trips. Consequently, around 23 millions trips are motorised trips, mainly car trips (67 % of the motorised trips).

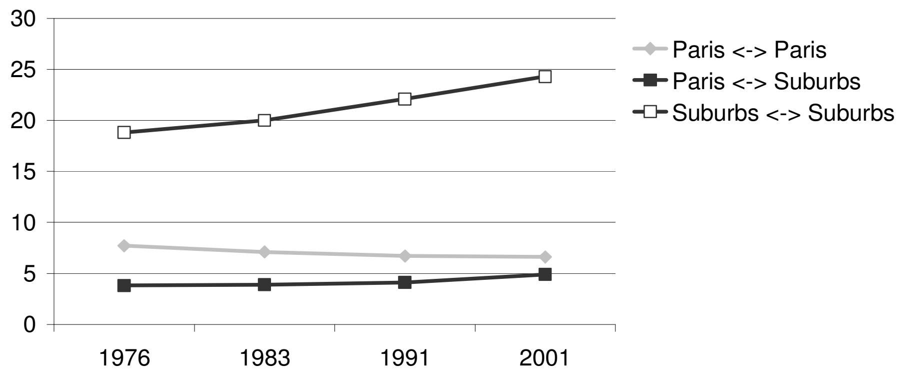
Since 1976, the use of car has increased considerably in Ile de France. However, between 1991 and 2001, this increase slowed down: just below +1 % per year compared to +2.6 % per year between 1976 and 1991.

Figure 1 – Number of trips per mode (trips per day in thousands) in Ile de France



Between 1991 and 2001, the number of trips with an origin and/or a destination in Paris has remained unchanged whilst it was growing steadily in the previous periods. In the same time, trips within the suburbs have increased substantially. These changes are of course the consequences of the emergence of a polycentric structure of the region, leading to an increased number of inhabitants and jobs outside Paris.

Figure 2 – Evolution of the origin and destination of trips in Ile de France (million per day)



Another major evolution concerns the increase of trips for leisure, shopping and personal purposes. Home to work or home to education trips remain stable in absolute value but decrease in proportion of the total.

1.3 History of STIF's transport model

In the early 90's, STIF decided to develop its own passenger trip forecasting system in order to improve its expertise. The main objective was to be able to have its own judgement about transport policies or new public transport schemes in order to be independent of transport operators or of other public bodies in Ile de France.

The model was designed to be able to produce strategic tests on transport policies, analyses of the evolution of mobility and traffic due to evolution of the population main characteristics and traffic forecasts on new public transport lines.

Hague Consulting Group (predecessor of Rand Europe) developed the whole framework in the 90's. Since the end of 90's, STIF achieves itself all the studies conducted with it.

Since its conception, ANTONIN (ANalyse des Transports et de l'Organisation de Nouvelles INfrastructures) has produced numerous traffic forecasts for projects concerning all type of public transport modes (suburban railway lines, metro lines, tramway lines, bus on dedicated lines) and concerning the creation of new lines or the enhancement of existing ones.

2. ANTONIN: MAIN CHARACTERISTICS OF THE MODEL SYSTEM

The structure of ANTONIN is based on that of widely accepted disaggregated transport models such as the Dutch National Model System. It is built under a disaggregate modelling framework using multinomial logit models to predict the various choices made by travellers in the Ile de France region. The data used in the model and the main points of its structure are described below.

2.1 The data

Zone system and data

ANTONIN was designed to be able to give results at different geographical scales, from the whole region to the local level. To achieve this goal, the region is split in 984 zones. Zones were defined to be pertinent with respect to the transport networks and the density of population and jobs. Paris comprises 300 zones because of its high number of metro stations, inhabitants and jobs. In the suburbs, major cities are split in a few zones whereas large zones are defined in rural areas.

For each zone, the model uses different variables to describe the population (number of inhabitants in total, by age, by sex…), employment (number of jobs), the number of places of students and parking costs. The data used to estimate the models were mainly obtained from the 1990 census. Forecasts are done using the results of last census (1999) or long term forecasts (for the year 2015).

Transport networks

ANTONIN is a multimodal modelling system, which means that both road and public transport networks are described in the model.

The public transport network in Ile de France offers a wide variety of modes and services. This is one of the reasons of the important use of public transport for trips within the region, especially in the dense part of the region (Paris and its near suburbs). The consequence of this complexity is that to go from an origin to a destination there can be many paths available to the traveller. Moreover, travellers can choose to walk or to drive to the first station/stop of their public transport trip, which adds even more complexity.

It was therefore chosen to use in the modelling 13 different travel modes instead of a basic choice between public transport, car and slow modes (cf. Figure 3).

Public transport networks are described for the morning peak and off peak hours. Rail, metro and bus services in Paris and its near suburbs are fully described. The description of bus services in the outer suburbs is simplified as well as the fare system which considers only full price tickets and monthly passes.

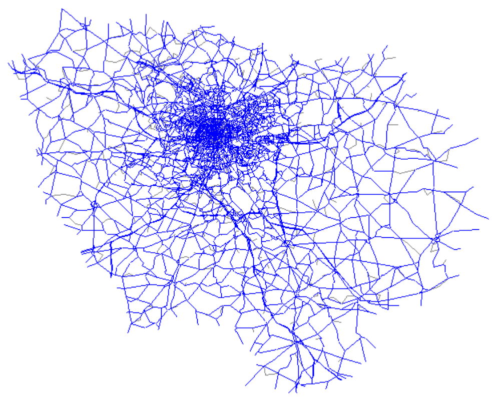
The road network comprises around 30 000 links describing mainly all major roads (including motorways of course) (cf. Figure 4).

Figure 3 –Transport mode combinations in ANTONIN

Individual modes
Car driver Car passenger Slow modes

Public transport modes	Walk access	Drive access	Walk or drive access
Rail only	x	x	
Metro only	x	x	
Bus only	x		
Rail/metro	x	x	
Rail/bus			X
Metro/bus			X
Rail/metro/bus			X

Figure 4 - ANTONIN car network



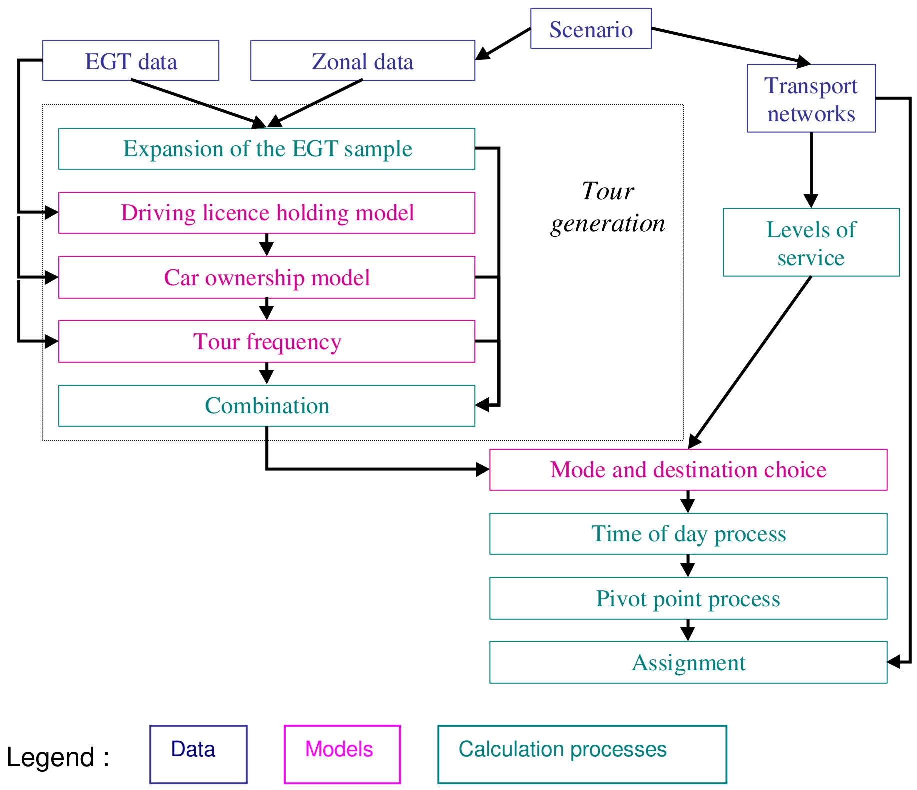
2.2 The architecture of ANTONIN

ANTONIN is based on four disaggregate models estimated on the results of the EGT 1991 (see Figure 5).

The ***driving licence holding model*** is applied to individuals and calculates the probability of having a driving licence whereas the total number of driving licences in Ile de France is generated by a cohort model to take into account the generations' effect.

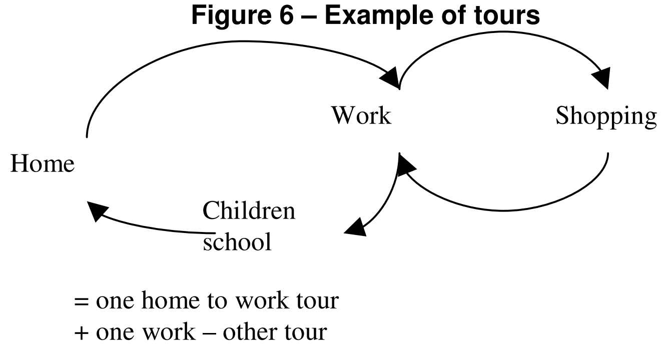
The results of the driving licence model are input to the ***car ownership model*** which is applied to households.

Figure 5 – The structure of ANTONIN



The ***tour frequency model*** is applied to individuals. Tours are defined as return trips based at home or at work/education place (cf. Figure 6). They are a simplification of reality as complex chains of trips are converted in go and return trips according to a hierarchy among trip purposes. (cf. Figure 7). Ten purposes of tours are considered. The model estimates the probability for a person to make one or more tours for each purpose during a day.

Figure 6 – Example of tours



These three models are applied to the individuals or households of the EGT 1991 sample, considering the fact that the EGT sample is statistically representative of the population of Ile de France. This ***"prototypical" sample*** is then expanded zone by zone in order to obtain the number of tours by purposes emitted by each zone. The expansion phase consists in the calculation of weights to apply to each individual of the sample in order to match targets value for each zone (such as the total number of inhabitants, the split by age / sex classes, the size of households etc.).

Figure 7 – Tour purposes

Home based tours	Non-home based tours
Home-work (high income/"cadre") Home-work (low income/"non-cadre") Home-business Home-school (children below 18 years) Home-education (students above 18 years) Home-regular shopping Home-other shopping Home-social activities	Work-business Work/education-others

The combination of the results of the first three models and their expansion for each zone gives a number of tours for each origin zone per day and purpose.

The last model is a ***combined mode and destination choice model***. It is applied to tours produced by each zone. It is assumed, at this stage, that the utility of a destination perceived by the traveller depends both on its characteristics with respect to the purpose of the trip (e.g. number of jobs available for home to work tours) but also on the characteristics of the trip itself according to the modes of transport. This means that a destination near the origin but offering limited opportunities can have a reduced probability to be compared to a destination far from the origin but offering a great number of opportunities.

The results of this last model are daily tour matrices for each of the ten purposes. The following steps of the forecasting system are calculation processes.

The ***time of day process*** convert the daily tour matrices in time period trip matrices (for instance peak period matrix covers the 7:30 – 9:30 time period). First of all, tours are split in two trips, the go and the return trips. Based on EGT observations, percentages are calculated to know which percentage of the first leg of a tour is made during a given time period. These percentages differ according to the purposes.

The ***pivot point process*** is then applied to obtain the final trip matrices for a scenario. The principle of the pivot point process is that only changes in results of the model between the base scenario and a future scenario should be applied to an existing origin – destination trip matrix to ensure the best quality of forecasts. Because of the unavailability of a detailed enough O-D matrix for Ile de France, ANTONIN uses the EGT matrix at the level of "department" (there are 8 "departments" in Ile de France) and splits these matrices on the basis of the results of the mode and destination choice model. Concretely, corrective factors are calculated by O-D, focused at the department level, for 3 modes and 3 purposes and are applied to ANTONIN results. These corrective factors are kept in memory and applied again to results of future scenarios.

The ***assignment phase*** is mainly focused on Public Transport (PT). A car traffic assignment module exists and uses a fairly classical iterative capacity constrained assignment technique. Public Transport (PT) assignment is based on a PT assignment technique (available in TP+) that spreads traffic over alternative competing PT paths that have a common stop node. PT assignment is done separately for the ten public transport modes / access combinations which means that if different combinations are available for an origin –destination pair, many paths can be used during the assignment.

3. THE OBJECTIVES OF THE UPDATE

Since its development in the 90's, STIF has regularly updated the database, zone data as well as network data (both public transport network and roads) of the model. But only limited enhancements have been implemented in ANTONIN, mainly in order to facilitate the analysis of results.

Of course, as executions of ANTONIN were conducted, some imperfections were detected. They mainly concerned the difficulty to use the model and the lack of detail in the description of Ile de France, due to the limits of the software used for the initial development of ANTONIN.

As explained above, ANTONIN is based on the results of EGT 1991-1992 and on the 1990 census. A new census was conducted in France in 1999 and the last EGT in 2001-2002. The re-estimation was therefore possible and appeared necessary to take into account the recent evolutions of individual mobility.

It was therefore decided to proceed to an update of the whole modelling framework in 2004. This update gave the opportunity to modify the structure of programs or models as required to solve the problems faced when using ANTONIN and to facilitate its use by the STIF's staff.

In this update study, it is important to note that STIF wanted to keep the general structure of modules as it is now. *Particularly, the main characteristics of the system, the disaggregate approach, the tour approach, the tour generation for a whole day, the mode and destination choice model and the pivot point procedure will remain unchanged*. The update concerns mainly the following points:

1. Computer environment and software
2. Updating to use recent survey and other data
3. Re-estimation using this new data

Each of these points are shortly described below.

3.1 Computer environment and software

The current version of ANTONIN works within DOS using programs written in 16-bit Borland Pascal as well as TP+ / VIPER batch scripts. However, especially the 16-bit DOS environment makes it difficult to add more complexity to the model system without compromising the speed of the model system too much. As a matter of fact, a complete run of the current version of ANTONIN takes around 13 hours on a standard PC.

To better handle this problem, a conversion of ANTONIN's programs is currently at work. DOS programs are being converted to Windows programs. Modules using TP + / Viper scripts are being transferred to CUBE Voyager. To speed up the runs, or more exactly to spare time to add new functionalities in the system, simplification of the calculation of level of services (by far the most time consuming module today) are being researched.

3.2 Data

The main difficulty in applying disaggregate models stays in the numerous variables that have to be used. Even when data can be found for the estimation of models in the base year, it is much more difficult to find the corresponding detailed data for forecast years especially if the term of the forecasts is 15 or 20 years ahead. Therefore, it was for instance decided to replace the number of jobs in shops used in the mode and destination choice model by shopping surfaces which is available at a detailed level.

Lot of work has been achieved at STIF to establish an automatic conversion of the description of public transport services used in the Internet traveller's information website (www.transport-idf.com) in a file with the required format to be used by the system.

Assuming that enhancements in the computer environment will provide some time savings, it was also decided to increase the number of zones to around 1300 and describe completely a selection of bus lines in the outer suburbs instead of a sketchy description. These changes are expected to enhance the quality of the results.

3.3 Re-estimation of the models

The EGT 2001/02 will be used as the base dataset on travel behaviour in the Ile-de-France. The disaggregate models that were estimated in the early 90's (see the Models as given in the diagram of Figure 5) will be re-estimated based on this new dataset. Additionally a new disaggregate model is estimated that describes ownership of Public Transport passes in the region.

The re-estimation of the models implies two major steps: first of all the existing model specification is used to estimate new models using the new EGT and other data. Secondly additional variables are sought to better explain travel behaviour for the current year.

The next section describes the re-estimation of the tour generation models (driving licence holding, car ownership and tour frequency) plus the estimation of the model for Public Transport passes.

4. THE RE-ESTIMATION OF THE MODELS

Using the EGT 2001/02 and including relevant additional data available, the existing ANTONIN tour generation models have been updated to reflect travel behaviour for that year. Given experience over the years and STIF's requirements, these models have been kept similar in structure to the existing ones.

The following models have been re-estimated so far:

1. Driving licence holding models for a) the head of household and/or his/her partner and b) other adults in the household;
2. Car Ownership models describing the choices a) between 0 and 1 car in a household with 1 licence and b) between 1 or more cars in a household with 2 or more licences;
3. Tour frequency models for each of the ten travel purposes identified describing on a given day whether a person would make a) any tour (0/1+ model) and b) another tour (Stop/Repeat model).

Additionally a new disaggregate model was estimated that describes ownership of Public Transport passes in the region. This model describes the choice between ownership of no PT pass, a *Carte Orange* pass (allowing for free travel between certain zones, mainly used by commuters) or an *Imagine'R* pass ("student pass").

Some more detail on each of the models is given below.

4.1 Driving licence models

To analyse the travel behaviour of persons, it is important to know what options travellers have in deciding on modes to travel. Therefore, it is necessary to estimate the number of licence holders within a transport zone and the characteristics of the households to which they belong.

An analysis of the changes in driver's licence holding between 1991 and 2001 demonstrates the requirement of a model to capture the various effects of age and sex as can be seen in this section. A simple tabulation is given below (based on EGT survey data), giving the percentage of people of 15 years and older that have a driver's licence in Ile-de-France.

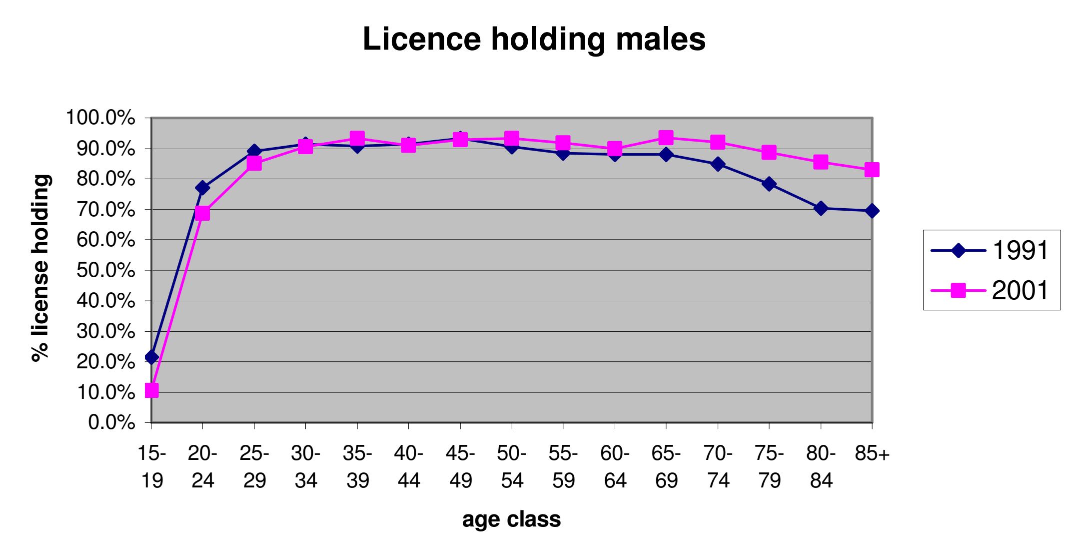
Gender	EGT91	EGT01
Male	83.1 %	83.4 %
Female	61.0 %	66.5 %
Total	71.5 %	74.6 %

Table 1 - Percentage of population over 15 years with a driver's licence

From Table 1 it is clear that the increase in driver's licence holding in general is almost entirely caused by changes in the licence holding by women. The difference in licence holding has substantially diminished.

A further analysis using the same data shows that (apart from gender effects) that licence holding is strongly related to age but also that the effect has declined over the period 1991-2001.

Figure 8 - Licence holding males for the two years.



From Figure 8 it is clear that for males the saturation level is at about 93% and remains stable at that level between 25 and 75 years of age. Looking at the development in the past years, it is clear that the only growth in male licence holding results from the men over 65 years becoming 10 years older and thus moving their higher licence holding rates into these age cohorts. Also interesting to note is that under 20 years the rate has halved, which has been seen in other countries too. Also for the 20-30-year-olds lower rates are visible as well.

Figure 9 - Licence holding females for the two years.

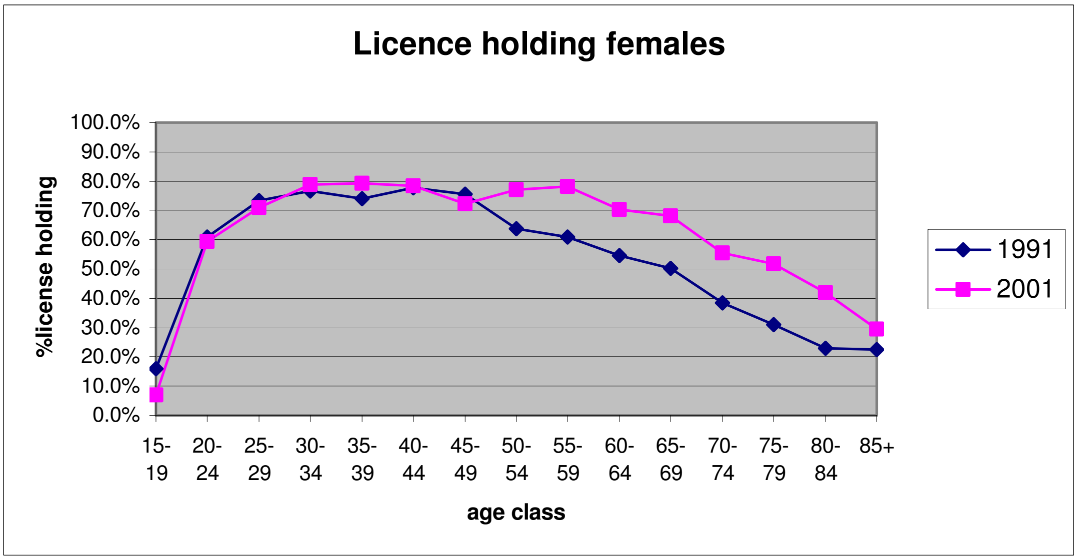


Figure 9 shows a different picture. For females it is obvious that (a) the saturation level is now at about 79% so some 14% points lower than for males, (b) for women over 50 years the licence holding rates are now some 10 to 20 points higher compared to the same cohorts in 1991 and (c) there is still a clear age effect visible, although less than in 1991.

While in 1991 female driver's licence holding was at about (a slightly lower) saturation level only for ages 25 through 49, ten years later these cohorts have also shifted ten years so that an age effect is only visible after 59 years of age (and of course an age effect is visible before 25 years but that is obviously the licence acquisition effect which we also see for males). Continuing this trend into the future, we expect that the overall rate will continue to increase for 20 years at a declining rate.

One would also notice that while the licence holding percentage for men goes some 10% down from the saturation level towards 83% for men over 85, for females this percentage drops with some 50% points down towards 29%! Looking at the licence holding rates for under 20 years, we observe (like we saw for males) that these rates are substantially lower than in 1991. However the change for 20-30-year-olds does not appear.

From these analyses it is clear that for the foreseeable future it is worth estimating disaggregate licence holding models to identify the differences in saturation level between males and females and also of the age effect for women over 60-65 years. These effects are taken into account in the logit models for licence holding.

From the figures presented it is also clear that cohort models are still useful in application, be it that this would be mainly necessary for women given that their licence holding rates in the older age groups are still catching up.

Our re-estimation results for the licence holding models confirm the developments we see in the data. Comparing the models with the same specification but with different EGT data, we see that the new models are generally less sensitive to changes in the independent variables (i.e. the coefficients estimated are closer to zero), this is as expected since licence holding over the total population has increased and shows a more homogeneous distribution over the population nowadays.

4.2 Car ownership models

To analyse the travel behaviour of persons, it is important to know what options travellers have in deciding on modes to travel. Therefore, it is necessary to estimate the number of cars within a transport zone and the characteristics of the households to which they belong.

The next table gives an overview (drawn from EGT 2001/02) of licence holding and car ownership in the Ile de France region, for those households that have one or more licences (our analysis confirms that almost all households with 0 licences have no cars, so we leave that out of our car ownership model). The table contains cell percentages, allowing analysis of the size of each of the combinations of household licence holding and car ownership in relation to the total number of households.

		HCA					Total
		0	1	2	3	4	
HLI	1	14.1%	30.8%	1.4%	0.1%	0.0%	46.3%
	2	2.2%	21.2%	21.4%	1.7%	0.2%	46.6%
	3	0.1%	1.1%	2.6%	1.9%	0.2%	5.9%
	4	0.0%	0.1%	0.4%	0.3%	0.2%	1.0%
	5	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%
Total		16.4%	53.1%	25.8%	4.0%	0.6%	100.0%

Table 2 - Percentage households by number of licences (HLI) and cars (HCA)

It is clear that households with one licence in majority have one car. Households with two licences have either one or two cars. The more licences the less apparent the relation becomes, but also the number of records reduces quickly with an increasing number of household licences and/or cars.

As in the previous Antonin model, two separate car ownerships models were estimated based on the number of cars and licences per household: one model for households with 0 or 1 car and only one licence and one model for households with 1 or more cars and two or more licenses. In Table 2 the cells have been marked that are modelled in these car ownership models. From the table it can be seen that the models describe the combination of licence holding and car ownership for 96.2% of all records, sufficiently high enough to maintain the existing car ownership modelling structure.

Our re-estimation results for both car ownership models show that the models are remarkably stable: Comparing the models with the same specification but with different EGT data (i.e. that of 1991/92 and 2001/02), we see that the new models are generally quite comparable to the previous ones, with for only a few variables a clearly lower sensitivity to changes in the independent variables (i.e. the coefficients estimated are closer to zero), reducing effects of e.g. gender as expected.

4.3 Tour frequency models

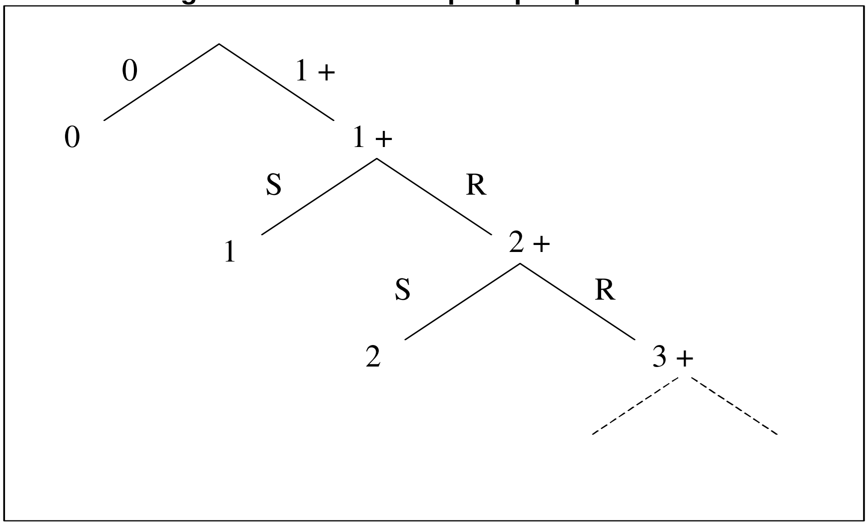
Tour frequency models in the Antonin model system are organised in two types:

- 0/1+: models to determine whether a person would make **any** tour
- S/R: Stop-Repeat models to determine whether a person would make **another** tour

Ten tour purposes (see Figure 7) were distinguished, for each purpose separate 0/1+ and Stop/Repeat models have been estimated.

Figure 10 gives the set-up of the tour generation model, linking the 0/1+ model (applied once) and the Stop-Repeat model (applied repeatedly to account for the situation where one or more tours are made).

Figure 10 - Model set-up Stop-Repeat Model



Our re-estimation results for the tour frequency models show that the models are quite stable over time: Comparing the models with the same specification but with different EGT data (i.e. that of 1991/92 and 2001/02), we see that the new models are generally quite comparable to the previous ones.

In the new models we have also tested the effect of including parameters for ownership of the Carte Orange and/or Imagine'R passes. Including these coefficients in some of the tour frequency models gave an important improvement in their explanatory power.

In four of the purposes (both Home-work purposes, Home-education and Work/education based other), owning a PT pass leads to a higher probability of making one or more tours on a day. For Home-business and Home-regularshopping a negative coefficient was found, which appears sensible, as travel for these purposes would generally be done by car and we can assume that the group with a PT pass is more likely not to own a car. For the remaining four purposes no effect could be found for making one or more tours on a day.

For only three purposes (both Home-Work purposes and Home-social) we could find an effect in the Stop/Repeat models, this effect was always negative. We consider these plausible given time budgets involved to make more than one tour on a given day (and the higher likelihood of making any tours on a given day for the 2 commute purposes).

More details on the PT pass model and considerations on the position of this model within the Antonin model can be found in the next section.

4.4 Public Transport Passes model

Public Transport Pass ownership is an important factor in mode and destination choice, especially in large metropolitan areas such as the Ile-de-France, which has an extensive public transport system. Owners of passes are much more likely to choose public transport than non-pass holders, and as pass ownership changes then so will the PT mode share.

An important consideration when specifying the pass ownership model was determining where in the choice hierarchy the pass ownership model should lie in relation to the other model components, namely models of driving licence holding, car ownership, tour frequency and mode and destination choice.

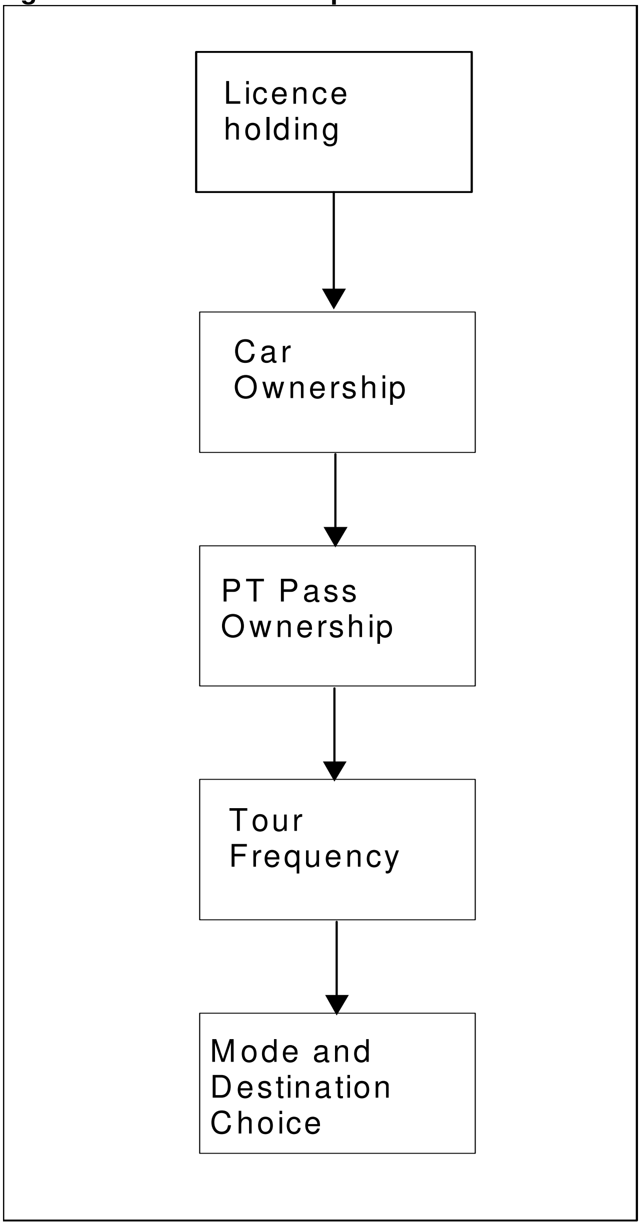
In this particular study -and similar to other studies by RAND Europe- the driving licence holding and car ownership model components lie at the top of the choice hierarchy. This means that pass ownership will not impact on licence holding nor on car ownership, but a person having a driving licence or a household owning a certain number of cars can be an explanatory variable in the pass ownership model.

The next issue is where the pass ownership model sits relative to mode and destination choice. It was decided to place pass ownership above mode and destination choice in the choice hierarchy, so that in the mode and destination choice models, PT costs will be specified as a function of predicted pass ownership. This choice was made since passes would generally be bought for travel over a longer period (and not just a single trip) and also to avoid complicated logsum terms representing the utility of the various PT pass alternatives in the mode choice utilities.

The final consideration was the placement of pass ownership choice relative to tour frequency. Overall it was decided to place pass ownership *above* tour frequency in the choice hierarchy for two reasons. Firstly, the decision to acquire a pass can be viewed as a medium term decision, made on a weekly, monthly or annual basis, whereas the decision as to how many tours to make is a shorter term decision. Typically shorter term decisions are represented lower in the choice structure. The second reason is more pragmatic: separate tour frequency models are used for different travel purposes, and a structure whereby tour frequency models for different travel purposes fed into a single pass ownership model would be cumbersome in implementation.

The final choice structure is given in Figure 11.

Figure 11 - Pass Ownership in the Choice Structure



In the structure implemented, the pass ownership model can contain parameters relating to licence holding and/or the car ownership of the individual, but not to their tour frequency or mode- and destination choice. All other explanatory variables for modelling pass ownership are person-specific attributes that are exogenous to the model system.

The frequency model is placed beneath pass ownership in the choice structure, and therefore the tour frequency models can incorporate parameters that reflect the pass ownership of the individual. Similarly the frequency models also incorporate licence holding, car ownership and availability parameters. The frequency models are placed above mode and destination choice (which is a standard model structure) and thus the mode and destination choice models can be sensitive to the results of all models placed above it.

Models of pass ownership choice have been estimated from the 2001 EGT Household Interview data. The models of pass ownership choice estimated from the EGT data predict changes in pass ownership at the individual level as a function of the socio-economic characteristics of the individual and their household. The models contain terms reflecting the impact of car ownership and availability on pass ownership. Alternatives identified are (note there are some 5% unknown/other passes):

1. No pass (70% of the individuals recorded in the EGT)
2. Carte Orange pass (20%)
3. Imagine'R pass (5%, only available to students under 26 years)

Both PT passes allow for free travel between certain zones of the region during weekdays.

The pass ownership model estimated shows the very significant impact of the variables included. Especially occupation of the respondent, "car competition" in the household (based on the number licences and cars in the household) and the dummy for households with a residence in Paris are amongst the highest significant variables, demonstrating the strong relation between these variables and the choice of PT passes in the region clearly.

5. NEXT STEPS

The work reported here is still ongoing as more models need to be re-estimated and those that have been re-estimated are in the process of implementation and validation. For the period ahead, the models will be put in a consistent framework so that the new model can be tested and used by the Syndicat d'Ile de France.