

---

## Analysis by panel data method estimation of car fleet models

Rachid Toumache and Khaled Rouaski

High National School of Statistics and Applied Economics, Algeria

Sabah Fadel

University of Algeria III, Algeria

---

### Key words

Car Fleet, Panel Data, Gompertz Function, Quasi-Logistic and Logistic Functions, Calibration of Econometric Models.

### Abstract

*The car fleet evolution in Algeria is due to the non-linear variation of the income represented by the national wealth (GDP) rather than other infrastructural factors such as car prices, fuel prices, the transport network, population density, and the extent of the country. This study predicts the future image of the Algerian car fleet, based on the technique of time-series cross-sectional data. The evolution of the car fleet is modeled using three utilities models provided by the literature namely the Gompertz function, the function Quasi-Logistics and Logistics function. In addition, these models were calibrated using panel data or pooled data. 46 countries (sections) were captured during 32 years from 1971 to 2002 (source: World Bank) to build four panels which include Algeria, China, India and the United States. For this choice, it was considered the international trend of the development of the park in terms of GDP and the countries with the same characteristics as ours. As a result of the work, a set of future Algerian car fleet scenarios were identified by the different models statistically significant.*

---

### 1. Introduction

Among the three Maghreb countries, Algeria owns the most important car fleet, it was stated that the car fleet rolling has 2 million tourism vehicles, 700,000 lightweight commercial vehicles, 536,000 heavyweight vehicles and 10,000 motorcycles. However, the average age of the car fleet is high, 55% of the vehicles have more than 20 years and 80% more than 10 years. The Algerian motorization rate is 71 vehicles for 1,000 inhabitants. Among the brands present, French carmakers are dominant; Peugeot and Renault represent half of the rolling car fleet (900,000 Peugeot vehicles and more than 600,000 of Renault vehicles). The total of other carmakers' cars is below 200,000 vehicles throughout Algeria. Recent years, the car market has seen the apparition of a growing number of new brands which constitutes new competitors, with the particularity of being aggressive on price segment, for French carmakers historically present in Algeria.

### 2. Identify, Research and Collect idea

The existing relationship between the vehicle ownership and the GDP per capita is represented by various non-linear models [1]:

– **Logistic model:** Tanner proposed a logistic model by adding more variables as the GDP per capita and the cost of vehicle. The utility form suggested commonly is :

$$Y = \frac{S}{1 + e^{at + b \ln(i) + \ln(p)}} \quad (1).$$

- **The quasi-logistic mode (1):** The same principle of the first model is applied in addition to a set of socioeconomic factors  $X_i$ .

$$Y = \frac{S}{1 + e^{-a X_1^{-b} X_2^{-c}}} \quad (2)$$

- Gompertz Function: Theoretically, the Gompertz function is written as :

$$Y = S \quad (3)$$

$i$  : represents the income per capita,  $S$  : saturation level,  $P$  : cost of the vehicle,  $b$ ,  $c$  and  $d$  are parameters of the mode (1),  $Y_i$  : the vehicle ownership level per capita, over the long run,  $GDP$  per capita income.

Concerning our application, the database is characterized by: The inclusion of two variables: the car fleet as a dependant variable and the GDP as an explanatory variable. It is a time series having 46 sections (countries) covering the period 1971-2002. This database is provided by the World Bank and the International Monetary Fund, gives us 1472 observations. In addition, the GDP is defined as the total of goods and services produced in the territory of any country during a given year, whichever the nationality of the producers; hence, it measures the wealth of a country.

The model in S curve «GOMPERTZ Quasi-logistic» allows establishing a unique trend based on one or many sections or countries with the goal to classify correctly the members of the population, they seek combinations between several countries. For each model, a saturation threshold, set in advance, is equal to the threshold of the country situated in the upper extreme of the S curve.

The utilities' models for the case of Algeria are written under the following linear form [1]:

- Logistic Model :  $Y_i = \frac{1}{1 + e^{-\alpha}}$  (4), Where the linear form is [3] :

$$\text{Log} ( \quad ) = \log(\alpha) + \beta * \log(\text{PIB}) \quad (5)$$

- Quasi- Logistic Model  $Y_i = \frac{1}{1 + e^{-\alpha X_1^{-b} X_2^{-c}}}$  (6) Where the linear form is [3] :

$$\text{Log} ( \quad ) = \log(\alpha) + \beta * \log(\text{PIB}) \quad (7)$$

- Gompertz Model:  $Y_i = S e^{-\alpha e^{\beta \text{PIB}}}$  (8), Where the linear form is [3] :

$$\text{Log} [\log(S/Y_i)] = \log(-\alpha) + \beta * \text{PIB} \quad (9)$$

$i$  : is the index of the individual, in this case: the country (Algeria)

Then in this study, the calibration of these models permits to forecast the size of the Algerian car fleet according to the GDP evolution, Concerning our application, the database is characterized by the inclusion of two variables, where the car fleet as a dependant variable and the GDP as an explanatory variable. It is a time series having 46 sections (countries) covering the period 1971-2002. This database is provided by the World Bank and the International Monetary Fund, gives us 1472 observations. In addition, the GDP is defined as the total of goods and services produced in the territory of any country during a given year, whichever the nationality of the producers; hence, it measures the wealth of a country. The two variables do not have a visible probability distribution. Given the fact of having a 46-country database, we judged useful to constitute the following different panels [7] :

- **P1**: World panel having the 46 countries.
- **P2**: panel includes: Turkey, Algeria, India, China and the United States.
- **P3**: panel includes: Algeria, India, China, United States, Syria and Egypt.
- **P4** : the following countries: United States, Italy, Mexico, Spain, Brazil, Egypt, Algeria, India, China and Syria.

### 3. Results

Now it is the time to articulate the research work with ideas gathered in above steps by adopting any of below suitable approaches. The aim of the S curve model consists in the determination, for different GDP levels, the areas where the fleet evolution changes its path. Indeed, once the model calibrated [1], Algerian car fleet scenarios will be established as a function of the selected countries choice. The first proposed model was assessed for all the available countries in our database, meaning the 46 countries with a long-term saturation of 0,850.

#### *a. Panel data estimation where all the coefficients are constant*

This result was obtained by EViews software using panel data regression with 46 countries. The precedent model's parameters are [2] :

$$Car = S * e^{\alpha * e^{\beta * PIB}} \quad (10)$$

$$\text{or } \log\left(\log\left(\frac{S}{Car}\right)\right) = \log(\alpha) + \beta * PIB \quad (11)$$

This model is a linear equation of type [3]

$$y = c + a * x \quad (12), \text{ with}$$

$$y = \log\left(\log\left(\frac{S}{Car}\right)\right) \text{ et } x = PIB \quad (13)$$

Finally, for all world countries, our model can be written as [3]:

$$\log\left(\log\left(\frac{0.850}{Car}\right)\right) = 1.539836 - 0.00010\epsilon \quad (14)$$

The aim of this equation is to determine the level of car fleet in function of GDP. The empirical application consists of proposing future values of the Algerian GDP, for example by running a linear regression of the curve and to calculate the car fleet level by equation (14). The statistical analysis of these parameters is: Student test is verified for the two parameters, the probability of each parameter equal zero is null, significantly inferior to 5%, commonly accepted as signification threshold. FISHER test has a probability equal to zero, i.e. the two parameters cannot equal zero simultaneously. The Coefficient of Determination Equals 0.86, considered as closer to the one of a linear relationship. In the contrary, The Durbin-Watson test confirms the existence of a positive correlation (its value equals 0.03, closer to zero) [2].

#### *b. Panel data estimation where the common coefficient is constant*

In this case, the model's equation is written as:

$$Car_i = S * e^{\alpha * e^{\beta_i * PIB_i}} \quad (15) \text{ Or}$$

$$\log\left(\log\left(\frac{S}{Car_i}\right)\right) = \log(\alpha) + \beta_i * PIB_i \quad (16)$$

With: The parameters  $S$  et  $\alpha$  are common for all the countries, by contrast of the parameter  $\beta_i$  which is estimated for every country. The obtained results are represented in the following table: For Algeria, the car fleet will be modeled by the equation:

$$\text{car} = \begin{cases} 0,850 * e^{\alpha * e^{-7,09 * 10^{-05} * \text{PIB}}} \\ \alpha = \exp(1.587995) = 4.894 \end{cases} \quad (17)$$

Where the  $\beta_i$  estimated by the model are included in the interval:  $-1,82 * 10^{-06}$  (Portugal) and  $-1,63 * 10^{-04}$  (China). The parameter of Algeria is closer to the Portugal one than China, which confirms the position of Algeria in the Gompertz curve. Moreover, the results of the model concerning Algeria with the two estimation methods are :

$$\text{Car} = \begin{cases} 0,850 * e^{4,894 * e^{-7,09 * 10^{-05} * \text{PIB}}} \\ 0,850 * e^{4,664 * e^{-10,3 * 10^{-05} * \text{PIB}}} \end{cases} \quad (18)$$

From this, we can draw the following remarks: The estimated parameters have closer values and  $\alpha$  et  $\beta$  have opposite evolutions. In addition, the model selection, for these two equations, is based upon the statistical significance of the parameters and the quadratic error square sum. The STUDENT test is verified for the two parameters concerning Algeria, we notice the probability of each parameter equals zero, so we accept the hypothesis of the two parameters being significant. The FISHER test has a null probability, meaning all the parameters cannot be null at once. Determination Coefficient Equals 0.965, obviously better than the first model (0.86). We notice a diminution in the DURBIN-WATSON test coefficient. It increased from 0.03 in the first case to 0.12 in the second one. The Error Squared Sum improved (diminished by 162 to 40.41). So the second model provides better significance than the first, and consequently, the second one will be compared with the following models.

### c. Panel data estimation with all coefficients varying

In this case, the model's equation is written as: In this case, the model's equation is written as:

$$\text{Car}_i = S * e^{\alpha_i * e^{\beta_i * \text{PIB}_i}} \quad (19) \text{ Or:}$$

$$\log\left(\log\left(\frac{S}{\text{Car}_i}\right)\right) = \log(\alpha_i) + \beta_i * \text{PIB}_i \quad (20)$$

With  $S$  is a common parameter for all the countries; however parameters  $\alpha_i$  et  $\beta_i$  are estimated for each country. For Algeria, the car fleet will be modeled by the equation:

$$\text{Car} = \begin{cases} 0,850 * e^{\alpha * e^{-10,1 * 10^{-05} * \text{PIB}}} \\ \alpha = \exp(1.784764) = 5.958 \end{cases} \quad (21)$$

With parameters included in the following intervals:

$\alpha_i$ : 1,3038 (Canada) et 3,75 (Greece).

$\beta_i$ :  $-33,5 * 10^{-05}$  (India) et  $-2,35 * 10^{-05}$  (Argentina). Algeria's parameter is closer to the Argentina one than the Indian, which confirms the position of Algeria in the Gompertz curve.

Moreover, the results concerning Algeria with the two last estimation methods are:

$$r = \begin{cases} 0,850 * e^{4,668 * e^{-10,3 * 10^{-05} * \text{PIB}}} \\ 0,850 * e^{5,958 * e^{-10,1 * 10^{-05} * \text{PIB}}} \end{cases} \quad (22)$$

Hence, we can conclude: The estimated parameters have closer values with a certain convergence of the parameter  $\beta$  toward a value of  $-10,1 * 10^{-05}$ . The parameter  $\alpha$  has a value

closer to 5 (5,958 et 4,664). Given the two equations having identical parameters, they will be used both to give two scenarios of the Algerian car fleet evolution in function of the GDP. The highlighted results show that estimations are better for the second and third model as the precedent section. In addition, we see evidence of  $R^2$  closer to 1 when passing from the first model to the third one. The sum of squared errors decrease, in its turn, for the third model. To conclude, we can say that the use of panel data for which are not common provides the best fit. Moreover, these four studied panels provide three evolutive scenarios of the Algerian car fleet. The graphical representation of these scenarios is illustrated in the (Appendices Fig.1).

*P1* offers a slightly remote curve relatively to other panels. This is due mostly in the country choice of the studied panels. The P1 panel comprises all the countries of the sample, by contrast of the other panels which have countries profile-similar to Algeria. The panel curves P2, P3 and P4 are almost identical.

The dashed lines delimit the three zones where the car fleet evolution changes its growth. That means when the GDP per capita exceeds 7,000 \$, the car fleet increase quickly than when it exceeds 30,000 \$, the car fleet will have a decreasing growth. Finally, the saturation is reached beyond 35,000 \$. Given the fact that Algeria is situated in the upper limit of the the second zone. This figure is of a high importance for the national institutions which manage the transportation sector, because it gives a future image on the size of the fleet when the country becomes richer. In order to establish a maximum scenarios, we opted to redo the same steps, but this time the curve not a Gompertz-type but a logistic and quasi-logistic ones defined by the equations:

$$\begin{cases} \text{Quasi - logitique} & : \text{Car} = \frac{S}{1 + \alpha \cdot \text{PIB}^\beta} \\ \text{Logitique} & : \text{Car} = \frac{S}{1 + \alpha \cdot e^{\beta \cdot \text{PIB}}} \end{cases} \quad (23)$$

$$\begin{cases} \text{Quasi - logitique} & : \log\left(\frac{S}{\text{Car}} - 1\right) = \log(\alpha) + \beta * \log \text{PIB} \\ \text{Logitique} & : \log\left(\frac{S}{\text{Car}} - 1\right) = \log(\alpha) + \beta * (\text{PIB}) \end{cases} \quad (24)$$

The results obtained by the application of these equations, depend on the other two cases, quasi-logistic and logistic respectively we chose the P2 and P4 for calibration. The same remarks made on the previous sections apply to the quasi-logistic model. This application gives the best statistical significance. Consequently they will be used for our car fleet forecasts.

The (Fig.2) represent the Algerian car fleet evolution scenarios. Whichever the panel chosen, we remark that all the curves are almost identical for GDP per capita level lower than 12,000 \$. Then the difference between the curves becomes obvious for high GDP per capital levels. For the case of Algeria (actual GDP per capita is about 7,700 \$), all the curves give the same future projection of the car fleet. We can assert that if our GDP doubles its value, the quasi-logistic model is insensitive in the choice of the panel.

By contrast, the inconvenient of this model lies on his inability to give a S-curve, meaning the saturation will not be reached for GDP per capita higher than 40,000 \$. Hence, the case of the United States is not well represented by the quasi-logistic model. the second and third models give the best statistical signification. Identically, they will be used in this study. The following curve (Figure.3) illustrates the application of these models for different chosen panels. We notice

that the P1 curve (all the world) is far from the others. The *P2*, *P3* and *P4* have their curves increasing quickly between lower GDP and higher GDP curvatures. Thus, there is a trend which all the countries (comprising Algeria) will reach once their GDP per capita passes from 10,000 to 17,000 \$. This trend is not verified for countries with high GDP levels and consequently these curves will be omitted in this study (Appendices Fig. 3).

Finally, the three models (Gompertz, Quasi-logistic and logistic) are reported in the same graphic. We notice the Gompertz curve has a median curve for the three models where the lowest GDP peak is identical for the two other models. For the case of Algeria and any other country in the first zone, the three models produce a spindle of curves that will be the next future possible scenarios for us (Appendices Fig. 4).

The difference between the three models becomes more stressed once we reach high GDP values. The quasi-logistic and logistic models become considerable compared to the reality and not significant. For this reason, we found that the literature advocates the use of the Gompertz model. Given our interest in low GDP countries, we extended the first zone of the following figure. Beyond the actual GDP per capita (7,700\$), the Algerian car fleet might evolve according to the different scenarios illustrated by the curve (Fig.4). We have two scenarios from the Gompertz curve representing the most probable evolution. The three scenarios representing the most pessimistic cases are coming from the logistic curve. Finally, the other scenarios from the quasi-logistic curve present the most optimistic ones (Fig. 4).

#### 4. Conclusions

The models provided by the academic literature related to the evolution of car fleet in the countries permitted to identified, in an acceptable manner, the most important parameters that affect, significantly, the ownership of vehicles. It shows that the wealth generated by the country as measured by GDP remains an important determinant of modeling car fleet at the expense of other parameters involving the price of vehicles, fuel, size of the transport network ... etc. This finding is reinforced by the fact that at the international level, the evolution of the car fleet in terms of GDP following the curves more or less identical across countries [5].

Excluding variables characterizing the behavior of individuals in car ownership, this study allowed us to highlight the influence of national wealth on the level of GDP over car fleet in Algeria. The objective is to analyze the distribution of vehicles according to GDP of different countries in which information is available on the World Bank and IMF databases. Thus, it will be possible to develop scenarios for the car fleet while considering the country as sections by panel data modeling.

We designed our calculation methodology by taking into account the available data from which this model was possible. We used the level of car park per 1000 inhabitants as the independent variable and GDP per 1000 inhabitants as the dependent variable. However, given the long period of study required by the modeling, we needed a 30 years history. Moreover, the theoretical models used in the study include the Gompertz equation, quasi-logistic and logistic. This choice is justified by the fact that the evolution of the car fleet in terms of GDP has a look of an S-curve with saturation at high GDP close to 1[6]. Since the panel data estimation is more appropriate in our case, the previous models were calibrated using the following assumptions: the parameters are constant for all panels, a certain number of parameters are constant for all countries and finally, all parameters are variable by country. The estimate of these last is by

transforming the models used in the linear form, without introducing changes in the equations. Models with constant parameters were statistically less significant than those defined by the other two hypotheses. In addition, we varied the countries studied in the different panels with the maintenance of the United States, India and China in all panels to see their influence on the estimated parameters. The conclusion obtained from our study is that the Gompertz curve is the most suitable for this model. With this curve, we were able to generate Algerian car fleet scenarios, which is a result of particular importance for the planning of transport infrastructures and refining capacity.

These scenarios were developed to forecast beyond 2002 [8]. At that time the national wealth was estimated at \$ 7,700 per capita and the fleet was 56 cars per thousand inhabitants. According to the current trend of GDP growth is 5%, we estimated the park in 2020 while admitting that Algeria's GDP would increase by 5% per year. In 2020, Algeria had a fleet size of 230 cars per thousand inhabitants and a GDP of \$ 18,000 per capita. Finally, the recommendations of this study are as follows [8] :

- The models used could be improved by incorporating other variables affecting the car fleet.
- The results could be compared with models of non-aggregated i.e. estimating each car type separately. The latter method requires the availability of data.

## References

- BOURBONNAIS. Régis, *"Econométrie,"* 6<sup>éd</sup> Edition, Dunod , Paris, 2005.
- CASIN. Philippe (2009) *"Econométrie, Methodes et Applications avec Eviews,"* 1<sup>st</sup> Edition, Technip, Paris.
- CHRISTOPHE. Hurlin, *"Econométrie des données de panel modèle linéaire simples"* Ecole Doctorale Edocif Méthodologie, Paris, 2004.
- DAMODER N.GUJARATI, *"Econométrie"*, 4<sup>rd</sup> éd , traduction par Bernard Bernier, Déboék , Paris, 2004.
- GREENE. William, *"Econométrie,"* 5<sup>éd</sup> , Edition Française dirigée par Didier Schlachter IEP, New York University, 2005 .
- AZAM Amjad Chaudhry "A panel data analysis of electricity demand in Pakistan," *In the labore Journal of Economics*, 2010.
- BOUACHERA. Taoufik, MAZRAATI. Mohamed "Fuel Demand and Car Ownership Modeling in India," *In: Algerian Petroleum Institute (Algeria)*,2006.
- BOURDEAU Béatrice, " Evolution du parc automobile Français entre 1970 et 2020 ," *thèse Doctorat et Ingénierie de l'environnement, à Université de Savoie-Chambéry, France Dundee*, 1996 .

Appendices

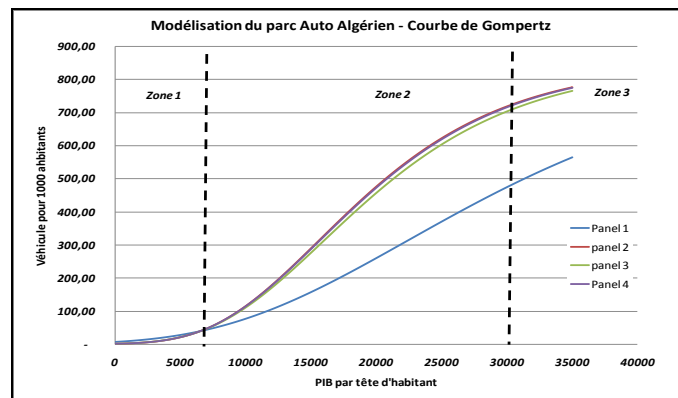


Fig.1 : Projection of the Algerian car fleet

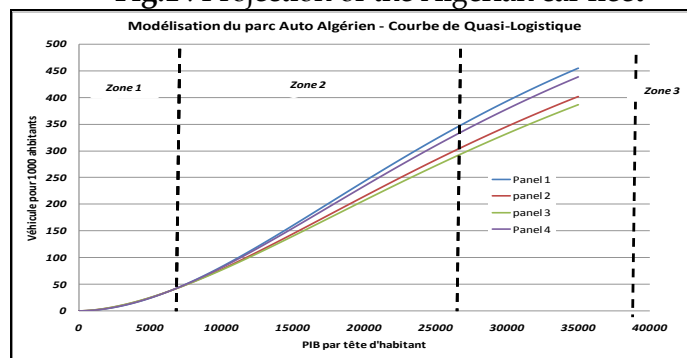


Fig.2: Algerian car fleet projection (Quasi-logistic)

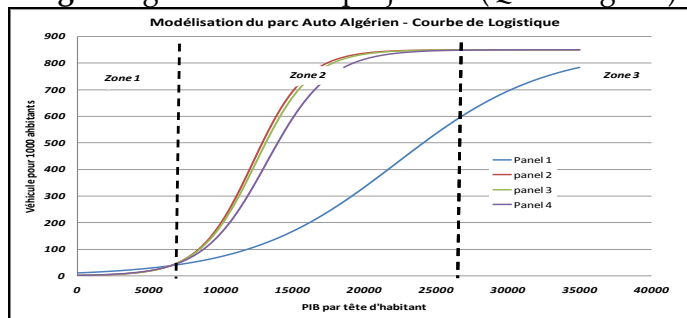


Fig.3 : Algerian car fleet projection (Logistic case)

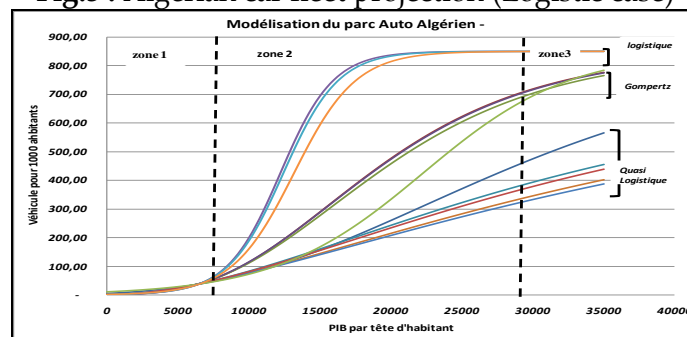


Fig. 4: Algerian car fleet projection (three models)