ATC economic modeling

Nathalie Lenoir, Jean-Claude Hustache

ATC Economic modeling

Nathalie Lenoir & Jean−Claude Hustache

Ecole Nationale de l'Aviation Civile
7 avenue Edouard Belin, B.P. 4005
31055 Toulouse Cedex
hustache@club−internet.fr

Introduction

Airspace congestion has become an increasingly significant phenomenon in Europe during the last decade. In several geographic areas, the air traffic control system is functioning at full capacity, causing congestion and resulting in delays difficulty bearable for the airlines.

Different approaches have been followed to solve the problem, but things have been mainly envisioned from a technical and operational point of view. On the operational side, several solutions have been brought into play : control sectors have been redesigned to optimize the flow of traffic, controllers work periods have been rescheduled to offer maximum capacity in periods of high demand, and the compatibility of national systems has been improved. On the technical side, new systems are being designed to help the controllers solve difficult conflict situations, and research is going on in several areas to enlarge control centers capacities in the future.

However, these efforts are costly, and attention is now focusing more and more on the economic aspects of air traffic control, particularly costs estimations, project evaluation and demand management. Some important questions are : What is the optimal organization in terms of costs? Is it possible to manage demand in order to have a more efficient use of existing capacity, considering the costs of dimensioning the system for peak periods?

Incentive theory and public economics enable to analyze the optimal setting in terms of structure and pricing. The structure of ATC services is however, a highly sensitive area, where economists can only provide guidance, for example by comparing the efficiency of different structures. ATC pricing is also a very important issue ; in many congested sectors, pricing is used to alleviate congestion and optimize the use of capacity, and this could be done in ATC.

Economic studies have to be supported by practical results, obtained through econometric studies. Those studies can lead to the definition of meaningful indicators on the basis of which management decisions can be taken. If those indicators were found to be consistent enough, they could provide guidance for investment and hiring policies.

So far, few econometric studies have been done, because meaningful data is hard to obtain, and work is only beginning in many areas, like cost estimations. Some results have been obtained on delays and on pricing (see Lenoir 1995) ; they are summarized in part 1. Another study (see Hustache 1996) focused on relations between traffic, delays, investments and capacity. The method and results are the subject of part 2.
The results obtained confirm the interest of further economic research and points to the directions it should follow. The estimation of a cost function is one of these. It gives the optimal way of using the production factors in order to reach a given capacity and a given quality. The study of the sensitivity of demand to prices and delays is another area of interest, and leads to examine the optimal structure of pricing and slot allocation. The directions of future research are exposed in part 3.

I. Delays and pricing

The estimation of the relation between delays and traffic is a prerequisite to the estimation of the cost of delays, which has to be taken into account for pricing purposes.

In a congested system, users do not have the right incentives to use the system : when entering the system, they take into account the delay they will have to bear, but do not consider the delay they impose on others, which can be important. Therefore they underestimate the « so called » social cost of their presence in the system. The social cost of the last user is the sum of the marginal cost of using the system (production cost) and of the cost of the marginal delay this user induces.

Usually, pricing is devised before congestion occur, and therefore, does not consider the delay cost. The user pays only the production cost, and thus does not consider the marginal delay cost (since this delay is bore by the subsequent users). He will therefore underestimate the real cost and this will result in too much congestion. In a congested system, pricing has to be changed, in order to give users the right incentives. Different pricing systems can be examined (see Lenoir 1996), but in any case (peak load pricing, priority pricing), the cost of delays has to be estimated.
This involves two steps : establishing the relationship between traffic and delays, and calculating the unit cost of delay (although since cost is not necessarily a linear function of delay, it can lead us to examine the relationship between delay and cost). Our work has concentrated on the first part, since data on delays is more readily available than data on airline delay costs.
Several methods can be used to link traffic to delays. The most accurate analysis would consist in linking traffic to delays at the sector level, since delays are dependent on the sectorisation and route configuration. However, this would not lead to a general relationship between global traffic and delays, which is what we are looking for.
Does such a relationship exist ? If ATC is well managed, one could expect that with minor variations, a given level of delays corresponds to a given level of traffic. The results will confirm this hypothesis (see Lenoir, 1995).
To estimate the relationship, econometric methods are well adapted, and their « simplicity » enables to recalibrate the model often (since capacity increases regularly, the model should be reestimated often). No hypotheses are made concerning the structure of ATC, therefore there are no biases due to a inadequate modeling of the industry, as could occur with estimations based on queuing theory (for an introduction to queuing theory, see Larson and Odoni, 1981).

Since traffic, delays, and capacity are linked, eliminating the effect of capacity changes implies studying traffic and delays at a given time. The data used consists in individual flight plans and delays, for a month (Sept. 1993). From this data we obtain average delays and traffic at a given time. We then look for relationships between delay and traffic at different times (delays can be dependent on present traffic but also on past traffic), using duration model (see Amemiya, 1985).

We find significant relationships between average delays and traffic at the same period, indicating no strong building up of delays. Estimation results are statistically better (R^2=0.86) for the same day of the week (4 Sundays of September, for example), showing that either capacity varies during the days of the week, or rather that traffic patterns are different.
The functional form is an exponential[1] showing that delays tend to increase exponentially after certain levels of traffic are reached. This enables us to calculate the marginal delay at a given level of traffic[2].
Other estimations, on individual delays, have failed, which indicates either that we do not have enough information to model individual delay, or that there is a strong stochastic component in individual delay which comes from the slot allocation rule (first come − first served).
It seems therefore more promising, for future work, to focus on the estimation of average delay, rather than individual delays.

[1].Average delay : R=exp (a + b sqrt (T) + c T + d T^2).
[2] Total marginal delay : TMD = T (T ($\partial R / \partial T$)

II. fundamental relations

Estimating the relations between traffic and delays at a given time is only part of what we should be interested in. In a subsequent work, we tried to see which dynamic relations exist between four key variables : traffic, delays, investments, and capacity.
Building four different models, the aim is to describe basically the ATM dynamic. At the starting point, hypothesis are the followings : Firstly, traffic data is exogenous, or more precisely depends on other variables like the growth national product, pricing policy, seasons, etc. Secondly, capacity is observed from a global and long term point of view, there is no flexible adjustment.
Then, for a traffic demand, the ATC production is adjusted by creating more or less delays. Once ATC administrators observe the delays, they can have an idea of the capacity shortfall. They will decide to invest, to hire controllers, or to use new technologies, regarding the present needs. Finally, investments should increase the capacity but it has to be checked that capacity improvement does not come too late regarding the traffic evolution.
Estimating each of the four models, we try to find functional forms in accordance with reality and to show the existence of time lags in certain models.
The following models have been built with yearly French data (see SCTA annual reports), covering a eleven year period (1984−1994).

II.1 delay = f (traffic)

Contrary to the first functional form exposed in part one, the focus is now on long term models. Moreover, different indicators are used for the delay measurement. None the less, one could not expect the relation to be linear. When a flight is delayed, any other flights should be rescheduled. As seen in part 1, a new entrant in the system causes a negative externality. This assumption is also confirmed in this model[3]. Delays[4] are increasing quicker than traffic[5] : a 20% increase in the traffic leads to an 100% increase in delays. This result could seem alarming but it is obtained without capacity variation. In fact, this scheme should be interpreted as a « do nothing case », or as a passive strategy in yearly capacity increase.
An economic interpretation leads to suppose two things :
1) There are decreasing returns to scale in the capacity use. So, we need to invest more and more or to find new technologies, if we want the ATC quality only to be constant.
2) There are negative externalities due to the lack of efficient incentives in the capacity use.

As the capacity is absolutely associated with the traffic−delay relation, another model is tested. A two capacity level model is built, based on monthly data. Actually, we suppose that ATC can provide two levels of capacity : a high capacity in summer, and a low capacity in winter. Unfortunately, the obtained results are not relevant. Statistically, regressions are not significant (R^2 = 30%) and moreover, the curves positions are not logical. With this last model, for the same traffic level, the high capacity system provides higher delays than the low capacity system.

Actually, this failure shows the limit of econometric studies in such a field. This kind of relation can not be modeled as soon as the high capacity is never supplied when the traffic is low. Building a traffic−delay relation with different capacity levels could only be made with simulation tools.

II.2 investment = f (delay)

The aim is to model the ATC administrators reactivity. How much should be invested to reduce very high delays? This link is absolutely « political », whereas the inverse relation is « technical » and tests the investment[6] efficiency (or its productivity). It will be treated in the third model.
A precise estimation of this relation should help to plan future investments, and could thus avoid a chaotic evolution due to short term motivations.
In a standard strategy, investments should be done before the lack of capacity occurs, in order to minimize the delays. After testing three different models, the final result is far from the ideal scheme. The 3 models were each based on one of the following assumptions :

1. Investments are determined in function of the future delays, there is an anticipation.
2. Investments are performed at the same time that delays happen
3. Investments are implemented one year after we observe delays.

---

[3]. Delay = − 1.26 + 2.2 $(\text{traffic})^2$
[4]. Delays are measured in hours of regulation.
[5]. The traffic is measured in kilometers controlled.
[6]. As investments amount were not available, depreciation and interest costs were used.

Both intuitive and statistical results lead to prefer the third solution[7]. In fact the correlation coefficient between data is higher (0.98% instead of 0.93%) when a time lag is introduced. This time lag may be the proof an imperfect strategy in ATC management. Actually, we suppose that a one year delay in the investment is a minimum limit. First, to obtain and analyze delay figures is time consuming. Secondly, it takes time to choose the right investment and to obtain the credits. as long term data are missing (more than 10 years) about delays and investments, higher time lags could not be tested(2−3 years).
A general critic which can be made about the ATC investments is a lack of reactivity, and an insufficient flexibility. Indeed, the fact that in Europe ATC centers depend on the Government budgets does not make the task easy. ATC investment projects can be in competition with other political plans, and they may not be decided strictly in function of the ATC needs and potential benefits.
Modeling this relation, we also try to learn about the function curvature. Is the relation linear ? As far as delays are low, there is probably no reason for investing. But, once delays become too costly for airlines, there is no doubt that they should claim for capacity, and would not accept to bear entirely the cost of delay. This results can confirm those intuitions, but statistical gains when using an exponential instead of a linear form are minors.

II.3 capacity = f (investment)

This relation measures the investment efficiency in terms of capacity[8]. As we proceed for the previous relations, two things have to be pointed out : What is the curvature ? Is there a time lag ?
About the curvature, different readings are possible.
If the curve is convex, it can mean that there are scope economics, coming from technological progress which allows a high performance level in terms of capacity and costs. Or, it can mean that capacity gains are obtained from other factors for which the cost is not part of the investments.
If the curve is linear, the scope economics are constant. An x% increase in investment grants an x% increase in capacity.
Finally, if the curve is concave, we can suppose the scope economics are decreasing.
Statistically, the two best functional forms are linear and logarithmic. Then, it is hard to give a conclusion about the investment efficiency in terms of capacity. None the less, a fundamental limitation of our model can explain this inaccuracy. This relation depends on the service quality : the more is lost in quality, the more is gained in capacity. As the ATC quality (in terms of delays) has decreased during the last 10 years, we can

imagine this model overestimates the investment productivity. This last consideration leads to prefer a logarithmic model[9].

It may seem surprising, but time lag between investments and capacity improvement was not found.

II.4 capacity = f (traffic)

The logical progression in this model would be to test the impact of the capacity on the traffic. If the capacity generates effects on the ATC demand, it is surely not on the traffic level but on the traffic distribution during the day. To test such relation annual figures are not appropriate. Hence, we test the impact of the traffic on the capacity.

As the first three models describe the indirect effect of the traffic on the capacity, estimating a direct relation should allow us to check the previous results. This last model provides two essential indications. As the traffic increases, the number of sectors increase, but the marginal productivity of a sector is decreasing. A 10% increase in traffic requires a 35% increase of sectors. Secondly the best model[10] is obtained by creating one year delay between traffic evolution and capacity adjustment. Then, if the capacity has to be adapted on time, investments should become more flexible, and decisions should be anticipated.

---

[7].Investments $_1$ = exp (18.04 + 0.5 E−4 delays $_0$).
[8]. The capacity is measured in simultaneously openable sectors.
[9]. Capacity $_0$ = −154 + 11.44 1n (investments $_0$)
[10]. Capacity $_1$ = 0.81 + 0.18 (traffic $_0$)^2

A general limitation of these models is that they focus on long term figures. If the relations are used to make forecasts, the models will only reflect the past decisions (and past errors). Thus, forecasts must be interpreted with care as the result of a « passive strategy »

Improvements could be done by using multiple regressions and by introducing the number of controllers in the capacity measurement.

III. Future directions of research

More econometric studies will be needed, in order to confirm the first results obtained. These results show that ATC, although a highly technical field, is no different from other industries, where econometric methods, without explicitly modeling the technology, can however find statistically meaningful results that offer new insights on the industry workings.

More ambitious results could be aimed at : the first study can be carried on and leads to investigate alternative pricing solutions for ATC, whereas the second study leads to cost evaluation.

As the first study shows, traffic and delays are linked, and eliminating delays may prove too costly as traffic goes on increasing. Dimensioning the ATC system for peaks may prove not only costly but also impossible. In this respect, the benefits of a pricing system well adapted to a congested ATC are not to be neglected. Thanks to a better attribution of capacity, delay can be significantly lowered. The utilization of capacity being thus optimized, any capacity increase can be weighted by a relevant cost−benefits analysis.

The second study, showing interesting relations between traffic, delays, capacity and investment, leads us to inquire about the optimal structure of ATC in terms of costs. Are production factors optimally used in ATC, or is there scope for capacity or cost improvement, through a better use of technology or staff, or a better timing of investments ?

The answers to these questions require the estimation of an ATC cost function, at country level, or better, at center level, based on historical cost data (if data permits such an estimation). This cost function would give us the cost of ATC production as a function of production factors (staff, investment ...)[11]. It could take into account delays, as an indicator of the quality of production, and indicate us what is the best ratio between more technology or more staff, depending on their respective cost.

The success of these studies will depend on the improvement of data collection and availability. The compatibility of data from different country (cost data) should be improved within the European zone. Data on airline costs (especially costs of delays) are also necessary to the improvement of pricing, and it is important to point out to them that it is in their interest to communicate more on this issue.

The conclusion to this paper could be that much needs to be done in the field of ATC by economists, but no valuable econometric study has ever been made without accurate data.

[11.] By definition a cost function is : $C = f(p,y)$ where C is total cost, y is production, and p is the vector of input prices.

REFERENCES

- Amemiya, T. (1985) : *Advanced Econometrics*, Basil Blackwell, Oxford.
- *Bilans annuels du SCTA* (1984 à 1995)
- Hustache, J.C. (1996) : *Econométrie du contrôle aérien*, Mémoire de DEA d'économie des transports, Université des sciences sociales, Toulouse. Note CEE No.32/96
- Larson, R. et A. Odoni (1981) : *Urban Operation Research*, Prentice Hall.
- Lenoir, N. (1995) : *Une approche économique de la congestion aérienne*, Ph'D thesis.
- Lenoir, N. (1996) : *Priority pricing in air traffic control*, presented at the IFAC international conference, June 1996.