Sétra

Service d'études sur les transports, les routes et leurs aménagements

A disagregate car ownership model based on French National Transport Survey

EUROPEAN TRANSPORT CONFERENCE

GLASGOW

12/10/2011

Richard GRIMAL

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE

Ministère de l'Écologie, de l'Énergie, du Développement durable et de la Mer

Ressources, territoires, habitats et logement
Énergie et climat Développement durable
Prévention des risques Infrastructures, transports et mer

Présent pour l'avenir

Sétra

Service d'études sur les transports, les routes et leurs aménagements

SOMMAIRE

Pages

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE

Ministère de l'Écologie, de l'Énergie, du Développement durable et de la Mer

Why modelling car ownership ?

- Car availability remains a determinant factor of mobility levels and modal shares, both for local and long-distance trips, most of them being realized by car. Mobility « cultures » are dependent on car ownership.

- Modelling approaches help identifying, de-correlating and comparing statistically significant influences, everything else being equal.

- Descriptive statistics but also econometric works question long-term stability of car ownership determinants. Dynamic elasticities differ from cross-sectional, themselves changing.

- Car ownership increases « everything else being equal ». Simultaneously, trends towards stabilization or decline of car-oriented mobility emerge among the young, city-dwellers and executives.

Methodology

A) Model development :

- We try to forecast household’s number of cars, from information about socio-demographics, household structure, land use patterns, conditions of access to mobility.

- We develop discrete choice models based on random utility maximization, where car ownership is given three modalities : 0 (car deprivation), 1 (single-ownership) and >=2 (multi-ownership).

- Models are specified and estimated from National Transport Survey 2007-2008, containing detailed information about 20 000 households.

- Starting from simple MNL models, more complex models such as segmented MNL and CNL are tested. Gains from sophistication are assessed.

Methodology

B) Assessment of model sensitivity and robustness :

- Model sensitivity and robustness are tested by simulation on specific subsamples and exogenous samples randomly drawn out from detailed census files.

- However, models first have to be adapted to take into account variable’s availability and their modalities in census files.

- Re-shaping car ownership distribution on census files also implies simulating driving license holding that is not available in census files. It implies developing a model for the number of driving licenses in households from National Transport Survey.

Methodology

C) Car ownership dynamics and long-term stability of explanatory factors :

- Simulation helps separating dynamics caused by changes in values of explanatory variables (socio-demographics, land use patterns, household structure, driving license dissemination…) from dynamics caused by changing effects.

- A model of car ownership is developed from NTS 1994, and applied to NTS 2008. Simulation results forecasts what car ownership would have been under the assumption of stable effects.

- Symetrically, a model is estimated on NTS 2008 with the same specification. Comparison with NTS 94 model allows identifying changes in strengths of respective effects.

Results part A - model development

- Some decisive variables were already identified in previous analysis, such as household size, income per consumption unit, density and centrality of housing area, number of working adults.

- Now, every plausible and available explanatory variable of car ownership is taken into account.

- Binary logistic regression with backward elimination is performed to remove irrelevant variables and select the most influential to keep into discrete choice models.

- Discrete choice models are then specified. Alternative-specific parameter values are defined to allow variability through alternatives, despite constant household's caracteristics.

Results part A – model development

- For instance, an MNL model was developed with the following specification :

*U_0 = B0_LICENSE * n_permis + B0_ABONTC * n_cartabon + B0_ICENTRE * i_centre + B0_IAPPT * i_appt + B0_IPARIS * i_paris + ε_0.*

*U_1 = C_1 + B1_ICENTRE * i_centre + B1_IPARIS * i_paris + B1_ABONTC * n_cartabon + B1_IAPPT * i_appt + B1_IACTIF * i_pr_actif + B1_CONJOINT * i_conjoint + B1_CONJACT * i_conjoint_actif + B1_IPROPRIO * i_proprio + B1_IQ4 * i_q4 + ε_1.*

*U_2 = C_2 + B2_LICENSES * n_permis + B2_NENFMAJ * n_enfants_maj + B2_IPERIURBAIN * i_periurbain + B2_IPROPRIO * i_proprio + B2_IQ3 * i_q3 + B2_IQ4 * i_q4 + B2_IACTIF * i_pr_actif + B2_CONJOINT * i_conjoint + ε_2.*

Where U_0, U_1 and U_2 represent respective utilities of « car deprivation » (no car), « single-ownership » (one car), and « multi-ownership » (two cars). Variable's meaning is given in next draft.

Results part A – model development

Variable	Meaning
N_permis	Number of driving licenses
N_cartabon	Number of transit seasonal tickets
I_centre	Indicator variable for home in central city
I_appt	Indicator variable for apartment
I_paris	Indicator variable for the urban area of Paris
I_pr_actif	Indicator variable for head of household's work participation
I_conjoint	Indicator variable for partner's existence
I_conjoint_actif	Indicator variable for partner's work participation (set to 0 if doesn't exist)
N_enfants_maj	Number of major children at home
I_periurbain	Indicator variable for periferic crowns of urban areas (zones outside built-up areas but many people commuting to them)
I_proprio	Indicator variable for property owners
I_q3	Indicator variable for third quartile of income per consumption unit
I_q4	Indicator variable for fourth quartile (the highest income group)

Variables : names and meanings in model 2008

Results part A – model development

Utility parameters

****

Name	Value	Std err	t-test
ASC_1CAR	-1.83	0.0857	-21.37
ASC_2CARS	-6.94	0.135	-51.38
ASC_NOCAR	0.00 --fixed--		
B0_ABONTC	0.734	0.0556	13.21
B0_IAPPT	0.910	0.0901	10.10
B0_ICENTRE	1.21	0.0876	13.82
B0_IPARIS	1.73	0.0851	20.29
B0_IQ3	-0.390	0.0754	-5.17
B0_LICENSE	-2.90	0.0720	-40.25
B1_ABONTC	0.184	0.0317	5.82
B1_CONJACT	-0.285	0.120	-2.37
B1_CONJOINT	0.257	0.0758	3.39
B1_IACTIF	1.09	0.0641	17.03

Utility parameters

****

Name	Value	Std err	t-test
B1_IAPPT	0.576	0.0631	9.13
B1_ICENTRE	0.453	0.0597	7.59
B1_IPARIS	0.610	0.0525	11.62
B1_IPROPRIO	0.465	0.0625	7.44
B1_IQ4	0.293	0.0785	3.73
B2_CONJOINT	0.968	0.108	9.00
B2_IACTIF	1.76	0.0793	22.19
B2_IPERIURBAIN	0.569	0.0968	5.88
B2_IPROPRIO	0.920	0.0780	11.79
B2_IQ3	0.412	0.0504	8.17
B2_IQ4	0.998	0.0911	10.96
B2_LICENSES	1.89	0.0627	30.11
B2_NENFMAJ	0.435	0.0925	4.71

Results part A – model development

Estimation results :

- L(0) = - 22 166.7 L(β) = -11 922.7 $\rho^2 = 0.461$

- Parameter signs are consistent with a priori expectations.

- Negative alternative-specific constants illustrate that, in the reference situation, car deprivation is more likely than car ownership (no driving license in household).

- The main explanatory variable is driving license holding, increasing probabilities of car ownership.

- Other variables are strongly influent : head of household's work participation, existence of partner, property owning, high financial resources, increasing multi-ownership probabilities,

- On the contrary, residential place in the urban area of Paris, in central city, apartment, are increasing car deprivation's probabilities.

Results part A – model development

- A very strong relationship between car ownership and driving license holding :

	No car	One car	Two cars	Three or more	Total
No license	97.9	2.0	0.1	0.0	100.0
One license	21.4	72.6	5.4	0.6	100.0
Two licenses	2.0	33.7	58.3	6.0	100.0

- Almost all license-deprived households are also car-deprived.

- The general case is that the number of cars equals the number of driving licenses. The minor case is that it equals the number of driving licenses minus one.

- Choice process dependence on driving license holding suggests developing alternative models. Model segmentation with segment-specific parameter values is tested against a simple MNL model for at least one-licensed households.

Sétra

Service d'études sur les transports, les routes et leurs aménagements

Results part A – model development

- Two segments are defined, one for single-licensed households, one for multi-licensed. The interest of segmentation is assessed through a likelihood ratio test : LR = - 2 (-11 406.6 + 4186.9 + 7171.4) = 96.6.

- LR is compared to a χ^2 distribution at the 0.05 level of significance with 24 degrees of freedom corresponding to supplementary degrees generated by segmentation.

- As χ^2 (0.05, 24) = 36.4 << 96.6, we can reject the null hypothesis and admit an interest for market segmentation. However, gain from segmentation is moderate.

- Alternative model forms can be tested, for instance CNL models. CNL models are equivalent to segmented MNL, suggesting there is little gain to be obtained from model sophistication. However, alternative specifications closer to the real choice process could give better results.

Liberté • Égalité • Fraternité
République Française

Ministère de l'Écologie, de l'Énergie, du Développement durable et de la Mer

Sétra

Service d'études sur les transports, les routes et leurs aménagements

Results part B – sensitivity and robustness

- Model sensitivity and robustness is tested through simulation on various subsamples.

Subsample	P(0)	P(1)	P(2)	E=0	E=1	E>=2
T_1	0.09	0.47	0.44	0.09	0.47	0.44
T_2	0.09	0.46	0.45	0.09	0.46	0.45
T_3	0.09	0.46	0.45	0.09	0.46	0.45
Urban area of Paris	0.19	0.48	0.33	0.18	0.49	0.33
Urban areas - > 500 000 dwellers	0.08	0.46	0.46	0.08	0.46	0.46
Midi-Pyrénées	0.06	0.48	0.46	0.05	0.47	0.48
Pays-de-la-Loire (id)	0.06	0.45	0.49	0.05	0.45	0.5
Rhône-Alpes (id)	0.06	0.45	0.49	0.06	0.44	0.5
Urban areas - 100 to 500 000 dwellers	0.06	0.46	0.48	0.05	0.45	0.5
City-centres	0.21	0.53	0.26	0.21	0.54	0.25

Table 2 : Simulated car ownership distribution on various subsamples of National Transport Survey, compared to real car ownership distribution

- Car ownership distribution, estimated by average probabilities, is correctly recreated on random subsamples from National Transport Survey. Simulation results are also sensitive to specific conditions. Car ownership distribution in highly-motorized and lowly-motorized areas is quite well recreated.

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE

Ministère de l'Écologie, de l'Énergie, du Développement durable et de la Mer

Sétra

Service d'études sur les transports, les routes et leurs aménagements

Results part B – sensitivity and robustness

- Model robustness is then tested by performing simulation on external data from detailed census files at the household level.

- Before, models have to be adapted to take into account variable's availability in census files. Notably, the most determinant variable of car ownership, driving license holding, is missing.

- Driving license holding is simulated thanks to another model developed from National Transport Survey.

- As treating all observations in census files would have mobilized too many computational resources (nearly 30 million lines), simulation is applied on random drawings.

Liberté • Égalité • Fraternité
République Française

Ministère de l'Écologie, de l'Énergie, du Développement durable et de la Mer

Results part B – sensitivity and robustness

- Differences between forecast distribution and real one are deriving from stratification bias of estimation subsamples, leading to under- or over-estimation of multi-ownership effects.

Subsample	P(0)	P(1)	P(2)	E=0	E=1	E>=2
T_1	0.206	0.476	0.318	0.193	0.475	0.332
Ile-de-France	0.335	0.408	0.257	0.316	0.476	0.207
Pays-de-la-Loire	0.137	0.495	0.368	0.120	0.472	0.408

- Results are sensitive to both highly- and poorly-motorized areas. Car ownership distortion is well recreated in sense and to a large amount in intensity.

- However, unobserved heterogeneity may still contain omitted factors, leading to endogeneity bias in parameter estimates. In particular, car ownership distortion is under-estimated.

Sétra

Service d'études sur les transports, les routes et leurs aménagements

Results part B – sensitivity and robustness

- Forecasting results can be improved by adding supplementary variables. In Ile-de-France, car-restraining behaviours among young adults, a high rate of tenants and limited parking facilities. In Pays-de-la-Loire, individual housing is widespread.

Subsample	P(0)	P(1)	P(2)	E=0	E=1	E>=2
Ile-de-France	0.334	0.436	0.230	0.316	0.476	0.207
Pays-de-la-Loire	0.137	0.482	0.381	0.120	0.472	0.408

- Enriching models with local variables leads to better forecasts, as a larger amount of heterogeneity was taken into account for model estimation.

Sétra

Service d'études sur les transports, les routes et leurs aménagements

Results part C – long-term car ownership dynamics and stability of explanatory variables

- Explanatory factors of car ownership, mobility, or transport demand, might not be stable through long-term periods. For instance, income elasticities are declining.

- New trends are arising, suggesting assumptions about behaviours (rationalization) among young adults, upper-class and city dwellers.

- Forecasts cannot always be derived from measured influences on cross-sectional data combined with assumptions about future values of explanatory variables.

Liberté • Égalité • Fraternité
République Française

Ministère de l'Écologie, de l'Énergie, du Développement durable et de la Mer

Sétra

Service d'études sur les transports, les routes et leurs aménagements

Results part C – long-term car ownership dynamics and stability of explanatory variables

- Behavioural change can be approached through model estimation. Another car ownership model was estimated from National Transport Survey in 1994. It was then applied to forecast car ownership distribution in 2008.

- Simulation results illustrate what car ownership distribution would have been under the assumption of stable effects, given changing values of explanatory variables.

	E=0	E=1	E>=2
Car ownership in 1994	0.251	0.488	0.262
Car ownership in 2008	0.195	0.473	0.332
Simulated car ownership in 2008 with NTS 94 model	0.216	0.497	0.286

- Part of car ownership increase is forecast by the model. However, a large amount of car ownership growth remains unexplained. In particular, multi-ownership is under-estimated.

Liberté • Égalité • Fraternité
République Française

Ministère de l'Écologie, de l'Énergie, du Développement durable et de la Mer

Results part C – long-term car ownership dynamics and stability of explanatory variables

- If estimators were unbiased and effects constant, multi-ownership should be correctly recreated by the model.
- Given observable effects in 1994, most trends (perpetuation of urban sprawl, increasing qualifications, driving license generalization…) contribute to increased car ownership, except for household structure.
- Differences between forecast distribution and real one might be caused either by estimation problems (endogeneity), or by changing strengths of effects.
- In order to reveal possibly changing effects, one can re-estimate a model with the same specification from NTS 2008, and compare parameter values.

Results part C – long-term car ownership dynamics and stability of explanatory variables

Parameter names	Parameter estimates 1994	Parameter estimates 2008
C1	-1,16	- 0,785
C2	-1,89	-1,64
B0_ABONTC	0,27	0,67
B0_PARIS	0,44	1,27
B0_ARTAGR	0,37	0,45
B0_APPT	0,56	0,68
B0_CENTRE	0,45	0,9
B0_Q1	1,28	0,91
B0_Q2	0,95	0,44
B1_ACTPR	0,53	0,83
B1_CONJOINT	0,55	0,26
B1_APPT	0,09	0,15
B1_Q3	-0,6	-0,13
B1_LICENSE	3,1	3,1
B1_NETUD	0,39	0,67
B1_NJEUNENF	0,26	0,21
B2_ACTPR	0,56	0,9
B2_CONJOINT	0,5	0,29
B2_Q3	-0,57	-0,08
B2_LICENSES	3,38	3,43
B2_ETUDIANTS	0,39	0,72
B2_NJEUNENF	0,27	0,22

Results part C – long-term car ownership dynamics and stability of explanatory variables

- Alternative-specific constants illustrate increased probability of car ownership, everything else (in observed heterogeneity) being equal.

- Influence of driving license holding on car ownership remains strong and stable.

- Income based-parameter values are strongly decreasing, probably illustrating decreasing prices with the development of second-hand market. Car would have come from a superior to usual good.

- Location effects are becoming more influent, especially in high-density areas (central cities, urban area of Paris…), where car deprivation is becoming usual among young city-dwellers.

- Head of household’s work participation and students living at home are becoming more influent (rationalization, autonomy ?)

Conclusion and implications

- The question remains of how interpreting changing effects on car ownership distribution, leading to wrong model forecasting from cross-sectional estimation.

- Interpreting them as changing behaviours might probably be too conclusive. They are probably three distinct dimensions to consider :

- The general dynamics of good diffusion, going through successive steps. Car market might have come to saturation, with decreasing marginal utility of supplementary cars among the well-off, and an adjustment process among low-income groups, leading to decreasing income elasticities. Homogenization also results from driving license diffusion.
- The adjustment process is facilitated by price dynamics with tbe second-hand market making car more affordable. Price dynamics, being temporal-dependent, couldn't be pictured in our modelling approach based on cross-sectional demand effects.
- Eventually, behavioural change : psychological latent factors such as desire for autonomy could explain increased multi-ownership, or, on the contrary, rationalization could explain car restriction and greater substitutability between alternatives among the young urban.

- The next step is to develop dynamic models from panel data.

Sétra
Service d'études sur les transports, les routes et leurs aménagements

Conclusion and implications

- One difficulty is about making distinction between underlying psychological attitudes and economic *inputs* as they could be correlated. For instance, the adjustment process might either be explained through psychological attitudes (rationality, need…) or price dynamics.

- Building dynamic models would require introducing price dynamics and collecting psychological information to help separating strictly economic from psychological factors in changing behaviours.

Liberté • Égalité • Fraternité
République Française

Ministère de l'Écologie, de l'Énergie, du Développement durable et de la Mer