Predicting the demand: Uncertainty analysis and prediction models in Spain

Anna Matas
José Luis Raymond
Mar González-Savignat
Adriana Ruiz
Project: SOCIO-ECONOMIC AND FINANCIAL EVALUATION OF TRANSPORT PROJECTS

Granted by Centro de Estudios y Experimentación de Obras Públicas (CEDEX)
Ministerio de Fomento
Ref. PT-2007-001-02IAPP

www.evaluaciondeproyectos.es
PREDICTING THE DEMAND: UNCERTAINTY ANALYSIS AND PREDICTION MODELS IN SPAIN

29 March 2009
Mar González - Savignat
Anna Matas
José Luis Raymond
Adriana Ruíz

Abstract

This working paper has two parts. In the first part, the main results from the literature on the level of error in predictions of the demand for transport will be presented. The second part is a summary of the prediction models applied in Spain by mode of transport.
# Tabla de contenidos

1. ANALYSIS OF THE UNCERTAINTY ASSOCIATED WITH PREDICTING THE DEMAND ...... 1
   1.1. LEVEL OF ERROR IN THE PREDICTIONS ................................................................. 1
   1.2. UNCERTAINTY IN PREDICTIONS OF THE DEMAND FOR TRANSPORT ................ 4

2. ANALYSIS OF THE PREDICTION MODELS BY MODE OF TRANSPORT .................. 6
   2.1. STATE PORTS .............................................................................................................. 6
   2.2. DIRECTORATE GENERAL OF MOTORWAYS .......................................................... 9
   2.3. DIRECTORATE GENERAL OF RAILWAYS ............................................................... 10
   2.4. AIR TRANSPORT ...................................................................................................... 10
   2.5. PLANS FOR TOLL ROADS AND MOTORWAYS ................................................... 12

3. DATABASES AVAILABLE IN SPAIN ........................................................................ 13

REFERENCES ................................................................................................................ 17
1. ANALYSIS OF THE UNCERTAINTY ASSOCIATED WITH PREDICTING THE DEMAND

Despite the thousands of studies published on predicting the demand in transport, there are actually very few articles that analyse the degree of error in these predictions. In this appendix we shall present the main results found in the literature, distinguishing between the studies that directly measure the level of error in the predictions and those that try to quantify the degree of uncertainty associated with both the explanatory variables and the model itself.

1.1. Level of error in the predictions

The survey of the literature focused on studies that examine a large number of projects and that therefore permit us to draw valid conclusions for our own study. This literature can be grouped around the studies by B. Flyvbjerg on the one hand and those by R. Bain on the other.

Bent Flyvbjerg, along with other authors, conducted a far-ranging study that yielded significant results on traffic predictions in transport infrastructure projects (Flyvbjerg, 2005; Flyvbjerg et al., 2005; Flyvbjerg et al., 2006; Flyvbjerg, 2008 and Naess et al., 2006).

These authors base their studies on 210 transport infrastructure projects in 14 different countries; 27 of them are related to railway investments, while the others are road projects. The period analysed spans from 1969 until 1998. The projects analyse the demand in urban railways, high-speed trains, conventional trains, tunnels, bridges and toll-free motorways.

The prediction error is defined as the percentage variation between the estimated traffic and the real traffic observed during the first year of operation. The estimated traffic corresponds to the prediction made when the decision was taken to build the project. Therefore, subsequent traffic estimates are not taken into consideration when one of the initial conditions has been modified or when additional information is available. Different results for the modes of transport analysed are derived from these studies.

In railways, the traffic is over-predicted in nine out of the ten projects studied. The average overestimate is 106%. The authors claim that there is a systematic bias in the traffic predictions in this mode of transport. The reasons for this deviation lie in the modal distribution model, which is excessively optimistic as to the railway’s capacity to capture a large share of the demand.
For *road* traffic, the prediction errors are also significant: in 50% of the projects the difference between the estimated and observed traffic is approximately ±20%. However, in this case there is no bias towards overestimation.

There is no significant difference for the standard deviation in the predictions of the demand for road and railway; in both cases these deviations are high, which indicates a high degree of uncertainty and risk.

On the other hand, there are no indications that traffic predictions have improved over time. Thus, according to these same authors, the improvements in the estimated demand models do not seem to have led to an improvement in the predictions themselves. Furthermore, it can be seen that the prediction errors have no relationship with the construction timeframe of the project and that there is only a weak relationship with the cost of the project.

Flyvbjerg (2005) holds that the choice of the project’s first year of operation to measure the prediction error does not influence the result and therefore that prediction errors remain steady over time. Finally, they confirm that the bias in predictions is related to the possibility of influencing the result of the investment.

As an alternative prediction method, the authors propose *Reference Class Forecasting*, which is based on the theory of decision-making under uncertainty developed by D. Kahneman. An application of this theory to the evaluation of investment projects can be found in Flyvbjerg (2008). The method consists of deriving predictions for a specific project by analysing other similar projects already carried out with the purpose of getting a distribution of probable results for the specific case being studied. Since this method is based on existing precedents, it is possible to obtain a higher degree of control over the systematic and deliberate errors or strategic biases that might conceal the political interests of certain investments.

For road investments, the studies conducted by Robert Bain investigate the prediction errors in toll roads and compare them to the prediction errors in toll-free motorways (Bain and Wilkins, 2002; Bain and Plantagie, 2004 and Bain and Polakovic, 2005). In these three articles, the authors examine the prediction errors for 32, 87 and 104 investment projects in toll roads, motorways, tunnels and bridges in numerous regions around the world. The prediction error is measured as the percentage difference between real and estimated traffic in the first year that the project is in operation.

Their results show that the estimates of the demand on toll roads tend to overestimate the amount of traffic they will have. Bain and Wilkins (2002) find that the demand is overestimated in 28 of the 32 cases. The percentage of overestimation is between 20% and
30%. On toll-free motorways, the prediction errors can be seen equally in both directions. Therefore, on toll-free motorways there is no systematic optimistic bias as there is with toll roads.

They also prove that the standard deviation is high and very similar in both toll roads and toll-free motorways. Once again, the high dispersion is indicative of the degree of uncertainty associated with traffic predictions. Eliminating the optimistic bias in motorway predictions shows a degree of error similar to that found for toll-free motorways.

Just as in the case of the study by Flyvbjerg, it can be seen that the accuracy of predictions does not improve over time. Therefore, neither the bias nor the prediction error can be attributed to the unforeseeable effects of the first year of operation. Finally, the variability in traffic predictions was higher for heavy vehicles than for cars.

Bain proposes an analytical tool, the Traffic Risk Index, which contributes to identifying and quantifying the main difficulties affecting traffic studies. This index is made up of a list of situations that influence errors in traffic predictions. The index can be used to examine the uncertainty in specific projects, while it also enables analysts to ask a series of questions that will help them to interpret the reliability of the predictions.

From another perspective in predicting the demand, Riddington’s study (2006) entails a criticism of the predictions of air traffic in the United Kingdom for the timeframe 2030. Even though this study does not analyse the degree of error in these predictions, it does shed light on the weaknesses in the models and variables used by the Department of Transport of the United Kingdom. His results point to the predictions’ lack of reliability due to methodological and econometric errors in the estimated model. Specifically, he points out that under the appearance of a multinomial logit model – taking population and generalised costs into account – which is apparently more complex, the predictions are adjusted by a relatively simple trend model, which furthermore is estimated incorrectly. The omission of relevant variables, the inadequate treatment of non-stationary series and ‘ad hoc’ assumptions about the GDP growth rate are just some of the problems analysed which lead the resulting predictions to lack reliability. Among other figures, Riddington points to the fact that the predicted 635% growth in the international market for the timeframe analysed is somewhat implausible. In general, the study concludes that the air traffic predictions are excessively optimistic and that no consideration is given to the existing restrictions on the capacity to handle this forecasted demand.

Likewise, in 2007 the USA’s Federal Aviation Administration conducted a programme that synthesises air traffic predictions. The report (Transportation Research Board, 2007) points out that few predictions in air transport have been subjected to an ex-post evaluation in
order to assess their degree of error. This report performs a general analysis of the relevant issues when predicting the demand, and it examines a range of issues like the information available and data gathering, the prediction methods, models and variables according to the type of predictions sought (macroeconomic, demographic, market, cost and technological, regulatory factors, restrictions on capacity, etc.) and their reliability.

The authors globally group these predictions into short-term and long-term, and those that consider the demand limited or not. Each of them can give rise to different predictions, and choosing the most appropriate one will depend on the objective being analysed.

They stress the importance of including uncertainty in the prediction in a systematic way and more formally than it is usually applied with the analysis of different growth hypotheses based on occasional estimates, which, the authors say, should be avoided. The following section contains a more detailed analysis of uncertainty in predictions.

1.2. Uncertainty in predictions of the demand for transport

Once the existence of significant prediction errors has been proven, we must inquire into the cause behind these errors and how they can be included in a project evaluation. Errors are essentially due to the uncertainty as to the evolution of the explanatory variables of traffic, or inputs, and the uncertainty associated with the estimated model. Below we discuss the results of the study on uncertainty conducted by Jong et al. (2007) based on a survey of 21 studies and on the development of their own application for an investment project in a Dutch motorway.

Given the fact that the future values of the exogenous variables are unknown, part of the prediction error is due to uncertainty as to the evolution of these variables. According to Jong et al (2007), the most common method for quantifying this type of uncertainty is some sort of Monte-Carlo simulation. Generally speaking, a statistical distribution is postulated for the variables and random samples are extracted; based on them, the traffic prediction is calculated.

The impact of the uncertainty of the inputs on the prediction is calculated based on the variance in these predictions. The studies surveyed by these authors find that the traffic flows are between 18% and 33% of the average for a confidence interval of 95%.

One of the chief criticisms levelled at this approach is that univariate distributions are postulated for the explanatory variables without taking the correlation among them into account.
A second source of uncertainty is associated with the specification of the demand equation – omitted variables, inappropriate assumptions about the functional form and about the statistic distribution of the random components – and/or error due to the use of estimated instead of real parameters.

A variety of approaches is used to approximate uncertainty. One possible way of proceeding is to use the Bootstrap and Jack-knife methods to create sub-samples, based on them the parameters of the model are estimated repeated times. Once the standard errors of the parameters have been properly calculated, they are used either to construct analytical expressions of the standard error of the output of the model or in information that enables a distribution function of the parameters to the constructed and Monte-Carlo simulations to be applied. The difficulty of constructing analytical expressions explains the use of simulations in the majority of studies.

According to the studies surveyed, the uncertainty derived from the parameters is lower than the uncertainty derived from the inputs.

De Jong et al. (2007) apply an analysis of uncertainty associated with both the inputs and the estimated model for predicting the extent of a motorway in the Rotterdam area. The model consists of a system of demand equations that includes different stages in the journey and is estimated using micro-data. It is a nationwide (LMS) model that includes the frequency of journeys according to purpose, the choice of destination and the mode of transport and time of day.

The uncertainty linked to the explanatory variables of the model is examined through time series analysis. This method enables us to construct a multivariate distribution for the following model inputs: available income, motorisation, costs per kilometre, population by age bracket and employment. The standard deviation and correlations are used for a 20-year moving average from the past 40 years. The idea is that the variation for the next 20 years (the timeframe of the prediction) can be determined by all the 20-year moving averages calculated for the past 20 years. The construction of the distribution function enables different values to be generated for the model inputs.

With regard to uncertainty related to the model, they use the Bootstrap technique to re-estimate its different stages. However, no significant differences were found with respect to the initial estimates. The analysis of uncertainty is conducted through multiple simulations that enable the estimated parameters, the inputs or both to be varied. The uncertainty related to inputs clearly dominated the uncertainty related to the estimated coefficients. Furthermore, the use of train/bus and underground was more uncertain than the use of the private car. For example, the uncertainty associated with the predicted number of car-
kilometres measured as the standard deviation from the average was 8.3%. In terms of the confidence interval of 95%, this entails a range of variation between 289 and 299 million car-kilometres.

As a general conclusion drawn from the literature examined, uncertainty in the demand prediction models is rarely dealt with in a statistically systematic, rigorous way. Hypotheses on high/medium/low scenarios are often put forth without proper justification for these scenarios. On the other hand, from the literature we can glean the need to pay more attention to ex-post evaluations of the predictions in order to analyse the degree of error in models and alternative models from a critical vantage point.

2. ANALYSIS OF THE PREDICTION MODELS BY MODE OF TRANSPORT

In this section, we shall analyse the demand prediction models that are used by the different public transport organisations in Spain. Specifically, we shall outline the models, if they exist, used by State Ports, Airports and Air Navigation (AENA), the Directorate General of Motorways and the Directorate General of Railways.

2.1. State Ports

Every four years, State Ports predicts the traffic for the Spanish Port System using a consolidated methodology that also undergoes constant improvements. The overall goal of this effort is to come up with a prediction of the traffic for each Port Authority and for each port in the Spanish Port System according to the type of freight and product which extends to the medium and short term. The last study, published in 2004, contains the predictions for the baseline year and the ensuing years 2007, 2010 and 2020.

The methodology used to predict port traffic consists of the following phases:

- Phase 1: Information gathering (both direct and indirect) and estimate of the models;
- Phase 2: Convergence of the procedures used;
- Phase 3: Comparisons with the Port Authorities;
- Phase 4: Drafting of the results.

Furthermore, due to the importance of port investments for the economy as a whole and the variability of the factors that influence the evolution of maritime traffic, two years after the drafting of each study an update is undertaken with the goal of assessing the degree of accuracy in the first year and consequently making the appropriate rectifications. This effort
entails choosing the flows which, according to given criteria (quantitative and direct), show significant deviations from the prognosis and coming up with a new prediction that includes the additional information obtained only for these flows. Below is an outline of each of the phases:

**Phase 1: Information gathering and estimate of models**

The methodology used stems from a combination of three approaches:

- **a.** Direct surveys of large freight carriers;
- **b.** Time series models that include economic variables;
- **c.** Casual process: Model based on input-output-tables.

The survey is targeted at clients and agents of the port system and provides both quantitative information (forecasted traffic, transport chain) and qualitative information (expected quality level, criteria of modal choice). The survey’s degree of reliability depends on the level of coverage it reaches, which in turn depends on how atomised each sector is. It is reasonable to assume that the survey reaches a very high level of significance in the energy sector given that this sector is generally concentrated in a handful of large companies. Indeed, this kind of information predominates when there are few importers/exporters of a given good, whilst when the level of coverage is so low that it does not allow for a sufficient degree of predictive reliability, the predictions obtained from the other procedures are accepted.

If on the one hand the information surveyed is important in that it reflects the companies’ individual problems and issues specific to each sector, on the other hand the companies do not tend to reveal their middle- or long-term projects, and in any case, many of them are reluctant to make long-term predictions owing to their lack of knowledge of future developments in the economy.

In the latest study performed, the direct information obtained in this way was integrated by the Delphi Consultancy, a panel of experts in economics and logistics assembled with the purpose of achieving the broadest consensus possible on the likely scenario that will characterise the logistics and transport sector in the long term.

The analysis of the time series based on historical information aims to adjust a curve that explains the evolution in traffic of each of the products considered by type of movement and by port (maximum breakdown possible), with the goal of detecting trends and cycles.
These two microeconomic approaches are accompanied by a third one defined as a casual process, which sets out to estimate the evolution in the different sectors from a macroeconomic standpoint. Basically, it is an approach that uses different sources of information (foreign trade, customs) to predict the evolution of the macroeconomic scenario for each sector into which the goods are grouped (in terms of gross added value or foreign trade, for example).

*Phase 2. Convergence of the procedures used*

Based on the predictions obtained from the three aforementioned procedures, convergence criteria are applied to ensure the overall coherence of the predictions. Generally speaking, the information obtained through the survey campaign prevails when the coverage level is high enough to ensure a sound level of reliability, while in the remainder of the cases the results from each procedure are compared and the hypotheses are revised until they are convergent.

Once convergence is achieved, there is an effort to fine-tune the results by explicitly bringing competition factors between ports into the analysis (distribution method). The so-called “context of transport” comes into play in this phase. Given the fact that in a context of port competition the concept of transport chain comes to the fore, the competing transport chains are chosen and an attraction coefficient is defined for each port based on its accessibility. Then the percentages of traffic distribution are calculated according to the attraction coefficient; a contrast with the actual distribution of traffic sheds light on the influence of other factors in the competition among different ports in terms of accessibility. The idea is that this “fine-tuning” of the analysis makes it possible to define a theoretical hinterland of the ports and ascertain each port’s limitations with respect to the others.

*Phase 3: Comparisons with the Port Authorities*

The results obtained in the previous stages are compared with the forecasts formulated by the Port Authorities, which are assumed to have a significant knowledge base. The Port Authorities can express their agreement with the results or put forth their own corresponding alternative forecasts.

*Phase 4: Results*

The result of the preceding phases is a series of forecasts made by each Port Authority and by each port in the Spanish Port System with the following level of disaggregation:

i. Navigation: high seas, coastal;
ii. Operation: loading, unloading

iii. Nature of the freight (9 sectors, 40 products)

iv. Form of presentation.

2.2. Directorate General of Motorways

The *Directorate General of Motorways* has a traffic prediction model for large corridors called the “Forecast System of Fleet, Traffic and Consumption”, which is owned by the *Ministry of Transport*. This model has been in operation since 1975. The latest document available is in the publication entitled “Forecasts of the traffic fleet and consumption for the 2005-2012 Sectoral Motorway Plan”.

The traffic predictions are made for three scenarios of economic growth: low, medium and high. The model encompasses the following stages:

1. Prediction of registrations by type of vehicle.

This figure is extrapolated based on the time series of vehicle registrations. Bearing in mind the average cost of each kind of vehicle, the percentage is calculated with respect to a macroeconomic variable, either consumption or GFCF of the expense in purchasing each type of vehicle for each year. By applying the extrapolated percentages of expense in each type of vehicle to the corresponding macroeconomic variables, and with a steady evolution in prices, estimates of annual registrations for each vehicle are obtained until 2020.

2. Estimates of the fleet of vehicles.

The *Directorate General of Motorways* figures on the fleet are corrected downward in order to obtain the fleet by age in each year and for each kind of vehicle. The fleet is predicted based on the information on registrations since 1945, and an estimated vehicle withdrawal trend is applied for each type of vehicle. The estimates are obtained by analysing the figures on withdrawals per vehicle provided by the *Directorate General of Motorways*. The system models these withdrawals each year according to a logistical function which takes vehicle replacement plans into account. Finally, the fleets are grouped together by age and type and compared with the available figures. The increases in the fleet go down as the traffic matures.

3. Traffic estimates each year according to age and type of vehicle.

This estimate is obtained based on the vehicle fleet, to which an average journey length is applied. This average is maximal in the first year of life and gradually goes down.
Extrapolations – both high and low – are made based on figures from 2005. In the first, the average journey length remains steady and in the second it gradually drops until reaching 10% lower in 2020. Furthermore, the system distributes the traffic into three different types of journeys: urban, periurban and interurban. The predictions are compared with the observed figures on the number of kilometres driven, and they are coherent.

The figures on traffic elasticities with regard to the GDP are for the different decades: 1960s: 1.77; 1970s: 1.27; 1980s: 1.45; 1990s: 1.47; and 2000s: 1.22.

4. Petrol consumption.

In the last stage, the petrol consumption by type of vehicle is predicted.

A comparison of the traffic predicted for the period 1989-2005 performed in 1989 with the observed real traffic shows that in aggregate terms the level of error in the prediction of interurban traffic is low.

2.3. Directorate General of Railways

The Directorate General of Railways does not have a demand prediction model. On the other hand, the predictions that appear in the informative studies of the projects submitted for evaluation by the Ministry of Transport significantly overestimate the future demand in most cases. In particular, predictions of the induced demand generated by new projects show systematic overestimates of the demand.

The conclusion is that there is no method for estimating the demand that is based on a study of the relevant traffic. Therefore, the figures on the future demand in railway transport which are provided in the informative studies must be dealt with and interpreted with limitations, as they are not derived from a rigorous prediction model that might be reliable. The Ministry of Transport works with figures provided by RENFE, highly aggregate figures like the ones included in the Statistical Yearbook. Ever since the liberalisation, access to information is more complicated and to date the new freight operators have not provided any kind of information.

2.4. Air transport

The traffic forecasts for the AENA airports are made by the Division of Development of Studies and Planning Techniques of the Directorate of Infrastructure Planning (abbreviated DPI).
In coordination with the Directorate of Operations and Network Systems (DOSR), the DPI is in charge of updating the forecasts on the evolution in passenger transport, airplanes and freight at the airports within the AENA network.

The traffic figures are updated twice a year, in March and October, and long-term projections are made for a 15-year period using results on the demand of passengers and airplanes, freight, passenger-times in arrivals and departures, airplane-times and airplane-times in arrivals and departures for different types of traffic. AENA’s overall forecast is the result of aggregating the disaggregated forecasts for each of the airports in its network, which are in turn segmented into domestic, national, commercial and total traffic.

In order to make the prediction, an econometric model is developed that relates the air traffic to observed variables whose correlation with the traffic values is plausible, such as demographic variables (civil population, active population, unemployed population, etc), tourism variables (entries into and departures from our country), microeconomic variables (elasticity-price, considering “ticket” price an aggregate of the cost of maintenance, taxes, fuel price, amortisation of assets, etc.) and macroeconomic variables (aggregates that measure tax and monetary policy such as interest rates, inflation, public spending, unemployment rate, salaries, production, volume of exports and imports, etc.) which come together in the GDP, the main variable with which air traffic is associated. The model introduces an average trend, yielding predicted values according to what has been observed in the past and what is forecasted for the associated variables.

One important factor in AENA’s predictions is that they also include changes in the transport supply, including the launch of the AVE high-speed trains in a given route, the entry of a new airport into direct competition and the opening of planned infrastructures.

They also make short-term predictions with studies of the most important airlines operating at the airport, current routes and possible short-term changes. To accomplish this, they check with the department of slot coordination and the airlines themselves, and they characterise the type of passengers through surveys, each airline’s fleet or their portfolio of orders. This enables them to estimate, for example, the average airplane size and the number of seats offered in the future. These particular studies are especially useful for more realistic short-term forecasts than the trend long-term model.

In conclusion, AENA’s overall forecasts are the aggregate result of short-term analyses, along with the long-term trends associated with the econometric model and the specific variations for each airport. The forecasted demand is always presented in three scenarios (high, medium and low).
Finally, it is important to point out that these predictions are subjected to consultation with each airport in question, as well as to the judgement of the division staff, which enables them to apply their experience and work performed in recent years.

2.5. Plans for toll roads and motorways

Any proposal for investment in Spain’s roads must submit an informative study that includes a detailed study of the traffic forecasts. To illustrate this, five investment projects in toll roads have been chosen and the predicted traffic has been compared to the observed traffic in the first few years the motorways were operating. For these projects, either the informative study or the pre-project have been consulted. The advantage of the former is that it includes a detailed study of the traffic. However, the pre-project already includes the motorway’s definitive route. In most cases, the pre-project maintains the same predictions as the informative study.

The traffic prediction models used differ according to the study and show varying degrees of complexity. Nevertheless, they all include the following steps:

1. Traffic prediction in the corridor analysed: The point of departure is the prediction of the traffic in the corridor for the first year after the motorway is opened. This prediction is based on historical figures from the car count station, on surveys of car counting and/or on mobility surveys. The projections from the baseline year until the first year in which the project is in operation (approximately ten years) are made according to the growth in GDP and/or the population. In some cases, the possible consequences of urban development changes in the region are also taken into account. Generally speaking, a distinction is made between long- vs. short-haul, light vs. heavy, and peak vs. off-peak traffic. Some studies base their predictions on estimated demand models, while others use standard values for the elasticities in demand with respect to the GDP or population.

2. Distribution of traffic between toll road and free motorways: The most common model is a logistic regression of the distribution of traffic among different roadways depending on their overall cost. This stage requires knowledge of the travel times and the monetary cost of each alternative. Some projects include more sophisticated studies based on individual data which enable time values to be estimated. This is the case, for example, of the study conducted by Steer Davies Gleave for the A-41 motorway based on an experiment of stated preferences. In other cases, the modal distribution is based on standard time values. For example, in the study of the Ocaña - La Roda motorway, the time value applied was between €9.60/hour and €15/hour according to the type of
traffic. The Modasign computer programme was used to assign the traffic among motorways.

3. Based on the prediction for $t_0$, growth rates in the traffic that decrease over time are assumed. Generally speaking, no rigorous justification for the rates chosen has been offered.

4. Some studies include an analysis of the traffic sensitivity for different assumptions included in the prediction.

One comparison of the predicted traffic compared to the observed traffic for the more recent toll roads in Spain shows a clear over-prediction of the traffic, although this seems to bear no relationship to the complexity of the model used. Specifically, the predictions in the projects examined are almost twice the amount of traffic actually observed.

A comparison was also made with the prediction errors for some of the toll-free roads. In the case of the motorways, the predictions’ deviation is less pronounced and no positive bias can be observed. These results coincide with the ones provided by Bain (2004).

Implications of the available information for the evaluation of projects

The demand models existing for ports, motorways and airports offer a prediction of the growth in traffic for a medium and long term which can serve as a point of departure for predicting the demand of the project to be evaluated. Given the fact that the demand for transport bears a close relationship with economic growth, aggregate predictions by traffic corridor are a sound point of departure for analysing the traffic of individual projects. In any case, they can be viewed as a benchmark to make comparisons and discard somewhat implausible variations in the demand.

Nevertheless, all three cases are relatively complex models, and for the ports, and to a lesser extent the airports, they combine predictions based on historical figures with qualitative opinions from the infrastructure managers. For this reason, it is difficult to systematically include these predictions in project evaluation.

3. DATABASES AVAILABLE IN SPAIN

Even though in order to conduct a particular study on an investment in transport it would likely be necessary to get disaggregated or detailed figures by requesting them from the competent organisations or the operators themselves, it is important to be familiar with the
databases available and accessible today. They can be very useful either for a preliminary examination or to supplement figures from other sources.

The databases on transport and mobility existing in Spain are neither systematised nor centralised in any single transport organisation or associated organisation that facilitates accessibility to them. A survey of these databases enables us to confirm that the type of information they provide is very limited if the goal is to use them to predict the demand. One of the greatest shortcomings is the absence of origin-destination matrices for the journeys, with both data on the aggregate traffic flows and individual data, for any mode of transport. With regard to the figures on flows of journeys, it is recommendable to present them with a higher level of disaggregation in origin-destination or kind of user. This level of disaggregation is fundamental for estimating demand models that are useful for predictions, and therefore for evaluating investments.

Thus, there is a significant need to improve and complete the databases in order to be able to more accurately estimate and assess the changes in the demand for transport, a crucial factor in when evaluating transport projects on which the profitability of the investment relies.

Below are the existing databases in Spain which contain information on transport and journeys, along with a brief description of each of them. You can see that in most of them, the figures available are too aggregate to be able to estimate and predict the demand accurately.

- **INE: National Statistical Institute**
  - Figures: Traffic, vehicles and drivers, companies, infrastructure and investment, transport safety, especially nationwide and some Europe-wide figures.
  - Modes: Road, railway, air and maritime.
  - Mobility: National and international, interurban and regional.
  - Web: [www.ine.es](http://www.ine.es)

- **DGT: Directorate General Of Traffic**
  - Figures: Vehicles and drivers, safety (series on accidents and victims).
  - Modes: Road.
  - Mobility: National, urban and interurban.
  - Web: [www.dgt.es](http://www.dgt.es)
• **Annual Report On Transport And Postal Services**
  - Detailed but aggregate information from the sector on all kinds of mobility and modes of transport.
  - Web: www.fomento.es

• **AENA: National Airports And Air Navigation**
  - Figures: Passenger and freight traffic according to origin and destination by airport (since 1990)
  - Modes: Air.
  - Web: www.aena.es

• **MOVILIA: survey on mobility of the Ministry of Promotion:**
  - *Movilia 2006 and Movilia 2000:* Figures on everyday mobility broken down by purpose for the journey, mode of transport, length, sex, educational level, timetable, etc., on the level of province, metropolitan area and size of municipality. The figures are available for both years that the survey has been conducted.

  - *Movilia 2007 and Movilia 2001:* Long-distance mobility (more than 50 km in the 2007 survey and more than 100 km in the 2001 survey), including the number of journeys by province of residence, by metropolitan area, by size of the municipality, by purpose for the journey, by educational level, by length, by modes of transport, by distance, etc.

  They also offer figures on flows of origin-destination journeys between autonomous communities by purpose for the journey and mode of transport. There are no flows at provincial level, and the inquiries posed to the experts in the Ministry of Transport in charge of the survey confirm that the sample is not representative for provincial flows.
  - Web: www.fomento.es

• **FAMILYTUR AND FRONTUR**

  These are surveys conducted by the Institute of Tourism Studies (Secretary of State for Trade and Tourism). The figures are periodic and by mode of transport.

  Frontour gathers information on journeys by citizens not living in Spain who enter Spain, and Familytour gathers information on journeys by residents of Spain who
leave Spain. The latter provides less biased information by mode of transport than the former.

- **Double Screen Road Surveys (1999, 2000)**
  - Well-gathered, highly disaggregated figures with origin-destination relations for a random sample on a workday.

- **ADIF: The Administrator of Railway Infrastructures**
  - This provides figures on the circulation of trains and travellers entering stations.
  - Web: [www.adif.es](http://www.adif.es)
REFERENCES


AENA (2008). *Metodología de actualización de las previsiones de tráfico de los aeropuertos de AENA*, AENA.


