Modeling of the behavior of alternative fuel vehicle buyers. A model for the location of alternative refueling stations

C. Fúnez Guerra a,*, Ricardo García-Ródenas b, E. Angulo Sánchez-Herrera c, Doroteo Verastegui Rayo c, C. Clemente-Jul d

a National Hydrogen Centre, Prolongación Fernando el Santo s/n, 13500, Puertollano, Spain
b Escuela Superior de Informática, Universidad de Castilla la Mancha, Paseo de la Universidad, 4, Ciudad Real, Spain
c Escuela de Ingeniería Minera e Industrial de Almadén, Universidad de Castilla-La Mancha, Plaza Manuel Meca S/N, Almadén, Ciudad Real, Spain
d Department of Energy and Fuels, School of Mining and Energy Technical University of Madrid (UPM), Rios Rosas, 21, Madrid, Spain

A B S T R A C T

This paper addresses the problem of estimating the infrastructure to be made available for refueling alternative fuel vehicles as a function of the profitability thresholds required by the investment. A methodology has been devised based on sales forecasts for alternative fuel vehicles. These methods use discrete choice models in which the factor of refueling infrastructure, rather than being considered simply as one more attribute of the model, acts as a constraint on the choice set for vehicle buyers. This methodology is used to estimate the infrastructure of hydrogen refueling stations and electricity charging stations for Spain (8,112 population centers) in 2030. Evolution of fuel cell vehicles over the years 2016 and 2030 is also estimated and compared with forecasts for countries such as France, Germany and the United Kingdom.

Introduction

The transportation sector is responsible for a large share of the European Union greenhouse gas (GHG) emissions, and consequently it is a central goal of the European Commission in sustainability development strategies. The European Union is attempting to replace 10% of conventional fuels with biofuels, hydrogen and ecologically-sourced electricity before the year 2020 [16] and for 2050 has set the much more ambitious target of 60% reduction in emissions of polluting gases [10].

The use of Alternative Fuel Vehicles (AFV) to replace vehicles powered by internal combustion, is an alternative form of road transport that may provide, in the long term, reduction in GHG emissions and improvement in air quality in cities [19,21].

European Union member countries have decided to implement programs to further accelerate the introduction of
AFV vehicles. The European Directive 2014/94/EU [17] establishes a common framework of measures for the deployment of alternative fuels infrastructure in the European Union. One of the main problems which must be addressed in the roll-out of the necessary infrastructure for the use of alternative fuels in transport is the so-called Chicken- or Egg dilemma [13, 21]. Consumers are not keen to buy alternative-fuel vehicles while the refueling infrastructure is hard to find, and refueling points will only become common when there are enough vehicles to make it worthwhile.

This study assumes that private enterprise will finance the roll-out of the refueling infrastructure if it considers that the investment required will reach a certain profitability threshold. This assumption removes the dilemma of precedence between supply and demand, and raises the question of whether the supply will generate sufficient demand, that is, whether the chicken will or will not lay an egg. Thus we call the approach the chicken or rooster problem as addressed in this paper.

The stated preference methods are the tools most widely used to analyze relevant factors in the market penetration of AFV. These studies use the marginal willingness to pay (WTP) and the analysis of market share in different scenarios to weight each of the factors. Martin Achtnicht and Hermeling [13], Hoen and Koetse [7] review these methods and show that availability of refueling stations is an essential factor. Hackbarth and Madlener [4]; Ito et al. [9] use this methodology to assess potential demand for infrastructure investment for AFV vehicles.

The literature considers the infrastructure as just another attribute of the utility of vehicles buyers. This paper, on the other hand, considers that the infrastructure limits the so-called universal choice set. We adopt a two-stage representation of decision making. In the first stage, the choice-set generation is simulated. The car buyers screen alternatives and eliminate those in which it is economically impossible to introduce the required infrastructure where they live. In the second stage, the buyers choose only from the alternatives remaining in the reduced choice set [12]. Finally, a sensitivity analysis of the profitability thresholds for the necessary investment in rolling out the infrastructure will allow the typology of the feasible infrastructure in the study area to be determined.

### Forecasting of alternative fuel vehicles sales by using discrete choice models

**A basic approach**

This section describes the basic method\(^1\) to predict the evolution of AFV as a function of the roll-out of infrastructure. This method assumes that vehicle sales satisfy the equation:

\[
 v_t^i = P_t \cdot \phi(\Delta GDP_t) \cdot \pi(i/t) \tag{1}
\]

where \( v_t^i \) is the number of vehicles of type \( i \) sold in period \( t \) (in thousands of units), \( \phi(\Delta GDP_t) \) is a multiplier which transforms the variation of GDP per capita in period \( t \) into vehicles sold per thousand inhabitants and \( \pi(i/t) \) represents the market share of vehicle type \( i \) in period \( t \).

This method has been applied to Spain. Three scenarios are considered (one baseline case, one optimistic and one pessimistic) consisting of a combination of forecasts of population development \( \{P_t\}_{t=2016}^{2030} \) and of the \( \{\Delta GDP_t\}_{t=2016}^{2030} \). The population estimate was obtained from the INE (Spanish Statistical Office). The baseline case considers a final population of 45.4 million, 44.3 in the pessimistic case and 46.5 in the optimistic case. The left of Fig. 1 shows the estimate of the multiplier \( \phi(\Delta GDP) \) using a Gaussian radial basis approximation and on the right of Fig. 1 the forecast for total vehicles sales \( \sum_i v_t^i = P_t \cdot \phi(\Delta GDP_t) \) for the three scenarios.

This study considers seven types of vehicle \( i \) that are already available on the Spanish market, or will be in the near future. We have considered gasoline and diesel (conventional technology, CT), biofuel vehicles (BV), natural gas vehicles (NGVs), hydrogen (fuel cell electric vehicles, FCEVs). There are

\(^1\) The data used in this paper can be found at the following web address http://cort.as/f5Vw.
also hybrid electric vehicles (HEVs), plug-in hybrid-electric vehicles (PHEVs), and fully battery-electric vehicles (BEVs).

The stated choice methods allow market share to be predicted \( P(i/t) \) as a function of a vector of attributes \( s \) of the set of vehicle types \( i \) considered in the study. Findings from the existing literature on AFV preferences show that next to purchase price and operating costs, driving range \([5,11,14]\) and fuel availability \([8,18]\) may have substantial effects on consumer preferences for AFVs. Emission reduction is also flagged as an important factor (see Refs. [6,18]). For these reasons, this study considers the vector of attributes \( s \) to consist of:

- **Purchase Price (PP)** is the money to be paid by the consumer to buy a vehicle (in €).
- **Fuel Cost (FC)** is money that a user spends to drive 100 km of distance (in €/100 km).
- **CO\(_2\) emissions (CO\(_2\))** is the amount of CO\(_2\) emitted by vehicles to travel 1 km of distance (in gr CO\(_2\)/km).
- **Driving Range (DR)** is the maximum distance that a vehicle can travel with a load to 100% of the fuel use (in km).
- **Fuel Availability (FA)** is the number of filling stations that the user has available to refuel (in % total existing filling stations).
- **Refueling Time (RT)** is the time that a vehicle takes to refuel from 0% to 100% its capacity (in minutes).

The most basic discrete choice model is the logit model, which expresses market share as a function of the attributes as:

\[
P(i|A, s) = \frac{\exp(V_i(s))}{\sum_{j\in A} \exp(V_j(s))}
\]

where \(V_i(s)\) is a function (linear in most studies) of the attributes, and \( A \) is the choice-set for a vehicles buyer (types of vehicle).

Most studies that use discrete choice models are cross-sectional studies. This paper deals with a longitudinal study and requires forecasts for the evolution of the different attributes with time, that is \(s_{t+1} = s_t + \Delta s_{t+1}\).

The forecast for the future of the attributes is based on Prospective Technological Studies\(^2\) for the different types of vehicles concerned shown in Fig. 2.

Hackbarth and Madlener [3] analyze the potential demand for AFV vehicles using stated-preference discrete-choice data and a mixed multinomial logit model (MMNL). This paper uses the model and parameters that the authors obtained with a sample of 711 respondents. The longitudinal market shares are shown in Fig. 3. The most important observation is that the attribute FA\(_t\) = 0 for the AFV in the year \( t = 2016 \) (it measures the roll-out of the infrastructure), nevertheless the market share is not zero. This effect will be dealt with in the next subsection using a constrained mixed multinomial logit model (CMMNL).

### A constrained approach

The set of alternatives \( A \) available to a buyer is constrained by the refueling infrastructure typology he/she can access.

For example, if a buyer has no possibility of refueling with hydrogen then the option of purchasing a FCEV is not feasible.

The paper considers that a buyer only has access to the infrastructure of the town where they live. This implies that the probability of purchasing a vehicle of type \( i \) depends on where the buyer lives.

The paper focuses on analyzing the roll-out of refueling infrastructure for hydrogen and electric vehicles. Thus, we assume the existence of a current infrastructure serving the other vehicle types and we wish to assess the introduction of refueling infrastructure for these two types of vehicle against the others. This premise results in four typologies of refueling infrastructure (henceforth TIR) which we shall call \( j \). We associate the value \( j = 1 \) with an infrastructure which has all possible fuels, the value \( j = 2 \) to hydrogen vehicles + current infrastructure, the value \( j = 3 \) to electric vehicles + current infrastructure, and finally \( j = 4 \) to the current infrastructure. The sets of alternatives associated with each typology are:

- \( A_1 = \{CT, NGV, HEV, PHEV, BEV, BV, FCEV\} \)
- \( A_2 = \{CT, NGV, HEV, BV, FCEV\} \)
- \( A_3 = \{CT, NGV, HEV, BEV, FCEV\} \)
- \( A_4 = \{CT, NGV, HEV, BV\} \)

Discrete choice models, when the attributes of all the alternatives present in the choice set are known, can be used to calculate the probabilities of acquiring a vehicle of type \( i \) for a TIR \( j \) in period \( t \),

\[
P(i|A_j, s^t) \quad \text{for all } i \in A_j
\]

We shall now determine which TIR are acceptable in town \( c \). A TIR \( j \) is feasible for a town \( c \) if it exceeds the profitability thresholds (break-even) for the types of fuel it includes. The criterion used to assess profitability of a refueling station is the number of vehicles per service station.

Currently in Spain, according to data from the Asociación Española de Operadores de Productos Petrolíferos, the number of service stations has grown from 8,622 in 2002 to 10,712 in 2014. Likewise, the number of cars in Spain, according to the INE dates, has increased from 18,688,320 in 2002 to 22,029,512 in 2014. From these data, it can be seen that each refueling station serves an average of 2.25 vehicles (in thousands of units). This quotient, number of vehicles per filling station, is indicated by \( b \) and is what we call breakeven. In towns where the number of vehicles per service station is higher than this breakeven value \( b \) it would be worth considering a service station if one does not exist.

Mathematically, this expressed as:

\[ \text{Feasibility } A_1 : \mathbf{x} \cdot \mathbf{u}_1^i \cdot P(\text{FCEV}|A_1, s^t) \geq b_{\text{FCEV}} \]  
\[ \text{Feasibility } A_2 : \mathbf{x} \cdot \mathbf{u}_1^i \cdot P(\text{PHEV}|A_1, s^t) + P(BEV|A_1, s^t) \geq b_{\text{PHEV+BEV}} \]  
\[ \text{Feasibility } A_3 : \mathbf{x} \cdot \mathbf{u}_1^i \cdot P(\text{PHEV}|A_3, s^t) + P(\text{BEV}|A_3, s^t) \geq b_{\text{PHEV+BEV}} \]  

\[ \text{Feasibility } A_4 : \mathbf{x} \cdot \mathbf{u}_1^i \cdot P(\text{BEV}|A_4, s^t) \geq b_{\text{BEV}} \]

\(^2\) More information at the link http://cort.as/f5Vw.
Vehicle Attributes

Fig. 2 – Attributes of vehicles between 2016 and 2030 (\(t\)).

Feasibility \(A_3\): always

Feasibility \(A_4\): always

where \(t\) is the vehicle life expectancy, \(\nu_c^t\) is the number of vehicles sold in town \(c\) in period \(t\). The left of the constraint (4) is an estimate of the number of vehicles of type FCEV which would exist in the equilibrium state. That is, the number of vehicles of type FCEV sold in a year multiplied by the average life of these vehicles. The constraint (5) considers that hybrid vehicles PHEV refill the fraction \(a\) of their fuel in the refueling infrastructure, and this constraint represents the economic viability of this type of infrastructure. In order for the TIR \(A_1\) to be viable both constraints (4) and (5) must be satisfied simultaneously.

The numerical trials found that the 8,112 Spanish towns are distributed among the following combinations of feasible TIR: \(\{A_1, A_2, A_3, A_4\}, \{A_2, A_3, A_4\}, \{A_2, A_4\}, \{A_4\}\).
shows that there are towns which could support any refueling infrastructure while in others only the current infrastructure is viable.

The next issue is to determine which of the feasible TIRs for a town c is introduced in a period t, this TIR is indicated by \( TIR_c \). To solve this requires defining the mechanisms operating in the companies (private initiative). We assume that companies, with or without public incentives, will create the infrastructure if the expected return to investment is positive (measured by the breakeven threshold b). A corollary of this mechanism is that companies will do business as long as it is possible to do so, and so the infrastructure in each town will grow to the maximum economically viable (feasible TIR).

Table 1 shows the application of this corollary to each situation. What is most noticeable is that there is a size of town for which it is only possible to introduce infrastructure for hydrogen or electric vehicles, and its introduction makes the other type economically non-viable. To remove this indeterminacy involves consideration of technological neutrality, chance, etc. We have simplified the question and in the numerical trials we have considered introducing typology \( TIR_3 \) in these cases.

Finally, the estimate of vehicle sales of type i for each period t is given by the expression:

\[
u^t_i = \sum_c u^t_c \cdot P(i|u^t_c, s^t) \quad (9)
\]

Note that for \( i \in TIR_3 \) gives \( P(i|u^t_c, s^t) = 0 \).

We have applied these two methodologies (MMNL and CMMNL) to Spain, and for the three scenarios of vehicle sales.

We have considered as a base discrete choice model the mixed logit model described in the article Hackbarth and Madlener [3]. The study considers \( t = 25 \) years and \( a = 0.5 \) and the breakeven threshold \( b_{FEV} = b_{PREV} = 7 \) (in thousands of vehicles). To get this we analyzed public-private partnership Mobility studies, such as for example Association Francaise pour l Hydrogene et les Piles a Combustible [1]; McKinsey Company [15]; Transnova [20]; and Germany Trade Invest [2]; which show the number of hydrogen vehicles, of hydrogen refueling stations, and hydrogen consumption. We have estimated the parameter \( u^t_i = \nu^t_i \cdot p_i \), where \( \nu^t_i \) is the set of total vehicle sales and \( p_i \) is the proportion that the population of town c in 2016 represents in the total. The results are shown in Fig. 4. It can be seen that the basic approach overestimates vehicle sales for which there is no TIR (FCEV, PHEV, BEV). The constrained approach corrects this.

Table 2 shows the number of FCEV vehicles estimated in 2030 for Germany, the United Kingdom and France and compares them with the estimates obtained for Spain using the CMMNL for the 3 scenarios. We have standardized the estimates obtained in the different studies by country population in 2016 (fourth column). Column 5 compares these standardized measures to the base estimate CMMNL. The estimates obtained with the CMMNL for Spain are seen to be similar to those for France, a country with a more moderate plan for the introduction of FCEV, which is in agreement with what is expected to happen in Spain.

The estimates for the number of FCEV in Spain in 2030 using the MMNL for the three scenarios are: 1.49, 1.63 and 1.83 million vehicles. These values lead to indices \( \mu = 0.0321 \), \( \mu = 0.0351 \) and \( \mu = 0.0394 \) well above the other countries. This shows the need to use the CMMNL to avoid overestimating with the MMNL.

In order for the figures for vehicles sold to agree with those predicted by the model, private initiative must be aware of business opportunities (the basic assumption of the approach) and the government must make introduction easier (by legally guaranteeing the investment through concessions, subsidies, etc.). The constrained MMNL overestimates AF vehicle sales for 2016 (see Fig. 4) because the refueling infrastructure for hydrogen and electric vehicles is currently negligible and so the basic assumption is not met.

### Table 1 – Typology of towns by feasible TIR.

<table>
<thead>
<tr>
<th>Feasible TIR's for town c</th>
<th>TIR's to be introduced in town c</th>
<th>TIR's introduced in town c</th>
</tr>
</thead>
<tbody>
<tr>
<td>( TIR_1, TIR_2, TIR_3, TIR_4 )</td>
<td>( TIR_1 )</td>
<td>( TIR_2 ) or ( TIR_3 )</td>
</tr>
<tr>
<td>( TIR_2, TIR_3, TIR_4 )</td>
<td>( TIR_2 )</td>
<td>( TIR_2 )</td>
</tr>
<tr>
<td>( TIR_3, TIR_4 )</td>
<td>( TIR_3 )</td>
<td>( TIR_3 )</td>
</tr>
<tr>
<td>( TIR_4 )</td>
<td>( TIR_4 )</td>
<td>( TIR_4 )</td>
</tr>
</tbody>
</table>

**Sensitivity analysis of breakeven parameter for the roll-out of alternative fuel infrastructure**

This section presents a methodology whose aim is to accelerate the introduction of alternative fuel infrastructures. Private initiative can determine in which towns, when and with which associated TIR to decide on their participation in this process, while the government can evaluate incentives permitting the development of a sustainable network which encourages the private sector to anticipate the investment.

This section analyzes both points of view. On the one hand, it is considered that the incentives allow variation in the breakeven value (decreasing it), while on the other hand the decision makers assess subsidy policy in terms of the capacity to reduce emissions, energy dependence, etc. These matters are closely related.
Table 2 – Comparison of the forecast for the number of fuel cell vehicles for the year 2030 in different EU countries.

<table>
<thead>
<tr>
<th>Country i</th>
<th>Program</th>
<th>FCEV in 2030 (in millions)</th>
<th>$\mu_i$</th>
<th>100 $\times \frac{\mu_i}{\mu}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germany</td>
<td>Now Germany Mobility</td>
<td>1.87</td>
<td>0.0228</td>
<td>121.35%</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>UK H2 Mobility</td>
<td>1.27</td>
<td>0.0194</td>
<td>88.34%</td>
</tr>
<tr>
<td>France</td>
<td>H2 Mobility France</td>
<td>0.773</td>
<td>0.0120</td>
<td>16.5%</td>
</tr>
<tr>
<td>Spain</td>
<td>Optimistic CMMNL</td>
<td>0.56903</td>
<td>0.0123</td>
<td>19.41%</td>
</tr>
<tr>
<td>Spain</td>
<td>Base CMMNL</td>
<td>0.47562</td>
<td>0.0103</td>
<td>0%</td>
</tr>
<tr>
<td>Spain</td>
<td>Pessimistic CMMNL</td>
<td>0.42353</td>
<td>0.0091</td>
<td>-11.65%</td>
</tr>
</tbody>
</table>

Where $\mu^*$ is the index $\mu$ for Base CMMNL.

A sensitivity analysis will be carried out on the breakeven parameter $b$ assuming that it has the same value in both types of infrastructure, that is: $b_{\text{FCEV}} = b_{\text{HEV+BEV}}$. Fig. 5 shows the TIR of Spanish towns in 2030 as a function of the number of inhabitants. Looking at the graph on the left, it can be seen that for a value $b = 7$ the TIR $\mathcal{A}_1$ is introduced in towns of more than 167,008 inhabitants, the TIR $\mathcal{A}_2$ between 84,789 – 138,048, the TIR $\mathcal{A}_3$ in the range 145,967 – 165,287. This supposes that in the case of Spain there would be 38 towns with infrastructure $\mathcal{A}_1$, 30 with $\mathcal{A}_2$, 7 for $\mathcal{A}_3$ and the rest 8037 keep the current infrastructure $\mathcal{A}_4$. The right of Fig. 5 shows sales of vehicles FCEV, BEV + PHEV as a function of the parameter $b$. When the parameter $b$ is increased the introduction of the infrastructure is reduced and so are the sales of those types of vehicle. Competition between alternative infrastructures can be seen for the value $b = 11$. An increase in this value of 10–11 (for the optimistic scenario) do not cause sales FCEV vehicles to reduce. In fact they rise because the electric vehicle charging infrastructure is removed from a town, and substituted for hydrogen fueling infrastructure.

To our knowledge there is no study in the literature which analyzes the introduction of refueling infrastructure (in Spain) in terms of thresholds of economic viability, and thus we have been unable to contrast our results.

From the point of view of private initiative it is vital to determine places, typologies and timings that are best for locating an alternative fuel station as a function of the its parameter $b$. Fig. 6 shows two situations of Spanish refueling infrastructures in 2030 for two values of the parameter $b$. 

Fig. 4 – Sales forecasts according to the two methodologies (MMNL and CMMNL) for the three scenarios analyzed.
Fig. 5 – TIR deployments versus city’s inhabitants (base scenario) and sales forecasts for the year 2030 as a function of the breakeven $b$ (for the three scenarios analyzed).

Note that it is possible to perform a sensitivity analysis not only the key parameter model breakeven ($b$) but any vehicle attributes considered. Table 3 shows the difference in total vehicle sales between the base case and a new scenario defined by the 50% trend increase in the fuel costs for CT, NGV, HEV and BV vehicles. It is noted: i) that although the fuel costs increased by 50% for CT, NGV, HEV and BV only CT and HEV vehicles decreased in sales; and ii) the MMNL constrained regarding MMNL reflects the lack of effect of this incentive to purchase an AF vehicle in small cities where the adequate infrastructure is not available.

### Table 3 – Differences in vehicle sales (in thousands) for the period 2016–2030 between the base case and the scenario defined by 50% the Fuel Cost for the vehicles CT, NGV, HEV, BV.

<table>
<thead>
<tr>
<th>Model</th>
<th>CT</th>
<th>NGV</th>
<th>HEV</th>
<th>PHEV</th>
<th>BEV</th>
<th>BV</th>
<th>FCEV</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMNL</td>
<td>-378.3</td>
<td>42.7</td>
<td>-178.7</td>
<td>184.3</td>
<td>65.1</td>
<td>42.0</td>
<td>222.9</td>
</tr>
<tr>
<td>CMMNL ($b = 3$)</td>
<td>-345.4</td>
<td>110.9</td>
<td>-154.8</td>
<td>114.0</td>
<td>42.2</td>
<td>89.5</td>
<td>143.6</td>
</tr>
<tr>
<td>CMMNL ($b = 7$)</td>
<td>-332.4</td>
<td>131.8</td>
<td>-166.1</td>
<td>90.7</td>
<td>33.1</td>
<td>103.9</td>
<td>119.0</td>
</tr>
<tr>
<td>CMMNL ($b = 20$)</td>
<td>-302.9</td>
<td>165.3</td>
<td>-127.4</td>
<td>46.9</td>
<td>15.8</td>
<td>126.7</td>
<td>75.6</td>
</tr>
</tbody>
</table>

Conclusions

This paper introduces a methodology for estimating sales of alternative fuel vehicles by stated choice methods. The key aspect of this approach is to introduce the infrastructure of the refueling network as a procedure which restricts the decision set for vehicle buyers and not as one more attribute of the basic discrete choice model. The sensitivity analysis of this model allows both private and public interests to assess their decisions about refueling infrastructure.
We have analyzed the situation in Spain and have obtained a typology of the existing infrastructure, in space and in time, and estimates for fuel cell electric vehicles in 2030, the results being comparable with those for France. We have proved numerically that the CMML corrects the overestimates of the MMNL.

Acknowledgments

This research was supported by the Spanish Ministry of Economy and Competitiveness of Spain-FEDER EU grant with number TRA2014-52530-C3-2-P and the Spanish Ministry of Economy and Competitiveness, Castilla-La Mancha Government and the European Regional Development Fund (ERDF) 2007–2013 included in the “Programa Operativo de I + D + I por y para el Beneficio de las Empresas - Fondo Tecnológico”.

REFERENCES