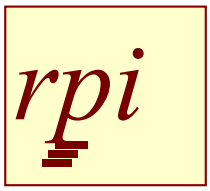


Regulatory **P**olicy **I**nstitute
31-33 Westgate
Oxford OX1 1NZ
UNITED KINGDOM
☎ + 44 (0)1865 792 858
Facs+ 44 (0)1865 241 885
http://www.rpieurope.org

STUDY ON THE IMPLEMENTATION RULES OF ECONOMIC REGULATION WITHIN THE FRAMEWORK OF THE IMPLEMENTATION OF THE SINGLE EUROPEAN SKY

European Commission Directorate-General of Energy and Transport

Tender No. TREN/f2/28-2002

FINAL REPORT

October 2003

REPORT PREPARED BY:
CHRISTOPHER DECKER, ALAIN JEUNEMAÎTRE , TIM KEYWORTH & GEORGE YARROW

PROJECT TEAM: CHRISTOPHER DECKER, HERVÉ DUMEZ, ALAIN JEUNEMAÎTRE, GEOFF JONES, TIM KEYWORTH & GEORGE YARROW

TABLE OF CONTENTS

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Introduction and background

1. Concerns have been expressed in recent years about the performance of the European ATM network. Notwithstanding the emergence of cross-border service provision -- Maastricht, and, for the future, CEATS and NUAC -- the structure of provision continues to be heavily influenced by existing national boundaries. The precise effect of the resulting fragmentation on the overall performance of the network remains a matter of dispute, but it is highly unlikely that an efficient structure of provision would correlate closely with national boundaries determined by political history.

2. The Single European Sky initiative provides an opportunity to address and to eliminate some of the barriers standing in the way of the evolution of the European ATM network toward more efficient structures and more effective performance. This Report is addressed to one part of the initiative, which is focused on issues of charging for service provision and of regulation of charging.

3. **To facilitate discussion of charging options, we introduce a distinction between operational airspace blocks (OABs) and airspace charging blocks (ACBs).** The former refer to blocks of airspace that are controlled by a single provider or a joint venture of providers. Following implementation of the Single European Sky regulations, functional airspace blocks (FABs), will have this characteristic. In contrast, an ACB is a block of airspace in which *en-route* charges are either identical or are set in a fully co-ordinated way, whether by a single service provider or by means of an agreement among the relevant service providers. OABs and ACBs may or may not coincide. For example, a two or more adjacent ANSPs may agree to set a common unit charge per service unit whilst each retains control of its own national airspace.

Review of the theory and practice of economic regulation of networks

4. Many, but not all, of the issues relevant to ATM have been addressed in other 'network' sectors of the European economy, including communications, energy, and rail, both in theory and practice We have therefore first reviewed this wider theory and practice with a view to formulating some of the principles of best-practice regulation which may be of relevance to ATM.

5. Points of importance that emerge from this work include:

   - neither full cost pass-through nor 'pure' price cap regulation is likely to offer the most effective way forward in relation to charging**: best-practice regulation is converging to 'hybrid' arrangements that allow for an appropriate balancing of risks and rewards**, taking **account of requirements to promote efficient investment and capacity expansion** ;

- liberalisation is leading network structures away from structures in which monopolistic suppliers deal with large numbers of small customers and towards structures in which monopolies deal with a much smaller number of commercially sophisticated, and sometimes large, network users: other networks are becoming more like ATM;

- these users, who stand on the demand-side of relevant markets, are playing an increasingly central role in determining the rules for the operation of networks, including charging rules: among other things, **user participation in governance and regulatory processes is serving greatly to improve the volume and quality of information available to policy makers**;

- following earlier stages of liberalisation, when the policy focus tended to be on promoting competition where feasible, and on reducing excessive costs, **issues of quality of service and of investment and network development tend to be acquiring more prominence**.

- **co-ordination functions, which are a key feature of network operations, are increasingly being unbundled and entrusted to specialist organisations, which go under the name of 'system operator' or 'network manager'**;

- in dealing with short-run matching of demand and supply/capacity, and in co-ordinating longer term investments across networks, the activities of system operators/network managers allow charging policy to be more targeted on the objectives of stimulating improved performance by service providers, providing longer-term signals to network users, and establishing an acceptable and efficient distribution of risk among providers and users.

Contractualisation

6. The growing importance of users in network governance is leading to a structure of relationships that is more 'contractual' in nature than in earlier periods. That is, increased obligations and commitments are being required of and accepted by the relevant parties in their dealings with one another. **In the second part of the Report, therefore, we first outline some of the ways in which this process can be taken forward in European ATM**. Among these are:

- enhanced disclosure, reporting and consultation requirements;

- the establishment of an independent European forum in which parties can engage on the issues;

- the development of European review bodies that can both scrutinise performance against commitments, and that can facilitate the translation of initial, more tentative commitments, into more formal ('contractualised') rights and obligations;

- the development of more explicit and more precise indicators of performance, to which commitments can be made and which can facilitate *ex post* evaluations of performance.

Risks and rewards

7. One of the most contentious issues in ATM is the allocation of risk, particularly of financial risk associated with traffic volatility. Current arrangements are manifestly deficient in this regard, and can give rise to significant short-term movements in charges that bear no relationship at all to movements in the costs of service provision. We have therefore considered some of the ways in which charges could be 'smoothed' or 'profiled' over time, so as to avoid erratic movements in charges and, hence, to avoid the financial difficulties to which they can give rise, and so as to establish less perverse, short-term correlations between charges and costs. Key points from this assessment include:

- There may good efficiency reasons for considering some degree of smoothing of user charges as desirable - in particular, the profiling of cost recovery may allow for a higher proportion of fixed costs to be recovered in periods when demand is less price sensitive and vice versa.

- We are skeptical of the efficiency benefits of funds (whether they be specified 'solidarity' funds or more general service provider reserves) where users are asked to pay in advance without receiving any firm and credible commitments in relation to what benefits they will receive in the future.

- Some ANSP exposure to traffic level changes in terms of revenue recovery allowances may be desirable in order to encourage flexible responses to changes in relevant circumstances. The potential for charge smoothing arrangements to weaken service provider incentives is an important factor to consider - that is, the extent to which smoothing arrangements can insulate ANSPs from desirable pressures that might otherwise be present to a greater extent.

8. We have proposed that the development of a 'Revenue Recovery Imbalance Account' could provide a framework within which alternative approaches to charge smoothing could be assessed and potentially implemented in a compatible manner. The aim of using such a framework is to seek to highlight the different dimensions – including in particular the range and extent of flexibility that is allowed for, and the constraints that are put on the usage of that flexibility - that impact on the desirability or otherwise of alternative approaches.

9. In our view, however, the most fruitful general approach to issues of risk is to go back to first principles of best-practice regulation and to recognise that there are close links between (a) establishing charging arrangements that encourage more efficient performance and (b) the resulting allocation of risk. **On this basis, we develop, in section 6, an over-arching framework -- labeled 'benefit sharing' -- for determining the *average en-route* charges of service providers.** The single, common framework allows for considerable flexibility in determining an appropriate

mix of encouragement and of risk sharing between providers and users. **The common framework encompasses cost-pass-through and price-cap regulation as special, limiting cases, and, via appropriate choice of parameters, it can be used to address issues that are usually associated with notions of 'solidarity', and that are therefore usually considered separately from other charging matters**. **It also provides for varying degrees of profiling or smoothing of charges over time, including via reliance on a Revenue Recovery Imbalance Account.** In our view, it is wrong in theory, and liable to lead to sub-optimal outcomes in practice, to address separately issues of encouragement/incentives, risk sharing, solidarity, and profiling/smoothing. They are best addressed together, and we have presented a framework in which this can be done.

Network management

10. One of the objectives of the Single European Sky initiative is to enhance ATM efficiency through the better use of capacity and improved congestion management. In this area, we believe that the enhancement of the roles of organisation entrusted with co-ordination activities is desirable. Specifically, **there is a strong case for the establishment of a European Network Manager** (or System Operator to use terminology from energy networks). The precise duties of such a Network Manager could lie along a spectrum that can roughly characterised as running from 'passive' to 'fully active', depending upon assessments of the likely effectiveness of national- and European-level contributions to co-ordination activities. Whereas in the rail and energy sectors co-ordination at European level remains fairly limited, the much greater relative importance of cross-border traffic suggests that a less passive European-level role would be more appropriate in ATFM.

11. In relation to the more specific questions concerning infrastructure projects, the Single European Sky initiative places emphasis on the promotion of greater levels of harmonisation, integration and interoperability in the development of infrastructure. The limitations of the current arrangements include difficulties arising from: lack of compliance and enforcement mechanisms at the European level; lack of accountability in relation to matters such as forecasts of future capacity requirements; and obstacles to the financing of collective or Pan-European infrastructure projects.

12. In our view the problems of the current system of ATM infrastructure interoperability and co-ordination at the European level could potentially best be addressed through the joint introduction of common Pan-European ATM infrastructure standards of interoperability and a European infrastructure oversight body responsible for identifying areas of potential improvement to the network as a whole and in encouraging greater levels of co-ordination with infrastructure projects.

Charging principles

13. The fourth part of the Report considers issues of charging structures, rather than the issues of average charge levels addressed via benefit sharing. It commences by setting out a number of high-level criteria that any system of charging will be expected to meet -- i.e. cost-reflectivity, non-discrimination, and transparency -- and

by discussing the meanings that can be attached to the relevant terms. We stress the importance of this discussion, because the relevant terms are open to a variety of interpretations, and they are sometimes used quite loosely to justify charging structures which, on closer inspection, turn out to be problematic or inefficient.

14. Key conclusions of this review included:

- We find no strong case for changing the weight factor used in the current *en route* charging formula, although some change in the exponent could be justified if carried out in conjunction with other changes that sought to capture cost differences in a different manner (for example, if there was an adjustment to the amount of revenue recovery linked to arrival/departure).

- The recovery of fixed cost components through non-distance related charges (for example, as part of a two-part tariff) can give rise to significant distortions of competition (in particular, through what is referred to in the energy sector as 'pancaking'). Such an approach would be highly undesirable for *en route* charging - other than as part of the development of an approach of the kind set out in Section 9 (ie. origin/destination/distance charging) – as the negative effects on competition are likely to be significant.

- The potential for the development of coherent and uniform congestion charging arrangements should be considered as part of a more general evaluation of potentially desirable longer term developments to air traffic flow management arrangements at the European level.

15. Following this general analysis, the possible roles of each of a number of specific factors that might feature as differentiating factors in any charging structure -- such as aircraft weight, time of day or year, distance travelled, altitude, etc. -- are explored. Our view is that use of any of a range of differentiating factors, and hence tariff structures, is compatible with high-level criteria such as cost-reflectivity, non-discrimination and transparency. **Determination of the most appropriate structure of charges must therefore rest upon their anticipated effects on system performance and on matters of practicality**. Included among the effects to be taken into account are the differential impacts of any particular charging structure on airlines, and the implications of these impacts for competition among airlines.

Charging options

16. Finally, we set out and discuss a number of different options for determining the structure of *en-route* charges: an origin/destination/distance approach; charge differentiation according to flight level; and charging by ACC. We note that the current arrangements, based on a division between terminal and *en-route* charges, are similar in structure to a more formalised origin/destination/distance structure, which in turn is not unlike the entry/exit/cross-border charge arrangements that are evolving in European energy networks. **The main problem with the current structure is, therefore, in our view, not that is inappropriate as a way forward, but rather that the division between terminal and *en-route* charges is not**

harmonised across EUROCONTROL member states. In consequence, and on the basis of some exploratory econometric analysis, we endorse the conclusion of the earlier PWC Report that the terminal charges set by some member states are too low; or put another way, that, for a number of ANSPs, the tariff structure is imbalanced.

17. In our view, **there are three major improvements that could be made to the current charging structure**:

- It should be formalised in terms of origin, destination, and distance charges.

- The origin and destination elements should not be set so as to reflect operations and costs that are specifically related to a particular flight phase. **Rather they should reflect the costs 'driven' by take-offs and landings, wherever those costs actually fall in operational terms.** Such costs include those arising from complexities in the control of airspace surrounding the point of origin or destination. Origin and destination charges might also be used as tariff components to which costs that are 'fixed' in nature could be allocated, efficiently and without discrimination.

- Constraints should be placed, either by agreement or by mandation, on the proportion of allowed revenues that service providers can recover from origin and destination charges on the one hand and from distance charges on the other hand, with the aim of achieving greater harmonisation in charging methodologies across European ANSPs. Such constraints could allow for the different mixes of traffic -- take-offs and landings, overflights, average distance controlled -- handled by different service providers

18. We have also analysed charge differentiation based upon a distinction between upper and lower airspace. The motivation behind this proposal is that control costs per service unit are lower in upper airspace, where flight paths are typically subject to fewer vertical movements than at lower levels. Whilst recognising this general point, we conclude that charge differentiation according to flight level is not likely to prove a satisfactory alternative to development of current arrangements into an origin/destination/distance structure. The most important reason for this is that, **on the basis of the available evidence, we can see no reason for believing that there is any very distinctive change in control costs that can be linked with a particular flight level**. Thus, whilst an upper/lower charging structure might be said to be cost reflective in some very broad and general way, it would not be cost-reflective within a range of altitudes around the designated boundary between upper and lower airspace. Substantial differentiation between upper and lower airspace *en-route* charges could therefore potentially be held to be discriminatory, and, because of this difficulty, the quantitative scope for charge differentiation would likely, in practice, to be limited. For this reason, **we conclude that upper/lower airspace charge differentiation is better seen as a potential supplement to the development of the current charging structure**. Whether or not there would be any material benefits from such further differentiation is far from obvious, but it is a matter that merits further investigation.

19. Finally, we consider possible charging options based upon each ACC setting its own *en-route* charges. The motivation for the proposal derives in part from the notion

that this would enable ACCs to compete with one another for traffic, and that the resulting competitive pressures would provide encouragement for improved performance by service providers. **Our judgement, however, is that the likely extent of inter-ACC competition can be expected to be generally weak and that, where greater competitive pressures are feasible (e,g, in upper airspace), it would be very easy for providers to take steps that would serve to mitigate those pressures (e.g. by co-ordination of pricing).** Put simply, the 'competitive paradigm' does not appear to be appropriate for ATM, although, where they exist, the possibilities for airlines to 'bypass' inefficient service providers can place incremental pressures on those providers. Again, therefore, **we conclude that ACC charging should be regarded as only a supplementary measure** (which is perhaps most useful in facilitating the restructuring of service provision in upper airspace), **not as an alternative to the further development of existing arrangements**.

General conclusion on charging

20. Pulling the various threads together, our overall conclusion is that, in seeking to promote the efficient development of the European ATM network, greatest reliance should be placed upon the pressures that can be brought to bear on the structure and performance of service providers from a combination of economic regulation and the more active involvement in network governance of users. In relation to charging, the key to progress most likely lies in the application of regulatory approaches such as that embodied in the benefit sharing proposals set out in this Report, which are targeted at influencing the average level of en-route charges, the incentives faced by service providers, and the distribution of risk between service providers and users. Relative to these matters, the determination of the precise charging structure to be implemented is of secondary importance.

21. That is, well constructed charging structures can be expected to have only a supporting (but nevertheless useful) role in guiding future developments in ATM. Our general view on charging structures is that there is merit in developing existing arrangements in ways that reflect costs of complexity by means of origin and destination charges (i.e. reformed terminal charges), which should be set according to principles that will ensure greater consistency throughout the European network.

SECTION 1 - INTRODUCTION

1.1 Background

(a) The Single European Sky draft Regulations

In 1999, the European Commission published a communication expressing its intention to introduce a package of measures to reorganise the air navigation services sector in Europe, resulting in the creation of a 'Single European Sky'.[1]

The introduction of this communication was followed by the establishment in 2000 of a 'high level group' of senior civil and military air traffic control authorities charged with examining the key elements of the Single European Sky proposal. In its final report tabled in November 2000 the high level group proposed (in relation to charging) that:

> *The charging mechanism should stimulate cost-effectiveness and include incentives for practices that increase capacity to enhance system-wide efficiency, whilst maintaining a high level of safety* [2]

In late 2001, the European Commission submitted its legislative proposals for the consideration of the European Parliament and the Council which included draft Regulations relating to: the framework for the creation of the Single European Sky[3]; the provision of air navigation services in the Single European Sky[4]; the organisation and use of the airspace in the Single European Sky[5]; and the interoperability of the European air traffic management network[6].

The draft directive of most direct relevance to the current study is focused *on the provision of air navigation services in the Single European Sky*. Chapter III (specifically Articles 13-15) of this draft Regulation establishes the general principles relating to air navigation charges within the framework of the Single European Sky, while also providing for the review of charges by the European Commission to ensure compliance.

(b) The current en-route charging system

The current *en-route* charging system used by the 15 Member states of the European Union is based on a multilateral agreement [7]. The key features of the *en-route* charging system established under this agreement are that:

[1] COM (1999) 614 final dated 1st December 1999
[2] *Single European Sky: Report of the high level group*, European Commission Directorate General of Energy & Transport, November 2000, page 30
[3] COM (2001)123 final/2 dated 30 November 2001 - 2001/0060 (COD)
[4] COM (2001)564 final/2 dated 11 December 2001 - 2001/0235 (COD)
[5] COM (2001)564 final/2 dated 11 December 2001 - 2001/0236 (COD)
[6] COM (2001)564 final/2 dated 11 December 2001 - 2001/0237 (COD)
[7] In addition to the 15 EU Member States, signatories to the Agreement include other countries, including ascension countries.

- The system has traditionally applied a full cost recovery principle to ATM service providers with pricing based on average costs. Under this full cost-recovery system, the unit rate for each country is calculated by dividing the forecast costs for the national service provider by the forecast number of service units. Differences between the charges collected and the actual costs in a given year are recorded and carried forward to the year n+2.
- Since July 1999, however, the option has also existed for the full cost recovery system to be replaced by the setting of price caps, providing certain measures are introduced (such as independent economic regulation and a commitment to introduce incentives to reduce cost and increase the efficiency of service provision).
- The calculation of charges levied to system users for *en-route* services employs a common methodology across the Member States, and other signatories to the Multilateral Agreement, that involves the multiplication of three basic elements: aircraft weight; a distance factor and a unit rate of charge.
- System users currently receive a single invoice for each flight that is the sum of the individual charges for each State that supplied *en-route* services during that flight.

The European Commission has indicated that it considers one of the principal weaknesses with the current charging system to be the absence of incentives that it provides for both system users and service providers to optimise the use of existing capacity, or to respond to signals to invest in new capacity. This is consistent with the findings of the Performance Review Commission that have noted that the current charging system does not create the appropriate incentives to encourage improved system effectiveness.[8]

In relation to terminal charges, most Member State currently levy a charge for terminal navigation services but unlike *en-route* charging there is currently not a common charging methodology across the Member States.[9]

An important aim of the Single European Sky package of proposals therefore is to examine how appropriate encouragement mechanisms might be developed under different charging systems to allow for better utilisation of existing capacity and to provide for efficient investment in new capacity.

[8] "The present full cost recovery regime does not provide incentives to deliver performance and to be responsive to user needs beyond levers normally available in the public sector. With the current system, airspace users are bearing most, if not all, of the business risks. On the one hand, if demand is higher than expected or if the planned capacity is not delivered, airspace users will incur higher delays. On the other hand, if demand is lower than expected or actual costs are higher than planned, the airspace users will incur higher charges." See: *An assessment of Air Traffic Management in Europe during the calendar year 2002* PRR 6 Version II, EUROCONTROL Performance Review Commission, May 2003, page 54

[9] In respect of terminal charges see *Study of the Terminal Charges for Air Traffic Control Services*, PriceWaterhouseCoopers, Final Report – March 2001

1.2 Scope of the current study

(a) Tender reference and terms

This study has been commissioned under tender contract TREN/f2/28-2002 to assist the European Commission Directorate General of Transport and Energy (DG TREN) in defining the rules required to implement the major provisions of the Single European Sky legislative package in relation to economic regulation.

Specifically the main objective of this study is to propose possible charging mechanisms that will promote the development of appropriate incentives on both airspace users and service providers to improve the system effectiveness.

Within the context of the study, improvements to system effectiveness include:

- Increases in capacity where needed
- Improved use of existing capacity
- Improvements to the sharing of risks between users and providers, while still providing for necessary investments to allow for additional capacity

For the purposes of this study the development of appropriate incentives for charging mechanisms have been assessed while bearing in mind the Commission's proposal to introduce functional blocks of airspace in Europe. In addition, we have been asked by the Commission to extend the study beyond just an examination of the development of appropriate incentives for *en-route* charging mechanisms to include an examination of terminal charging where it was considered relevant.

Other issues we have been asked to consider as part of the study include[10]:

(i) **To better define financial incentive mechanisms that will encourage increases in capacity, while improving the use of existing capacity.** This is addressed throughout the report, but specifically in sections 4, 5 and 6.

(ii) **The creation of a solidarity mechanism to provide for: the temporary compensation for loss of income by ANSPs in the event of an important fall in traffic and for the financing of collective projects.** This issue is addressed in two sections of the report. The potential for a solidarity mechanism to provide for temporary loss of income is addressed in section 5. The issue of the financing of collective projects and the role for a solidarity mechanism in that context is discussed in section 7.

(iii) **Specific measures in relation to the concept of functional blocks of airspace.** In section 2 we discuss the concept of FABs and how they conceptually relate to charging, while section 9 examines the incentives that different charging schemes might provide for the creation of FABs. Sections 4, 5 and 6 can also be

[10] Although we were initially asked to include an examination of the potential for internalising external costs as part of the original tender, this element of the study was later removed.

understood to be of direct relevance to FABs where they are concerned with the incentive arrangements faced by service providers.

(b) The study's approach

As noted above, the high-level objective of this study is to propose a set of possible charging mechanisms *that will provide the* appropriate incentives *upon both* system users *and* service providers to *increase* system effectiveness.

It is important to note that this study is focused on the *charging* issues related to the Single European Sky package and not on *operational* issues related to the Single European Sky package. Thus, this study is focused on the development of arrangements which will increase system effectiveness from an economic/financial standpoint – that is, provide the desirable conditions to ensure that system capacity is both developed and used efficiently, and that the financial risks associated with ATM are shared efficiently between system users and service providers.

Our approach to examining and assessing potential charging mechanisms has therefore been influenced by the expected responses that the introduction of different potential charging mechanisms will have on both system users and providers and, more specifically, whether these responses will be consistent with increased system effectiveness. Put another way, **we have examined potential charging mechanisms that are likely to encourage behaviour in both system users and service providers which will either: increase system capacity; allow for better usage of existing capacity; or improve the sharing of risks between users and service providers.**

In examining these issues this study draws upon a range of source materials including reports, position papers, studies and consultations. These materials have been submitted or prepared by organisations such as: the European Commission; various Member states, EUROCONTROL; national service providers both individually and through representative bodies such as CANSO; commercial airlines; airline user groups such as IATA, AEA and ERA; the ETC ATC; and from other ATM systems organisations such as NAV CANADA, the FAA, Airservices Australia and Airways New Zealand. The study also draws upon the experience of other European regulated sectors such as the rail, electricity and gas sectors, as well as the broader economic and regulatory literature relating to economic incentives and network industries.

1.3 Structure of the report

This report is divided into nine sections and two appendices. The first part of this report (sections 2 & 3) provides the context for the study and begins with a review, in **Section 2,** of the restructuring issues in ATM. This is followed in **Section 3** with a review of the theory and practice of price regulation in network industries, highlighting the areas that are likely to be of greatest use for ATM.

The second part of this report (sections 4 –6) begins in **Section 4** with an examination of the scope for contractualisation in ATM and in particular seeks to highlight ways in

which more formalised arrangements in ATM may be developed. **Section 5** considers some of the specific issues raised by traffic volatility in ATM and examines potential for smoothing mechanisms for charging. **Section 6** develops a general charging framework that can allow a considerable degree of flexibility in implementation, taking account of both allocation of risk and 'encouragement' factors.

The third part of this report (**Section 7)** discusses potential approaches to improving coordination in both flow management and infrastructure investment at the European level.

The final part of this report begins in **Section 8** by outlining and reviewing some general criteria against which charging arrangements in ATM can be assessed, and by assessing a number of factors that are, or might be, used when defining charging structures. This is followed in **Section 9** by an evaluation of some alternative options for charging structures that might be promoted via the Single European Sky implementation rules.

SECTION 2
REVIEW OF RESTRUCTURING ISSUES

2.1 Introduction

Currently, the European ATM system is built around national ANSPs that each enjoy a monopolistic position in their respective parts of airspace. Although the arrangements have, in practice, performed adequately, such fragmentation of provision can give rise to a number of problems that can cause performance to be less good than it could be under alternative arrangements. Thus, there is strong evidence indicating that European ATM costs are high, at least in relation to costs in the US[11], although it is an open question whether this is largely attributable to factors directly linked to fragmentation (e.g. loss of scale economies, inefficient duplication, higher costs of co-ordination, inefficient sectorisation, etc.) or to other factors such as management effectiveness or weaknesses in co-ordination.

Issues surrounding fragmentation of provision are not unique to ATM networks. Thus, the European electricity network exhibits some features that are similar, and even within single national jurisdictions there can be substantial structural fragmentation of electricity systems – as has historically been the case in Germany and the USA. However, compared with a network such as electricity, the air transport network is characterised by a very high level of 'cross-border flows' (i.e. international flights).

Addressing potential performance problems arising from fragmentation in ATM is one of the main objectives of the Single European Sky initiative. In this context, it is to be stressed that 'decentralised' provision of ATM is not itself a problem. What is at issue is the achievement of a structure of provision that will be most effective in serving the requirements of airlines. *A priori*, it is highly unlikely that a structure of provision built around historic political boundaries will satisfy this condition, and an aim of the Single European Sky initiative is therefore to remove existing barriers and obstacles to the development of more effective systems of provision. It is anticipated that structural adjustments will occur by means of co-operative arrangements among existing service providers, facilitated by Single European Sky regulations. The introduction of the concept of Functional Airspace Blocks (FABs) is a specific example of such 'facilitating regulation'.

2.2 Dimensions of restructuring: operations/service provision vs. charging

Before reviewing current restructuring issues in European ATM, it will be useful first to introduce a crucial distinction that is central to much of the rest of this Report. **It is the distinction between restructuring of ATM provision (operations) and restructuring of ATM charges.**

[11] "A comparison of performance in selected US and European en-route centres", EUROCONTROL Performance Review Commission, May 2003

Much of the past debate on ATM restructuring in Europe has focused on service provision. That is, attention has been centred on how national providers might co-operate in the design and management of airspace so as to improve operational effectiveness. The implications of FABs have, for example, often been analysed in these terms.

The drafting of regulations for the Single European Sky requires, however, that attention also be devoted to issues of charging for ATM services. **This raises important questions concerning the relationship between airspace blocks defined for operational purposes and airspace blocks defined for charging purposes.**

The first point to make is that there is no *necessary* co-incidence between these two types of airspace block: they may or they may not be the same. In order to maintain the conceptual distinction, and to facilitate later analysis of the relationships between the two, we will make the following distinction:

An Operational Airspace Block (OAB) is a block of airspace in which control is entrusted to a single ATM provider, whether that provider is one part of an existing organisation, a stand-alone organisation or a co-operative joint-venture among several organisations. Currently, the architecture of OABs is defined by ACCs, and OABs may be smaller than national airspace (e.g. in larger EU member states), the same as national airspace (e.g. in smaller countries), or trans-national (e.g. Maastricht). An OAB corresponds with the Single European Sky concept of a FAB, but we have not used the latter terminology here so as not to confuse the pre- and post-implementation periods.

An Airspace Charging Block (ACB) is a block of airspace in which en-route charges are fully harmonised. By this is meant that charges are either the same or are set jointly, either by a single organisation or collectively by a group of co-operating organisations. For example, charges might differ according to location in order to reflect congestion, but if the charges are determined by one organisation or by collective agreement, the relevant airspace would still be regarded as a ACB. Currently, ACBs are defined by national boundaries, illustrating the point that OABs and ACBs need not be the same.

Given this distinction, it can be seen that restructuring issues involve questions about not only the evolution of OABs, but also the evolution of ACBs and the relationships between the two.

2.3 Recent developments in European ATM and their implications for charging

Before considering some more specific aspects of the restructuring issues, it can be noted that there have been two major strands of recent developments that may have significant consequences for the issues of economic regulation and charging that will be considered in later sections of this Report. The first relates to the governance of service provision, the second to EUROCONTROL charging procedures and initiatives.

2.3.1 Governance in service provision

In relation to governance in service provision, there are two developments that are of particular significance when assessing various, alternative options for ATM charging.

(a) Commercialisation of service provision

First, the draft Regulations has endorsed the dynamics of increased autonomy of ATM providers, [12] based upon the formal separation of economic regulation from service provision, which has traditionally been vertically integrated to governmental administration. The draft Regulations also address the development of working relationships between service providers, [13] and it is of particular relevance in the case of the formation of cross border FABs where national member states will have to agree the designation of one or more service providers to manage the block of airspace [14].

Thus, the new environment, more commercialised in orientation, may facilitate contractual financial arrangements in service provision and revenue sharing. Given this, one particular issue is to establish whether ANSPs will be granted increased freedom in deciding their own tariff structure and how charging alternatives would best serve the objectives of the Single European Sky.

(b) Unbundling services

Second, there is a more general trend toward separation of activities in ATM provision. Moving towards corporatisation has meant creating national ATM monopolies, initially based upon integrated provision of the different service components, such as: infrastructure, R&D, operations, support services (e.g. meteorological services), etc. Some of these services can, however, be outsourced and/or may be organised in different ways. By the same token, separation of provision from administration introduces a distinction between regulatory costs and cost of service provision. One particular set of question that arises, therefore, concerns the extent to which the various service components should be unbundled, whether from the perspective of service provision or of charging, and how the relevant costs should be recovered.

In summary, the increasing commercial autonomy of ANSPs, coupled with the potential unbundling of service components (e.g. infrastructure, operations, supports services) and of costs (e.g. regulation, service provision), presents new challenges for the future development of charging structures.

2.3.2 EUROCONTROL task forces

EUROCONTROL has been developing proposals for new approaches to charging along two dimensions, namely: changes in the airspace architecture for charging purposes, and the possible alternative charging formulae.

[12] Council of the European Union. 15853/02, 11 March 2003, (1) p.2.
[13] Ibid Article 9 (2) p.14.
[14] Ibid Article 7 (4) p.13

(a) The current airspace architecture charging model

As already noted, the current charging regime is governed by the pattern of national boundaries. To each national airspace territory there corresponds an *en-route* charging block/zone (ACB), to which is attached a national unit rate calculated according to common principles – i.e. recovery of the costs of ATM provided in the *en-route* charging block/zone.

Each charging block/zone can cover a number of national Flight Information Regions[15]. FIR boundaries do not change easily: they need to be agreed at the ICAO level between countries.

The delineation of FIRs does not always correspond to operational requirements and therefore the costs levied in a charging zone may not reflect the underlying costs of service provision. The service provider of a member state may also partly control an airspace portion of another Member State. In general, this occurs largely at the fringes of the relevant airspace blocks, and it can work both ways (a member state may receive and/or provide service to another member state). Therefore, the member states tend to consider that the reciprocal services are sufficiently balanced, and that no financial transfers are appropriate. Since each provider's costs are ultimately recovered, and since entering into bilateral agreement can be a cumbersome process, there is little incentive to align the unit rate more closely with the costs of service provision in the relevant airspace.

Since these issues currently largely arise along the fringes of airspace blocks, the resulting economic effects are likely to be only limited in extent. However, in cases where the implications of reciprocity become more significant, albeit still of limited geographical scope, formal delegations of service provision may be agreed (for example, as between France and Germany with Switzerland). [16]

The current FIR architecture model of charging - computation of the distance flown by FIR and allocation of charges according to the limits of established FIRs, rather than by airspace controlled by providers - is enshrined in the EUROCONTROL multilateral agreement and in the legislation of some countries. To date, EUROCONTROL member states have tended to prefer to stick to the collection of revenues accruing within their own, national FIRs, but, in principle, other ways of charging and billing can be envisaged. To change the current regime, however, would require unanimous agreement at the EUROCONTROL Commission level (Ministers of the Member States).

In this respect, the Single European Sky introduces a significant change. It envisages the creation of a single FIR in the European upper airspace independent of national boundaries. Therefore, at least in the upper airspace there will be immediate scope for defining *en-route* charging blocks/zones independent of national boundaries.

[15] An airspace of defined dimensions within which flight information services and alerting services are provided.

[16] For a comprehensive analysis of delegations, see Francis Schubert: *The Financing of Cross Border Air Traffic Services, a legal perspective,* Skyguide, 2003.

More generally, the current architecture of FIRs does not necessarily limit the development of new ACBs and associated charging mechanisms in other parts of airspace. It is potentially open to ANSPs to harmonise charges and create an ACB whilst maintaining initial revenue allocations based on the existing FIR arrangements. Any unwanted effects on the financial positions of individual ANSPs could then be addressed via side agreements (i.e. on compensatory financial flows).

The EUROCONTROL Central Route Charges Office (CRCO) has studied options of adapting its procedures for calculating charges so as to be able to apply them to any controlled airspace, irrespective of existing FIR boundaries. Using data from CFMU together with the RSO tool [17], it would appear that the CRCO would be able to define a flexible method of charging based upon consideration of the point of entry to and the point of exit from an airspace block, coupled with the distance flown according to the preferred route in the flight plan. Such an approach could be adapted to any of a range of different ACB configurations, including those outlined in Section 9 below. [18]

(b) EUROCONTROL task forces on charging

A considerable amount of work has already been done on various options for the development of charging arrangements, whether based upon the current architecture of ACBs or on alternative configurations that might emerge in the future.

The EUROCONTROL CRCO conducted a series of simulations in 2000 and 2001 to assess the extent to which changes in the current charging formula might help to promote a better use of airspace [19]. Among the possibilities for price differentiation under consideration were charging: according to peak/off-peak periods; by ACC; by sector; and by overflights/landings-and-take-off as a proxy for an upper/lower airspace distinction. In addition the CRCO has put forward additional proposals on the possibility of two-part tariff arrangements. [20]

The task force also took stock of proposals in relation to terminal charges that had been put forward in the Commission's study on terminal charges [21], and in particular of the possibilities of extending the 20km rule to a 40 or 80 km rule, and of creating separate charges for particular services (AIS, MET, CNS Services).

The principles that guided the CRCO work included improving cost reflectivity in *en-route* charges, defined in a broad sense, and as such were not focused on the wider policy issue of cost effectiveness. The principle of cost reflectivity will be discussed more fully in Section 8 below, but here it can simply be noted that, at least in one of its

[17] The Route per State Overflown tool allows a measure of the distance flown according to the description of the route indicated in the last filed flight plan.
[18] See internal CRCO/EUROCONTROL memo: *Preliminary study: differentiation between upper and lower airspace –charging area defined independently from UIR/FIR boundaries.*
[19] Enlarged Committee for Route Charges, minutes of meetings of the Task Force on Possible Pricing Mechanisms (covering year 2000 and 2001).
[20] See PPM Task force minutes of 13 February 2001.
[21] PriceWaterhouseCoopers: *Study on terminal charges for Air Traffic Control services,* EC, final report March 2001.

interpretations, it can be understood to incorporate costs imposed on other users [22], via delays for example, as well costs borne by ANSPs.

Consideration of two-part tariff schemes by the CRCO was based on the distinction between fixed and variable costs in ATM provision. Given this distinction, it is possible to envisage a charging scheme based, say, on a per flight or per weight factor component linked to fixed costs and a distance-related component linked to variable costs. It can be noted, however, that any split between fixed and variable costs is highly dependent on the time period over which such costs are determined – the proportion of costs that are deemed variable will increase with this time period – and that this (time-period) issue needs to be explicitly addressed before cost allocations can be determined.

Another aspect of this strand of work was evaluation of the potential financial impacts of alternative charging arrangements on the various categories of users (regional airlines, international carriers etc.). This is important because any change will tend to have consequences that fall unevenly on airlines. Among other things, such effects can, if substantial, have material effects on the state of competition in airline markets, as well as on factors such as service frequencies to and from less densely populated areas.

There is, then, already a solid foundation of work on which assessments of options for the development of charging structures can be based. This analysis needs, however, to be developed in a wider framework of evaluation, for a number of reasons. **First and foremost, reform of charging structures alone might, by and of itself, do little to encourage operational restructuring or to promote operational efficiency. Second, the relatively static conceptual frameworks in which much economic analysis addresses charging structure options need to be developed and refined to address investment and capacity issues (the 'time-period' questions).** Third, there is a range of different approaches to charging that can be said to be 'cost-reflective', which have differing implications for users and for overall efficiency. There is a need, therefore, to have recourse to a wider analytic framework that considers how users will react to the alternatives and how those reactions will in turn affect ANSPs. Put simply, effective charging structures will reflect demand-side considerations as well as supply-side cost considerations.

In summary, recent developments in European ATM provide scope for charging restructuring which, among other things, moves away from ACBs that are determined by national boundaries. The future relationships between evolving ACBs and airspace controlled (OABs) are matters of some importance that will be examined in later sections of this Report.

2.4 Operational restructuring

In this sub-section we will review first some key characteristics of ATM service provision, and second the experience of the Maastricht ACC and the planned new ACCs: NUAC and CEATS. This is to provide both a guide to the ATM structure with

[22] In economic terminology, these are a type of *opportunity cost*.

which regulators, including the Commission, initially have to deal, and an indication of the operational restructuring that has either already occurred or is in the process of implementation.

2.4.1 Service provision characteristics and the Single European Sky

Potentially adverse consequences of fragmentation in service provision in Europe have been highlighted in many past documents. Of particular note is the PRC Report that compared, according to various measures, the performance of a sample of ACCs in Europe with a sample of ACCs in the USA[23]. The PRC conclusion was that, although there was significant variation among ACCs, the average US performance was substantially higher than the average European performance. In consequence of the way in which the samples were selected, this difference could not be attributed to differences in traffic complexity.

As explained in the introduction to this section, the contributions of various 'explanatory' factors responsible for the observed performance difference have not been fully evaluated (and the PRC study did not seek to do so, being focused on measuring and comparing performance, rather than on explaining performance). Further, it would be much too simplistic to jump to a conclusion that a large part of the difference is attributable to 'fragmentation' and that consolidation will necessarily lead to significant improvements in efficiency in Europe. It is well known from studies of consolidation in other sectors -- which have been stimulated in part by an interest in assessing, for merger control purposes, the likely economies of scale and scope to be expected from increased concentration in supply – that the mere fact of putting two organisations together is no guarantee of improved performance.

Notwithstanding these cautionary points, the current structure of ATM in Europe must, whether independently or considered in conjunction with other factors, be potentially considered to be a significant factor in accounting for certain performance weaknesses in the European system.

(a) Characteristics of the current structure of service provision

The following review of the current structure of service provision is based on data prepared by the EUROCONTROL Experimental Centre [24]. Data have been collected for 34 ANSPs (considering Maastricht UAC as an independent entity), covering 64 ACCs. The source data refer to eight days (4 days in summer, 4 days in winter) in 2001. The categories applied to flight phases are approach (<FL95), lower airspace (FL95-FL285), and upper airspace (FL285-FL470). Data were also assessed using a FL245 threshold for upper airspace. Using FL245 or FL285 does not give rise to any significant difference in conclusions, at least at the broad level with which we are concerned here.

In relation to ANSPs and ACCs, the points that can be made include:

[23] "A comparison of performance in selected US and European en-route centres", EUROCONTROL Performance Review Commission, May 2003

[24] We would like to thank here Patrick Ky and Claire Leleu for the help they provided.

- Of the 34 ANSPs, 19 operate only one ACC Centre, and 8 two ACCs. Only the largest European countries have a portfolio of ACCs, but the numbers are relatively small, falling in the range 3 to 5.

- Of the 64 ACCs, only 3 are clearly specialised upper airspace centres (Maastricht, Karlsruhe, London ACC).

- Corresponding to the three specialist upper airspace centres are a number of ACCs that specialise in the management of the lower airspace (and approach) lying below the upper blocks. These are:

  For Maastricht: Amsterdam, Brussels
  For Karlsruhe: Frankfurt, Bremen
  For London ACC: London TC, Manchester.

  In addition, one other European ACC, Dublin, is specialised in managing lower airspace.

- There have, therefore, been some, albeit still limited, departures from an operational airspace architecture based upon single OABs defined purely by national or sub-national boundaries National boundaries have been traversed in one of the three cases, Maastricht, where the OAB lies over the Netherlands, Belgium, Luxembourg, and part of Germany.

- The remaining ACCs, which comprise the large majority, control all airspace above the relevant territories in which they are located (although ATC can feasibly be de-localised). This structure appears to reflect the history of European systems: the location of ACCs dates back to thirty to forty years ago when ACCs were built where radar was installed, and national sovereignty obligations covering small airspace blocks likely explains the lack of specialisation and de-localisation. Thus, particularly in small countries, the national ANSPs have had to control simultaneously the lower airspace around their main airports and a relatively small volume of upper airspace.

In relation to air traffic control operations, it can be noted that:

- The number of controlled flights is higher in the lower than in the upper airspace. One reason is that some intra-European flights do not reach FL245 or FL285. However the specialised upper airspace centres (Maastricht, Karlsruhe, London ACC) display similar numbers of controlled flights to the most busy lower airspace locations. In other words, fewer planes are flying in the upper airspace, but the upper flight control is concentrated in a smaller number of Air Traffic Control Centres.

- The average mean route length (MRL) in the lower airspace is fairly even among ACCs, at around 100km per flight. This distance corresponds, at least to a reasonable approximation, to the final descending and initial ascending phases, the procedures for handling which are relatively similar at most airports. In the

European upper airspace, above FL285, the mean route length is in most cases twice as great (i.e. around 200km), but for some ACCs is four times or more times as great. Examples include the Brest, Shannon, and Scottish ACCs, which control oceanic traffic.

- airspace density (total controlled kilometres by adjusted surface), which provides an indicator of the intensity of control, does not vary greatly between lower and upper airspace. But in terms of the average number of flight level changes (vertical movements), the data indicate that there are around three times the number of movements in lower airspace than in upper airspace. Taken together with the fact that the MRL is shorter in lower airspace than in upper airspace, this is consistent with the view that lower airspace is the more complex to manage/operate.

Given these characteristics of the current situation in relation to service provision, we now consider various routes via which restructuring might occur.

(b) Division of service provision according to flight phases

Phases of flight divide airspace into broad categories[25]. Broadly speaking, the categories in use are *en-route* (covering both upper airspace > FL245, where most flights are level flying, and lower airspace, between FL95 – FL245, where many flights are in ascending or descending phases after taking off or before landing); approach control, which is the terminal phase of the flight between *en-route* and aerodrome control; and aerodrome control, the phase directly associated with take-offs and landings.

Whilst these categories are widely accepted, the resulting division of flight phases does not map automatically on to operational divisions. An airport may have a control tower to which may or may not be attached an approach control centre with sectors and controllers. The approach control itself varies in terms of flight level and distance. In other cases approach control may be provided from an ACC that also controls part of the *en-route* phase of flights. Such an ACC may be specialised in lower airspace and approach control, or may deal with *en-route* control at both lower and upper levels as well as approach.

The 'non correspondence' between operational terms and airspace categories raises two important issues. First, the use of distances to/from airports and/or flight levels to determine ACBs and charging structures will not lead to charging structures that correlate exactly with operational units whose costs might be reasonably straightforward to estimate. The question is: does this matter? And the response, which we will consider in Section 9, has obvious implications for restructuring, whether of operational units or charging blocks or both. Second, from a user perspective, it has to be considered whether being charged according to the flight phase or flight level will have significant implications for the conduct of airline businesses.[26]

[25] Ibid PWC study, pp.2-4
[26] Ibid PWC study, p.36

(c) Flight efficiency and route choice

Flight efficiency can be approximated by direct routing and great circle trajectories [27]. The most efficient way of flying – taking account of gravity, air friction, and the resulting fuel consumption – is to fly to higher altitudes relatively quickly and to reach an optimised level apex before the final descending phase of the flight. Fuel consumption is also affected by the number of level changes during the flight [28].

Without being able to provide hard quantitative evidence, interviews with personnel in airline operational centres, with pilots, and with ANSPs seem to indicate that airline companies select flight plans (flight level, route) primarily on the basis of timing, delay effects and of impacts on connected flights (particularly when hub-and-spoke strategies are being adopted). Thus, at peak hours when congestion gives rise to delays, new last minute flight plans may be filed to avoid congestion. Among other behaviour patterns that have been reported and acknowledged by ANSPs, is the avoidance by airlines of particular routes or sectors (e.g. Belgian and Swiss airspace).

In principle, therefore, users may exert some pressure, through options for substitution among flight plans, on ANSPs. Such pressures could, again in principle, encourage ANSPs to provide services that meet user requirements in terms of costs, capacity required to avoid undue congestion, efficient route provision (according to length, number of vertical movements, etc.), and so on. **In practice, however, the evidence that we have seen does not suggest that these pressures are likely to be very material in quantitative terms, and this is important in that it has direct implications for the evaluation of those potential options for restructuring and improved ATM performance that are based upon the notion of 'increasing competition'.**

(d) Two views of the Single European Sky and its implications for operational restructuring

From one perspective, in order to address issues of fragmentation of service provision, the envisaged Single European Sky regulations can be seen as resting on three important changes:

- the delineation of a European upper airspace above FL285 combined with the creation of a European FIR;
- the anticipated introduction of FABs, in upper airspace, which are not constrained by national boundaries; and
- the possible management of each FAB by one provider or by an alliance of providers.

[27] See "Study for the European Commission on the Regulation of Airspace Management and Design", Wilmer, Cutler & Pickering, 14 May 2001

[28] Usually any flight conforms to predefined routes. When flying time is less than 50 minutes, it is uncommon for the aircraft to go above FL245. The pilot can make a request to be allowed to proceed to a higher level, depending upon the availability of capacity, but any such excursion would typically last for only a few minutes.

According to this perspective, the approach might be interpreted as one in which lower airspace is considered to be largely 'national', encompassing airspace blocks close to airports and with ascending and descending phases, and in which the upper airspace is considered to be largely 'European', encompassing the control of overflights or cruising flights travelling from one country to another.

Continuing with the argument, the introduction of the division between upper and lower airspace - with the provision for separate upper and lower airspace charges - would promote a reconfiguration of the service provision in the upper airspace, which in turn would require at least some reconfiguration of lower airspace management. Member states will agree on FABs and designate the accountable providers, subject to a recognition process that requires approval from the 'Single European Sky Committee' and the fulfillment of particular, designated conditions. Among other things, the draft airspace regulation package requires FABs to be supported by a safety case. The FAB is also required to: enable optimum use of airspace taking into account traffic flows; be justified by overall added value; ensure transfer of responsibility for air traffic control between ATS units; and ensure compatibility between the configuration of upper and lower airspace. [29]

Broadly speaking, these changes might be viewed as leading to the following 'virtuous circle'. Gains in merging upper airspace blocks will be substantial, arising from factors such as the ease of innovation and the benefits of adopting common systems, including reduced maintenance costs. **Charges for the use of upper airspace will fall substantially, providing incentives for smaller countries, operating independently, to move out of upper airspace ATM and to specialise in lower airspace ATM. Over time, large upper airspace ATM providers would offer services at lower levels, thus disseminating the efficiency gains more widely.**

We will call this the "upper airspace restructuring is key" (UARIK) view.

Against this can be set a perspective which sees the Single European Sky initiative as both less programmatic, and more diffuse and wider in its intended effects. It can be summarised as follows.

The Single European Sky legislation will permit and facilitate the creation of FABs and, subject to the necessary implementation rules, ACBs in upper airspace. However, on the alternative view, the Commission has proceeded in this way not so much because this is seen as the key that unlocks the door of comprehensive restructuring, but more because it is simply one of the things that it was both feasible to do and that, considered on its own and without any further, far reaching implications, it was desirable to do.

The evolution of the architecture of airspace, and of the structure of ACCs, will, however, be determined by national authorities. There is no guarantee that large FABs and ACBs will be created without further encouragement and pressures, not least because of the political influences that will continue to be at work (see below on the positions of Norway and Finland in relation to the creation of NUAC).

[29] Article 5,COM (2001)564 final/2 dated 11 December 2001 - 2001/0236 (COD)

Even if such large upper airspace blocks are created, and if *en-route* charges in upper airspace are substantially lowered, there is the further question of whether this will, in reality, lead on to effective and efficient restructuring of ATM at lower levels of altitude. There is, for example, the possibility that the creation of new, upper airspace centres will simply give rise to extra assets and extra costs, without leading to offsetting reductions in assets in lower airspace provision (which is one of the concerns surrounding CEATS, about which see below). **In that case, although *en-route* unit costs and charges in upper airspace can be expected to fall, unit costs and charges for ATM in lower space will simply rise, and, on average, network users may be no better off, and possibly worse off, than before.**

On this alternative view, the key to efficient restructuring and improved ATM performance lies not in legislation facilitating the creation of cross-border FABs and ACBs in upper airspace – although that is welcome as a development in its own right – but more in the incentives, encouragement, and pressures exerted on ANSPs and on bodies responsible for overall network co-ordination and development, by means of regulation and supported by the involvement of airlines in network governance, to improve network performance. That is, the central task is not to create particular airspace structures, whether in terms of particular configurations of FABs and ACBs, in order to drive performance improvements, but rather to put in place regulatory arrangements and processes that will provide the necessary encouragement and pressures (the carrots and the sticks) for restructuring (where appropriate) and improved performance *in relation to all aspects of ATM: upper airspace en-route, lower airspace en-route, and terminal.* This last point is of particular importance since much of the overall cost of ATM, and hence much of the available opportunity to improve performance, lies at the lower airspace and terminal levels.

For obvious reasons, we will call this the "regulation is key" (RIK) view.

Adjudication between these alternative views, UARIK and RIK, is critical to the future development of European ATM policy. If UARIK is correct, then much of the relevant legislative task is done, and what remains is simply to ensure that the implementation rules for charging allow for the creation of appropriate ACBs in upper airspace, to accompany the prospective upper airspace FABs. If the RIK view is correct, then the issues to be settled concerning ATM regulation, including the process of regulation, are both wider in scope and more similar in nature to those to be found in other network sectors of the European economy.

More immediately, the view taken affects the way in which the subsequent sections of this Report can be read. Specifically, on the UARIK view Section 3, and much of the following sections is unimportant, and attention can be focused on Sections 8 and 9; whereas on the RIK view, Section 3 and its successors contain highly important lessons that can be drawn from regulatory experience gained in other contexts.

2.4.2 Review of European service provision restructuring

So far only one cross-border European upper *en-route* centre is in operation, Maastricht. Two additional ones are planned in the coming years, CEATS covering the parts of South Central Europe and NUAC in the Nordic region. It is to be emphasised that the Single European Sky has not yet come into force. Therefore none of these existing or planned trans-national ACCs can claim to have met the criteria of the FAB recognition test.

(a) Maastricht UAC

Belgium, in 1964, was the first Member State to delegate the *en-route* control of a part of its airspace –above FL195 - to EUROCONTROL, initially based at Brussels National Airport and later, in 1972, at the Maastricht Centre. Two years later, the Centre was entrusted with control of the Northern German upper airspace, above FL245.

In 1986, the Netherlands agreed to delegate to Maastricht provision of service for its upper airspace, in this case above FL300. In 1993, the Centre was given the responsibility for *en-route* control of traffic flows above FL245, a move that was accompanied by some reorganisation of the Belgium upper and lower airspace. In addition the Maastricht Centre hosts a DFS military control unit.

Initially dealing with complex flows – including ascending and descending phases -- Maastricht has gradually evolved towards more of an upper airspace control centre, but without having reached the Single European Sky upper airspace FL285 division [30].

However, Maastricht as an organisation and in relation to charging, cannot be considered as a fully independent European ACC. The airspace division tends to replicate national boundaries. Maastricht is divided up into the Hanover sector group, the Brussels sector group, and the Delta-Coastal sector group. In a way, control is based on national considerations, and revenues are apportioned on the basis of the number of controllers allocated to the relevant sectors. Currently the division of revenues is: Luxembourg (1.06%); Belgium (34.39%); Germany (43.78%); The Netherlands (20.77%) [31].

Maastricht illustrates the dynamics of separation between upper and lower airspace for operational purposes in circumstances where national upper airspace blocks are relatively small, and where scale economy effects, to the extent that they exist, might be expected to be greatest. It is a multilateral arrangement among parties that might be justified on a 'minimum efficient scale' rationale, rather than as an attempt to creating a particular centre dedicated to managing an extensive European airspace zone and/or specialising in level flight service provision.

From a charging perspective, Maastricht does not exist. That is, there is no ACB corresponding to the OAB. No differentiation has been introduced.

[30] The PWC Report places the division between upper and lower at between FL245 and FL295. The European regulation sets a level in between.
[31] Maastricht Report 2001

(b) Central European Air Traffic Services UAC (CEATS)

CEATS was made official by the signature of a multilateral agreement on 27 June 1997. It is a new, planned Upper Area Control Centre which would cover a large part of the South Central European airspace: Austria, Bosnia and Herzegovina, Croatia, the Czech Republic, Hungary, Italy (the Northern part which consists of the sectors of Padova ACC), Slovakia and Slovenia. The regional airspace block is defined vertically above FL285/FL290. CEATS is scheduled to be operational by 2007-2010.

The motive underlying the creation of CEATS and the introduction of the division between upper (European regional) and lower (national) has been summarised as: enabling an increase in capacity and a decrease in the cost of services per unit, whilst maintaining flight safety. To achieve that end, CEATS will introduce [32]:

- uniform airspace procedures
- optimal sectorisation irrespective of national border
- direct routings
- uniform and cost recovery implementation of new technologies and concepts
- sharing of resources.

As in the Maastricht example, one driving force in CEATS appears to have been the small sizes of upper airspace controlled by national ACCs and the possibilities of achieving economies of scale in provision. However, it can be noted that the options for, and the consequences of, flight bypass of the upper airspace of any national ANSP whose unit rates rose out of line with those of nearby territories are somewhat greater than in other parts of European airspace. **Consolidation of the airspace will reduce such options for airlines – in effect, they will limit 'competition' – and this was likely another motivation for the development.**

One interesting aspect of the formation of CEATS is the way the restructuring process has been planned. As indicated earlier, the creation of such a centre involves some restructuring in relation to both upper and lower airspace, giving rise to a requirement for accompanying transfers of assets and responsibilities. Rather than a introducing 'big bang' with a sudden switch to the new centre, a step by step approach has been preferred. It consists in gradually transforming the current upper national ACCs sectorisation into a virtual upper centre, and then relocating resources to rationalise the number of centres in the region [33]. **One concern about this restructuring path is how far it will bring reductions in costs in the lower national airspaces. This will only occur if the resources devoted to national centres are reduced, possibly via closure of one or more national ACCs. But if commitments to restructure national ACCs are weak, there is at least a possibility that the new upper-airspace ACC will simply serve to increase overall costs, upper and lower combined.**

[32] CEATS operational concept document, 2002
[33] CEATS presentation background, 2002

(c) The Nordic Upper Airspace Centre (NUAC)

The NUAC project involves the creation of a quasi-European Nordic upper airspace control centre. It is planned to cover the national airspace of Denmark, Finland, Norway, and Sweden. Although it is gaining momentum, the adhesion of Norway and Finland is still uncertain, not least because participation will imply commitments to close parts of their national ACCs.

The project aims to combine, in one upper ACC, service provision above FL285 - which is currently split among 10 different ACCs - with the ambition to become a competitive provider of ATM for neighbouring countries [34].

There are a number of particular features of the NUAC project that are worth noting. First, one Nordic airline company (SAS) is a very substantial user of ATM in the Nordic airspace. SAS is the leading buyer of ATM in Sweden (44%), Norway (42%), and Denmark (35%). It has hubs in Copenhagen and Stockholm and is the main airline company at Oslo airport. Arguably, therefore, there is a degree of balance between the monopoly power of the ANSPs, the suppliers, and the purchasing power of the main user of service, the buyer. Under such circumstances, the emergence of a collaborative environment between suppliers and the buyer is more likely. The situation is in this regard unique on a regional (although not on a national) scale in Europe.

Second, the civil-military interface has distinctive features. The Nordic region has sufficient airspace availability to render the military issue less complex than in many other parts of Europe when mapping the Nordic airspace. The staffing of Luftfartsverket in Sweden (LFV) numbers 1200, with 20% of the controllers working at military airports and all controllers having the ability to handle both civil and military traffic. The military itself makes use of civil airports. Hence controllers are familiar with the control of both military and civil flights.

A third notable feature of NUAC concerns the restructuring process, which illustrates some of the political dimensions of creating FABs. The NUAC project is not so much the instrument for change as the consequence of the need for change. Sweden provided the impetus when, in 1998, new direct routes received political support. Then attempts were made to determine how to increase and better manage capacity. To date, Sweden has had three ACCs, Malmö, Stockholm and Sundswall, each responsible for its own FIR.

Different options were debated so as to decide whether one or two centres would be most appropriate. From interviews at the LFV, the choice of two centres appears to have been based more on political than on technical grounds, leading to one in Malmö for *en-route* upper airspace control, and the other one in Stockholm for lower airspace.

But while NUAC introduces a clear separation of airspace at FL285, the Malmö ACC will also be a lower airspace centre handling a small part of the overall traffic, in a few lower sectors, in order to optimise operations. The NUAC project anticipates

[34] NUAC project phase 1 report, from project vision, September 2002, p.14.

prospective efficiency gains as a result of better direct routing and resectorisation [35]. Thus, it is expected to benefit users through a reduction of flying time and fuel consumption. The project also envisages an increase in controllers' productivity by facilitating the implementation of new concepts and technology.

Finally, NUAC does not intend to replicate the Maastricht form of organisation in which, broadly speaking, member state nationals control their respective parts of the combined airspace. The aim is rather to introduce multinational management of upper airspace in the Nordic Region, with revenues being allocated to a NUAC private entity, of which member states will be shareholders. In that respect, one issue for NUAC is how to comply with different national regulatory frameworks. While a joint regional regulatory committee is a possibility, the establishment of a regional regulator may be more appropriate.

2.5 Summary

In summary the following points can be made regarding restructuring issues in ATM:

- Currently, the European ATM system is built around national ANSPs that each enjoy a monopolistic position in their respective parts of airspace. This fragmentation of service provision can give rise to a number of problems that can cause performance to be less good than it could be under alternative arrangements. Addressing potential performance problems arising from fragmentation in ATM is one of the main objectives of the Single European Sky initiative.

- When thinking about restructuring it is useful to draw a distinction between restructuring of ATM provision (operations) and restructuring of ATM charges, as restructuring issues involve questions about not only the evolution of Operational Airspace Blocks (OABs), but also the evolution of Airspace Charging Blocks (ACBs) and the relationships between the two.

- Recent developments in European ATM provide scope for charging restructuring which, among other things, moves away from ACBs that are determined by national boundaries.

- It is critical to the future development of European ATM policy that the UARIK and RIK approaches to restructuring are adjudicated between. If UARIK is correct, then much of the relevant legislative task is done, and what remains is simply to ensure that the implementation rules for charging allow for the creation of appropriate ACBs in upper airspace, to accompany the prospective upper airspace FABs. If the RIK view is correct, then the issues to be settled concerning ATM regulation, including the process of regulation, are both wider in scope and more similar in nature to those to be found in other network sectors of the European economy.

[35] Ibid p.43, reduction of sectors from 32 to 21. Estimated average reduction in average flight time of 2 to 3 minutes.

- To date only one cross-border European upper *en-route* centre is in operation, Maastricht, although both CEATS and NUAC are planned for the coming years. It is to be emphasised that the Single European Sky has not yet come into force. Therefore none of these existing or planned trans-national ACCs can claim to have met the criteria of the FAB recognition test.

SECTION 3
RECENT DEVELOPMENTS IN THE PRACTICE OF REGULATION

The objective of this section is to note current trends in the development of best-practice regulation. These developments build upon earlier experience, but their focus is often on wider issues – such as the effective and efficient *development* of the *structure* of networks, regulatory governance, etc. – rather than narrower issues of charging/pricing. A number of these wider issues are relevant to the Single European Sky programme.

The discussion in this section is complemented by the review of some of the major general themes in price regulation, as they have developed in theory and practice across various regulated segments of the communications, energy, transport and water sectors contained in Appendix 1.

3.1 Initial clarification of concepts

It will be useful first to clarify a number of points that will emerge and re-emerge throughout the following discussions.

(a) Definition of the relevant product/service

Any price/charge is a price/charge *for something*. It is therefore meaningless to discuss prices/charges without first defining the product or service being supplied in a precise manner.

This small point gives rise to a number of preliminary observations:

- The services supplied may be complex, comprising a set of different components.
- The components may be supplied in different mixes or *bundles*.
- Different customers may value the individual components differently, and may wish to purchase different mixes.
- There has been a general trend in network industries toward unbundling of services offered, so that, in effect, greater variety and choice is offered to network users/customers.
- For any given service or bundle of services, quality of provision is frequently a major issue. Services of different qualities are not economically equivalent.

It is to be expected therefore that, when setting charges or determining charging principles, suppliers will define, with an acceptable degree of precision, what it is that will be provided to the customer in terms both of the scope of products/services and of their qualities/service standards. This is in accordance with normal commercial practice

in competitive markets, and it provides a benchmark 'contract' (whether explicit or implicit) against which performance and compliance can be assessed.

In relation to the detail of unbundling, there will typically be a number of different ways in which services and their associated charges can be disaggregated. Choices among the various alternatives should, in principle, be based upon a mix of demand-side (what customers want) and supply-side (cost) factors.

Although the point is an obvious one, it is worth making explicitly because, **in monopolistic sectors, there is often a tendency to think that service menus and charging structures should be determined on the basis of cost information alone. That is, there is often a preconception that, when unbundling products/services, the component parts should be disaggregated in a way determined exclusively by a concern to reflect cost differences among those components. Such neglect of the demand side is, of course, simply a perpetuation of 'monopolistic' thinking. Effective unbundling and disaggregation of charges requires that careful attention be paid to the pattern of *demand-side substitutability*.**

(b) Revenue requirements, the level of charges/prices, and the structure of charges/prices

In order to provide incentives for commercial undertakings to continue to maintain supply, overall revenues must remunerate overall costs, suitably defined. Or put another way, the average *level* of prices/charges must 'reflect' average costs. One regulatory task is, therefore, to determine the overall revenue requirement (alternatively, the average price/charge level). This principle is well recognised in ATM, and the substantive questions to be addressed lie in specifying more precisely the costs to be included and the detail of how charges are to be calculated from the cost data.

A number of issues tend to arise in this area. **Charges could obviously be set on the basis of average costs, whatever those average costs may be shown to be. Such an approach is, however, known to encourage inefficiency in provision.** Expenditures incurred may therefore be subject to some sort of efficiency test, which may be more or less stringently applied. The normal formulation of this requirement is that *charges should reflect efficiently incurred costs*, and expenditures that, in the light of the information available to the provider at the relevant time, are *manifestly* excessive are disallowed. With this adjustment, pressures toward efficient provision, including by the restructuring of service provision when existing structures are manifestly inefficient, can be established, whilst accepting the principle that charges should allow for the recovery of average costs.

Most regulated undertakings supply a range of products/services that differ in their economic characteristics. Thus, even in cases such as electricity supply, where the final product may be almost homogeneous in terms of physical characteristics (e.g. kWhs supplied), there are often very major *economic* differences in the service provided due to factors such as location and time. Put very simply, it may be much more costly to supply 1 kWh in a peak period than in an off peak period or to supply 1kWh to an isolated rural location than to a medium sized city.

For any allowable revenue requirement, when provision involves more than just one, simple service, a variety of different *structures* for prices/charges will be feasible. In general, it is desirable on grounds of efficient use of economic resources that the structure of prices/charges reflect the structure of costs. The rationale is that, when customers (airlines in the case of ATM) are determining the mix of services that they wish to procure, cost-reflective charges will imply that they implicitly take account of the effects that their decisions will have on providers.

It follows that the importance attaching to issues concerning charging structures tends to depend upon the ease with which customers can substitute between different service mixes. If the behaviour of customers is relatively insensitive to the relative charges, there is little gain to be had from fine tuning price structures. Where demand substitutability is higher, the potential efficiency gains will likewise be greater. In this latter case (higher demand substitutability), however, it is necessary also to take into account the potential inefficiencies caused by changes in customer conduct that result from design weaknesses in charging structures.

One source of such weaknesses is inadequate cost information. Whilst it might be considered desirable to unbundle services previously provided jointly, so as to reflect differences in the costs of supply of service components, the ability accurately to reflect costs depends upon the quality of the relevant information. If demand substitutability is high, there can be significant, *unwanted* customer response to price differences that exist only as a result of cost estimation errors.

Another possible source of weakness in tariff design are the 'discontinuities' in charges that result from efforts to simplify the charging structure. For example, in network industries, tariff boundaries or thresholds are sometimes defined in terms of some numerical characteristic such as the size of the load (in electricity and gas). At the boundary/threshold there may be a sharp discontinuity in charges (e.g. lower unit prices at higher volumes), which may not correspond to a similar discontinuity in costs. That is, costs may vary more smoothly with the relevant variable (size of load), but the discontinuity in charging is introduced so as to simplify the tariff structure. If substitutability is high -- which is when well-designed, cost-reflective charging has most potentially to offer -- **non-cost reflectivity around boundary/threshold points can cause significant, unwanted and inefficient substitution.**

To illustrate, a gas customer with load a little below a tariff threshold may find that purchase costs can be reduced by wastefully consuming extra gas in order to qualify for a different tariff that has lower unit rates. Although the tariff structure may appear broadly cost-reflective – there are lower prices at higher volumes – it is not at all cost-reflective around the tariff boundary.

The 'discontinuity' problem potentially arises in ATM whenever a relatively small movement in aircraft position, whether horizontally or vertically, gives rise to a step-change in charges that is not matched by a step-change in costs caused by the movement.

(c) Duration of charging/pricing periods

Pricing/charging constraints, whether set in terms of an allowed overall revenue or an average per-unit charge or some mixture of the two, have to be determined for a prescribed period. **This charging period may be short or long, and the choice of its duration is one of the most important decisions to be made by the competent authority.**

The significance of the decision can be illustrated by considering its effects under an arrangement in which average prices are set so as to be equal to average cost. If prices are set for only a short period, and are continuously adjusted as costs change, the result is a simple form of *cost-plus pricing*. If, on the other hand, prices, although initially set on the basis of costs, are then held constant for an extended duration, the outcome is very different. For at least the relevant period, prices are independent of costs, and cost changes within this period are not immediately reflected in prices (although, of course, they may, and likely will, be so reflected at some later time).

The time that elapses between sequential adjustments to charges is often referred to as *regulatory lag*, since, in the US system of utility regulation, allowed prices were traditionally held constant until subsequently changed by a further regulatory review at some unknown future date. In a European context, and in the light of more recent developments in regulatory thinking and practice, it is probably better described as the *charge duration*, since such a term can encompass both US-type arrangements, where charges are set for an indefinite period but reviewed at some later (uncertain) date to be determined by events, and other arrangements in which either the charges themselves or, alternatively, a charging formula are set for a *predefined* duration.

(d) Ex ante versus ex post assessments

In assessing costs for the purpose of setting charges, the focus may be on measurement of costs actually incurred or on the expected future evolution of costs, including capital costs. This *ex post/ex ante* distinction is often linked to the differences between accounting and economic approaches to cost measurement. In terms of general principle, the forward-looking (*ex ante*) approach is usually considered to be preferable in terms of implications for economic efficiency. However, forward-looking cost assessments tend to be more subjective and to give rise to more difficult monitoring problems.

At one extreme of a set of regulatory approaches we can imagine arrangements in which charges/prices are set entirely on the basis of *ex post* cost assessments. If *the charge duration* is set to be short, charges will track measured cost changes quite quickly; if the *charge duration* is set to be long, the 'tracking' will potentially allow for more persistent and larger deviations between prices and costs.

Whilst longer charge durations tend to provide greater encouragement for suppliers to improve performance, it can be seen that they will give rise to problems when they are associated with *ex post* approaches. If, for example, there is a systematic trend in costs

that is beyond the control of the supplier, charges will tend to be persistently below or above costs for no very good reason, and in the former case will potentially give rise to financing problems for suppliers.

(e) Regulatory credibility/certainty/stability

A final, critically important concept that merits brief discussion at the outset is that of regulatory *credibility*. Particularly where commercial undertakings are expected to make substantial investments in assets that are specific to the supply of the relevant products/services, there is a very real issue in regulated industries arising from the dependence of future revenue streams on regulatory decisions. An undertaking may be reluctant to invest if it believes that there is a significant likelihood that, once investments are made, the undertaking's inability to redeploy assets to other purposes will be exploited by the relevant authorities through decisions that lead to unexpectedly low revenues.

The credibility issue is an underlying theme of the economics of regulation, and we will return to it at several points below. Here we simply point out that, although most discussions of the issue focus on the implications of lack of commitment to future prices/charges, similar types of considerations can also apply to matters of investment. For example, an ATM provider might embark upon a programme to expand capacity or reorganise its activities and then find, if it is dependent upon the state for finance, that the promised or expected funds are not forthcoming. Appreciating the possibility of such events, airlines may in turn be a little less willing to make investments complementary to the ATM investments.

Although the extent of the problem has varied from country to country, instabilities and uncertainties in the flow of funds available to state-owned enterprises have been a widespread phenomenon. Access to finance for expansion was arguably the most important factor in stimulating the launch of the wave of utility privatisations in the UK, and there is every indication that it is an important issue in developing strategies for the future development of European ATM.

The potential role of independent economic regulation in this area cannot be overstated. For so long as the businesses of suppliers/providers are subject to direct political influence, credibility is likely to remain an issue, and private sector investment will be impaired -- either funds will not be available or will be available at an unnecessarily high cost of capital. This follows from instabilities, over time, in the agendas/preoccupations/priorities of national governments, such as when investment funds are cut in response to deterioration in the general (macro) public finances (which becomes the political priority).

Although it is not a panacea, the development of independent regulators (entrusted with limited, focused and well defined duties) operating within a stable legal framework, can substantially mitigate the credibility problem, and open the way for more reliable, lower cost sources of private finance. To date, the international experience with independent regulation has been positive, although it is equally clear that the way in which the

institutional framework is structured is a matter of great importance. This is because independent regulation tends to be under constant pressure from two sources:

- Political agendas that frequently are pre-occupied with issues, often of a short-term nature, that lie beyond the scope of the established regulatory duties. Specifically, political actors tend to put pressure on regulators to use their delegated powers for purposes other than those for which those powers were intended.

- Regulatory *capture*, by which is meant a position in which one or other potential interest group gains a disproportionate influence on regulatory decisions. Such capture can develop in a number of ways: bribery; consideration of future career prospects by regulators; political control (when regulatory authorities or commissions are elected); bureaucratic control (e.g. when there is an over-close relationship with another public agency, with a different agenda); and intellectual influence (e.g. when 'single-issue' ideas come to predominate).

3.2 Key trends in best practice regulation

An account of how regulators have, in practice, addressed the various issues described in Appendix 1 would be fascinating, but long and complex. For current purposes, therefore, this Report simply sets out a few of the more obvious trends.

One over-arching message is clear, however: the terms on which regulatory choices are frequently debated are, in the light of the regulatory histories that can be observed, often simplistic, misleading and unhelpful. The UK provides a good illustration of this point. The UK system is often characterised as one in which an independent regulator sets high-powered incentives for a privately owned monopoly via an RPI-X pricing formula, and this approach is sometimes contrasted with the cost-of-service approach that was traditionally favoured in the USA, but which was explicitly rejected as an option when BT was privatised in 1984.

There are some elements of truth in this characterisation, but they are only elements. Those elements were even a reasonable, first approximation to the wider picture in 1984, but, in 2003, the characterisation would be misleading in relation to 95% or more of what it is that most regulators actually do. Over a period of twenty years, things have moved on a good deal, and regulatory arrangements have evolved considerably. The original RPI-X philosophy probably stands further away from current UK approaches than it did from cost-of-service regulation in 1984, a fact that is often not recognised, including in the UK itself. **We therefore strongly recommend that discussions concerning future charging regimes for ATM should seek to avoid getting bogged down in arguments between false choices between limited and artificial options (e.g. cost-of-service vs. RPI-X).**

Given this, we can identify some of the actual tendencies in regulatory activities that can be widely observed

(a) Unbundling and the introduction of competition

There has been a strong trend toward the explicit unbundling of services previously provided by monopolistic network operators. In part, this has been stimulated by the options for introducing competition in some of the services. Competition in this context can take one of two forms:

- Service-to-service competition, in which different suppliers compete for the business of customers in a normal way, and
- Where this is not feasible, and it remains preferable to retain one service provider, competition *for* the franchise to provide the service (e.g. provision is *procured* via competitive tender).

Even in cases where either form of competition is infeasible, there has been some tendency toward unbundling, albeit the pressures here have been weaker. One stimulant for this development is recognition that different users may wish to purchase different combinations of network services. Unbundling, coupled with separate pricing of the various network components, potentially leads to more efficient use of the network, and, perhaps more importantly, helps communicate to providers just what it is that their customers (system users) want to purchase.

(b) Unbundling of 'co-ordination' activities

One of the most important developments in network regulation has been the recognition that 'co-ordination' activities can be unbundled from other aspects of service provision. The requirement for co-ordination arises from two factors:

- The difficulties in devising pricing arrangements that can, *over very short time periods*, efficiently balance demand and supply/capacity.
- The existence of network externalities, which likewise raise problems for fully decentralised, 'market-based' approaches to co-ordination.

In the absence of such co-ordination, decentralised networks can be prone to both short-term and long-term inefficiencies (e.g. excessive congestion, under-utilised capacity). Effective provision of co-ordination services is therefore of the highest importance for network performance.

When networks were dominated by vertically integrated monopolies, the relevant co-ordination activities were typically entrusted to incumbent enterprises. Liberalisation has, however, led, at least in some sectors (e.g. energy), to the separation of these activities, under the heading 'systems operations' – or, as it might be termed in relation to ATM, 'network management'. The organisation entrusted with the relevant tasks can either be an existing service provider or a distinct entity (an 'independent system operator' (ISO) or 'independent network manager'). Either way, the relevant undertaking is properly regarded as a 'service provider', although the services provided (co-ordination) are quite distinct in nature and are supplied both to network users and to

other network service providers (each set of parties benefits from improved co-ordination).

The relevant co-ordination services may be short-term in nature, most typically involving the matching of demand to available supply/capacity on a minute-by-minute basis (flight management in ATM); or they may be of a longer-term type, such as helping to secure co-ordination in investment across the network (infrastructure management); or they may be both. Like other service providers, system operators/network managers need to be remunerated for their activities, and, since co-ordination is monopolistic in nature, they need to be regulated in one way or another. Thus, regulation may be based upon a cost-of-service approach, a price-cap approach, or some hybrid, raising the kinds of trade-offs discussed above.

The roles entrusted to system operators/network managers have very major implications for pricing/charging arrangements in networks. Put simply, for a limited and well-defined set of activities (congestion management, investment co-ordination in the presence of externalities), bilateral transactions between the system operator(s)/network manager(s) and other service providers on the one hand and system users on the other hand *replace* bilateral transactions between other service providers and system users. **This reduces the weight placed on prices/charges (which relate to bilateral transactions between other service providers and users) in securing network efficiency. In consequence, charging structures can be greatly simplified** (since they are not relied upon to address very short-run congestion issues or longer-term externality problems). **In turn, this enables some of the perverse effects of short-run marginal cost pricing on investment incentives to be avoided** (see Section 8 and Appendix A1.7).

(c) A shift away from cost-of-service regulation

Rate-of-return regulation in its pure form, if it ever existed, has long-been abandoned, by virtue of factors such as regulatory/charging lag and audits that lead to potential disallowance of inefficiently incurred costs and investments. Whilst the approach may not be overly inefficient in some circumstances and over relatively limited periods of time, at some point there generally develops a tendency for costs to be inflated and padded in a way that is perceived to be 'out of control'. The shift away from cost-of-service regulation implies a shift *towards* some or other form of incentive regulation, broadly defined (see Appendix 1).

(d) The evolution of hybrids

There has not, however, been a general shift to price cap regulation. The UK experimented with pure price cap regulation in 1984, but, in practice, moved away from the approach fairly quickly thereafter. Rather, the tendency has been toward hybrid forms of incentive regulation, in which there are explicit or implicit arrangements for providers and their customers to share the benefits of any performance improvements (and to share the costs of any deteriorations in performance). A wide variety of 'sharing rules' have been adopted, in an attempt to set the 'parameters' of those rules in a way that appropriately reflects the relevant economic circumstances.

(e) The emergence of more explicit, ex ante regulatory 'contracts'

Analysis of regulation has long made reference to the notion of an implicit *regulatory contract*, even under the cost-of-service approach. **There has, however, been a tendency toward more explicit definition of what is expected of the regulated undertaking and of the regulator.** For example, in contrast to the uncertain regulatory lag associated with traditional rate-of-return regulation in the US, more recent developments have tended to pre-specify the timing of regulatory reviews, albeit with 'get-out' provisions to allow for flexibility in the light of radically altered circumstances.

An associated, notable development has been the move, in some cases, to define the 'outputs' that the firm is expected to deliver over the relevant charging period. This makes the arrangements much more like a normal commercial contract, in which it would be expected that there would be references to outputs as well as to prices.

In this context, 'output' is to be interpreted broadly. It can include specification of service quality, where such quality can be measured, with providers and users sharing in the benefits of any over-achievement of contracted quality, and likewise sharing in the costs of any under-achievement of quality standards – for example, by means of adjustment of charges or by compensation payments. **Likewise, 'output' might in some circumstances be defined in terms of capacity provided, rather than the actual usage of capacity.** In this way, payments to providers can be linked to the actual output of investment programmes rather than the inputs in to those programmes (e.g. investment expenditure).

(f) Incorporating users/customers into the 'contractual' framework

Whereas traditional views of regulation have interpreted arrangements in terms of hypothetical 'contracts' between regulators and suppliers/providers (and much principal-agent theory has developed in this way), there is increasing emphasis on the contractual obligations of suppliers to users/customers, and *vice versa.*

These arrangements effectively take the form of collective, multi-party agreements between providers and users. They do not preclude individual service agreements between a particular provider and a particular user, but rather lie 'above' such individual agreements and, in some cases, govern the latter. Such multi-party agreements are generally subject to some form of regulatory oversight and supervision. Both multi-party and individual level agreements are usually subject to strict requirements that they be non-discriminatory in their effects.

(g) Growing recognition of the implications of regulatory contracts and structures for private capital markets

Provided that they are credible, regulatory 'contracts' that link the allowed revenues, and hence charges, of providers to capacity have the effect of facilitating the use of private sector capital in the development of network infrastructure. At the same time,

they thereby ease concerns that users may have about 'pre-funding' investment in infrastructure. Put simply, **the regulatory guarantees make it easier for providers to raise finance for investment projects, against a more secure prospect of future revenues.** From the system user's perspective, such an arrangement means that they only contribute toward such investment once the benefits of that investment are available (i.e. once there is an output), as would be expected to be the case in most normal markets.

The regulatory commitments do, however, need to be credible, and have to extend over a duration commensurate with the nature of the relevant investment. Thus, for example, in the UK gas industry, which has until recently worked with a five-year contract period, we have recently seen the emergence of regulatory commitments in relation to gas terminal capacity that extend to as much as fifteen years into the future.[36] In the case of investment in the London underground transport system, contractual arrangements extend for a duration of 25 years.

Credibility is enhanced not only by making more explicit, *ex ante* commitments, but also by the structure of the regulatory system itself. There is increasing acknowledgement of the importance of the independence of the regulatory arrangements from other parts of the political system, which can be driven by shifting and unstable objectives, sensitive as they are to the preoccupations of electoral politics.

That said, the lip service paid to 'independent' regulation is not always reflected in political practice. Well-designed regulatory systems therefore need to be characterised by institutional barriers to political 'capture'.

(h) Increased focus on investment issues

When networks are first commercialised or privatised, there is in many, but not, all cases, a legacy of gold-plating in capacity provision. For this reason, investment issues are not necessarily at the top of regulatory agendas in the first stages of such development. However, a variety of factors – technological change, environmental problems, demand growth -- have tended to push issues of investment in network infrastructure ever higher up the list of policy priorities.

As explained in Appendix 1, the potentially distorting effects of price controls on investment incentives have long been recognised, but, in networks where little investment is required, or where there is some possibility of *network competition*[37], these distortions may be damped. Where substantial investment is warranted, and the prospects for network competition are limited, regulatory agencies have sought, and are seeking, to address infrastructure investment issues more directly and more explicitly.

[36] This does not mean that charges are set for fifteen years, since charges reflect costs other than capacity. Rather it means that the network owner is given commitments in relation to one specific investment activity, a further example of unbundling, in this case the unbundling of provision of a certain type of network capacity.

[37] By network competition is meant a situation in which there can be direct competition among providers of network capacity. Thus, in telecoms, there can be a variety of forms of network competition: fixed link transmission, fixed vs radio transmission, telecoms lines vs TV cables in the local loop, etc.

Linking allowed charges to capacity made available is one way to do this, but it has the disadvantage of being a form of central planning in that the regulatory authority is required to take a view on what levels of capacity are likely to be required, and where and when. In order not to revert to arrangements that, in many countries, have not worked well, either for the user or for the taxpayer or for both, the most recent regulatory developments in this area have sought to introduce what might be called 'deep' incentives.

The idea here is that the regulatory contract specifies a base-line plan for the development of the system, based upon best-available information at the time of the contract. The provider commits to this plan in return for commitments on the levels of charges that will be allowed. However, the provider is free to depart from the plan if, in the light of evolving information, such adjustments are indicated as being warranted *on the basis of the conduct of system users*. Specifically, charges are adjusted in such a way as to give the provider extra financial returns if less investment than planned is made in circumstances in which users indicate lower demand than originally anticipated, or if more investment than planned is made in circumstances in which users indicate more demand than was originally anticipated.

Very roughly, this approach can be characterised as:

- Traditional output/capacity based regulation for the baseline plan, plus
- Incentive regulation, based on market signals, for deviations from the baseline.

As always, the incentive elements can be made more or less high-powered. There is, however, flexibility to combine low-powered incentives for the base-line with high powered incentives for deviations from the base-line (i.e. for flexibility), which is a potentially attractive way of simultaneously addressing the concerns of providers and users.

(i) Increased role of users

One way of interpreting the traditional role of regulatory authorities in network industries is to see them as acting on behalf of consumers, and, in some cases, this is reflected in the explicit duties given to the relevant agencies (e.g. a duty to protect the interests of consumers). Such an interpretation makes obvious sense when there is a monopoly provider – such as an electricity, gas, telecoms, rail, or water utility – serving many millions of small customers. In effect, the regulatory agency exists to redress an imbalance of economic power between the monopoly supplier and its customers.

As networks have been unbundled, however, and as 'final service' markets have been liberalised, regulatory attention has come to focus ever more on the provision of core, monopolistic, infrastructure services, not to millions of small final customers but rather to a relatively small number of companies that use the networks to render services to final customers. ATM networks have always been like this – services are supplied to

airlines who serve millions of travellers – but the development is relatively new, and still far from complete, in most energy, communications and rail networks.

There is still, of course, an imbalance of power between providers and users, but, in network services markets, it is much less stark. Given the fewness and the commercial sophistication of network users, there is much less of a requirement for regulators to take the role of always 'acting on behalf of the customer'. Accordingly, it is possible to discern a strong trend towards increasing the *direct* contacts between providers and users, since channeling all issues of concern to network users through a regulatory agency – which, at the end of the day, is a public, administrative body, rather than a commercial organisation – may not be the most effective way of resolving outstanding problems. **In consequence, system users – on the demand side of the relevant service markets – are coming to play an increasing role in the more formal aspects of regulatory processes.**

The increasing role played by system users in influencing the development of networks has a potentially very large benefit in terms of the extra information that becomes available. Much economic analysis of network regulation emphasises the asymmetry of information between regulated suppliers and the supervising authority. In our view, a more important problem is the poor quality of information available to both sides. In either case, increased user participation can substantially improve matters: more demand-side information is brought to the process, the extra information can ease problems of asymmetry, and there are increased pressures for discovery and revelation of information from the supply side.

(j) The development of contractual networks

Where network services have been unbundled and have been made capable of being offered under competitive conditions, users and providers can contract directly for service in a standard, commercial way. However, during the unbundling process and, in the longer term because of network effects (potential externalities), there have developed a variety of agreements and 'network' codes that are multilateral in scope. These are effectively 'rule-books' for the operation of networks in situations where there are many users and more than one network services provider. To paraphrase the title of a study of comparative deregulation in North America, Japan and Europe, *freer markets mean more rules*. [38]

Thus, alongside the physical network itself, there has increasingly developed a *contractual* network. Further, where contractual commitments are sufficiently specific, secondary trading in contracts is also developing, creating, for example, markets in rights to the use of network capacity. Slot trading in air transport is, where it exists, one particular example.

[38] *Freer Markets, More Rules: Regulatory Reform in Advanced Industrial Countries* (Cornell Studies in Political Economy), S. K. Vogel, Cornell Univ Press, (September 1996)

(k) The emergence of more inclusive governance arrangements

Since networks develop and change over time, any set of contracts and rules must necessarily have a degree of flexibility in order to support continuing, efficient physical operations. The emerging contracts and rules therefore include provisions for their own modification. This might involve reference to an arbitration process (for contracts), and, for network codes, the rule change provisions are built around more formalised arrangements for governance.

Different governance arrangements have developed to supervise the multilateral contracts and codes for network use. One option is to allocate rule-making powers to the relevant regulatory authority, after a process involving consultation and hearings with all interested parties (including, but not restricted to, providers and users). Another is to place greater weight on formalised industry panels, populated with experts from both providers and users (possibly supplemented with 'outside' experts), to propose developments to the network codes, with regulators playing a less central role.

3.3 Summary: Reassessing the role of regulation

Much thinking about economic regulation is still dominated by ideas developed for traditional, pre-liberalised networks. As the above summary of developments indicates, the role of regulation in network sectors such as energy, communications, and transport, has shifted, and continues to shift, in fundamental ways. The traditional role of correcting imbalances of market power between the supply and demand sides of the relevant markets remains, but in a context in which the underlying imbalance is much less stark than in traditional models. The challenge is less to 'represent consumers', on the ground that it is infeasible to give 'voice' to consumers in any other way when each consumer is small and there are many millions of them, than to develop and supervise institutional arrangements in which the 'voice' of network users – relatively few in number, frequently large, and commercially sophisticated – can be heard directly by providers and can not be ignored.

Increasingly, regulators of networks are involved less in telling providers what they should and should not do, and on what terms, and much more in:

- Facilitating the development of contractual networks among providers and users.
- Acting as adjudicators when differences emerge.
- 'Prodding' the various parties to act, when opportunities for improvement arise but no 'first-mover' appears.
- Promoting the discovery and revelation of information in formats that will be of assistance to all parties in seeking efficient operation and development of the network.

To date, air transport networks have exhibited some marked differences from other networks in communications, energy and transport, but unbundling/business separation,

liberalisation and the growing significance of cross-border transactions in those other sectors have led to a significant degree of convergence. Today, the regulatory challenges are much more similar across sectors. In electricity, for example, the European network is characterised by a multiplicity of transmission undertakings, some privatised, some public owned, providing transport services (transmission of electricity) for a relatively small number of commercial customers (electricity generators and suppliers). Notwithstanding the fact that the relative importance of cross-border traffic is much higher in ATM than in electricity, there are increasing opportunities for each sector to learn from the other, and for common principles to applied across these, and other, European networks.

SECTION 4
CONTRACTUALISATION AND SERVICE PROVISION ARRANGEMENTS

4.1 Introduction

A key trend in the regulation of network industries has been the development of increased levels of 'contractualisation'. This can be broadly understood as relating to a movement to improve and formalise definitions of the rights and obligations of users and providers in relation to a given set of services. An important part of this process, then, involves better defining what it is that is being paid for.

There has been considerable and growing interest in the potential for developing increased levels of contractualisation in relation to ATM from range of parties - including user and service provider industry bodies[39] - but a number of key issues in terms of how that process can be expected to operate have remained unresolved. So, for example, questions are raised concerning: the desirability and feasibility of developing contractual arrangements between different parties (eg between individual airlines and individual ANSPs, between industry bodies, etc); the extent to which 'contracts' could be commercial agreements as opposed to being effectively some form of regulatory arrangement; related to this, the extent to which different forms of agreements could be made enforceable; and, more generally, the processes by which such contractual arrangements could be most usefully developed.

This section of the report is concerned with this set of issues, and in particular seeks to highlight ways in which regulatory activity (understood broadly) can facilitate and underpin the development of more formalised arrangements in ATM.

The Common Council position contains a number of provisions of direct relevance to the consideration of issues of contractualisation. In particular, Article 14, Section 2(e) states that:

> *Transparency of the cost-base for charges shall be provided. Implementing rules for the provision of information by the service providers shall be adopted in order to permit reviews of the provider's forecasts, actual costs and revenues. Information shall be regularly exchanged between the national supervisory authorities, service providers, airspace users, the Commission and EUROCONTROL.*[40]

Section 3(e) of Article 14 states that:

> *Charges shall allow for the safe and effective provision of air navigation services with a view to a high level of safety and to cost efficiency and shall*

[39] '*Quality of ATM Service including service level agreements'*, IATA Draft Policy, Technical Operations Policy manual – Part B resolutions 13.04.04

[40] *Common Position adopted by the Council with a view to the adoption of a Regulation of the European Parliament and of the Council on the provision of air navigation services in the Single European Sky*/dated 11 March 2003 - 2001/0235 (COD)

stimulate integrated service provision. To that effect, such charges may be used to provide:

– mechanisms to encourage air navigation service providers and/or airspace users to support improvements in air traffic flow management such as increased capacity and reduction of delays, while maintaining an optimum safety level. The decision as to whether to apply such mechanisms remains within the sole responsibility of each Member State; [41]

The following considers issues of relevance to contractualisation in the context of these Articles of the Draft Regulation. Sub-section 4.2 provides a brief overview of the current institutional arrangements for charge setting, and highlights a number of identified weaknesses in these arrangements. Sub-section 4.3 considers the potential for contractual arrangements to be developed through commercial contracting. Sub-section 4.4 considers ways in which regulatory arrangements can facilitate the development of increased levels of contractualisation and presents a stylized set of steps in order to provide a framework for considering different implementation options. Sub-section 4.5 turns to a consideration of specific modifications of the current institutional arrangements that may be desirable. Sub-section 4.6 examines how contractualisation relates to EUROCONTROL.

4.2 The current institutional arrangements for charge setting

4.2.1 Multilateral Arrangements[42]

The EUROCONTROL Route Charges System was established by a Multilateral Agreement by which Contracting States agreed to adopt a common policy and create a joint system for the calculation, billing and recovery of route charges. The Enlarged Committee for Route Charges is responsible for supervising the operation of the system. The Enlarged Committee holds at least three sessions each year, and represents the main forum for multilateral consultation with users at present. User organisations are currently able to attend the meetings of the Enlarged Committee with observer status.[43]

The route charging principles stipulate that preliminary data submissions should be made by 1st June each year, and that this data should be provided to users at least ten calendar days before the June meeting of the enlarged Committee. This submission should provide actual cost and traffic levels[44] for the previous calendar year (referred to as year n), and forecast levels for year n+1 (i.e. the current year) to n+5. Cost forecasts for year n+2, therefore, form the basis for estimating the charge level for the next calendar year. For all years (i.e. n to n+5), costs should be disaggregated according to the route charging principles (this disaggregation is discussed below). A finalised version of this data should then be submitted in November of each year.

[41] ibid. Article 14
[42] The details provided in the following section draw heavily on the CRCO (July 1999) document: "Guidance on the Rules and Procedures of the Route Charges System"
[43] With the exception of a few items specifically earmarked as being confidential
[44] In addition to this, some ancillary data is provided including the number of air traffic controllers.

The Enlarged Commission of EUROCONTROL consists of Ministers of the Member States and is the legislative body governing the common Route Charging System. The Enlarged Committee is the body in charge of supervising the operation of the system, preparing the decisions of the Commission, and constitutes the main forum for multilateral consultation with user organisations. The Provisional Council was established as part of a process of early implementation of the Revised EUROCONTROL Convention, and is made up of representatives of Member States at Director General of Civil Aviation level.

The Enlarged Committee is assisted by the Central Route Charges Office (CRCO) in terms of the definition and application of the principles for establishing route charges by existing Member States (and their service providers), and the compliance by new Member States with the principles. However, the CRCO does **not**:

- audit the accounts of States or service providers
- assess the cost-efficiency of States or service providers
- set targets
- enforce compliance with route charging pricinples

In cases where there are concerns of non-compliance with the route charging principles that are raised by either users or service providers, the CRCO can provide a source of expert input into bilateral meetings on request. Matters can also be addressed in writing to the CRCO, in which case it will raise the matter bilaterally with the relevant party and will inform the party that expressed the initial concerns of any agreement reached and/or of reasons for disagreement. If this process is considered unsuccessful, matters can be presented in writing to the Enlarged Committee, who may formally task the CRCO to analyse the situation and submit a report. If these processes have not resulted in a satisfactory outcome for the parties involved, the enlarged Committee may inform the Provisional Council of the matter. In addition to concern-led activity of this kind, the CRCO will also seek to clarify apparent misinterpretations/misapplications of the route charging principles by Member States that it finds during the course of its routine analysis.

In our consultations with users, significant concerns were expressed concerning the adequacy of the Enlarged Committee in terms of providing a forum for user representation. Significant weaknesses of this process have also been highlighted recently by the EUROCONTROL Performance Review Commission, and its most recent performance review – PRR6[45] – provides a useful account of some of the problems with current multilateral arrangements.

A key issue for the development of more formalised arrangements concerning the provision of ATM services is the generation of forward looking target values for key variables which can provide a baseline against which actual outcomes can be evaluated. With respect to the provision of forward-looking forecasts by ANSPs, the PRC made a number of observations:

[45] *An assessment of Air Traffic Management in Europe during the calendar year 2002* PRR 6 Version II, EUROCONTROL Performance Review Commission, May 2003

- whilst the current route charging arrangements require States to provide forward-looking projections of costs and traffic levels for the coming five years, only half of the relevant countries actually complied with this requirement in 2001;
- Capacity and cost management processes remain disconnected: States are not required to present justifications for proposed cost increases, and there is no link with explicit performance objectives;
- It is also stated that capacity management has so far been conducted without reference to corresponding costs.

The PRC highlighted a range of specific weaknesses that it found to characterise the current arrangements of the Enlarged Committee:

- forward-looking information is virtually not discussed in the enlarged Committee;
- a club/cartel syndrome where it is not in the States' interests to challenge each others cost bases;
- conflicts of interest arise given that contracting states are often represented by ANSP representatives – it was noted that whilst this gives ANSPs a strong voice in the enlarged Committee, user representatives have only observer status;
- the consensus decision-making process hampers progress.

In the context of the importance of information flows, and in particular of user involvement in regulatory processes highlighted in the review of best practice regulation in Section 3, these represent substantial areas of concern.

4.2.2 Member State institutions and processes

It is notable that the CRCO have developed a standard agenda and guidelines that can be used in order to structure bilateral consultations at the EUROCONTROL member state level that include issues such as: a review of the previous year's results; an update on the latest estimate for the current and the coming year, and forecasts for years beyond N+2; and a brief summary of the capital expenditure programme in terms of the expected timing and level of its impact on charges[46]. In practice, the approach taken varies between States both in terms of the level of information disclosure and the consultation processes in place.

Whilst these processes clearly provide some basis for user involvement at the State level, significant weaknesses have been identified. In particular, it is notable that the PRC have argued that:

> *Bilateral consultation meetings, when they take place, provide the only opportunity for the users to seek explanations for the level of charges that will be imposed. They seldom take the form of proper consultations on the required*

[46] Although it is noted that the CRCO guidance should only represent a minimum entitlement and that States should normally give more.

> *level of performance and the opportunity for new investments. There is little transparency on the outcome of these bilateral meetings* [47]

They go on to say that:

> *In principle, proper consultations should result in agreements between ANSPs and users, whereby the former commit themselves to deliver service (e.g. capacity) and the latter accept to pay the corresponding costs* [48]

This is clearly consistent with a more general emphasis on the desirability of increased levels of contractualisation (referred to above), with the processes of charge setting and service provision resulting from a joint process that involves commitments. The remainder of this section is focused on ways in which such arrangements could develop and could be facilitated by modifications to the current arrangements.

4.3 The potential for commercial contracting arrangements outside of regulatory processes

Before considering potential regulatory responses to some of the weaknesses highlighted above, it is useful to consider the extent to which cost and quality of service issues could be effectively addressed through 'service level agreements', or other forms of private contractual arrangements, between individual users and individual ANSPs. It is notable that this matter was raised on a number of occasions during our consultations with users and providers.

Important practical difficulties arise when considering the potential for commercial contracting on service levels between service providers and users. The key difficulty is that the relationship between an individual service provider's performance and the actual service received by airlines is not straightforward - it is likely to be subject to significant and complex network effects. Furthermore, the recognition of the importance of network effects has resulted in institutional responses that significantly complicate private contracting approaches focused on individual ANSPs.

A particularly important issue here is that whilst individual ANSP's can determine the amount of capacity that they make available, they do not determine which parties get access to that capacity. **Rather, the current system can be understood as effectively treating the provision of capacity by ANSPs as an 'input' into the effective capacity of the European network.** The allocation mechanism for this capacity is then substantially[49] carried out by the CFMU. This is a substantial barrier to commercial contracting between individual airlines and ANSPs for *en-route* capacity, as in practice the ANSP's cannot actually provide for delivery to a specified customer.

---

[47] *An assessment of Air Traffic Management in Europe during the calendar year 2002* PRR 6 Version II, EUROCONTROL Performance Review Commission, May 2003, page 53

[48] ibid. page 54

[49] ANSP's have some flexibility in relation to divergences from the last filed flight plan.

The existence of network effects is therefore a significant complicating factor but it does not preclude private contracting approaches to service provision levels (recognising the monopoly issues that arise). One could envisage, for example, the potential for individual airlines to contract for long-term capacity rights with a European network manager (perhaps a modified form of CFMU) on a gate-to-gate basis. Such contracts could define the rights and obligations of both parties including in relation to charge levels (with perhaps some minimum level of reservation charge that was not related to usage) and compensation arrangements for non-delivery. Under such arrangements, the European network manager could contract separately with individual ANSPs for the local capacity rights necessary to provide the gate-to-gate service.

Clearly, **this is some distance from the current situation, and at present the fragmented nature of the system, and the relatively passive role played by CFMU when allocating capacity, makes a commercial contracting approach extremely difficult to envisage in the short term**[50].

4.4 Contractualisation within a Regulatory Framework: Identifying a Stylised set of Regulatory Requirements

The following maps out a series of steps that a process of contractualisation can take in a regulatory context. The aim here is to provide stylised framework within which to consider the potential to formalise a set of compatible implementation options in relation to ATM. The compatibility of a framework with a range of different implementation options is of particular importance as significant differences in regulatory approach between countries can be expected to continue to coexist, and are explicitly allowed for in the Draft Regulations.

The development of contractual arrangements within a regulatory context can be usefully described in terms of a set of steps that involve different degrees of formalised regulatory requirements, and the different powers conferred on particular types of institutions. Table 4.1 sets out a stylised set of steps, where the regulatory intervention becomes increasing more onerous for service providers in terms of requirements. In broad terms, the list moves from a situation where service providers are faced with requirements in terms of process, to one where they may be faced with externally determined financial incentive arrangements linked to delivering specific levels of performance. The list is intended to be indicative, and to highlight a set of potential steps that could be compatible, but that could also be unbundled and implemented in different combinations. An important aspect of identifying steps separately in this manner is to provide a basis for considering those aspects that could be most effectively implemented at the European level, and those should remain at the discretion of individual States.

Steps 1 – 7 are clearly focused on process, and involve no explicit formal commitments being put into place in terms of performance (other than, of course, in terms of

[50] The potential for more active approaches to 'system operation' at the European level is discussed in Section 7 below.

compliance with the process requirements). They involve the definition of a set of indicators (step 1), and requirements to publish target levels of those indicators for a defined period (step 4). These steps can be understood as generating information flows in a common format. Steps 2 & 3 specify requirements in terms of formal user involvement in the target setting process, with a requirement for users to be consulted (the institutional arrangements for this consultation process are discussed further below), but also a requirement for user views to be reported on and responded to. That is, service providers could be required to explicitly address concerns and comments that are addressed to them. This, process, in itself, provides a potential mechanism that may generate the publication of new and useful information.

Steps 6 & 7 provide for third party scrutiny of the service providers proposals in the light of the consultation process and report, and for subsequent service provider response to this. Whilst no adjudicating powers are assumed in these steps, the review body could have some role in terms of 'sunshine' regulation, and may act as a form of 'prodder' in relation to the process. A potentially desirable extension here could involve consultation reports having to be 'signed off' by a review body. This could then provide a basis for ensuring that ANSPs provide adequate responses to issues raised by users, and provide a potential basis for generating additional relevant information flows over time.

Steps 8 & 9 are concerned with translating a set of targets into a formalised incentive contract that provides for enforceable commitments. Step 8 is intended to provide a potential mechanism for implementing agreed positions that have emerged from the consultation process, by means of a tightly defined regulatory role. This role then provides a potential mechanism for giving form to what would otherwise be informal agreements between users and providers, without the introduction of a regulator that can develop its own proposals. Thus, whilst some adjudication is introduced, it is basically passive.

It is notable that this, and the potential extension highlighted in step 7 (where a review body has to 'sign off' consultation reports), introduce bodies with very narrowly defined functions when compared to what is normally understood to be an economic regulator. In both cases, the focus is principally on process requirements with the review body faced with a simple binary decision (yes/no) in relation to what could be well defined criteria.

Step 9 goes beyond this and instead effectively involves the introduction of an economic regulator. That is, the review body has powers to propose alternatives, and acts to some extent as a user representative. Where agreement cannot be reached between the regulatory body and service providers, the matter could then be referred to an adjudicating body.

TABLE 4.1: THE DEVELOPMENT OF CONTRACTUAL ARRANGEMENTS

1. A set of desirable performance indicators are defined in relation to service provider activity;

2. Service providers required to consult with users and other industry participants on appropriate target levels for each performance measure over a defined period (for example five years);

3. Service providers required to publish a written report on the consultation process and this should include recognition of and responses to arguments and issues raised in submissions by other parties;

4. Service providers required to publish target levels for a defined period following consultation;

5. Service provider consultation (referred to in 2) to include explicit proposals on the compensation/incentive arrangements that should apply in cases of under/over-performance relative to target levels (to allow for generalised application, these proposals could include zero compensation/incentive factors);

6. Service provider's report on the consultation process (including submitted responses), and service provider proposed target levels and incentive/compensation arrangements, to be submitted to a designated review body. Review body has no powers but is required to publish its assessment which is made subject to pre-determined criteria (appropriate criteria are discussed below);

7. Service providers required to respond to review body report within a designated time period, and may publish modified target levels taking into account review body findings;

8. Designated review body has narrowly defined powers that could involve a right of veto in relation to the service provider proposal, or the power to choose between service provider proposed options on the basis of pre-determined criteria, but not to put in place alternative options (other than a 'zero compensation/incentive factors' option coupled with a proposed set of target levels). Thus the review body, subject to a check, could provide a means of formalising the service provider proposed options for contractual form to the ANSP proposal. [As an alternative, the review body could have a duty to make a recommendation concerning which proposal to adopt to a political body that would have determinative powers (which again could be narrowly defined).

9. Designated review body has powers to put forward its own proposals for target levels and incentive/compensation arrangements to be put into place, although its decisions would be subject to appeal by the service provider. [As in (7) above, the review body could instead recommend to a political body which would be determinative. In this case, the recommendation could be subject to an appeal]

4.5 Identifying potentially desirable institutional developments

The above provides a stylised account of a set of steps that include differing levels of obligation placed on service providers, and different levels of powers given to supervisory bodies. The current route charging arrangements would allow for all of these steps to be adopted at the national level if considered appropriate by a given State. In particular, the arrangements allow for a departure from the full-cost recovery system where a system of economic regulation is introduced, and this could (but need not necessarily) include steps of the kind set out in the above table. At present, however, only the UK have adopted an economic regulation approach with UK NATS regulated by the UK CAA.

A key focus in the context developing economic regulation arrangements as part of the Single European Sky initiative, however, is the extent to which the arrangements can be modified at the European level.

4.5.1 Disclosure of Information and other process requirements

As was noted earlier, at the European level, the Route Charging Arrangements currently require forward-looking forecasts of costs and traffic levels to be provided each year by States. However, as was also noted, compliance with these disclosure arrangements was limited to only half of the EUROCONTROL members for 2001. There are also disclosure arrangements under the ECIP and LCIP processes that generate planned levels of capacity enhancements. However, these processes are substantially disconnected from the cost reporting and forecasting process developed for route charging purposes, and the relationship between movements in cost and capacity levels is typically not made explicit.

A key area where the current arrangements could be usefully developed therefore relates to the form and extent of ANSP disclosure of information requirements and the enforcement of compliance with these requirements.

(a) Compliance with Disclosure Requirements

Compliance with information disclosure requirements in relation to forward-looking cost levels was highlighted as a significant problem area above. **One approach to this issue would be to make the eligibility of cost submissions for charging purposes in a given year dependent, or partially dependent[51], on a specified level of forward looking information being provided.** A similar approach could be adopted in relation to other process requirements, such as providing for a defined period of consultation (for

[51] We note that the PRC have suggested that the allowance of charge increases could be made dependent on the provision of required information (see PRR 6)

example, proposed levels published so as to allow one month for users to respond), and reporting on the consultation process[52].

(b) State/ANSP/ACC Disclosure Arrangements

It is important to note that the focus of disclosure requirements under the route charges system relates to States, rather than specifically to ANSPs. This only generates a significant issue at present in relation to MUAC, since the costs of MUAC have to be allocated between Belgium/Luxembourg, Germany and The Netherlands for charging submission purposes. However, over time, the relationship between individual state cost-bases and the costs associated with specific ANSPs may become less closely aligned with restructuring developments. If information disclosure in relation to the route charges system is to be the primary focus of dialogue between users and providers on cost and service provision levels, it is desirable to ensure that information on different ANSPs is presented in such a way that consistent comparisons and evaluations can be made. In particular, it would seem desirable to ensure that ANSP's are required to meet the same information disclosure requirements as States.

A separate issue that arises concerns the potential benefit of generating information flows that relate to ACCs in addition to information relating to ANSPs. Clearly, information on ACCs would be necessary in situations where it was considered appropriate to implement an approach to charging based on ACCs[53]. However, more detailed information provision by ACC could also be of value from a regulatory perspective – in the sense that it provides additional information that can inform cost and performance comparisons and assessments – and, may be useful in the context of signaling and informing potential restructuring opportunities. It is notable that the disclosure of cost information for a number of ACCs has resulted from the benchmarking work of the PRC[54] and the PRU[55], and the development of increased disclosure of information by ACCs over time – in particular, in a manner that can be reconciled with ANSP and State charging data provision - appears a desirable objective. However, given the potential for difficult allocation issues to arise, consideration would need to be given to time period over which any disclosure requirements could reasonably be implemented.

(c) Disaggregation of Cost Data

In their annual route charging submissions, States are currently required to provide a breakdown of costs for years n (previous completed year) to n+5 by type and by category as follows.

[52] Clearly it may be desirable to evaluate the extent to which a report on the consultation process has satisfied a number of criteria, such as adequately responding to user comments. This is discussed in terms of review processes below.
[53] The desirability of adopting such an approach is discussed in Section 9.
[54] "A comparison of performance in selected US and European en-route centres", EUROCONTROL Performance Review Commission, May 2003
[55] *Year 2000 European ANS Providers Benchmarking Report*, EUROCONTROL Performance Review Unit, ACE/2000 Data Analysis Working Group, September 2002

Costs by type:

- Staff
- Operating Costs
- Amortisation
- Interest
- Other – if necessary

Costs by category:

- ATM/CNS
- Training
- Studies/Tests and Trials
- Administration
- AIS
- MET
- Search and Rescue
- Other – if necessary

A key area where a more detailed breakdown of costs would be desirable concerns communications, navigation and surveillance costs (CNS). The Draft Service Provision Regulation envisages that the provision of communication, navigation and surveillance services (as well as aeronautical information services) should be organised under market conditions whilst taking into account the special features of such services and maintaining a high level of safety[56], and provides for the certification of separate provision of these services[57]. A robust separation of CNS costs can be understood as an important factor in ensuring that procurement processes are not unduly distorted by inappropriate cost allocations. For example, an ANSP could seek to guard against alternative CNS providers – and thus potentially hamper the development of alternative more efficient arrangements - by artificially depressing the levels of cost allocated to these services, if the shortfall could be otherwise recovered through route charges. It is also notable that if there were to be a development of separate upper airspace charges[58], then this would raise issues concerning the way in which CNS costs for a given area should be procured and charged for.

A requirement to disclose CNS costs separately would represent a first step in addressing these issues, and given the different characteristics of the individual components, separate disclosure of the costs of each component (i.e. communications separate from navigation separate from surveillance) would appear desirable unless a clear case to the contrary can be provided. Separate disclosure would ideally involve the breakdown of costs by type being shown separately for CNS in order to highlight the cost structures of these services.

[56] Preamble point 12 of draft service provision Regulation COM (2001)564 final/2 dated 11 December 2001 - 2001/0235 (COD)
[57] Article 6(3) of draft service provision Regulation
[58] This matter is discussed in detail in Section 9

In terms of providing conditions that are likely to facilitate competition for the provision of communication, navigation and surveillance services, however, separate disclosure of cost data can be understood at a first step. As with unbundling in other areas, disclosure requirements can be backed by requirements for differing degrees of separation such as internal accounting separation or full financial and organizational separation. In addition to initiatives focused on cost data, the development of market-based arrangements for CNS provision could be facilitated through more specific requirements in terms of the definition of service requirements for the provision of each component (with the potential for the subsequent use of benchmarking in order to assess declared levels), and the definition of procurement processes (for example, requirements to undertake market testing, or some form of tender processes to be undertaken).

(d) The quality of cost data

Clearly the value of cost data is dependent on its quality, and a number of concerns with respect to data quality have been raised over time. In 2000, the PRC recommended that the Director of EUROCONTROL should be given the right of initiative to investigate States' compliance with the EUROCONTROL Principles for Establishing the Cost-base for Route Facility Charges and the Calculation of the Unit Rates.[59] In response to the PRC's recommendation, the Provisional Council requested that the Enlarged Committee for route charges investigate the issue of compliance, and they in turn passed the matter over to the 'FIFU task force' who have been considering these issues.

Importantly, significant provisions are included in the Draft Service Provision Regulation that should impact on data quality. In particular, Article 11, which is concerned with 'Transparency of Accounts' requires that ANSPs draw up, submit to audit and publish their financial accounts, and that these accounts should comply with the International Accounting Standards adopted by the Community[60]. It also requires that ANSPs publish an annual report and regularly undergo an independent audit. Importantly, Article 11(3) directly addresses accounting issues in relation to route charging, and states that:

> *When providing a bundle of services, air navigation service providers shall, in their internal accounting, identify the relevant costs and income for air navigation services, broken down in accordance with EUROCONTROL's principles for establishing the cost-base for route facility charges and the calculation of unit rates and, where appropriate, shall keep consolidated accounts for other, non-air-navigation services, as they would be required to do if the services in question were provided by separate undertakings.*

A potential area of ambiguity that arises here concerns the relationship between the internal accounting separation envisaged by the Draft Regulation and the financial accounts to be audited.

[59] *"Fourth Performance Review Report",* PRR4, 2000, Recommendation 33

[60] The regulation requires states that 'Where, owing to the legal status of the service provider, full compliance with the International Accounting Standards is not possible, the provider shall endeavour to achieve such compliance to the maximum possible extent'.

In order to maximise the effectiveness of the auditing processes set out in the Draft Service Provision Regulation, it would be desirable if ANSPs were required to present their financial accounts in such a way that their *en-route* business accounts are shown separately, and to provide a reconciliation between data presented in their audited accounts and that submitted as part of the route charges system.

4.5.2 A forum for dialogue between service providers and users

As was noted earlier, there are significant weaknesses in the current arrangements in terms of the adequacy of fora for dialogue between users and providers on key cost and service quality issues. Some of the weaknesses of the Enlarged Committee on Route Charges were highlighted above. The approaches to consultation taken by different ANSPs vary considerably, but as the PRC's review indicated, these processes are frequently not transparent in terms of both the processes involved and their impact on decisions.

More generally, there would seem to be significant benefits, given the interrelatedness of many of the issues involved, for consultation processes to be organised in a coordinated manner. The use of a common European consultation framework could also be expected to reduce the costs of user involvement in a number of ways, including through the use of standard formats and timetables for information flows. Given the importance of network effects on ANSP activities and costs, and particularly in the light of the potential development of functional blocks of airspace, there would seem considerable benefit for consultation to be organised in relation to a European Forum, at least as a central point of focus.

Given that relevant issues are most closely interrelated over more narrowly defined geographical areas (and that this may form the basis for the development of Functional Blocks of Airspace), and in order to provide more practical fora for discussion with other providers and with users, there may be significant benefits in a European forum developing regional sub-groups.

We would note that there are plans to reform the arrangements of the Enlarged Committee on Route Charges, and given this, a modified Enlarged Committee could potentially provide an adequate European forum. However, we would note that many of the current perceived weaknesses of the Enlarged Committee are likely to relate, at least in part, to a lack of clarity in terms of what the role of the Enlarged Committee should be. In particular, if the Enlarged Committee is understood to be principally a political or a regulatory decision-making body, then it may be considered that a main source of weakness is the fact that some states are represented by ANSPs. The key problem here is that there is not a clear separation between service provision and regulation (and as the PRC have noted, this can give rise to well recognised conflicts of interest).

A more general question then concerns what role political and/or regulatory decision-making bodies can most effectively adopt in processes of this kind, and more specifically how they should be involved in consultation processes between providers and users, and in review processes. Consistent with our review of best practice regulation, we would argue that are significant benefits in consultation fora that are

clearly separated from political decision-making processes. A modified Enlarged Committee may be able to provide such a forum, but this would, of course, raise the question of which body undertakes the political decision making functions of the current Enlarged Committee (and in particular the final approval of charges).

More generally, we note the comments of CANSO, IATA and others who have argued for the importance of a consultation body that is independent of EUROCONTROL in the context of the development of implementation rules for the Single European Sky package[61]. A key contention here is that as an inter-governmental organisation, EUROCONTROL is not best placed to facilitate what is principally part of a commercial contracting process.

Whilst the body responsible for organising a forum of this kind can be expected to have some bearing on the manner of its workings, a key issue will in any case be the basis of user involvement. As noted above, **users can only attend meetings of the Enlarged Committee on an observer basis. Given the importance of user involvement, a substantial change in this position is justified.** As was indicated above, movements to strengthen the user voice in consultation processes could include the introduction of requirements on ANSPs to report on the consultation process, and to explicitly respond to issues raised by users in that process.

4.5.3 Review/Advisory Bodies

The PRC currently provides a European review body with no explicit adjudicative or enforcement powers. In terms of the regulatory roles identified in the review section, the PRC can be said to actively engage in development and formatting of information flows, and in 'prodding' or disturbance activity. It is also important to recognise the PRC's advisory role for the EUROCONTROL Provisional Council. The PRC has made recommendations to the EUROCONTROL Provisional Council on a wide range of issues including in relation to the safety reporting system, delay targets, airspace design, and information provision from ANSPs. Some of these recommendations are fairly general in nature, but increasingly over time there have also been more specific recommendations. The PRC publishes a log of its recommendations in its annual performance review reports, showing the response of the Provisional Council and the actions to follow. All of the PRC's recommendations appear to have been endorsed, although clearly the practical effect of this depends to a significant extent on how specific the recommendation was.

The PRC, then, provides an existing example in ATM of where an advisory body, through producing authoritative reports and recommendations, can allow for a process where political decision making bodies (in this case the EUROCONTROL Provisional

[61] *An Industry Consultation Body for the Single European Sky,* An Industry-wide Joint Proposal by the European Association of Aerospace Industries (AECMA), Airports Council International–Europe (ACI-EUROPE), the Civil Air Navigation Services Organisation (CANSO), Association of European Airlines (AEA), European Business Aviation Association (EBAA), European Regions Airline Association (ERA), International Air Carrier Association (IACA) and the International Air Transport Association (IATA), April 2003.

Council) can adopt a position of routinely signing-off on their advice in relation to a set of issues. Thus, determinative powers are retained by the political decision making body, and could be invoked to 'over-rule' the advisory body's recommendations in specific circumstances. This raises the question of the extent to which the PRC, or some other advisory body, could usefully adopt a similar role in relation to more specific matters of target setting and encouragement mechanisms.

The steps outlined in table 4.1 above envisage a range of different roles that could be undertaken by a review body (or set of bodies). In step 6 above, it is envisaged that a review body would evaluate the consultation report submitted by each ANSP – which would include proposed target levels for identified performance indicators. If ANSPs were required to submit a report on the consultation process that responded to issues raised by users, an important role for a review body could be to evaluate whether such a report had been adequately completed. As noted above, this could be formalised as a 'process requirement' that required 'sign off' by a European level body, other than in cases where a Member State has introduced economic regulation.

More generally, the extent to which matters can be determined by decisions at the European level is clearly a central issue. The Draft Service Provision Regulation[62] explicitly states that the introduction of encouragement mechanisms will be at the discretion of member states. Given this, the role of economic regulator, as commonly understood, is clearly envisaged as one that could only be imposed at the State level. However, as was stressed in Section 3, regulation – particularly in situations where as in ATM the demand side includes a range of large sophisticated users - encompasses a range of activities, and the fact that the decision of whether or not to introduce encouragement mechanisms is envisaged as a Member State level decision, does not rule out a range of potentially desirable regulatory activity at the European level.

One potentially such desirable role could relate to situations where an effective consensus between users and service providers on an appropriate encouragement mechanism emerged from consultation processes. In such a situation, it would appear desirable if there were a mechanism at the European level whereby the proposed 'contract' could be examined, and subject to specified criteria, could be implemented. In particular, a review body could be required to assess the proposal in terms of a relatively well-defined criterion such as: whether the introduction of the agreement could be expected to better facilitate the operation of an efficient and coordinated European network. A clearly specified role for EUROCONTROL in such a process could involve the provision of expert advice to the review body on network implications. By defining the assessment in terms of a simple yes/no type of question, it could be ensured that only proposals presented by the ANSP can be agreed. It therefore ensures that individual States – through the supervisory arrangements that they have in place for relevant ANSPs – can effectively determine the flexibility that relevant ANSPs have to enter into such agreements. It does, however, provide a potential mechanism

[62] Article 14(3e) of draft service provision Regulation COM (2001)564 final/2 dated 11 December 2001 - 2001/0235 (COD)

whereby forward-looking agreements can be given a contractual form through a process of regulatory oversight at the European level.

4.5.3 Investment Planning Issues

A preliminary step in developing appropriate incentives for efficient investment involves the clear identification of the parameters that are relevant for each investment project. At a broad level this involves identifying and defining:

- what *outputs* are to be delivered as a result of ATM investment projects;
- the *timing* of the delivery of these specific investment outputs;
- what are the *benefits* attached to these outputs;
- the *costs* associated with the investment project

The identification and definition of the outputs of ATM investment projects and the timing over which these outputs would be delivered, or will become operational, typically form part of the investment planning process that is currently undertaken by individual ANSP's.

The report of the EUROCONTROL workshop of 30 September 2002 outlines certain broad criteria that can be used to check whether the investment project could be considered economically viable, these include whether the:[63]

1. cost per unit of aircraft operations decreases as a result of the investment
2. the most cost-effective option has been considered in the planning process
3. investment is economically viable for the users
4. costs and benefit monitoring and reporting mechanisms are implemented

The workshop concluded that it was important to refer to the above assessment criteria at different phases of the investment process. In the next section we examine potential forums or mechanisms through which this assessment of investments project can be conducted.

In two forthcoming reports on the economic evaluation of the benefits and costs of ATM investments, EUROCONTROL seeks to develop a consistent approach and standard methodology for the economic evaluation of investment projects by service providers.[64]

[63] EUROCONTROL (2002) Report of the Economics Workshop -30 September 2002 EUROCONTROL HQ, Brussels

[64] "Description of the Process for Economic Evaluation of ATM investments", EUROCONTROL Economics, Issue 1.0/Document 12, 10 February 2003 (unpublished); and "Approach to Assess the Benefits and Costs of ATM investments", EUROCONTROL Economics, Issue 1.0/Document 20, 12 June 2003 (unpublished). These studies also draw upon the discussion papers of the Economics workshop held at EUROCONTROL on the 30th September 2002

(a) Processes for assessing ATM investments

Once the outputs and benefits/costs of ATM investments have been identified a separate issue relates to the appropriate forum and process through which this information can be assessed and examined. The need for a transparent and inclusive forum for the assessment of ATM investments is emphasised in the forthcoming EUROCONTROL report on the investment process which notes that:

> *The economic evaluation approach should be developed in a transparent way involving and informing stakeholders from the beginning to ensure an as wide as possible ownership of the results.*[65]

The following considers two potentially complementary approaches that can provide useful inputs into investment assessment processes: investment plans; and the use of an economic evaluations group.

(b) ATM investment plans

One method by which ATM investment proposals can be documented in order to facilitate assessment is through the use of investment or business plans.

The use of investment plans or statements has been introduced in a number of other regulated sectors and typically involves the publication of forward looking statements covering five to ten years that outline in detail the planned system investments over that period. In some industries the company responsible for the investment plan must consult with interested parties prior to the publication of the investment plan which can allow for third party scrutiny of the investment programme prior to it being finalised and undertaken.

The development of plans relating to future capacity expansions, or future investments, is not uncommon amongst ANSPs in Europe and elsewhere in the world, such as in Australia, New Zealand and Canada. Below we highlight a number of examples of how ATM investment and business plans are used in different countries.[66]

(c) The use of investment plans by ANSPs

In France an annual investment plan is prepared and submitted by the General Directorate for Civil Aviation to the French Finance Minister for approval in the year preceding its introduction. The annual investment plan details information on all investments categorised into specific sections and also provides a list of authorisations in the coming year. The investment plan is reviewed by both DNA and endorsed by the head of the General Directorate for Civil Aviation prior to being submitted to the Finance Minister. The investment plan is also presented to users for consultation in the year preceding its introduction, and user consultation on the investment programme is also conducted as part of the air navigation charges consultation.

[65] EUROCONTROL Economics, Issue 1.0/Document 12, 10 February 2003 (unpublished), page 5

[66] We are grateful to Avinor , DNA, DFS, Skyguide, Austrocontrol and Luftfartsverket Sweden for their assistance in helping us to understand this process

In Norway, on the other hand, annual or periodic investment reports detailing estimates of completion dates and comments on deviations and corrective actions are prepared for internal monitoring and control purposes. Investment plans are not submitted to either a regulator or the Government for approval. However, user consultation occurs in respect of both specific investment projects and on total investment plan levels.

Investment plans covering three-year periods are prepared by Luftfartsverket in Sweden as part of its annual business planning process, which include information of investment objects from a five-year plan and is approved by the Director of the Air Navigation Services Division. At the same time, each year an investment plan outlining the total amount of planned investments for each of the next three years is sent to the Government as well as to the Ministry of Industry, Employment and Communications for approval. Consultation with users on investment plans occurs if the investment is more or less initiated for or dimensioned for one specific customer, such as the Military, whilst other users are informed about our major investments at regular meetings.

In both Austria and Switzerland there are no specific requirements to consult with users regarding investment projects. While in Austria consultation generally occurs on major investment projects between Austrocontrol and the Austrian Airlines Group (their biggest user), in Switzerland some projects may be discussed in collaboration with airports, but not airline users. In Switzerland, an annual investment plan is prepared internally but is not submitted to a Regulator or Government for approval, but rather is approved internally.

In the UK, NATS is required to prepare three documents relating to investment. These include a full *ten-year business plan*; an *annual service and investment plan* and *an annual business plan report*. The ten year business plan details forecasts of future demand, the standards of service (volume, delays) that NATS plans to meet, the capacity to provide and capital investment plans. As a condition of the business plan's approval by the regulator, NATS is required to consult with users and other interested parties on the business plan, including through general consultation meetings and workshops. In addition to the requirement to produce a *ten-year business plan*, NATS is also required under its licence to produce an annual *service and investment plan* and a *business plan report*.

In Canada, NAV CANADA produces and publishes a three-year business plan which details current and proposed capital investments over the period.[67] In addition to the publication of the three-year business plan, NAV CANADA also publishes an *annual information form* that contains a brief update on recent developments and current trends, including the development of ANS infrastructure.

Similarly, Airways New Zealand also produces an *Air Navigation Plan* which provides an overview of the current system and looks at the future of air navigation systems in New Zealand.[68] It is a 15-year plan which describes Airways intentions for new air

[67] "NAV CANADA business plan 2002-2005", NAV CANADA, 2002

[68] "Airways' Air Navigation Plan for New Zealand", Airways Corporation of New Zealand Limited, May 1999

traffic control and navigation systems implementation, and the plan is updated every two years.

Air Services Australia has also recently published its first *Air Traffic Management Strategic Plan* which identifies strategies to achieve key operational improvements over the 15 year period from 2000 to 2015.[69] The Strategic plan provides for regular review of this process and details a cost-benefit methodology which should be employed in developing projects under the Strategic Plan.

(d) Potential Developments

While the majority of ANSPs in Europe we have spoken to have stressed that they produce investment plans and programmes for internal purposes, there is in many cases no formal requirement for them to consult or publish this information more widely. In practice this significantly limits the potential for external assessment or scrutiny of the investments, and associated costs, that are being undertaken by the individual ANSPs. **Perhaps more significantly, and as discussed elsewhere in this report, this could result in failures of coordination within the European ATM network, both because individual ANSP's may be potentially unaware of relevant investments that are being undertaken in neighbouring regions, and, perhaps more importantly, because ANSPs are not required to account for their investment decisions within a common framework and forum**. Issues of overlapping investments at some points and underinvestment in other parts of the European network were highlighted as an area of concern during discussions with various stakeholders during the course of this study.

Requirements to publish forward looking business plans, that specify the content to be included – which could potentially include a detailed specification of each of the items listed under the benefit and cost categories noted above – provide a potentially important contribution to investment appraisal processes in European ATM, and could provide a common basis upon which proposed investments could be reviewed and assessed by stakeholders.[70]

(e) The creation of an economic evaluations group

An approach to the assessment of investment projects which was proposed in a discussion paper for the recent EUROCONTROL workshop on the investment process focused on the creation of a specific economic evaluations group who would be given responsibility for: [71]

- conducting, supervising and reviewing economic evaluations
- monitoring and ensure consistency
- maintaining the economic evaluation approach

[69] "Australian ATM Strategic Plan", Ed. 1.01, Airservices Australia, November 2001

[70] One of the recommendations of EUROCONTROL Document 12 is that "the approach should provide a consistent cost and benefit typology, standardised inputs and outputs and a controlled set of incremental baselines", page 6

[71] EUROCONTROL Economics Workshop -30 September 2002 EUROCONTROL HQ, Brussels 30 September 2002 – Discussion paper on stakeholder involvement, page 4

- advising the ATM/CNS consultancy group and other bodies on economic evaluations

The perceived benefits of introducing an economic evaluations group noted in the discussion paper is that it will provide for continuity and consistency of and between economic evaluations.

Whilst some of the roles highlighted above - reviewing, monitoring, and advising on the adequacy of economic evaluations - can clearly be understood as desirable, questions clearly arise in relation to which institutions are best placed to undertake such activities, the scope of activities that they should be engaged in, and, more generally, their relationship of these roles to consultation processes between users and providers. There would also be issues about the extent of accountability of such an economic evaluations group for the investments that are approved through this process.

4.6 Contractualisation and the roles of EUROCONTROL

The above discussion highlights a more general set of issues concerning the impact that developing processes of contractualisation and economic regulation should properly have on the activities of EUROCONTROL. These issues are largely beyond the scope of this study, although given the range of activities that EUROCONTROL engages in, the issues cannot be wholly avoided. We would note, that in line with best practice, and consistent with the planned approach for ANSPs, a clear separation of service provision and those activities that can be understood more properly as regulatory activities would seem highly desirable.

We would also note the comments by the Performance Review Commission in relation to weaknesses of the current Air Traffic Flow Management arrangements, including from their comparisons with US operations, and that the PRC plans to review EUROCONTROL activities more thoroughly in developing its next performance review report. As a general point, it seems entirely appropriate that similar processes of scrutiny should be directed at EUROCONTROL's service provision activities as are being proposed in relation to ANSPs.

4.7 Summary

To summarise, the following points can be made regarding contractualisation and service provision arrangements:

- A trend in the regulation of network industries has been the development of increased levels of 'contractualisation'. In ATM in particular there has been considerable and growing interest in the potential for developing increased levels of contractualisation in relation to ATM from range of parties - including user and service provider industry bodies

- In particular, under the current arrangements significant concerns have been expressed concerning the adequacy of the Enlarged Committee in terms of providing a multilateral forum for user representation. Given the importance of user involvement, a substantial change in the current position is justified

- Important practical difficulties arise when considering the potential for commercial contracting on service levels between service providers and users. The key difficulty is that the relationship between an individual service provider's performance and the actual service received by airlines is not straightforward - it is likely to be subject to significant and complex network effects.

- The development of contractual arrangements within a regulatory context could involve a set of stages that require different degrees of formalised regulatory requirements, and the different powers conferred on particular types of institutions. In broad terms, this process moves from a situation where service providers are faced with requirements in terms of process, to one where the may be faced with externally determined financial incentive arrangements linked to delivering specific levels of performance. An important aspect of identifying stages separately in this manner is to provide a basis for assessing which steps could be more effectively implemented at the European level, and those should remain at the discretion of individual States.

- Consistent with this increasing focus on contractualisation, the current institutional arrangements could be usefully developed to require greater levels of disclosure of information by ANSPs and greater enforcement of compliance with these disclosure requirements.

- Investment planning raises a number of specific issues in terms of contractualisation which concern, amongst other things, the uses that can be made of investment plans and forecasts, and how they might relate to allowable revenue assessments

SECTION 5
RISK SHARING & TRAFFIC VOLATILITY IN ATM

Section 4 of this report was principally focused on the processes by which the development of increased levels of contractualisation can be facilitated through modifications to current institutional arrangements. Before considering the specific forms that incentive, and more generally risk and benefit sharing, contracts could take (in Section 6), we first consider some of the specific issues raised by traffic volatility in ATM, and potential responses to the associated risks that this volatility can generate for service providers and airlines.

We begin this section with a discussion of the effects of traffic volatility on both service providers and airlines and why the introduction of a potential smoothing mechanism is seen as desirable by some stakeholders. This is followed by a discussion of broad categories of traffic volatility faced by system users and service providers and a consideration of the extent to which risks associated with this volatility might be partly controllable by service providers. We then examine potential ways in which the smoothing of user charges can be financed, including through the introduction of a solidarity mechanism, drawing upon the experiences of such mechanisms in practice such as the NAVCAN rate stabilisation fund and the NATS exceptional user contribution. This includes a discussion of the potential for the development of a 'Revenue Recovery Imbalance Account' as a framework within which alternative approaches to charge smoothing can be assessed and potentially implemented in a compatible manner.

5.1 The impact of traffic volatility on service providers and users

Following the downturn in traffic volumes after 9/11, much attention has been focused on the impact of changing traffic volumes from unexpected exogenous shocks on the level of user charges for *en-route* services and on the impact that this has on service provider revenue. This has lead to calls from some stakeholders (including both ANSPs and airlines) for the development of mechanisms that could be used to smooth that impact of changing traffic volumes on user charges and service provider revenue.

The following general concerns are typically cited as reasons why the development of a smoothing mechanism in ATM is required:

(a) The timing of charge increases given airline financial difficulties

As has been seen over the past couple of years, the current charging arrangements can result in significant increases in ATM charges at times when the economic conditions facing airlines are at their most severe.

While the current arrangements have some degree of formalised adjustment incorporated into the charge calculation process - with over or under-recovery - of revenue for a given year (n) adjusted for in charges two years later (n+2) - questions

remain as to whether the degree and form of charge provided for formally in the current arrangements is sufficiently responsive to economic circumstances, particularly in situations where there has been a significant fall in traffic levels.[72]

(b) Financing difficulties faced by ANSPs

Where actual traffic levels differ from those that were forecast for the calculation of unit rates, service provider revenues will differ from their declared level of cost base. In particular, where traffic levels are lower than forecast, an under-recovery of revenue will result. To the extent that there is a time lag before this is returned to users through charges, and the shortfall is significant, this can give rise to financing difficulties for service providers.

(c) Spiraling charges

An additional form of concern relates to a potential spiraling effect that can result from significant user charge increases by particular ANSPs, the basic process of which is described in the box below.

It is notable that these situations highlight that in practice ANSPs and their owners – typically, state governments – do in fact bear some degree of financial risk for falling traffic under the current charging arrangements. Furthermore, pressures for cost reduction that can arise from ANSP financial difficulty can be of particular relevance from an airline perspective. In particular, in recent years airlines have put considerable emphasis on the arguments concerning ANSP cost levels and the extent to which ANSPs have reduced their cost levels in response to a substantial fall in traffic. The airlines have also put considerable focus on the extent of their own cost-cutting efforts, and highlighted the need for significant attention to be paid to ANSP cost levels.

(d) Assessment

In principal, given the characteristics of the aviation sector, there may good efficiency reasons for considering some degree of smoothing of user charges as desirable. In particular, the profiling of cost recovery may allow for a higher proportion of fixed costs to be recovered in periods when demand is less price sensitive and vice versa. However, the assessment of the desirability of different levels and forms of charge smoothing is complicated by a range of factors related to views on the efficiency and flexibility of service providers, and the financing arrangements that might be used. These issues are considered in turn in the next two sub-sections.

[72] In this respect, we note that as an exceptional measure, 2002 unit rate increases were deferred for 3 months.

Spiraling Charges

The basic process of concern is as follows:

- a significant fall in traffic results subsequently in a significant increase in the relevant unit rate;
- the significant increase in the unit rate results in a significant degree of avoidance of the ANSP's airspace (to take advantage of the lower charges offered in neighbouring airspace);
- this avoidance activity results in a further significant fall in traffic levels which, in turn, subsequently results in a higher unit rate, which encourages more avoidance activity, and so on.

This effect relates to situations where the relative price of a specific ANSP increases, and airlines find it profitable and practical to substitute away from using that particular ANSP's airspace. This effect is, therefore, most significant for nations that have a relatively small amount of airspace (for ANSPs with larger airspace, substitution is more likely to be a relatively marginal issue).

It can be noted in relation to this spiralling effect that substitution by airlines to other airspace is not costless. It is to be expected that the avoidance activity could result in airlines using less efficient routes that involve more fuel usage, with its associated costs (including externalities).

5.2 Traffic volatility and ANSP revenue allowances

The discussion thus far reflects debates about risk sharing in ATM that have tended to see this issue as largely separable from other aspects of charging arrangements. However, as the summary of regulatory principles in Section 3 makes clear, the incidence of risk is heavily influenced by the general approach that is taken to the regulation of charges. For example, and subject to qualifications concerning the effects of lags in charge adjustment that may be associated with charge smoothing arrangements, cost-of-service regulation tends to allocate financial risks associated with traffic volatility to users, whereas price-cap regulation tends to shift such risks much more towards service providers.

A key issue for evaluation concerns the extent to which it is indeed desirable to seek to shift some of the risk associated with traffic volatility to service providers. In order to assess this issue, it is useful to seek to identify some of the different potential sources of traffic volatility. There are many potential sources of traffic volatility in ATM and any

examination of the most efficient allocation of this risk should begin by developing a better understanding the various categories.

Broadly speaking, in considering the allocation of risk a desirable balance, in efficiency terms, will depend on a number of factors including:

i) the extent to which the likelihood of particular outcomes can be influenced by the behaviour/conduct of service providers and/or users;

ii) the extent to which the physical and/or financial risk mitigation mechanisms available to different parties can be expected to differ significantly in terms of their efficiency.

A useful distinction to employ when thinking about who should bear the different risks associated with traffic volatility is that of 'partly controllable' and 'uncontrollable' risk.[73] In the context of ATM, 'partly controllable' risk can be defined as those financial risks associated with traffic volatility that can be materially influenced by either service providers and/or system users. Conversely, 'uncontrollable' risk would be the financial risks associated with traffic volatility that are common to the entire ATM sector and which cannot materially be influenced by system users and/or service providers.

(a) Supply-side volatility and risk

The potential sources of traffic volatility generated by supply-side factors in ATM can be broadly categorised as being the result of either:

(i) *Technical failures*: which can be caused by a variety of sources such as unexpected equipment or system failures or by the poor management and maintenance of systems and equipment.

(ii) *Industrial action*: such as strikes either by controllers or by related services (support staff etc.).

(iii) *Exogenous factors*: such as weather

It is arguable in relation to both (i) and (ii) above, in certain circumstances service providers should be able mitigate the expected traffic volatility arising from these events. In these circumstances, this suggests that the risk associated with traffic volatility caused by either technical failures (for example, where it is related to poor up-keep or maintenance of equipment) or industrial action (where it relates to poor

[73] This distinction is closely related to the concept of moral hazard, where economic actors may not bear the full consequences of their actions due to uncertainty and incomplete or uncertain contracts which prevents the full assignment of damages for their actions. An example, in the case of insurance markets is that people who take out insurance may take greater risks than they would do if they did not have insurance because they know they are protected, which results in the insurer may getting more claims than it expected. The difficulty arises because the conduct of the insured party affects risks (i.e. risks are partly controllable) in ways that are not easily specified and monitored.

management or arbitration practices) is to some degree controllable by service providers. It follows that, to the extent that these risks are potentially controllable by service providers, these risks should be borne to a greater extent by service providers.

(b) Demand-side volatility and risk

Similarly, a list of broad sources of demand-side traffic volatility can also be compiled which might include:

(i) *Business cycle effects*: such as volatility in demand related to economic booms and busts

(ii) *Airline market developments*: changing mix and pattern of demand through developments such as the emergence of low-cost airlines

(iii) *Exogenous factors*: eg. 9/11, SARS, Iraq war

It is arguable that to some extent the risks associated with (ii) above may be partly controllable by the system users, and thus these risks should be borne by users. That is, the traffic volatility associated with changes in the structure of the airline market feature some elements that, in certain circumstances, could reasonably mitigated by system users. For example, a fall in demand on specific routes for large national airlines following the introduction of low budget competitor on alternative routes is at least partly a reflection of business conduct, and it follows that the risks associated with this source of demand volatility should lie with those system users.

The above discussion is clearly a very preliminary discussion of relevant issues, and does not suggest clear-cut divisions. What the discussion does suggest, however, is that at least some part of the risk associated with traffic volatility should be understood as partly controllable by service providers. It can be noted that previous analyses of ATM cost structures[74] indicate that a very high percentage of ANSP costs are fixed in the short term[75]. However, the fact that it may be difficult to respond flexibly to traffic level variations does not mean that flexibility is not important factor. Indeed, a greater focus on flexibility to respond to alternative situations in the face of significant uncertainties concerning future forecasts has been a very common feature of many network industries. As indicated above, these uncertainties can relate not simply to the overall volumes of traffic, but also to the composition of traffic over time in terms of relevant temporal (e.g. peak/off-peak) and spatial (routes, flight levels and associated distances) parameters. It is notable that the importance of service provider flexibility in relation to performance levels was emphasized by the PRC in its recent report, PRR6:

[74] PPM Task Force
[75] In the longer term, a higher proportion of costs will, of course, be variable.

> *Flexibility of resource allocation clearly appears as an important factor for both delays and cost-effectiveness. It was identified as a main factor in the productivity difference between US and European ACCs.*[76]

These factors imply that some ANSP exposure to traffic level changes in terms of revenue recovery allowances may be desirable in order to encourage flexible responses to changes in relevant circumstances[77]**.**

5.3 Potential options for financing charge smoothing mechanisms

As highlighted in Section 5.1 above, another important issue when considering the potential development of alternative charge smoothing arrangements concerns the mechanisms by which they can be financed. It can be noted that this issue is of central relevance to (at least one part of) discussions concerning the potential development of some form of solidarity mechanism. In this section we examine a number of potential sources of financing for smoothing user charges that could be developed to seek to lessen the impact of traffic volatility on both users and service providers.

We note that the Enlarged Committee on Route Charges has considered some of these possibilities,[78] including pluri-annual application of the adjustment mechanism, and the creation of a stabilisation fund with a credit line facility. However, none of the measures considered gained sufficient support by the Study Group. It was noted that smoothing mechanisms in themselves do not encourage either cost efficiency or investment in improved performance.

(a) *The establishment of a 'rate stabilisation fund'*

One potential mechanism for smoothing user charges is through the creation of a 'fund' specifically established to stabilise user charges and ANSP revenues by providing for the additional financing in times of falling traffic through contributions from user funds that have been accumulated in a fund over time. As can be seen from the box below this approach to smoothing user charges has been adopted in Canada since 1998 in the form of a rate stabilisation account.

In the process of our consultation, some stakeholders have indicated considerable doubt in relation to the development of this type of fund. In particular, they have questioned the efficiency of holding money in a fund, and have also raised significant concerns about the credibility of how the fund would be managed – in particular, whether the funds would in fact be diverted to another use at a later date.

In considering the desirability and efficiency of a rate stabilisation fund in a European context, a number of issues would need to be addressed, amongst which would include:

[76] *An assessment of Air Traffic Management in Europe during the calendar year 2002* PRR 6 Version II, EUROCONTROL Performance Review Commission, May 2003, page 65

[77] The question of when revenue should/could be recovered from users, as opposed to how much should be recoverable, is discussed below

[78] SGR/02/32/2886; 2876; 2879, considered at the meeting on 24-25th October 2002, Prague.

- how the mechanism would be managed, and by whom
- whether the fund would be European wide or separate funds would be maintained on a regional or national basis
- the specific terms of access to the fund – ie: how much would traffic have to fall before it would trigger access to the fund
- whether the fund is permanently established or established for only a temporary, and revisable, period

However, to reiterate the point made in Section 7 in relation to the creation of an infrastructure development fund, we are skeptical of the efficiency benefits of these types of funds where users are asked to pay in advance without receiving any firm commitments in relation to what benefits they will receive in the future.

(b) The establishment of pre-approved credit lines

An alternative option for smoothing user charges in times of falling traffic might be through the establishment of pre-approved lines of credit by service providers.[79] The establishment of credit lines would allow service providers to access additional financing in times when traffic falls below forecast levels in order to mitigate the impact on user charges.

An advantage of this mechanism as compared to the establishment of a solidarity mechanism is that it gives the service providers the flexibility to determine, on the basis of their own assessments, when they need to access additional financing. It also has the advantage of only allowing service providers to access funds when they actually require them and therefore avoids the concerns of airlines about 'pre-financing'. However, the viability of this type of smoothing mechanism depends greatly on the access to, and terms on which, service providers are able to obtain financing. For example before providing finance, capital markets may require some form of institutional changes or guarantees possibly in the form of independent regulation or government security.[80] In addition, the use of credit to finance their activities exposes service providers to the typical risks associated with capital markets such as unexpected interest rates movements, which potentially introduces another element of financial instability to the sector. However, in this regard it has been suggested that better financial rates may be made available to ANSP's if credit lines are provided at a regional, rather than national, level.

[79] It is our understanding that this mechanism is already being used by some service providers.

[80] In this regard, note the experience of NATS following September 11 when NATS bankers indicated that they would not allow drawings under the Capital Loan Facility and NATS was unable to finance all of its planned expenditure. See "NATS' Application to Re-open the EUROCONTROL Charge Control" CAA Decision, March 2003, page 10

NAV CANADA rate stabilization account

In 1998 NAV CANADA established a 'rate stabilization account' for the purposes of mitigating; "the effect on its operations of unpredictable and uncontrollable factors, principally fluctuations in air traffic volumes resulting from the cyclical nature of the commercial air carrier industry".[a]

In terms of the mechanics of the operation of the account, if the actual revenues in the rate stabilization account are greater than the Company's forecast financial requirements, then any excess will be reflected as a liability against the account and will be returned to customers through reduced service charges in the future. However, if actual revenues in the stabilization account are insufficient to cover forecast future financial requirements this shortfall will be recorded as an asset in the account and will be recovered through higher future customer charges.

Deposits or withdrawals from the rate stabilization account need to either be approved by the NAV CANADA Board or be based upon variations from amounts used in establishing the service charges. The account was initially endowed with $50 million CAD which rose to a peak of $58m CAD by the fiscal year 2000. The contributions made to the account in 2001 were $16m compared to total ANS charges of some $908m (1.76%).

Following the fall in traffic post 9/11, a further program of cost reductions and revenue improvement known as the Mitigation Plan was undertaken to which all stakeholders were asked to contribute.[b] The plan included a temporary payment deferral plan to help users' cash flow. The rate stabilisation fund was fully exhausted and in August 2002 had a deficit of $19m.

Since that time, the NAV CANADA Board has determined to re-establish the fund account to $50m within five years, and in its most recent review of service charges effective from 1 August 2003, NAV CANADA have increased the average service charge by 6.9%. According to the press release the revised rates comprise two elements: an increase in base rates which are set to allow for a breakeven position for the fiscal year 2003/04; and "an additional rate adjustment related to the recovery of the past shortfall and replenishment of the rate stabilization account". [c]

(a) "Backgrounder on rate stabilization account", Available on NAV CANADA website
(b) "NAV CANADA business Plan 2002-2005"NAV CANADA, page 17-18
(c) "Announcement of revised service charges"NAV CANADA press release of July 21 2003

(c) Allowing for ANSPs to hold Reserves through 'over-recovery' in good times

A further mechanism for smoothing user charges in periods of falling traffic involves allowing ANSPs to build up financial reserves through 'over-recovery' during periods of high traffic to be used to finance the 'under-recovery' of revenue in times of low demand.[81] In many ways this is similar to the development of a number of national based solidarity mechanisms whereby each country establishes it's own fund which it manages over time, building up reserves through 'over-recovery' in times of high demand and using these reserves in periods of low traffic and ANSP 'under-recovery'.[82]

An obvious advantage of this type of 'over-recovery' mechanism is that it can potentially introduce stability to both user charges and service provider revenues over time. However, as was noted in regard to the solidarity mechanism above, questions still surround the efficiency of holding money in these types of funds, and whether sufficient oversight and regulatory mechanisms will exist in the different member states to ensure that the funds are appropriately and credibly managed.

(d) Allow for the re-negotiation of forecast traffic levels and allowable revenue

A further mechanism for smoothing user charges might involve introducing processes that provide for the re-negotiation and reassessment of forecast traffic levels in specific circumstances, such as when the actual traffic level falls below the forecast traffic level by a particular amount. If a formalised process is introduced which could trigger such a renegotiation this approach might allow for user charges and the allowable revenue of service providers to be adjusted with immediate effect following any significant fall in traffic levels.

Where incentive arrangement are introduced a fall in traffic levels can give rise to a different form of concern – in particular, that the financial position of an ANSP may be heavily undermined as a result of the arrangements. This raises the question as to whether explicit recognition of a potential need to renegotiate arrangements under extreme circumstances is desirable. The NATS application to re-open its EUROCONTROL price control is an example of such a situation and is described in the box below.

An obvious problem with this approach, and one that was highlighted recently by the NATS application, is that it may prove cumbersome and costly to administer. In addition, it may not be desirable for service providers as it introduces greater uncertainty regarding their longer-term financial stability, which could impact their decisions regarding the financing of investments and other activities.

[81] This mechanism is particularly favoured by CANSO. See 'Position Paper on the Communication from the European Commission to the Council and the European Parliament on the Implementation of the Single European Sky', February 2002.page 5

[82] The issues of the incompatibility of this type of mechanism with national laws, particularly taxation laws, have been raised by a number of stakeholders

The NATS Exceptional User Contribution

In March 2003, the UK CAA confirmed its decision to vary the charge control condition and future regulatory policy by introducing a range of measures – known as the exceptional user contribution - to help boost NATS long-term financial stability.

The specific changes to NATS licence approved by the CAA included that[1]:

- the price cap be adjusted from RPI-4% in 2003 and RPI-5% in 2004 and 2005, to RPI-2% for 2003, 2004 and 2005;
- until 2005 volume risk will be shared by NATS and users, with NATS facing no more than half of this risk; and
- enhancements to the financial ring-fencing of NERL to protect users.

In order for the exceptional user contribution to be approved by the CAA, NATS agreed to implement a series of measures – known as the composite solution - relating to the restructuring of its financial structure, which included :

- securing of £130 million of new investment by BAA and the UK Government which would allow NATS to reduce its senior debt from the outset;
- restructuring its debt to introduce greater flexibility for NATS to borrow from a broader range of sources
- the re-instatement of a facility providing funds for NATS' investment plan.
- the introduction of a Standby Facility (provided by NATS' bankers) and a Liquidity Reserve Account (from NATS' internal resources) totaling £37.5 million to be drawn upon in the event of financial pressures;
- allowing NATS to raise additional borrowings subject to satisfying normal commercial criteria, thus increasing its financial flexibility;
- a dilution of the relative shareholding of the Airline Group reflecting the new equity investment by BAA and the UK Government and a restriction on dividend distributions by NERL until 2008 to retain cash within NERL to fund debt service obligations and its investment plan.

(a) *"NATS' Application to Re-open the EUROCONTROL Charge Control",* CAA Decision, March 2003, page viii
(b) *"NATS' Application to Re-open the EUROCONTROL Charge Control",* CAA Decision, March 2003, page 8

5.4 Development of a Revenue Recovery Imbalance Account

The above discussion has indicated a number of approaches that have been adopted to charge smoothing, and highlighted a number of key concerns and issues that arise in relation to those approaches. This section provides a potential way forward for the

consideration of these issues. The aim is not to seek to present a 'solution', but rather to provide a framework within which alternative approaches could be assessed and potentially implemented.

A key feature of this framework is that it highlights ways in which different degrees of financial flexibility could be allowed for, identifies specific constraints that could be put in place to limit the extent of this flexibility, and provides for a set of arrangements within which different implementation options could be adopted in a mutually compatible manner. As is noted below, the current adjustment mechanism can be understood as a specific case within this framework.

The central feature of this approach involves the development of a Revenue Recovery Imbalance Account (RRIA) for each ANSP. The RRIA would record the imbalance for a given ANSP between the revenue that it was allowed to recover from users, and the revenue that it has actually recovered from users. In many respects, this is a very familiar concept in relation to route charging, as the CRCO routinely calculates this difference for a given year as part of the existing over/under-recovery arrangements. An important weakness of the current regime (when applied in the standard manner), however, is that any deficit or surplus in year *n* is automatically adjusted for in year *n+2* irrespective of circumstances at that time.

An alternative approach would be to allow some flexibility in the adjustment process, such that instead of effectively having separate accounts for each year that have to be 'balanced' after a fixed delay period (i.e. in n+2), an ongoing account was introduced with the 'imbalance' position on that account recording the cumulative position over time. It is notable that an approach of this kind could be compatible with any set of arrangements that generates an 'allowable' revenue for each year that can be compared with actual revenue. That is, it is compatible with a full cost recovery system, and with any alternative mechanisms that result in an allowable revenue figure being generated for each year (as opposed to a purely price-based approaches that do not provide revenue commitments).

The precise form of arrangements would depend on the approach taken to a number of key issues. In particular:

- what rules should govern the generation of imbalances?
- what rules should be in place to govern the timing of disbursement/recovery of positive and negative imbalances?
- what limits should be placed on the level of the aggregate imbalance that can be generated or carried forward from one year to the next?
- whether separate limits should be placed on the allowable level of change in the imbalance position from one year to the next?
- whether or not a symmetrical approach should be adopted to positive and negative imbalances
- what procedures should operate in situations where limits are exceeded?
- what conditions should be placed on the eligibility of ANSPs to operate an imbalance account with specified properties?.

A key aim of a framework of this kind is to seek to focus attention on the range of relevant dimensions that can make up any specific approach, and thus the range of alternative combinations that can be considered. This can be an important, as it is not obviously the case that the view that might be taken on, say, the development of reserves, would be the same given a different combination of other conditions. However, it need not be the case that a single uniform approach would have to be adopted. Rather, it could be envisaged that a menu of options could made available. Indeed the current arrangements could be described in terms of the above as a specific option where the generation of imbalances can only result from differences between the forecast and actual number of service units, a disbursement/recovery rule is clearly defined and fixed, there are no limits on imbalances (and thus no issue of exceeding such limits), and the only conditions placed on ANSPs are those that are generally applicable under the route charging system.

The above list indicates some central factors that can be varied in order to develop potential options to the current system. The following considers some potential options and issues.

(a) The Generation of Imbalances

An important first distinction concerns the flexibility that ANSPs have in terms of the generation of imbalances. It is useful here to differentiate between two sources of imbalance:

- deviations between the actual number of service units for a given year and the number forecast in the charge setting process for that year;
- deviations between the expected level of revenue recovery (given the forecast of the number of service units) and the allowable level of revenue recovery;

This provides a basis to distinguish between unplanned and planned deviations. This is useful as it allows for limits to be set for planned deviations. Thus one approach could be to allow ANSPs flexibility to plan for a lower or higher expected level of revenue recovery than their allowable level of revenue recovery for the coming year, within explicit limits on the maximum expected levels of over/under-recovery that can be planned for in a given year. Since the desirability of planned over/under-recoveries will depend on the opening imbalance position of an ANSP (that is, the cumulative sum of past imbalances) then it may also be desirable to have an additional limit introduced that restricted ANSP flexibility such that the effect of a planned over/under-recovery could not be to take the cumulative imbalance position outside of pre-defined boundary levels. That is, pre-determined limits could be put both on the magnitude of expected imbalance that can be planned for in a given year, and on the expected imbalance position that related to any planned for deviations.

A difficulty that arises when making the distinction between planned and unplanned deviations, however, is that to some extent, by varying the forecast number of service units, planned deviations can be made to appear to be unplanned. This could be addressed by either ensuring the use of third party forecasts of the number of service units (CRCO provide the forecast for a number of States), or perhaps requiring some

reporting process to explain situations where an ANSP's forecast differ from CRCO forecasts by more than a defined margin.

(b) The Disbursement/Recovery of Imbalances

A similar approach as described in terms of the planned generation of imbalances – that is, of ANSP flexibility within defined limits – could be adopted in relation to questions concerning what rules should govern disbursement/recovery of imbalances. Thus, explicit limits could be put on the imbalance position that could be carried forward from one year to the next[83]. Such an approach would essentially introduce a disbursement/recovery rule for any cumulative imbalance above defined levels, such that the excess would have to be adjusted for in the following year's charges.

An alternative, less mechanistic approach, could involve the setting of specific levels of imbalance position that trigger process requirements. Thus, for example, this could allow for a specific case to be made for not making the mechanistic correction given prevailing circumstances.

(c) Conditions for Flexibility

An important factor in terms of the potential for allowing flexibility of the kind described above concerns the robustness of arrangements over time. In particular, it is important that users and providers consider that imbalances will over time be adjusted for appropriately, and that flexibility in the management of imbalances will not influence incentives in relation to underlying costs (for example, in a situation where ANSPs have generated surpluses). Key issues here are the regulatory framework that the ANSP operates within, and the form of contractual arrangements that the ANSP is subject to in relation to revenue recovery and performance.

Regulatory scrutiny can provide an important source of credibility in relation to the monitoring of imbalances, and the underlying funding position. More generally, process requirements in relation to the generation and disbursement of imbalances, and the management of the imbalance positions, can allow for third party scrutiny. For example, flexibility of the kind discussed above could be made conditional on ANSPs consulting on proposed approaches, and these consultation processes including clear statements of the financial position that underpins a proposal.

The desirability of allowing different levels of ANSP flexibility may depend on the extent to which ANSPs are subject to credible commitments in terms of performance more generally. Thus, it may be considered appropriate to make the allowance of flexibility conditional on cost and output commitments having been formalised either in terms of the route charging system, or a national regulator[84].

[83] Clearly these limits would need to be consistent with any limits put in place in relation to the planned generation of imbalances.

[84] This refers to potential approaches for contractualisation of cost and service commitments discussed in Section 4.

It should be noted that the approach set out above could be unbundled in a number of ways, such that, for example, flexibility (within defined limits) in relation to the disbursement/recovery of imbalances could be implemented without a similar form of flexibility being introduced in relation to the generation of planned imbalances. Different approaches to the scope of flexibility allowed for, and the size of limits set, could justify different approaches in terms of conditions that may be appropriate to put in place.

5.5 Summary

This section has assessed a number of issues concerning traffic volatility and risk sharing in ATM, including the extent to which the risks associated with traffic volatility might be partly controllable by service providers, and potential arrangements for financing charge smoothing. Key points from this assessment include:

- There may good efficiency reasons for considering some degree of smoothing of user charges as desirable - in particular, the profiling of cost recovery may allow for a higher proportion of fixed costs to be recovered in periods when demand is less price sensitive and vice versa.

- We are skeptical of the efficiency benefits of funds (whether they be specified 'solidarity' funds or more general service provider reserves) where users are asked to pay in advance without receiving any firm and credible commitments in relation to what benefits they will receive in the future.

- Some ANSP exposure to traffic level changes in terms of revenue recovery allowances may be desirable in order to encourage flexible responses to changes in relevant circumstances

- The potential for charge smoothing arrangements to weaken service provider incentives is an important factor to consider - that is, the extent to which smoothing arrangements can insulate ANSPs from desirable pressures that might otherwise be present to a greater extent.

- The above points strongly suggest that discussions of potential charge smoothing arrangements should take place as part of the consideration of the general approach to charging to be adopted, since the resulting incidence of risk will be likely to be heavily dependent on this general approach.

We have proposed that the development of a 'Revenue Recovery Imbalance Account' could provide a framework within which alternative approaches to charge smoothing could be assessed and potentially implemented in a compatible manner. The aim of using such a framework is to seek to highlight the different dimensions – including in particular the range and extent of flexibility that is allowed for, and the constraints that are put on the usage of that flexibility - that impact on the desirability or otherwise of alternative approaches. The desirability of a particular course of action will depend on the specific combination of choices made in relation to these dimensions, and the trade-

offs underpinning them. The points made in the above bullets highlight our views on a number of issues that are of central relevance to assessing these trade-offs. The RRIA provides a framework for assessing these issues as part of more general charging questions, and is compatible with the framework presented in the next section.

SECTION 6
A GENERAL CHARGING FRAMEWORK TO DETERMINE RISKS AND REWARDS IN ATM

In this section we develop a general charging framework that brings two key strands of analysis – the approach to charging and the distribution of risks – together and that can allow for a considerable degree of flexibility in determining the most appropriate outcomes, taking account of both the allocation of risk and 'encouragement' factors. In the terminology of Section 3, the framework can reasonably be described as a 'hybrid' approach, and it has the following characteristics:

- It provides for the sharing of benefits from good performance, and the disbenefits from poor performance, between ANSPs and airlines.
- Simultaneously, it provides for the sharing of risks between service providers and airlines.
- It provides encouragement for improved performance.
- The precise mix of encouragement and risk sharing is flexible, being determined by choice of the relevant parameters. A variety of preferences on this matter, which might reflect the views and circumstances of national regulators, can be encompassed by the same, general formula.
- It allows for the profiling of financial revenues and payments over time.
- Traditional cost-of-service regulation (cost plus) and price cap regulation (RPI - X) are special limiting cases.
- It can be extended to incorporate considerations of solidarity among service providers.
- It can be extended to incorporate factors such as delays and capacity provision, provided the latter can be suitably measured.
- Similar approaches have been implemented successfully in other contexts.[85]
- A practical example of a hybrid approach in ATM is the case of Airways New Zealand. Under this scheme, a portion of any profit above a defined target level is shared with airlines on an agreed basis.

For simplicity, the approach will be referred to as "benefit sharing". Before setting out this framework in detail, we consider a number of more general issues that arise when

[85] Approaches of this kind have been implemented in US telecommunications and electricity transmission in the UK.

seeking to specify which variables could most appropriately be incorporated into incentive formulae.

6.1 Developing a benefit sharing framework for ATM: some general issues

6.1.1 Service Quality/Safety Issues

As emphasised in section 3, when incentives to reduce costs are introduced, it is important to recognise the potential for unwanted effects to be generated. In particular, costs can be reduced not only by becoming more efficient, but also by, in effect, offering lower quality of services to customers. In practice, therefore, implementations of incentive regulation tend to pay close attention to quality of service issues.

More generally, a focus on the definition of quality of service issues – that is, what level of service is actually delivered to users - in conjunction with a consideration of charge levels – what price is charged for a given level of service – is central to processes of contractualisation. Central factors from a user perspective include the price/quality bundles that are actually delivered, and the profiles of alternative bundles that are available at specific points in time. This can be of particular importance in relation to new investment issues – which are of considerable importance in ATM - where a clear specification of the costs and timing of outputs to be delivered can allow for explicit recognition of new capacity commitments within an incentive contract framework. As was discussed in section 3, such approaches can provide for clarity over future cash flows (and in doing so facilitate private financing arrangements), whilst at the same time guarding against concerns of pre-financing. Also, by focusing attention on 'outputs', and linking financial rewards to their delivery, such approaches can also help to focus attention on alternative and potentially more efficient ways in which these outputs can be delivered (including through restructuring and the development of FABs).

(a) Safety

Consistent with the above comments, in practice implementations of 'high-powered' incentive arrangements such as price cap regulation tend to pay close attention to quality of service issues and, speaking broadly, such approaches are only likely to be attractive where quality of service can be clearly defined and monitored. This is one of the reasons why a pure form of price-cap regulation is problematic for ATS. Safety is clearly a highly important, though by no means the only, dimension of quality of service, and it can be difficult to monitor. For example, it might be difficult to detect, *ex ante*, changes in arrangements that led to small, but potentially very harmful, degradations in safety. The PRC's comments on the absence of adequate performance indicators for safety at the European level is notable in this respect[86]. In particular, the PRC note that insufficient compliance of states with EUROCONTROL safety reporting standards is a major obstacle to measuring matching indicators, identifying key risk areas and managing safety. We would note that, as with reporting processes more generally, proper compliance with safety reporting processes can be made a condition that is

[86] See PRR 6 Version II, EUROCONTROL, May 2003

linked explicitly with the allowance of specified forms of charging activity. This could involve linking process compliance to the eligibility of charges or charge increases (consistent with the discussion of information disclosure above), or alternatively to the eligibility of incentive contract approaches to charging.

(b) Delays and Capacity

Unsurprisingly, discussions of service quality in relation to ATM are typically heavily focused on delays. After safety – which is taken to be a necessary condition – delays are the primary focus in considerations of ATM performance and quality of service provision. However, when seeking to assess the performance of specific providers, it is important to try to identify those aspects of service provision which are – at least to some extent - within the control of that provider over any given time period.

When considering delays, immediate issues arise concerning the extent to which the cause of a given delay can be adequately isolated, and the extent to which focusing regulatory attention in general, and financial incentives in particular, on an identified delay category might give rise to unwanted distortions elsewhere – for example, if *en-route* delays can be passed between control areas (for example, from one ANSP to another, or between *en-route* and terminal).

Key issues here include the potential negative network effects associated with particular behaviour, but also the potentially wasted resources associated with seeking to manage particular indicators at the expense of core service provision tasks. A related issue concerns the impact of movements in traffic levels on delay statistics, and the potential for this to distort the effectiveness of delay indicators as measures of ANSP performance.

Whilst delays are clearly an important measure of the quality of service experienced by users (and passengers) when using the network, their relationship to the performance of specific ANSPs is more difficult to interpret. When seeking to assess ANSP performance, a useful distinction can be drawn between delivered 'outputs' and actual 'outcomes'. Of course, the value of a given level of an 'output' being delivered will vary depending on actual conditions, and – as emphasised in relation to cost levels – the flexibility with which available resources are used to provide for actual fluctuations is an important facet of the quality of service provision.

(c) Baseline service standards

One approach that has been developed in order to seek to address this type of situation is the use of baseline commitments, with incentives focused on deviations from these commitments, with, in some cases, the impact of the deviation taken into account when determining the resulting level of incentive payment/cost.

Thus, if more than the baseline level of a given output was provided and this resulted in better 'system' performance than would otherwise have prevailed, then this additional provision is rewarded under the incentive arrangements. If the additional provision of the output did not enhance system performance, then it would not be rewarded.

Similarly, provision of a lower level of capacity than the agreed baseline level would only give rise to a loss under the incentive arrangements where this resulted in a deterioration of system performance compared with what could have been expected under the baseline case. It is notable that in situations where there is a case for specific short-term capacity reductions, this case could be explicitly assessed up-front, and associated reductions could be explicitly taken into account in determining the baseline capacity profile so as not to result in an unduly negative assessment in terms of performance assessment and any associated incentive arrangements. Such adjustments may be of relevance where it can be demonstrated that some short-term reduction in capacity levels is an efficient part of a restructuring process.

For ATM, this form of approach could take a number of different forms. At present the PRC advises the EUROCONTROL Provisional Council on average system delay targets, and as a result of this process a Key Performance Indicator is agreed upon for ATFM delays. EUROCONTROL translates delay target into capacity targets for individual ACCs, using the FAP methodology. ANSPs then indicate capacity increase commitments for each of their ACCs, which are recorded in Local Convergence and Implementation Plans (LCIPs). These commitments may or may not equal the EUROCONTROL targets, and remain indicative and non-binding.

As was highlighted above, the consideration of cost issues in the generation of the EUROCONTROL targets is currently very limited, and an important first step that appears desirable whether or not specific incentive arrangements were to be developed on this basis, would be the specification of these capacity commitments through a process that involved the simultaneous consideration of cost forecasts.

Nevertheless, the current EUROCONTROL capacity planning process provides a potential framework within which committed baseline capacity levels could be linked to some form of incentive arrangements. Where medium term capacity enhancement plans were determined, these could allow flexibility in terms of the manner in which capacity was provided, but provide for an incentivised commitment to deliver the agreed outputs by a given point in time.

If capacity provision were used as an ‘output’ in incentive arrangements, it would be important to determine the time periods over which committed levels of capacity should apply, and what the relevant area that they would apply to should be. In terms of temporal variations, since traffic flows can vary significantly within a day, between week days and week-ends, and throughout the year, it is important to determine whether a single baseline level of capacity is desirable or whether some form of capacity provision ‘profile’ that recognised these factors would generate better incentive effects. A profile of this kind could be developed on the basis of recent experience (for example, the previous year flows and capacity levels declared to CFMU) adjusted for special events (etc) and the expected impact of any agreed increase in the baseline level.

It is notable that ACCs will typically provide for a set of different capacity provision levels in relation to a given sector reflecting expected variations in traffic over time. The demand for capacity can therefore exceed available supply either because it exceeds the peak available capacity level (i.e. the highest of the capacity provision levels), or,

despite the fact that it is lower than the peak level, because it exceed the actual capacity level made available at a given point in time. The relative importance of these situations as sources of delay is clearly an empirical matter, and one that could be of some importance in determining the most appropriate development in the form of output commitments to be given over time.

Currently the EUROCONTROL capacity planning process defines a single peak level of capacity as a target level for each ACC. In addition to the temporal issues highlighted above, this raises potential issues concerning the extent to which ACC's represent the most appropriate basis upon which to define capacity commitments, compared, in particular, with a more operational focus on sectors. An important issue here is that the control area of a specific ACC may include sectors/control areas that vary considerably in terms of the significance of capacity constraints. If an ACC-based measure of capacity were used as part of a set of incentive arrangements, this could potentially give rise to incentives to focus on areas where capacity delivery was least expensive, which may not coincide with identified scarcities.

Whilst these temporal and spatial factors clearly raise potential issues that, (unsurprisingly) indicate that care must be taken when seeking to develop capacity-based output measures, they do not present an insurmountable barrier. They do indicate that a relatively cautious approach is likely to be desirable as a starting point, in terms of the financial stakes that should be linked to such scheme. However, this is typical of most new incentive arrangements, with the financial exposures generated by a given scheme typically growing over time as more confidence develops in the measures used, and appropriate modifications have been made to address identified weaknesses.

The comments above also indicate information disclosure arrangements and consultation processes of the kind discussed earlier in this section are likely to be of considerable ongoing importance even where the information disclosed does not form a specified part of an agreement. This, again, is typical of incentive arrangements. That is, whilst the incentive contract itself is typically defined in terms of set of formulae with specified variables, the monitoring of performance by regulators and third parties takes account of a considerably wider set of variables. A particularly important factor here is that a specific incentive contract represents but one part of a broader and ongoing relationship between service providers and users, and this can have a significant dampening effect on incentives to engage in opportunistic behaviour that seeks to 'cash-in' on the limitations of the explicit contract that are understood as violating less explicit understandings of appropriate conduct. That is, there can be significant reputational consequences of such conduct.

These points indicate that whilst there are likely to significant limitations in terms of the usefulness – in terms of developing incentive mechanisms - of the planned capacity enhancement numbers that emerge from the LCIP process, they may provide a reasonable starting point. More particularly, if equivalent capacity figures were developed in the context of more integrated processes of consultation with enhanced levels of user involvement, an incentive formula that focused on these peak capacity numbers could provide a focal point around which broader assessments of performance

could be based, with a view to greater degrees of formalization over time that could seek to incorporate other relevant factors such as flight efficiency[87].

An approach of this kind provides for a baseline set of capacity provision commitments, but does not, of itself, provide for desirable incentives to respond flexibly to variations in actual traffic levels and patterns. A potential way of linking a capacity-based approach to actual traffic variations could involve focusing more directly on delays. In particular, incentive arrangements could be linked to the difference between actual delays allocated to a given ANSP, and the level of delays that would have been expected given the actual level of traffic and committed baseline levels of capacity. Such an approach would have the advantage that it presents an assessment of ANSP outputs in relation to the primary quality variable (after safety) that is typically of concern to users (i.e. in terms of actual delays). The analysis of the difference between levels of actual and expected delays presented in the PRC reports PRR5 and PRR6 would appear to provide a useful basis for such an approach.

6.1.2 Incentives for Cost Efficiency

As a general presumption, best practice regulation indicates that relatively unfocused approaches to incentives for cost efficiency are likely to be desirable other than in situations where there is a relatively high degree of confidence in relation to the robustness of relevant information flows. Two factors are of particular importance here. Firstly, detailed specifications can generate incentives for gaming, with for example, resources expended on seeking to reclassify cost components from more onerously to less onerously regulated categories. Secondly, there may be significant asymmetries between the position of service providers and third parties (including regulators and users) in terms of both access to existing information, and discovery of new relevant information. The importance of these factors can change considerably over time, and confidence in the robustness of specific measures can underpin the development of separate or sub-schemes related to particular components. However, given the current position in relation to information flows in ATM, an unfocused approach would seem most appropriate at this stage.

Such an approach could be implemented by means of a control focused on average or total 'allowable' revenue. In the review of price regulation in Appendix 1, the relationship between the recovery of revenue and the level of demand/output was discussed. Specifically, it was noted that if the level of *total* revenue that a service provider is allowed to recover is fixed then system users carry most of the risk associated with fluctuations in demand/output, whereas if the level of *average* revenue is capped this exposes service providers to greater risk from fluctuations in demand/output and may lead to them being unable to cover their fixed costs if demand falls to less than expected levels.

The distinction between revenue requirements set on the basis of the level of *total* revenue or *average* revenue therefore has obvious implications for the allocation of the risks associated with traffic volatility between service providers and system users in

[87] Issues of flight efficiency were considered in the PRC's most recent performance review (PRR6), but are a relatively early stage of development compared with capacity/delay issues.

ATM. However, in practice, the resulting level of risk sharing under either a *total* or an *average* revenue approach can be modified by the introduction of a specific adjustment factor that adjusts the level of allowable revenue in response to deviations between actual and forecast traffic levels. It was argued in Section 5.2 that some degree of ANSP exposure to traffic level variations was likely to be desirable. The approach taken below is to base the principal formula on unit cost levels – that is, on the level of average revenues – and then to use a traffic volume adjustment factor to 'unwind' some of the impact of traffic volumne movements on ANSP allowable revenues (with the amount of unwinding determined by the decision taken on the value of the relevant parameter)[88].

Given the structure of costs in ATM provision, with for example a high percentage of ANSP costs accounted for by labour costs, the use of a consumer price index to adjust allowable revenue levels provides a relatively straightforward and well understood way of tracking generalized and uncontrollable increases in prices. For greater clarity, the formulae presented below do no include an inflation adjustment mechanism (although adding such an adjustment mechanism would be a straightforward matter).

6.2 A general benefit sharing framework

6.2.1 The basic formula

As is generally the case, the ultimate aim is to determine an allowable revenue (RA) for the year, which is translated into a unit rate by dividing by an estimated traffic volume.

Let UC be the relevant level of unit costs in a given year, UC* be the target or benchmark level of unit costs (the determination of this latter number will be considered later), and UCA be the 'allowable' unit costs for the year. Allowable costs -- CA = UCA.X, where X is the number of service units -- are those costs that the ANSP will be able to recover through charges, either immediately or at some point in the future. The difference between allowable costs and actual costs (C = UC.X), which may be negative (allowable costs may be below actual costs), can therefore be interpreted as a performance 'benefit', B, that accrues to the ANSP.

Define UCA as follows:

$$UCA = \min \{UC^* + \alpha [UC - UC^*]; UC - F\} \text{ when } UC > UC^*$$

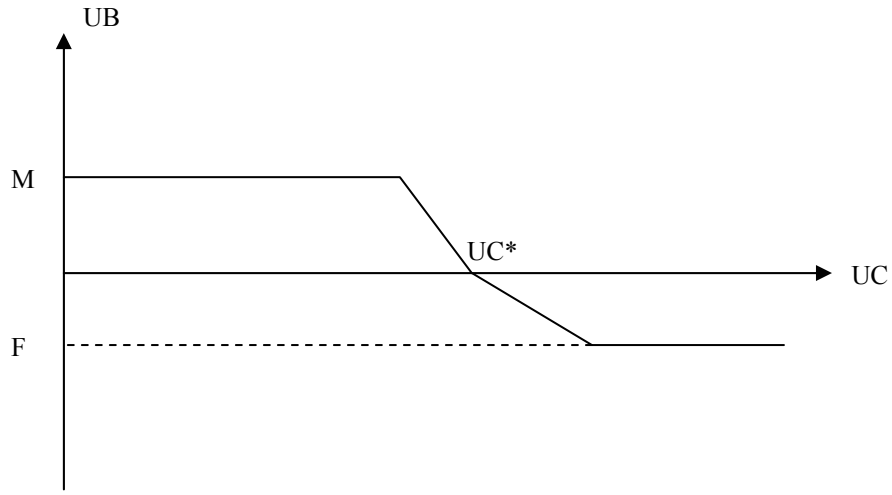
$$UCA = \max \{UC^* + \beta [UC - UC^*]; UC + M\} \text{ when } UC < UC^*.$$

[88] As indicated in the above discussion, an alternative approach that could generate similar outcomes would be to define the principle formula in terms of a *total* revenue allowance, and to use a traffic volume adjustment factor instead to introduce a degree of exposure to deviations between forecast and actual traffic levels in terms of ANSP allowable revenues. As above, the degree of exposure would be determined by the choice of parameter values.

This looks complicated, but the relationships are, in fact, relatively simple, as can be seen by charting the link between ANSP financial benefit per service unit (UB = UCA - UC) and actual unit costs.

As can be seen, allowable unit costs are equal to actual costs (benefit is zero) if actual unit cost exactly hits the benchmark level.

If the ANSP does better, and costs are below the target level, the ANSP gets to keep a fraction β, up to a maximum of M.



If the ANSP does worse, and costs are above the target/benchmark level, the ANSP absorbs a fraction α of the performance difference, up to a maximum of F (the 'floor' on the ANSP's financial exposure per service unit).

Thus:

- The (non-negative) parameters α and β determine the rewards and penalties of, and hence the risks borne by, the ANSP for better or worse than benchmark performance.
- The lower are the values of these parameters the greater the incentives and the greater the risks falling on the ANSP in relation to cost performance.
- The two parameters need not be the same. For example, if α > β (as shown in the diagram) the upside rewards and risk are greater for the ANSP than their downside counterparts.

- The ceiling/maximum and floor parameters place bounds on the financial exposure of ANSPs, thus providing extra protection against extreme events. By the same token, the risks arising from extreme events fall more on airlines.

- Cost-of-service regulation is a special case, in which either M = F = 0, or $\alpha = \beta = 1$ (allowable unit costs are then equal to actual unit costs).

- Price cap regulation is a special case in which $\alpha = \beta = 0$, and M and F are both large.

The approach is capable of further refinement, for example by creating a 'deadband' around the target/benchmark, such that all unit cost outcomes within this range are fully recoverable. Such refinements are not pursued here, since they are only of marginal interest in the present context.

The important point to note about the approach is that, within a common framework, different mixes of encouragement for improved performance and risk sharing can be implemented. Where the approach has been used, regulators and regulatees have, when first moving away from full cost pass-through, tended to take a cautious view of the trade-offs, and then become bolder as experience is gathered and confidence in the arrangements grows. Thus, when first moving away from cost of service regulation, values of α and β have been set relatively close to one, and small maximum and floor levels of financial exposure have been selected. Later α and β have been reduced, and M and F have been increased.

6.2.2 Setting the target/benchmark level of costs

Normal regulatory practice is to make determinations of the relevant target/benchmark level of costs at periodic reviews, which is somewhat different from the semi-automatic cost pass-through mechanism currently used in ATM. It is, however, possible to move away from simple cost pass-through whilst still maintaining a formulaic approach to the setting of benchmarks, as will now be illustrated. The key concept is that of regulatory or contract lag, as discussed in Section 3.

A simple mechanism that fulfills this purpose is:

$$UC^*(t) = UC^*(t-n) + \gamma\,[UC(t-n) - UC^*(t-n)],$$

where t-n is the latest year for which reliable cost information is available. The lower the value of γ, the slower the rate at which target/benchmark unit cost tracks toward actual unit cost (the longer is the regulatory/contract lag). When $\gamma = 1$, target/benchmark costs are simply set at the last recorded, actual unit cost level, which maps across into the current ATM position.

The only loose end is that some initial target benchmark unit cost has to be set (thereafter the target/benchmark is adjusted automatically).

6.2.3 Descriptive outline

The charging arrangements set out above work as follows. Suppose an ANSP manages to achieve a reduction in unit costs to below the relevant benchmark level. The ANSP will get to keep a fraction of the difference, up to some predefined maximum. The rest will be passed through into lower charges (via a process such as that described in 6.2.4 below). Future cost targets/benchmarks will, however, be reduced somewhat, although by an amount less than the level of out-performance recorded in year t. If, in future years, the ANSP achieves no further cost reductions, it will nevertheless tend to continue to out-perform relative to the target/benchmark, but by smaller and smaller amounts. In this way, the target/benchmark unit cost will gradually track toward actual unit cost, at a rate determined by the parameter γ, which is a parameter that can be set by regulation. Unit rates will likewise fall over time, toward the lower unit cost.

A similar process occurs in the reverse direction. Target/benchmark unit costs and unit rates will track upwards, over time, to a higher, actual cost level. In both cases, it can be noted that the adjustment of unit rates will tend to be smoothed.

6.2.4 Revenue imbalances

In any given year, actual revenues can be expected to differ from allowed revenues, giving rise to revenue *imbalances*, an issue that was addressed in Section 5.4 above. Revenues may be higher than allowable revenues, or they may be lower. The question is: how are such over-recoveries and under-recoveries to be handled?

As suggested in Section 5.4, there is merit in smoothing out the adjustments that are associated with this factor. The following provides a formulaic approach that is consistent with that framework, and that allows for different views to be taken on the desired level of smoothing. If RI(t-n) is the accumulated revenue imbalance, compounded at an appropriate interest rate, in the latest year (before year t) at which the numbers can be accurately estimated, and if RA(t) = UCA(t).X(t) is the allowed revenue in year t, the following mechanism would complete the circle:

$$RA(t) = CA(t) - \lambda RI(t-n)\,[1 + r]^n,$$

where r is the interest rate. That is, a designated fraction of the cumulated imbalance is passed through to users in the year.

Again, the parameter λ, which can be determined as a matter of policy, is a smoothing factor: the lower its value the more gradual the rate at which imbalances are passed through into unit rates.

6.2.5 Setting the unit rate

Since the allowed revenue for year t is only calculable *ex post* – it depends on the number of service units supplied in year t – it cannot be used directly to set the unit rate for the year. It is, however, possible to forecast its value, on the basis of forecast unit allowable costs UCAF(t) and forecast service units XF(t). That is:

$$ARF(t) = UCAF(t).XF(t) - \lambda RI(t-n)\,[1 + r]^n.$$

The unit allowable cost forecast would be based on latest available cost information, taking account of the adjustment process by which unit allowable costs are determined, as specified above. The service unit forecast could be generated as under the current arrangements.

The unit rate in year t would then be:

$$UR(t) = ARF(t)/XF(t).$$

6.2.6 Adjusting for short-run traffic volatility

As discussed earlier, one of the key problem areas in ATM is that, in the short-run (e.g. within a year), a substantial fraction of service providers' costs can be expected to be fixed. Hence, within relatively short periods, unit costs tend to fall significantly when traffic volume rises faster than is projected, and to increase significantly when traffic volume falls relative to projections. Since most of the short-run volatility in traffic volume is beyond the control of service providers, it can be argued that the approach described above would place excessive risks on service providers, since the cost targets/benchmarks would be particularly stringent in periods of depressed demand.

To the extent that this is an issue, it can be addressed by means of a 'traffic volume adjustment factor' in the determination of allowable revenues.[89] For example, the adjustment factor in year t could be defined as:

$$TAVF(t) = v.[X^*(t) - X(t)],$$

where $X^*(t)$ is a benchmark traffic volume for the year, set by regulation or by agreement. $X^*(t)$ could be based upon agreed, trend projections of traffic volume. The TAVF will be positive when traffic volumes are below benchmark/trend levels, and negative when traffic volumes are above benchmark/trend levels.

Allowable revenue in year t is now defined as:

$$RA(t) = CA(t) + TAVF(t) - \lambda RI(t-n)\,[1 + r]^n,$$

and forecast allowable revenue can be adjusted accordingly, taking account of the forecast traffic adjustment volume factor.

The parameter v, which again would need to be determined by regulation or by agreement, can be interpreted as the unit, short-run fixed cost, at the benchmark traffic volume, that the service provider will be allowed to recover in a given year. If the great bulk of short-run costs are deemed to be fixed, it might be set at a level a little below unit costs calculated at the benchmark traffic volume. Lower values of v have the effect

[89] The underlying economics here is similar to the disaggregation of allowed revenue into a total revenue allowance and an average revenue allowance, as described in Appendix 1.

of placing more of the shorter-term risk associated with traffic volume volatility on providers and less on users.

6.2.7 Extensions

The benefit sharing approach can be developed in a number of ways. Whilst omitting detailed discussion, we note some of the possibilities:

- Target/benchmark unit costs could be set on a 'yardstick' basis. That is, rather than adjusting the target/benchmark for a particular service provider simply on the basis of that provider's own cost performance, the target/benchmark could be set on the basis of cost performance across the network.

- The benefit sharing approach can be used to determine the allowable revenues of system operators/network managers. In this case, the equivalent of the unit rate might be a charge per service unit made by the system operator/network manager to other service providers and users.

- Allowable revenues can be adjusted to reflect factors such as delays and capacity provision, provided that the latter can be suitably measured and attributed.

- Allowable costs can be predetermined on the basis of a baseline 'agreement', specifying such matters as capacity to be made available, service standards, etc. The benefit sharing arrangements could then apply to *deviations* from the baseline, and would incorporate downward adjustments in allowable revenues in the event of failure to meet the agreed commitments. A discussion of a potential framework for developing such arrangements was provided earlier in this section

6.3 Summary

In this section we develop a general charging framework that allows for "benefit sharing" and that can allow for a considerable degree of flexibility in determining the most appropriate outcomes, taking account of both the allocation of risk and 'encouragement' factors. Key points include:

- When incentives to reduce costs are introduced, it is important to recognise the potential for unwanted effects to be generated, and in practice, therefore, implementations of incentive regulation tend to pay close attention to quality of service issues – the context of ATM this may refer to effects on safety and delays and capacity. Given the potential for these issues to arise the use of baseline commitments, with incentives focused on deviations from these commitments, have been developed in a number of regulatory contexts.

- The ultimate aim of the general benefit-sharing framework is to determine an allowable revenue for the year, which is translated into a unit rate by dividing by an estimated traffic volume. The approach is capable of further refinement which

allows for different mixes of encouragement factors and for improved performance and risk sharing to be implemented within a common framework.

- Where the approach has been used, regulators and regulatees have, when first moving away from full cost pass-through, tended to take a cautious view of the trade-offs, and then become bolder as experience is gathered and confidence in the arrangements grows.

- In the context of ATM the approach can be refined and developed in a number of ways, including to incorporate an allowance for short-run traffic volatility.

SECTION 7
ATM CO-ORDINATION AND INFRASTRUCTURE MANAGEMENT

This section of the report builds upon the observation noted in section 3 that in the absence of co-ordination, decentralised networks can be prone to both short-term and long-term inefficiencies (e.g. excessive congestion, under-utilised capacity). Of particular relevance here are two of the high-level aims of the Single European Sky initiative: to improve system effectiveness through improving the use of existing capacity, while also ensuring that capacity is developed where needed at different points of the network.

The first part of this section examines potential mechanisms and rules that may promote greater efficiency in Air Traffic Flow Management (ATFM) in the short term – that is, better use of existing capacity. The second part of this section focuses on potential approaches that could promote greater levels of harmonisation in respect of infrastructure and investment decisions across the European ATM network. This is followed by a discussion of the potential financing options for collective projects, including the issues surrounding the establishment of a solidarity mechanism to finance collective infrastructure projects.

7.1 European ATM System Operation

One of the primary objectives of the Single European Sky initiative is to realise the efficiency gains of more harmonised and interoperable European airspace through the more efficient management of traffic flows. The Single European Sky proposal specifically notes that:

> *The Single European Sky aims to enhance air traffic management efficiency, to make better use of capacity and to improve the handling of congestion by creating mechanisms for setting priorities and dealing with crises. As part of this initiative, comprehensive common European rules for scheduling aircraft departures and arrivals will be drawn up and enforced*[90]

This sub-section of the report draws upon studies undertaken in regard to system and congestion management in ATFM, as well as on the experiences of other European network industries. It begins by briefly examining the current ATFM system in Europe and detailing some of its limitations. This is followed by an assessment of the potential for an enhanced, or more active, system operation/network management role at the European level. Finally, we examine the approaches adopted to system operation in the European networks for electricity and rail and explore the potential to apply some of these approaches to the European ATM network.

[90] "A Single European Sky: Broadening horizons for Air Travel" European Commission (2002), page 12

7.1.1 ATFM in Europe

The process of ATFM in Europe currently occurs through a combination of co-operation between the airline operators, ANSPs in the ECAC member states, and Central Flow Management Unit (CFMU) of EUROCONTROL. Since 1995, the CFMU has acted as the central co-ordinator for air traffic flow management in the 36 European Civil Aviation Conference (ECAC) member states.

(a) The ATFM Phases

In operational terms there are three phases involved in the flow management process, which are:

- The *strategic planning* phase which occurs between six months and seven days ahead of operations. The flow management functions undertaken at this phase involve preparing a strategic demand forecast based on an assessment of the number of planned flights. This allows any forecast potential capacity and demand imbalances to be identified and for potential options to be identified.

- The *pre-tactical planning* phase which occurs between a few days and a few hours before take-off and involves actions undertaken to attempt to alleviate any identified potential capacity/demand imbalances. This typically involves the introduction of restrictions on traffic movement through the introduction of regulations in specific sectors.

- The *tactical* phase which refers to the period just before take-off and during a flight. The flow management functions undertaken at this stage typically involve applying the regulations that were developed in the *pre-tactical* phase and real-time monitoring of the flow management of the system.

Each of these phases represents an input into the flow management process and represents a potential opportunity for capacity and demand imbalances to be addressed.

(b) Interaction between ANSPs and CFMU

The primary interaction between the CFMU and each ACC is through a flow management position (FMP), which is located in each ACC and is responsible for flow management co-ordination and interfacing with the CFMU. The interaction between individual ACCs and the CFMU at the different operational phases of the ATFM process primarily occurs through the FMP. At the strategic phase of the process the Flight Data Operations Division of the CFMU produces a set of traffic demand forecasts for each ACC in Europe, based on the expected level of traffic demand, which is transmitted to the FMP in each ACC.

At the pre-tactical stage the FMP is responsible for checking whether the available capacity in the ACC, based on available staff, is sufficient to match the forecast traffic demand previously set in the strategic phase. This information is then passed on to the operational rooms in the individual ACCs who are responsible for organising available

capacity into sector groups. If there is a mismatch between available capacity and expected traffic flow the operational supervisor might at this stage restrict the amount of flow in a particular sector. In some ACCs this process of sectorisation is done in co-operation with the CFMU who may suggest that a specific sectorisation is developed to cope with the daily demand pattern. According to a recent PRC report, however, the role of the CFMU in this process is still weak.[91] Finally, the ACC submits to the CFMU its 'declared capacity', that is the number of aircraft per hour their ATC sectors can safely accept.

Based on the declared capacity and the different sectorisations provided by individual ACCs, the CFMU then develops a regulation plan that includes the series of regulations that are to apply on the following day across the network. A regulation plan can be seen as a series of capacity constraints that is imposed on the network. The primary form of regulation currently imposed in Europe is that of ground delay, although level-capping, tunnelling and advisory re-routings may also be used.

In developing the regulation plan the CFMU incorporate a slot allocation rule which applies a 'first planned–first served' principle which presumes that flights should arrive over the restricted location in the same order they would have arrived had there been no ATFM measures. According to the CFMU the slot allocation rules typically result in a large number of small delays rather than imposing large delays for a small number of flights.

Finally, the regulation plan is then submitted to the ACC's who have the opportunity to request a change to some of the regulations on the day prior to its introduction. The regulation plan is also submitted to the aircraft operators, which in certain circumstances allow the aircraft operators to be involved in the process to address ad hoc problems.

7.1.2 Problems with the current ATFM process

The European Commission has estimated that some 350,000 aircraft flight hours a year are wasted because of a combination of air traffic management delays and airport delays,[92] while the CFMU has estimated that some 60.1% of *en-route* delays in 2002 were caused by insufficient ATC capacity. The most recent PRC assessment of Air Traffic Management for 2002 estimates that the costs of *en-route* ATFM delays to airspace users was between 700-1,000 million euros and to airspace passengers was 900-1,100 million euros.[93]

These findings are consistent with studies such as the *'Independent study for the improvement of ATFM'* (the Jaquard report),[94] and other studies undertaken by different stakeholders including the aircraft operators, aircraft traffic controllers and

[91] "A comparison of performance in selected US and European en-route centres", EUROCONTROL Performance Review Commission, May 2003, page 80
[92] "A Single European Sky: Broadening horizons for Air Travel" European Commission (2002), page 5
[93] "CFMU Annual Summary 2002", Version 1.2, February 2003, page 12; and "An assessment of Air Traffic Management in Europe during the calendar year 2002" PRR 6 Version II, EUROCONTROL Performance Review Commission, May 2003, page 30
[94] "Independent Study for the Improvement of ATFM", Philippe Jaquard (IGACEM), 11 September 2000

EUROCONTROL, all of which see the current ATFM process in Europe as limiting the efficient use of available capacity.

(a) EUROCONTROL Studies

In its special report on air transport delays, the Performance Review Commission (PRC) of EUROCONTROL identified ATFM delays as a major cause of air transport delays in Europe.[95] The PRC report finds that "the present ATFM system is somewhat inflexible and stretched to its limits"[96] and identifies the following limitations of the current ATFM process:

- ATC sectors were overloaded (that is, the level of actual traffic frequently exceeded the level of declared capacity)
- Capacity is wasted due to unused ATFM slots
- The equity criterion – first planned, first served – is not optimal
- Conflicts between airport operations and ATFM slots

In terms of the outcome of these limitations on the ATFM system, the PRR2 report found that some 44% of ATFM delays originated from a demand/capacity mismatch in 30 sectors out of 468 sectors in 1999. The report also found that just a 1% increase in the amount of declared capacities (as compared to sustainable capacities) would reduce delays by approximately 6%, which translates into a saving of 300 million euros per annum. This is of particular relevance, as the expectation of a significant problem of sector overloading can be expected to result in some slack being incorporated into declared capacity levels.

One of the recommendations of the PRC report to the Provisional Council of EUROCONTROL was that the use of existing capacity be improved, and that an independent study be undertaken "into how to optimise the use of existing capacity, and to improve ATFM principles, processes and operation, in order to reduce delay"[97].

As noted above, the PRC study identified unused ATFM slots as a limitation of the current ATFM system. This issue was investigated further by the EUROCONTROL Experimental Centre (EEC) in a separate study.[98] In this study the EEC estimated that recovering just 3.6% of unused slots would result in a reduction of 21% in total delays – where an unused slot is defined as a slot not used by a flight which could have been assigned to an already delayed flight to reduce its delay. The study also found that if unused ATFM slots were recovered then the *maximum* flight delay would be reduced by an estimated 27%.

[95] "Special Performance Review Report on Delays (January – September 1999)" PPR2, EUROCONTROL Performance Review Commission, November 1999
[96] ibid, page 5
[97] ibid, Recommendation 5 page 28
[98] "Analysis of unused ATFM slots 2000", EUROCONTROL Experimental Centre

In a separate study comparing the performance of US and European *en-route* centres the PRC again identified the current ATFM system in Europe as contributing to system inefficiency and delays.[99] Specifically the study noted that:

- In the US, ATFM activities are focused on efficiently feeding airports to maximise their throughput, airport capacity being seen as the scarce resource. Conversely, in Europe the *en-route* capacity constraints are imposed to satisfy the strategically capped traffic at different European airports[100]
- There is only one data system used in the US, whilst different systems are used in Europe which limits the possibility for more dynamic ATFM activities
- In the US, aircraft operators in non-core areas may choose their routes, which 'dilutes' the traffic flows in these areas. In Europe, all flight plans are filed according to published routes.
- The distinction between the three ATFM phases in Europe is not made in the US, with most ATFM actions occurring in what would be the tactical phase in Europe – that is, less than a few hours before the operation.
- In the US there are a variety of possible ATFM procedures used to manage potential capacity problems. The most preferred ATFM action employed in the US is the Miles in Trail (MIT) measure and the least preferred ATFM action in the US is ground delay programmes (currently the most preferred in Europe).
- In the US there is an on-going collaborative decision making process regarding ATFM, and coordination with users is given a high priority
- The study found that the ATFM process in the US is decentralised but is more synchronised than in Europe.

A conclusion of the PRC comparative study was that the ATFM system in the US better utilised airspace capacity than the European ATFM system, which it suggests has impacts for both controller workload and productivity.

(b) Other studies

Aircraft operators have also identified the current European ATFM system as contributing to inefficiency and delays in Europe.[101] IATA argue that it is the *structure* of the current ATFM system in Europe which is at the core of the problem as it allows the CFMU to effectively protect individual ANSPs from overloads, with the costs associated with this protection being borne by the aircraft operators. The IATA action plan proposes that a permanent European wide capacity planning process be introduced to ensure that capacity is added to the network where and when it is needed. IATA also propose that the maximum use of direct routes through free routing should be encouraged in wide areas of airspace (which is consistent with current practice in non-core areas in the US as noted above).

[99] "A comparison of performance in selected US and European en-route centres", EUROCONTROL Performance Review Commission, May 2003

[100] In this regard the CFMU annual report for 2002 notes that the ratio between the Airport delay and the total delay, which remained around 22% during the previous 5 years, increased significantly in 2002 to reach 33.6%

[101] "ATC in Europe: An IATA Action Plan for the New Millenium", International Air Transport Association CO219/9/99

The need for a new approach to flow management is also highlighted as one of the ATC/EUC proposals and they note that "improving flow management techniques will greatly improve the European results as far as departures are concerned" [102]. The ATC/EUC call for new ATFM measures designed on the basis of overall system efficiency, which will allow for a better use of the available capacity. The specific proposals of the EUC to improve flow management include:

- the use of only one statutory flight plan which "should be taken as a mandatory contract between AO's, pilots, Air Traffic Controllers and the CFMU".
- the abolishment of the current "first planned-first served" principle and a move toward a system which would allocate slots on the basis of a "wider global interest framework".
- that the CFMU should be moving toward 'controlling' rather than 'regulating' demand and that it should be able to "propose or impose re-routings and flight level changes should it be useful to ensure an efficient and expeditious treatment of traffic flows. This should be done in co-ordination with FMP, FMD and Aircraft Operators, the latter filing one unique flight plan"

The proposals of the ATC/EUC call for a more global approach to ATFM and that flow management rules be contained in a charter to be developed between all ATC participants. This charter should provide rules for all ATFM related behaviours and the measures that should be introduced in the event of non-compliance.

7.1.3 The potential for a European ATFM Network Manager

The role of network manager (or 'system operator' in energy markets) is common to most network industries such as electricity and gas networks, and the rail sector. Here we focus on that aspect of the role concerned with ensuring that demand is allocated to available supply on the most efficient basis, and with those functions of a system operator that involve the various activities undertaken to manage a network close to real time.

As in other network industries, the role of the system operator – or flow manager in ATM terms – is to safely manage and co-ordinate the European ATM network and facilitate greater interoperability between individual ANSPs. Specifically, the objective of the ATFM network manager is to balance demand and capacity to protect ATC sectors from overload so as to keep delays to a minimum and to avoid congestion in different parts of the network.

As discussed earlier, the system operation role for ATFM in Europe is currently performed through the co-operation of individual ANSPs and the CFMU. An issue that immediately arises in thinking about a dedicated European ATFM network manager involves the specific functions and responsibilities that it could perform and how they might relate to the functions performed by individual ACCs at different points on the network. Put another way, a key question to be addressed when thinking about a

[102] "ATFM improvements", Air Traffic Controllers European Unions Coordination

European network manager involves thinking about how 'active' or 'passive' a role the European network manager could perform.

At one end of the spectrum a dedicated European ATFM network manager can be seen in a relatively *passive* role, where its functions are limited to collecting, formatting and transmitting information regarding capacities and projected flight plans. The European ATFM network manager of this type would be responsible for collecting information and formatting/processing it in a particular way (i.e.: through a model or algorithm) but would not be responsible for 'actively' managing constraints in the network.[103]

At the other end of the spectrum there is potential for the network manager to play a more *active* role. In addition to the responsibilities for collecting and formatting information this might also involve giving the network manager greater responsibility for managing constraints in the European network. In a recent report the PRC identify two effects that an ATFM measure can have on capacity allocation:

> *It may increase global throughput by rerouting aircraft from overloaded to less loaded sectors or by reducing complexity for a given throughput by sequencing the aircraft on a flow. It may also lead to less severe restrictions on traffic if compliance with these restrictions is better, that is, if there are fewer overdeliveries, or if dynamic ATFM actions are available to handle such overdelivieries* [104]

There are obviously many potential forms that an *active* European ATFM network manager might take. For example, within the current ATFM structure, an *active* European ATFM network manager might be responsible for ensuring compliance with submitted flight plans and capacities. This might be achieved through the introduction of an automatic penalty system for breaches of flight plan that is associated with the negative network effects generated, for example, where a lodged flight plan is unused. An active network manager might also have a role in the identification of alternative routings which might relieve network congestion.

Additionally the European ATFM network manager might be responsible for the development of new procedures that allow for greater levels of congestion management in the system. There are potentially both demand and supply side responses which could be considered, and which would empower the network manager to interact with users and providers close to real time to manage the network more efficiently. On the supply side this could involve the introduction of specific procedures to encourage service providers to make available more capacity during congested periods. Alternatively on the demand side this might involve the recognition of the different costs associated with delay by different users. In other words, as occurs in other network industries, the network manager might have the ability to relieve congestion through either 'buying' extra capacity from ANSPs or 'buying' delays from users.

[103] This type of network manager is obviously very closely related, but may not necessarily be identical to, the current functions performed by the CFMU

[104] "A comparison of performance in selected US and European en-route centres", EUROCONTROL Performance Review Commission, May 2003 page 79

As is the case with all issues relating to the European network, the introduction of a more active European network manager will involve some reallocation of roles and responsibilities between national ANSPs and the European level. In particular, the specifics of the relationship between the individual ANSPs and the European network manager, along with an analysis of the magnitude of the network externality effects, would be major factors in the consideration of how active or passive a role a network manager should have at the European level.

7.1.4 Network management/System operation in other European networks

In considering potential ways of addressing the problems with the current ATFM process in Europe identified above, it is useful to consider the approaches to network management/system operation adopted in other networks, such as electricity and rail. A review of the approaches to system operation in these networks is contained in Appendix 2 to this report.

This review of approaches to system operation in the energy and rail sectors suggests that the use of existing capacity of European ATFM may potentially be improved if:

- A 'flow product' relating to the *en-route* slot is created, which could potentially incorporate rights and obligations on behalf of both users and ANSPs.
- The possibility of allowing users to trade in the 'flow product' could be considered in order to efficiently allocate capacity within the system
- Appropriate incentives and/or compliance measures could be introduced that reward system users who comply with their commitments under the 'flow' product, while conversely penalising those system users who are not in compliance.
- The introduction of a 'use-it-or-lose-it' rule for booked capacity might ensure that potential unused slots are reallocated within the system close to real time.
- An ATFM network manager could be empowered to engage in certain system balancing activities close to real time such as the authority to contract with system users for additional capacity (through, for example, entering into the market to re-purchase additional capacity from system users) or alternatively requiring them to compensate constrained system users for the capacity that it cannot provide.
- Congestion charges might provide a mechanism for the more efficient operation of the ATFM network, as may the introduction of charges for those who book system capacity but do not use it

Similarly, if we apply some of the approaches adopted in other networks to congestion management to ATM, it suggest that:

- While the use of capacity curtailment methods for congestion management in ATFM may have the advantage of not requiring a capacity or cost allocation rule, it is unlikely to be the most economically efficient method of congestion management. This finding is to some extent supported by the recent comparison of US and European ATFM performance which highlighted the difference between the US approach to congestion management of using Miles in Trail and the European approach of imposing *ex ante* regulations.

- The use of capacity auctions could potentially promote a more economically efficient system of capacity allocation. However, this needs to be balanced against the increased complexity that auctions may introduce, and issues that may arise in relation to the specific conditions of ATM.
- Capacity redispatch presents an alternative method for congestion management based on operational co-ordination between different system operators to ensure that capacity imbalances are managed as far as possible within a system, so as to only apply minimal pressure on the congested parts of the network.

The above points are intended to be suggestive and potential forward ways of thinking about how the European ATFM process might be improved on the basis of developments that are occurring in other sectors.

7.2 Improving European ATM infrastructure interoperability and development

A key element of the Single European Sky package is to improve the interoperability of the European ATM network through promoting a greater level of harmonization in the infrastructure decisions made by individual member states.[105] In part this is to address the problem that has been outlined by the European Commission as follows:

> *Some national authorities have made vigorous attempts to boost capacity. Unfortunately, these have not always been co-ordinated with neighbors, whose problems they have sometimes exacerbated.*[106]

This is a problem that ATM shares with other European transport networks and which was highlighted in the recent Van Miert High Level Group report which notes that:

> *...the implementation of cross-border projects is hindered by specific factors such as different political agendas, the lack of coordination of administrative procedures on either side of the border and the difficulty to agree on a sufficient amount of public contribution to make projects bankable. The political decision-makers are sometimes inclined to sacrifice cross-border projects for the benefit of national projects.*
>
> *It requires a long-term vision in order to avoid, as is often the case today, short-term decisions on financing infrastructure - according to the political priorities of the day. It also requires a Community vision, on the level of the enlarged Europe, for the planning of major infrastructure.*[107]

[105] See in particular "Amended proposal for a Regulation of the European Parliament and of the Council on the interoperability of the European Air Traffic Management network" - 2001/0237 (COD), 11 March 2003

[106] "A Single European Sky: Broadening horizons for Air Travel" European Commission (2002), page 6

[107] "High Level Group on the Trans-European Transport Network Report", European Commission, 27 June 2003, page 21

In this section we examine potential methods through which greater co-ordination could be developed to promote:

- improved interoperability between existing ATM infrastructure; and
- more efficiently planned and network-focused development of new ATM infrastructure [108]

The first part of this sub-section examines the potential for the development and introduction of a common series of common system wide procedures, mechanisms and standards. This is followed by an assessment of the potential role for a European ATM oversight body or a European infrastructure manager.

7.2.1 Existing European Infrastructure Co-ordination Measures

There are currently a number of initiatives undertaken by EUROCONTROL in co-operation with ANSP's which seek to better coordinate and harmonise ATM infrastructure across Europe. At the core of these initiatives is the European ATM capacity planning process which is a 5 year rolling plan with the objective of coordinating capacity enhancement at the European level. This Europe wide capacity planning process is complemented by the European Convergence and Implementation Programme (ECIP) which outlines and quantifies performance targets that should be achieved throughout the European network in the medium-term.

The medium term performance targets established by ECIP are used to develop Local Convergence and Implementation Plans (LCIP) which identify local measures to enhance capacity in line with ECIP targets for the medium term. Each ANSP details in its LCIP the activities they plan to undertake to enhance capacity for each of their ATC centres. Once agreed, LCIP documents are endorsed by both State authorities and EUROCONTROL.

Alongside the network wide planning measures, individual ANSP's may also produce local medium-term capacity plans that in most cases need to be ultimately co-ordinated with the LCIP measures.

7.2.2 Limitations of Current ATM Infrastructure Co-ordination Measures

In the preamble to the Single European Sky proposal it is noted that:

> *...even though progress has been achieved during the last few years towards seamless operation of the EATMN [European Air Traffic Management Network], the situation still remains unsatisfactory, with a low level of integration between national air traffic management systems and a slow pace in*

[108] Consistent with the use of the term ' infrastructure' in other European networks and the draft regulations for the Single European Sky, the use of the term 'infrastructure' in this section refers not only to ATM equipment and assets but also to traffic management and navigation systems.

> *the introduction of new concepts of operation and technology necessary to deliver the additional required capacity.*[109]

Thus while recognising that programmes such as ECIP/LCIP have resulted in some greater level of integration, the Single European Sky proposal recognises that there are still a number of limitations with these processes.

The first of the limitations is that the ECIP/LCIP processes rely heavily on a 'co-operative planning' process between EUROCONTROL and the individual ANSP's. While this has resulted in "a substantial increase in the capacity of the European Air Traffic Management network" according to the 2002 ECIP status report, the success of these processes are still subject to the discretionary actions of individual ANSP's in implementing suggested capacity harmonisation or improvements. The vulnerability of the ECIP/LCIP processes to local ANSP discretion in implementation is recognised in the 2002 ECIP status report which notes that:

> *the number of pan-European objectives in the ECIP continues to increase….however, the progress made by states in implementing Pan-European objectives was slow when compared to previous years and a number now run the risk of being delayed*[110]

In short, a major limitation of the European infrastructure management process is the lack of sufficient compliance or enforcement mechanisms to ensure that infrastructure improvements or developments are actually undertaken by individual ANSP's. The converse of this problem is the difficulty that some ANSP's experience in reconciling the capacity enhancement forecasts underlying the ECIP/LCIP with their own medium term capacity forecasts for specific ACC centres. It has been noted that the inability to reconcile the 'high-level' forecasts with local forecasts per ACC has lead to some delay or postponement of particular infrastructure improvements measures.

This raises a further limitation of the current infrastructure management process which is that it is unclear who 'owns' or is 'accountable' for the forecasts of capacity which underlie the infrastructure projects. While it is recognised that both representatives from each nation state and EUROCONTROL endorse the LCIP for each ANSP, the question has been raised in practice of who precisely is accountable when the actual capacity does not correspond to the planned capacity forecast included in the LCIP.

A further, and related, limitation of the current infrastructure management process is related to the financing of infrastructure projects. This is a particular issue for those infrastructure projects that are of marginal benefit to a particular ANSP but will have positive European network externalities. Similarly as noted in the preceding discussion, within the current framework there may be reluctance on the part of an individual ANSP to finance a specific infrastructure enhancement in its area on the basis of forecasts which it does not agree with and yet for which it could be held financially accountable.

[109] Point 4 of Preamble to Amended proposal for a Regulation on interoperability - 2001/0237 (COD), 11 March 2003

[110] "ECIP status report 2002", EUROCONTROL (2002), Page (iii)

7.2.3 The development of common European ATM Infrastructure Standards

One potential approach to improve infrastructure management at the European level might be to introduce a more formal and detailed set of rules and standards that could be used to harmonise and integrate existing infrastructure across the network and to assess future infrastructure developments. The development of a common set of European ATM infrastructure standards is consistent with the approach adopted to provide for greater levels of interoperability and harmonisation in the other European networks, particularly the conventional and high speed European rail networks.[111]

Article 3 of the draft Regulation on the interoperability of the European ATM network sets out the form and content of the implementing rules for interoperability.[112] The draft Regulation proposes that the European interoperability standards be drawn up through the cooperation of the European standardisation bodies, EUROCONTROL and Eurocae (a non-profit technical standards body). It is proposed that responsibility for the assessment of conformity with the interoperability standards be undertaken by designated 'notified bodies' in each member state.

This approach does to some extent already occur in European ATM following the introduction of Council Decision 93/465/EEC which sets out a series of common technical specifications that govern the procurement of ATM equipment and systems, however, the Single European Sky proposal seeks to extend the application of these technical standards to impose obligations on all ATM actors.[113]

An important feature of this approach is that it allows for a decentralised decision-making process which could include all stakeholders and provide for the development of a common set of European standards. As is the case in other network industries, the setting of these standards could be overseen by a joint body comprising technical and operational representatives. [114] This approach is also consistent with the Single European Sky package, which notes in the preamble:

> *It is therefore in the interest of all those involved in air traffic management to develop a new partnership approach allowing the balanced involvement of all parties and stimulating creativity and the sharing of knowledge, experience and risks; such partnership should aim at defining, in cooperation with industry, a*

[111] In the case of the European high speed and conventional rail networks the infrastructure is governed by a series of technical standards for interoperability (TSI's). TSI's are defined to be the specifications by which each sub-system or part subsystem is covered in order to meet the essential requirements and ensure the interoperability of the trans-European rail system. The TSI's are drawn up to the order of the European Commission by a joint body representing the national infrastructure managers, the railways companies and the industry . See "Directive 96/48/EC of 23 July 1996 of the European Parliament and of the Council of 19 March 2001 on the interoperability of the trans-European high-speed rail network" OJ L235, 17.9.1996; and "Directive 2001/16/EC of the European Parliament and of the Council of 19 March 2001 on the interoperability of the trans-European conventional rail system" OJ L110/1, 20.4.2001

[112] Amended proposal for a Regulation on interoperability - 2001/0237 (COD), 11 March 2003 , Article 3

[113] Point 18 of Preamble to Amended proposal for a Regulation on ATM interoperability - 2001/0237 (COD), 11 March 2003

[114] It has been suggested that one possibility might be that the Single Sky Committee might perform this function.

coherent set of Community specifications that can fulfill the widest possible range of needs.[115]

However a key question concerning the introduction of common European standards of interoperability is whether this measure on its own will be sufficient to remedy the problems outlined in section 7.2.2 above in relation to the European infrastructure coordination and development process.

7.2.4 A Role for a European Infrastructure Oversight body or Manager ?

One possible method of providing greater levels of harmonisation and interoperability for infrastructure projects might be through the creation of a European infrastructure oversight body or a European infrastructure manager. The key distinction between a European oversight body and a European infrastructure manager lies in the potential roles that they undertake and the resources which they would require.

(a) Infrastructure managers in Railways

The recent Directive of the European Parliament on the railways infrastructure provides some basis for thinking about the roles, remit and financing abilities for infrastructure managers.[116] While it should be stressed that the Railways infrastructure directive does not provide for an infrastructure manager at the European level (but rather at different parts of the network, such as in different member states), the directive is useful in outlining some of the potential functions and tasks that an infrastructure manager in ATM could fulfill at the European level.[117]

Specifically, the functions of the infrastructure manager(s) identified in the Railways directive include:

- At points of congestion on the network the infrastructure manager is required to undertake a capacity analysis study within six months of the identification of the infrastructure as congested,
- Within six months of the 'capacity analysis' study the infrastructure manager must produce a 'capacity enhancement plan' which shall identify: the reasons for congestion; the likely future development of traffic; the constraints on infrastructure development and the options for capacity enhancement. The capacity enhancement

[115] Point 7 of Preamble to Amended proposal for a Regulation on interoperability - 2001/0237 (COD), 11 March 2003

[116] "Directive 2001/14/EC of the European Parliament and of the Council of 26 February 2001 on the allocation of railway infrastructure capacity and the levying of charges for the use of railway infrastructure and safety certification", OJ L75/39, 15.3.2001

[117] Specifically 'infrastructure manager' is defined in Article 2(h) of the Directive to mean: "any body or undertaking that is responsible in particular for establishing and maintaining railway infrastructure. This may also include the management of infrastructure control and safety systems. The functions of the infrastructure manager on a network or part of a network may be allocated to different bodies or undertakings". Note also Article 4(3) which requires that "Infrastructure managers shall cooperate to achieve the efficient operation of train services which cross more than one infrastructure network" and that "They may establish such joint organisations as are appropriate to enable this to take place"

plan shall also include a cost-benefit analysis for each of the possible measures identified, and
- Should the infrastructure manager not 'make progress' with the action plan identified in the capacity enhancement plan then it may be prevented from levying fees for the relevant infrastructure.

Again it should be noted that the Railways infrastructure directive does not refer to the creation of a European wide infrastructure manager.[118] Rather, it focuses on co-operation between different infrastructure managers to enable the "efficient creation and allocation of infrastructure capacity which crosses more than one network".[119] However, the discussion of the role of infrastructure manager(s) in the Directive is useful in identifying potential functions that a European wide infrastructure manager in ATM may be created to undertake. In particular, it is useful in that it provides a role for infrastructure manager(s) in identifying, investigating and consulting on potential infrastructure improvements to the network.

(b) A European Infrastructure Oversight Body

One possible way of addressing the problems outlined in section 7.2.2 above is through the introduction of a European infrastructure oversight body. This type of European infrastructure body would principally be a co-ordinating body that might be responsible for performing the following roles:

- *An analytical function:* collecting, formatting and analysing information at the European level;
- *An identification function:* identifying potential infrastructure projects that may be beneficial to the network as a whole;
- *A compliance function:* ensuring that infrastructure improvements or developments are actually undertaken.

The European infrastructure oversight body might, for example, be responsible for undertaking an analysis of proposals for alternative routing, airspace design and re-sectorisation improvements that would be beneficial to the network as a whole.

The introduction of a European infrastructure oversight body of this type could be understood as being compatible in general terms with the decentralised approach to the development of common interoperability standards discussed above, and will provide a relatively modest institution with well-defined roles. In particular, the function of the European infrastructure oversight body in generating better information flows could be seen as complementing this otherwise decentralised process. In addition, as the European infrastructure oversight body would be focused on exploring, identifying and highlighting co-ordination failures that might provide a constructive basis for addressing these failures within a largely decentralised process.

[118] Note however the "Proposal for a Regulation of the European parliament and of the Council establishing a European Railway Agency" COM (2002) 23 final 23 January 2002 which would be responsible for safety and interoperability.
[119] Article 15(1) of the Directive 2001/14/EC

The improvement in information regarding investments across the network projects that would be collected, analysed assessed by the European infrastructure oversight body could also prove useful for Member states and other economic regulators in assessing the individual ANSPs investment allowance as part of the rate setting process outlined in the previous chapter. That is, the investment information produced by the European infrastructure oversight body could provide a useful comparative benchmark for assessing the level and efficiency of investment spending of individual ANSPs.

As a co-ordinating body the European infrastructure oversight body might also have a potentially useful role in co-ordinating finance from a number of ANSPs for specific projects that are beneficial to their region as a whole. For example, the European infrastructure oversight body might, on the basis of cost-benefit analyses, approach individual ANSPs and obtain finance for specific cross-border projects. Similarly, the European infrastructure oversight body might also be encouraged to actively pursue general funds for transport infrastructure investment at the European level for specific projects. However, unlike the European infrastructure manager discussed below, the role of the European infrastructure oversight body would necessarily extend to include responsibility for contracting or actually undertaking the network enhancements that it identifies, which would remain the responsibility of individual ANSPs. This has important implications for the extent of resources and funding that such a European infrastructure body would require to operate.

It is our view that in light of the problems of the current system outlined in Section 7.2.2, the introduction of this type of European infrastructure oversight body may result in significant benefits in terms of co-ordination of investment decisions across the European network. Crucially, however, this view is subject to the assumptions that the roles of European infrastructure oversight body are well defined and that it performs these roles efficiently. The example of the PRC/PRU – which are relatively modest in size but have access to external resources – could be a possible model upon which such a European infrastructure oversight body might be based. Similarly, experience from other network industries suggests that a network co-ordination body with relatively modest resources can bring significant benefits to a network.

(c) A European Infrastructure Manager

Alternatively there may be potential for the introduction of a more active European infrastructure manager. The roles of such a European infrastructure manager might extend beyond the roles performed by the European infrastructure oversight body to include responsibility for contracting to facilitate the development of investment projects. This 'active' European infrastructure manager would therefore act to some degree as a centralised infrastructure investment body responsible for identifying, financing and potentially owning future investments in the European network.

Two observations can be made in relation to this type of European infrastructure manager. The first point concerns the access it would to finance for its activities, particularly if it is to have a role in funding infrastructure projects. One possibility which is discussed below in Section 7.3 could be that the European infrastructure manager approaches investment banks for funding in the normal project finance fashion

to fund a specific investment. This might be the case for example where individual ANSPs are not willing to finance cross-border projects on the basis of projected benefits *ex ante*. Alternatively, as is discussed below, in the event that a dedicated European infrastructure fund is created the infrastructure manager might have access these funds to undertake these cross border investments.

A second point is that the introduction of a European infrastructure manager will have obvious implications for how individual ANSP's infrastructure projects relate to those undertaken by the European infrastructure manager. This is an issue that is beyond the scope of this report, however, it is worth noting that the introduction of a centralised infrastructure role will need to be accompanied by a clear division of roles and responsibilities for infrastructure management between the national and European level to prevent inefficient duplication of activities.

Finally, it is notable that there was a broad reluctance on the part of stakeholders for the introduction of this type of European infrastructure manager. This reluctance was based upon the view that such a body would prove to be inefficient for the system as a whole, and that the introduction of such a body will result in a duplication of roles with other organisations.

(d) Assessment

In our view the problems with the current system of ATM infrastructure interoperability and co-ordination at the European level are significant and could potentially best be addressed through the joint introduction of two measures. The first is through the introduction of common Pan-European ATM infrastructure standards of interoperability. Similarly, in our view there is a strong argument that these measures should be accompanied by the introduction of a European infrastructure oversight body responsible for identifying areas of potential improvement to the network as a whole, and in encouraging greater levels of co-ordination with respect to infrastructure projects

7.3 Infrastructure financing options

In the previous discussion we examined potential approaches to identifying, prioritising and potentially undertaking European infrastructure projects. **An important related issue, and one which is central to one aspect of the creation of a solidarity fund, involves the financing of these infrastructure projects.**

Put simply the current possibilities for financing potential infrastructure projects are limited to:

(i) User charges
(ii) Financing from individual Member states,
(iii) Joint financing from individual Member states
(iv) External financing (including project finance)

7.3.1 The creation of an Infrastructure fund from User Charges and Member States

The issues relating to potential options for developing user charges and the contractual relationship between users, ANSPs and member states are discussed elsewhere in this report. However, in the context of the discussion in this section it is important to note three points relating to the use of user charges or Member State financing to create a collective infrastructure fund.

Firstly, Article 14 (3(e)) of the Common Position of the Council in relation to user charges states that:

> *"charges may be used to provide: revenues to benefit projects designed to assist specific categories of airspace users and/or air navigation service providers in order to improve collective air navigation infrastructures, the provision of air navigation services and the use of airspace."*[120]

While this Article clearly allows for user charges to be levied to finance collective infrastructure projects, it does not specify the *way in which* these charges can be levied, or how the *timing* of these charges relates to collective infrastructure projects. That is, the Article does not specify whether there needs to be an additional charge (such as a surcharge) to fund infrastructure projects, or whether existing user charges can be re-allocated to fund collective infrastructure projects. Similarly, the Article does not say anything specific regarding the timing of user charges, which could be either *before* or *after* the collective investment has been undertaken. These issues are of great importance to many participants in the ATM sector and are discussed in more detail below.

(a) User Charges

In relation to the creation of a collective European infrastructure fund it should be noted at the outset that there is broad reluctance on the part of users to create any form of European infrastructure fund through levying *additional* user charges. In our view, this reluctance is entirely rational: to be asked to pay without receiving any firm commitments in relation to benefits would, in economic terms, be even worse than cost-plus provision (which at least delivers the service ahead of the bill).

An alternative possibility for the financing of collective infrastructure projects, which may be more acceptable to system users, might involve a *re-allocation* of the existing contributions from user charges toward infrastructure projects at the national level to a dedicated European collective infrastructure fund. However, before any such reallocation of existing charges can be seriously considered it would be necessary to undertake a more detailed analysis of the current user contributions to the financing of infrastructure projects at different parts of the network, including through EUROCONTROL charges. The potential re-allocation of user charges toward a

[120] *Common Position adopted by the Council with a view to the adoption of a Regulation of the European Parliament and of the Council on the provision of air navigation services in the Single European Sky*/dated 11 March 2003 - 2001/0235 (COD)

European infrastructure fund would crucially also need to be accompanied with the introduction of new processes or institutions to oversee the financial management of these infrastructure projects.

(b) Member State financing

A second possibility for the financing of infrastructure projects at the European level is through financing by Member states. **Here too, the potential for the creation of a European infrastructure fund, which pools the contributions of Member States, might also exist.** However, in its recent report, the Van Miert High-Level Group on the Trans-European Transport Network was critical of the contributions currently being made by Member states toward trans-European networks and estimated that:

> *The Member States are currently investing less than 1% of their gross domestic product in building transport infrastructure and devoting only one third of this investment to achieving the trans-European network*[121]

As discussed above, one possible explanation for this reluctance on the part of Member states to invest in the trans-European network is because of a lack of incentive that may exist to undertake and finance infrastructure projects unilaterally which are only of marginal benefit to their part of the network, but may be highly beneficial to the network as a whole. Put another way, the issue surrounds the extent of willingness, and extent of responsibility, of each Member State to finance infrastructure projects which provide positive externalities to the network as a whole.[122]

One possible method for partially addressing the problem outlined above may be through the joint financing by individual member states of specific infrastructure projects that may be beneficial to both states. This process might occur through, for example, bilateral negotiations between neighboring member states. **Indeed, as noted above, one of the potential functions of a European infrastructure oversight body might be to identify projects of collective interest and then either approach, or act as a forum for, different member states to establish the terms for financing arrangements.** However, an obvious problem with focusing too heavily on this approach to financing European infrastructure projects is that it is still necessarily subject to the initiative and discretion of Member states. Although should this prove too significant a problem this could potentially be addressed through the introduction of a European infrastructure manager who could either obtain finance to fund these projects up-front and then receive payment from Member states once the benefits of the investment are realised, or fund the infrastructure investment through its own resources.

[121] "High Level Group on the Trans-European Transport Network Report", European Commission, 27 June 2003, page 7

[122] Page 8 of High Level Group report above notes that "By their very nature, trans-European network projects benefit the whole of the Union. Consequently, Member States should go beyond a purely national logic which has led - apart from a few, all too rare exceptions - to their excluding funding for any infrastructure outside their territory" and they draw the attention of "economic policy decision-makers to the incongruity in the long term between what is at stake in carrying out these projects and the constraints curbing public funding".

In summary, the creation of a European infrastructure fund through either a surcharge on users or through contributions from Member states, is likely to be unpopular not least because of the danger of the potential inefficiency it introduces, and because it potentially detaches the funding of collective projects from the beneficiaries of these projects. In our view, empowering a European infrastructure oversight body or European infrastructure manager to identify and co-ordinate finance for specific cross-border projects as they arise may provide for the more efficient development of the network.

7.3.2 External Financing for European ATM Infrastructure Projects

Consistent with the subsidiarity principle enshrined in the EC treaty, issues surrounding infrastructure investments have typically been primarily the responsibility of each Member state, with the European Community's role being limited to acting as a coordinating body.

Article 129(c) of the EC Treaty allows the Community to support the financial efforts of the Member States for projects of common interest, and since 1990 the European Council has provided for a European infrastructure action plan for the development of projects of European interest, and at that time funds were first set aside to fund some of these projects. [123] Specifically the form in which this support can take is through: [124]

- financing feasibility studies;
- loan guarantees;
- interest-rate subsidies; and
- in the case of transport networks, subsidies through the Cohesion Fund

Council regulation 2236/95 outlines the general rules for the granting of Community financial aid in the field on trans-European networks. Projects that are eligible for assistance under this Regulation include:

- projects of common interest
- that contribute to achieving the purposes of the networks,
- that contribute to the economic objectives of the Commission White Paper on Growth, Competitiveness and Employment,
- that are accompanied by an environmental impact assessment,
- that are intrinsically viable, with solid financial backing, participation by the private sector and a need for Community support.

[123] Article 129c of the EC Treaty that: "In order to achieve the objectives referred to in Article 129b, the Community may support the financial efforts made by the Member States for projects of common interest financed by Member States, which are identified in the framework of the guidelines referred to in the first indent, particularly through feasibility studies, loan guarantees or interest rate subsidies; the Community may also contribute, through the Cohesion Fund to be set up no later than 31 December 1993 pursuant to Article 130d, to the financing of specific projects in Member States in the area of transport infrastructure. The Community's activities shall take into account the potential economic viability of the projects."

[124] These were laid down by Council Regulation 2236/95 of 18 September 1995 on general rules for granting Community financial support (amended by Regulation 1655/99 of 19 July 1999).

Despite the introduction of Regulation 2236/95 and subsequent initiatives at the European level to allocate specific fund for transport infrastructure projects, the Van Miert High Level report recently observed that:

> *Both the commitments entered into at the European Council of Essen and the recommendations of the new priorities made by the Group risk remaining a dead letter if the European Community does not release new financial resources. In particular, the Group recommends to the budgetary authorities that they should positively consider an appropriate allocation of funds, and one which truly acts as an inducement, be set aside for the trans-European transport network within the forthcoming financial perspectives*[125]

Finally, it should be noted that the ability to access this form of financing to fund Pan-European infrastructure improvements in ATM will be subject to competition for these funds from the other European transport sectors. Although the funds are intended to apply to all European transport networks, it appears as though an increasing proportion of these funds are being directed toward infrastructure investments in the rail and road sectors.[126]

7.3.3 Other forms of external finance

Other forms of external financing that may be relevant in thinking about potential funding for Pan-European infrastructure projects in ATM might also include support from either the European Investment Bank (EIB) or the European Investment Fund (EIF), or through private sector support in the form of a public-private partnership.

In regard to each of these the following brief comments can be made:

- Loans from the EIB are granted on the basis of banking and capital markets criteria, particularly the financial feasibility of a project, that is the ability to repay, and technical and environmental suitability. Similarly the EIF uses the same conditions to provide its guarantees.[127]

- External funding from private sources is specifically encouraged in Regulation 2236/95 through the development of partnerships between the public sector and private operators. One specific form in which these private-public arrangements can be facilitated include through the use of financing instruments. The Commission has stated that it wishes to see wider use to be made of structurally subordinated loans and early operational stage loans.

[125] "High Level Group on the Trans-European Transport Network Report", European Commission, 27 June 2003, page 8

[126] A notable exception to this is the financial support for the Gailleo project

[127] "High Level Group on the Trans-European Transport Network Report", European Commission, 27 June 2003, page 8 point 23. In this respect it should be noted that one of the key proposals of the Van Miert report is that "The Group is keen to stress the crucial role of the European Investment Bank (EIB) through its loan policy. It suggests to develop the financing capacity of the bank through various financial engineering techniques in particular for cross-border projects. Moreover, it suggests that the EIB strengthen its links with the European Commission"

The development of private financing for infrastructure projects is also encouraged in the Van Miert report which notes that:

> *Finally, given the extent of the financial requirements, the Group is calling for initiatives to promote public-private partnerships. An appropriate legal framework, particularly as regards concession rights and charging for infrastructure use, must be introduced at Community level. Such partnerships must also be based on a distribution of risks which is acceptable for the private sector. New guarantee mechanisms ought to be set up, such as in the context of a mutual risk fund, in order to cover, inter alia, the risks of delays or failures to complete certain sections which could jeopardise the viability of a project*[128]

7.4 Summary

In conclusion the following points can be made in relation to potential ways of improving ATM flow management co-ordination and infrastructure management and development:

European ATM System Operation

- The current ATFM system in Europe involves a combination of activities by aircraft operators, individual ANSPs and the CFMU. There are number of problems identified with the current European ATFM system in various studies and by stakeholders

- One possibility for the improvement of European ATFM is through the creation of a dedicated European network manager. The European ATFM network manager could perform either 'active' or 'passive' functions, depending upon the extent to which some functions relating to system operation could be better performed at the local or European level

- An examination of the approaches to system operation and congestion management adopted in the European internal markets for electricity and railways reveal that they are also considering alternative mechanisms for the allocation of capacity close to real time, while also ensuring that certain safety and interoperability requirements are met.

European ATM infrastructure interoperability and development

- The Single European Sky initiative aims to promote greater levels of harmonisation, integration and interoperability between European ATM infrastructure. Currently infrastructure projects are identified, undertaken and commissioned through a combination of EUROCONTROL and individual ANSPs programmes.

[128] "High Level Group on the Trans-European Transport Network Report", European Commission, 27 June 2003, page 8 and point 26

- The limitations of the current system include a lack of compliance or enforcement mechanism at the European level; questions of ownership and accountability for forecasts; and issues relating to the financing of Pan-European infrastructure projects

- In our view the problems of the current system of ATM infrastructure interoperability and co-ordination at the European level could potentially best be addressed through the joint introduction of common Pan-European ATM infrastructure standards of interoperability and a European infrastructure oversight body responsible for identifying areas of potential improvement to the network and encouraging greater levels of co-ordination with infrastructure projects.

Infrastructure Financing Options

- The options for financing European ATM infrastructure projects are limited to either: user charges; contributions from Member states or to external sources of financing such as European transport infrastructure funds or through loans from the European investment bank or through other project finance avenues.

- In our view, the creation of a European infrastructure fund through either a surcharge on users or through contributions from Member states, is likely be inefficient because it potentially detaches the funding of collective projects from the beneficiaries of these projects.

- A better way forward might be to allow the European oversight body or infrastructure manager, to identify and obtain finance for specific cross-border projects as they arise including through either European Commission funds or possibly through obtaining project finance.

SECTION 8
THE STRUCTURE OF ROUTE CHARGES: GENERAL PRINCIPLES AND SOME GENERIC ISSUES

8.1 Introduction

In this section we first consider one or two of the most relevant implications of the earlier review of best-practice regulation for the relationship between charges and costs, then we discuss some general criteria against which charging arrangements can be assessed and which are referred to in Article 14 of the Service Provision Regulation[129]. The remainder of the section considers a set of generic issues that arise in the consideration of alternative charging parameters, and in particular considers the approach to weight and distance in the current charging formula, and the potential for the useful development of two-part tariffs and congestion-based charging.

8.2 Charging principles: overview of the issues

8.2.1 Charges and costs

As outlined in the earlier review of best-practice regulation, in considering the appropriate relationships between prices and costs in any charging structure there is a crucial trade-off between promoting dynamic efficiency (innovation, restructuring, improving service quality, cost-reduction, etc.) and promoting static efficiency (making the best use of currently available resources and assets). This distinction is not always made in discussions of charging principles for networks, with the result that the resulting analysis is often excessively static in nature. That is, too much emphasis is frequently placed on setting charges that closely reflect *prevailing* levels and structures of costs, with the result that, where charges are set in this way, there is little encouragement for dynamic improvements in performance.

A well-designed charging structure will, in terms of economic efficiency, have a dual purpose:

- It will signal *costs* in a way that encourages customers to use existing network assets and resources in the most effective way, for example by making it cheaper for customers to avoid congested areas (spatial differentiation) or congested times (peak/off-peak differentiation).

- It will signal *opportunities* for improvements in supply-side performance in a way that encourages suppliers to discover and implement those opportunities.

[129] *Common Position adopted by the Council with a view to the adoption of a Regulation of the European Parliament and of the Council on the provision of air navigation services in the Single European Sky*/dated 11 March 2003 - 2001/0235 (COD)

The general approach adopted by best-practice regulation in seeking to resolve the trade-off between static and dynamic efficiency is now well established across many network sectors and across many countries, although the detail of implementation, which reflects particular circumstances, tends to differ. It can be summarised as follows:

> *Prices/charges should be set so as to be under constant pressure to move towards the relevant cost-based levels, such that the rate of convergence reflects the balance of static and dynamic opportunities for improved efficiency.*

In relation to ATM in Europe, there is a clear consensus that the potential dynamic improvements in supply-side performance are particularly significant, and this has major implications for the future development of the charging structure.

(a) Marginal cost pricing

The issues raised by the static/dynamic efficiency trade-off can be briefly illustrated by reference to the principle of marginal cost (MC) pricing. Movement toward social marginal cost pricing is a stated aim of EC transport policy, and the potential for implementing MC pricing in respect of terminal charges was considered in detail by PWC.

It is often assumed that setting charges equal to short-run marginal costs (SRMC) is optimal from the point of view of promoting economic efficiency, and this is a view to be found in many introductory economics textbooks. However, the optimality result only holds in certain, well-defined circumstances, which are unlikely to be found in most realistic situations. Two examples illustrate the point.

First-best and second-best. Implemented in a particular economic sector, SRMC may cause distortions because:

- It leads to financial deficits, which must be covered by taxes that cause distortions in other parts of the economy, or

- Complementary or substitute activities are not priced at SRMC, in which case SRMC pricing applied in one sector will mean that *relative* prices might not accurately reflect *relative* costs, thereby distorting customers' decisions.

When such problems arise, it is generally efficient for charges to reflect factors other than SRMC. For example, when there is need to cover a financial deficit, deviations from SRMC might best be set to reflect willingness to pay (Ramsey pricing). In other cases, deviations from SRMC might reflect the pricing of complementary activities, at least for so long as pricing is not harmonised across all sectors.

In short, charges might be set to *reflect*, but not to equal, marginal costs, since other factors are also important.

Dynamic efficiency. SRMC pricing may have adverse effects for dynamic efficiency. For example:

- SRMC tends to be higher when demand is pressing hard against capacity. SRMC pricing can therefore provide incentives for suppliers to hold back on investment in incremental capacity, so as to maintain or increase prices and revenues.

- Innovation is often discouraged, since charges may not enable fixed costs of R&D to be recovered.

(b) Long-run marginal cost pricing

In practice, long-run marginal cost (LRMC) may sometimes provide a better first-base for setting charges than SRMC, for reasons that have little to do with static efficiency in resource allocation. Long-run approaches often yield charges that are more stable over time, and that avoid perverse incentive effects in relation to investment. Thus, for example, suppliers are not rewarded with higher prices for creating congestion by under-investing in capacity.

In effect, LRMC can work better than its short-run counterpart precisely because it breaks any very close link between charges and short-term cost movements, as is required to encourage dynamic efficiency. This is the most important consideration, not the marginal or incremental nature of the cost calculation.

8.2.2 Charge differentiation and unbundling

Determination of the most appropriate *structure* of charges, with differentiation among the different components of that structure, is an exercise in unbundling. Distinct service components are identified and priced, and network users can choose among alternative combinations of service components.

As indicated in the earlier review, the most appropriate form of unbundling can be expected to depend upon a mix of demand-side and supply-side factors. The 'dual purpose' test described above can be applied, for example by asking of any proposed, 'differentiated' charging structure, does it:

- signal *costs* in a way that encourages customers to use network assets and resources in the most effective way?

- signal *opportunities* for improvements in supply-side performance in a way that encourages suppliers to discover and implement those opportunities?

8.2.3 Charging principles

Before considering the detail of possible charging differentiation arrangements, it will be useful to clarify the meaning of three, 'high-level' criteria that are generally used to assess alternative options: cost-reflectivity, non-discrimination, and transparency. These three criteria are all explicitly referred to in Article 14 of the Draft Service

Provision Regulation[130]. Precisely because these are high-level criteria, they are open to differing interpretations, and this can have an important bearing on the relevant assessments.

(a) Cost-reflectivity

Cost-reflectivity can be interpreted narrowly – e.g. by requiring that prices be set equal to SRMC, LRMC, allocated costs, etc. – but this is to miss the point of the principle. What matters is that prices/charges be set in ways that best guide service users and service providers to more efficient outcomes. Costs are, of course, very important factors in this exercise, but, particularly when dynamic efficiency issues (innovation, restructuring, investment) are of importance, the linkages between prices and costs are not mechanistic.

Efficiency is not the only aspect of cost-reflectivity that may be relevant. It might be held to be 'fair' that an individual or undertaking responsible for particular decision, which leads to extra costs that can be attributed to that decision, should bear those costs, and should not bear the costs attributable to decisions of others. The approach is, however, silent on what to do about non-attributable costs, and it is therefore limited as a general principle of equity/fairness/solidarity. In practice, it may also come into conflict with other interpretations of equity/fairness/solidarity.

Equity, as well as efficiency, rationales for cost-reflectivity are, nevertheless, not to be discounted entirely, particularly when the relevant costs are measured broadly (in line with EC Transport Policy, which refers to social costs, not just to direct costs).

(b) Non-discrimination

Similar points apply in relation to non-discrimination. Interpreted narrowly, discrimination might be said to occur whenever relative prices differ from relative marginal costs. However, such a narrow interpretation risks the introduction of charging structures that are inefficient, because they do not serve the purposes described above.

In practice, the principle of *non-discrimination* tends to be interpreted in a way that excludes certain pricing structures that are judged to have adverse effects by virtue of being inefficient, anti-competitive, or manifestly in violation of some relevant notion of equity. It is, therefore, in relation to effects on efficiency, competition and (clearly defined) equity that the non-discrimination principle should be applied.

The efficiency and competition issues in ANS are broadly similar to those that arise in other network sectors (and more generally throughout the European economy). Equity

[130] Article 14(2e) of the Draft Service Provision Regulation states that: "Transparency of the cost-base for charges shall be provided". Article 14(3a) states that: "Charges shall be set for the availability of air navigation services under non-discriminatory conditions". Article 14(3d) states that: "Charges shall reflect the cost of air navigation services and facilities made available to airspace users taking into account the relative productive capacities of the different aircraft types concerned".

considerations are, however, more specific than elsewhere, as reflected in the ICAO principles. In particular, these principles indicate that:

- There should be no discrimination against airlines on the basis of nationality: two airlines requiring similar services should be charged the same amount, irrespective of nationality.

- Charges should not be such as to give rise to what might be called 'pure scarcity rents' (e.g. a charge for use of airspace that is completely unrelated to service provision, whether past, present or future).[131]

The latter does not appear to mean that there can be *no* deviations between charges and costs, suitably defined, but rather that any deviations need to be justified in terms of their function in encouraging more effective service provision

(c) Transparency

Transparency is slightly different from the other two principles, in that it refers less to the final charging outcomes and more to the process by which charges are determined. Transparency has a number of potential benefits, including:

- Encouraging revelation of information, particularly cost information, which may assist in monitoring and control of supply-side efficiency and commitments.

- Improving forecasts of the future evolution of charges, which may be particularly useful to network users in making future plans.

8.3 Charging structure parameters: some generic issues

The following provides an assessment of some generic issues concerning the charging structures in ATM. These issues concern the manner in which a given level of charges is recovered from users, and as such can be understood as having relevance to all of the charging options discussed later in the section.

(a) Weight and 'willingness to pay' [132]

The current use of a weight factor in the charging formula seeks to reflect willingness to pay in line with Ramsey pricing principles for the efficient recovery of fixed costs. However, the extent to which the current weight factor generates efficient prices is clearly a matter that can be questioned.

[131] Since charging scarcity rents may be efficient, as it encourages optimal use of a scarce resource, we interpret this aspect of non-discrimination as being, at heart, a matter of international equity.

[132] In our view, the relevant factor to be taken into account is 'willingness' to pay rather than 'ability' to pay. The former will be affected by the availability of substitutes, as well as to the income position of the end user. Ability to substitute is clearly relevant to the determination of the demand elasticities, that are relevant for Ramsey pricing.

The weight factor in the current *en-route* charging formula is calculated by taking the square root of the result obtained by dividing the maximum take-off weight (in metric tons) of the aircraft by 50. That is, charges are currently recovered on the basis of differences in the MTOW raised to the power of 0.5. In its report on Terminal charges, PWC argued that the best metric for ability (better, 'willingness') to pay is the standard seating capacity of an aircraft[133]. Whilst the current weight factor results in significantly higher charges for a given route for heavier aircraft, the per seat costs of the flight for heavier aircraft remain significantly lower. When considering the relationship for most aircraft types currently operated in Europe, PWC found seating capacity to be closely correlated with MTOW subject to an exponent of 0.7.

This suggests that the current formula may not fully reflect differences in willingness to pay, and thus that there may be a case for efficiency benefits arising from increasing the level of charge differentiation related to weight through an increase in the exponent. However, when considering the likely efficiency effects of any change to the weight factor, a number of factors need to be considered. Firstly, Ramsey prices are a function of willingness to pay **and** marginal costs. If there are significant differences in the marginal costs of service provision between large and small planes then these differences would justify a difference in the per seat costs. Thus, the fact that heavier planes will on average spend more time flying at level altitude, and that given this they may on average require less controller time per km than lighter planes could justify a difference of the kind observed in terms of Ramsey pricing.

Secondly, it has been argued[134] that it may be desirable for there to be appropriate incentives for airlines to use larger planes, in order to reflect the fact that the use of larger planes can reduce the number of movements to be controlled for a given number of passengers and thus alleviate the costs of delays. The growing importance of airport capacity constraints in Europe has been referred to in support of this view.

On this latter point, we would note that airport capacity constraints will clearly vary between locations, and seeking to address this issue through variations to the weight factor used for *en-route* charging would represent very poor policy targeting. On the more general point, we would note that the current weight factor does, on average, continue to result in lower costs per seat for heavier planes, and thus does provide for an incentive related to size of plane. In order to determine whether differences in marginal cost and/or the potential impact of plane size on the resulting costs of delay would justify a change to the current weight factor, detailed analysis of the likely effects of a given change on airline behaviour would be necessary. This is likely to be a relatively complex exercise given the wide range of factors that are likely to impact on choice of aircraft type. It is also notable that the results of such analysis are likely to be indicative, rather than clearly suggesting a specific weight factor.

We further note that (a) the suggested PWC exponent (0.7) was derived from a statistical analysis, and is therefore itself a point estimate subject to estimation errors (implying that the number could, with reasonable probability, be a little higher or a little

[133] PWC Study on Terminal Charges, page 54
[134] Issue raised during consultation process, for discussion see section 8.4

lower) and (b) PWC indicated that the current exponent (0.5) might in any case provide an acceptable approximation to willingness to pay.

Given these points, our preliminary assessment of these issues is that there is not a strong case at present for changing the exponent used in the current charging formula, although as noted in the discussion of specific options later in the report, there may be grounds for some increase in the exponent if carried out in conjunction with other changes that sought to capture differences in marginal cost in a different manner (for example, if there was an adjustment to the amount of revenue recovery linked to arrival/departure).

(b) Using specified intervals for weight rather than a continuous function

The current charging formula takes account of differences of MTOW by using a continuous function $[(MTOW/50)^{0.5}]$. One alternative approach that has been proposed[135] would involve instead defining the impact of weight in terms of a set of intervals, with different factors applied depending on whether the MTOW of the aircraft was, for example, less than 25 tons, 25 to 100 tons, 100 to 200 tons, 200 to 300 tons, or greater than 300 tons.

If the aim is to incorporate both willingness to pay and encouragement of larger aircraft as factors relevant to charging, we are unclear as to why this would provide a better balance between the two elements. PWC suggested an increase in the exponent, from 0.5 to 0.7, which might lead to a closer reflection of willingness to pay, but simultaneously reduce incentives to increase aircraft size. For reasons given above, we are not convinced that this would be a significant improvement on the status quo (other things equal). Likewise, however, we do not believe that the evidence suggests any obvious improvement from movements in the opposite direction, which would strengthen incentives towards use of larger aircraft but would likely weaken the linkage between charges and willingness to pay.

In addition, there is also the point that the proposal could, by its nature, introduce substantial discontinuities into the relationship between charges and costs. For example, an aircraft of 201 tons might be required to pay a significantly higher charge than an aircraft of 199 tons, when both the costs of providing the service and the willingness of the airline to pay for the service were little different between the two cases. This could distort airline choices, leading to otherwise avoidable inefficiencies.

(c) Charging by passengers

Another alternative approach that was raised during our consultation process was for charging by passenger, with the potential for *en-route* costs to be shown directly on the ticket price (in the tax box). As discussed in the PWC report, this raises questions of how to deal with cargo (including questions of the extent to which some passenger planes are used for cargo purposes). If *en-route* costs were to be put onto the ticket price, this may also result in a dampening of incentives on airlines to put pressure on

[135] Issue raised during consultation process

ANSP costs. Given that charging by passenger would represent a significant change to the system of charging, with associated costs and practical difficulties for application, and that there are no clear efficiency benefits associated (and indeed the change could dampen incentives), we find no grounds to consider that such a change would be justified.

8.4 Two-part tariffs – introducing a non-distance related component

As an alternative to introducing willingness to pay factors via the weight factor, consideration could also be given to the option of relying on two-part tariffs which, for each aircraft movement, introduced a non-distance related element into *en-route* charges. In the limit, two-part tariffs could entirely replace the weight element, but a more realistic option is partial substitution between the two, with the introduction of non-distance related charge leading to a lesser emphasis being placed on weight of aircraft.

There are, however, a number of difficulties with the proposal that likely limit the extent to which non-distance related components could usefully be used, at least for the moment. Consider, for example, a flight path that passes through the airspace of several member states. The airline would be charged a 'fixed' cost by each Member State. In contrast, a flight path that travelled a similar distance, but only within one member state, would pay fixed charges only once.[136] This effect – which in the energy sector is referred to as pancaking – arguably discriminates in favour of internal flights relative to cross-border flights. Since this is in direct contradiction to the principles of the EC Treaty, it is unsurprising that the Commission has been highly negative in its views of pancaking in the energy sector, and it would clearly be inconsistent to take a different view of the phenomenon in relation to ANS charges.

A further difficulty is that heavy reliance on non-distance related factors might, depending upon how it is implemented, be discriminatory in ways that had significantly adverse effects for both economic efficiency and competition. Suppose, for example, that a fixed charge is levied on each movement, irrespective of aircraft size. This would clearly discourage the use of smaller aircraft, since the charge-per-passenger could be reduced by increasing size. If the incentive effect exceeded the cost reduction benefit to the ANSPs, the result could be inefficiently over-sized aircraft, flying at lower load factors.

Other effects that would need to be considered include:

- possible effects on competition, including on barriers to entry, that would result from any increase in the per-unit passenger costs of using small, rather than large aircraft, and

[136] Note that, the issue is not one of allocating and charging seperately for fixed costs. Rather it is whether or not such separate charges are consolidated into a per kilometre rate

- possible impacts on flight frequency, which is a matter of direct concern to end users (and possibly business users in particular).

In relation to this second point, although lower frequency might, for a given volume of passenger traffic, reduce ANS costs (by reducing the number of flight movements), the other side of the same coin is that any such reduction in frequency will have negative impacts on end-users, since, other things being equal, end-users can be expected to benefit from higher frequency of service.

Notwithstanding these points, it is possible to envisage some potential role for non-distance related charges when two conditions are fulfilled:

- fixed charges are levied on a Europe-wide basis, so as to eliminate pancaking effects (which otherwise, in our view, would rule out the option), and

- there are some costs that can reasonably be attributed on a per-flight basis, rather than a per-km basis (e.g. the costs of processing flight plans).

In relation to the second of these, it can be noted that such costs are not strictly fixed (the cost driver is simply number of flights, rather than some distance-related cost driver) and are, in any event, likely to be relatively small, implying that impacts on the conduct of airlines, efficiency, and competition can be expected to be highly limited.

In addition to the above situations, it can be noted that Terminal charges provide for an effectively fixed cost component to be recovered in a manner that does not give rise to pancaking effects. Issues related to potential variations in Terminal charges are discussed in Section 9.

8.4.1 The Distance Factor

Distance is a key parameter in the current route charging formula and represents an indicator of controller time occupied (which is recognised as an important cost driver). For a given flight, distance is currently measured using the CRCO's Route per State Overflown (RSO) system. This system calculates the great circle distance in kilometres, between the aerodrome of departure within, or the point of 'entry' into, the airspace of the Flight Information Region (FIR) of a given state, and the aerodrome of first destination within, or the point of exit from, that airspace. The entry and exit points are defined as the points at which the lateral limits of the airspace are crossed by the route described in the last filed flight plan[137].

The distance taken into account is equal to the distance calculated on the basis described above less twenty kilometres for each take-off and each landing in a given state. The deduction of 20 km represents that part of the flight that – for charging purposes – is currently allocated to terminal charges. The distance factor used in the charging formula is equal to one hundredth of the adjusted distance measure.

[137] This incorporates changes made to the initial flight plan filed made by the operator, as well as any changes approved by the operator as a result of air traffic flow management measures.

A particular problem with the current use of the distance factor concerns the undesirable incentives it generates in relation the avoidance of congestion[138]. In particular, if an airline seeks to re-route (either independently or in response to an offer from CFMU) in order to seek to avoid a congested area, the distance travelled will be likely to be greater than would have been the case if no avoidance activity was undertaken. As a result, the charges paid by an airline that re-routes to avoid a congested area will be likely to be higher than they would have otherwise been. It is notable that this could be in addition to extra fuel costs that may be associated with traveling a longer distance.

In principle, a system that instead used fixed distance components would has a number of desirable properties. Most notably, users would not pay more in terms of direct charges when congested areas are avoided by choosing an alternative route; and ANSP's do not face a reduction in revenues when they engage in activities that result in more direct routing. However, it can be noted that fixed distance terms have been used in the past, before the introduction of the current RSO system in January 1998. The previous system of measuring distance, which had been used from 1971 when the Route Charges System became operational, was based on the Most Frequently Used Route (MFUR) between two aerodromes. The main rationale presented for the movement to the RSO system was a desire to find a more cost-related way of calculating route charges[139]. In particular, the new system was intended to better reflect the actual traffic in the airspace of any given State. Our understanding is that the primary complaints concerning the fixed distance system were based on the fact that significant divergences could arise between actual chosen routes and the fixed routes assumed for billing purposes. The key issue underpinning the change in the system, then, appears to have been the question of the allocation of revenues.

We would note that in principle this allocation issue could be addressed whilst maintaining fixed distance factors for charging. Indeed, one could envisage a system where fixed distances were used for the purposes of charging users, but where the allocation of collected charges was managed on the basis of the RSO system (i.e. on the basis of the route set out in the last filed flight plan. Whilst there would clearly be a number of detailed issues to assess in terms of the development of arrangements of this kind, we consider that the potential use of fixed distance terms – with appropriate allocation arrangements – merits further consideration.

8.4.2 Congestion-related charging mechanisms

Attempts to develop a set of arrangements aimed at reflecting congestion costs in *en-route* charges can be understood as broadly consistent with a movement toward social marginal cost pricing, which as noted earlier, is a stated aim of EC transport policy. As the earlier discussion of marginal cost pricing indicated, however, the desirability of any specific mechanism will be heavily dependent on the relevant circumstances.

For *en-route* services, an important factor includes the fact that capacity allocation – as opposed to capacity provision - is predominantly managed by the Central Flow

[138] The distance factor also generates potential incentive issues in relation to the directness of routes offered by ANSPs given that revenues depend on distance.

[139] *Off to a successful start: CRCO implementation of Route per State Overflown (RSO),* Skyway 1998

Management Unit (CFMU) of EUROCONTROL. In particular, CFMU processes flight plans in order to identify situations whereby expected traffic flows would exceed forecast available capacity, and acts to match supply and demand where an imbalance is identified. That is, the identification and management of constraints is currently managed at a European level.

There would seem to be a strong case for considering developments focused on reflecting congestion costs in pricing in the context of a more general evaluation of potentially desirable developments to the central flow management arrangements. Two factors that were raised earlier in this section in the context of SRMC pricing are of importance here. Firstly, a key issue in terms of the efficiency of congestion-based charging arrangements concerns the extent to which the resulting relative prices of substitute activities adequately reflect underlying variations in cost conditions (taking into account congestion related costs). In a European ATM context, this strongly implies that a consistent European approach to congestion charging would be desirable. Secondly, by focusing congestion charging efforts at the European level, issues of ANSP revenue recovery can be substantially isolated from the introduction of congestion charging, thereby avoiding potentially perverse incentives being generated for ANSPs to generate capacity constraints.

Consistent with this latter point, a useful distinction can be drawn between 'baseline' charging arrangements focused on long-term variables, and charging arrangements designed to reflect more short-term variations. The *en-route* charging arrangements – which deal with ANSP cost recovery - can be understood here as representing baseline charging arrangements. Congestion charging arrangements could be developed separately to this at the European level in conjunction with other developments in system operation arrangements[140].

8.5 Summary

The first part of this section provided a discussion of some general criteria against which charging arrangements can be assessed. Key points highlighted included that:

- Cost-reflectivity can be interpreted narrowly – e.g. by requiring that prices be set equal to SRMC, LRMC, allocated costs, etc. – but this is to miss the point of the principle. What matters is that prices/charges be set in ways that best guide service users and service providers to more efficient outcomes.

- Costs are, of course, very important factors in this exercise, but, particularly when dynamic efficiency issues (innovation, restructuring, investment) are of importance, the linkages between prices and costs are not mechanistic. Too much emphasis is frequently placed on setting charges that closely reflect *prevailing* levels and structures of costs, with the result that, where charges are set in this way, there is little encouragement for dynamic improvements in performance.

[140] System operation issues were discussed in further detail in Section 7

The remainder of the section provided a review of a number of generic issues that arise in the consideration of a range of alternative charging parameters. Key conclusions of this review included:

- We find no strong case for changing the weight factor used in the current *en route* charging formula, although some change in the exponent could be justified if carried out in conjunction with other changes that sought to capture cost differences in a different manner (for example, if there was an adjustment to the amount of revenue recovery linked to arrival/departure).

- The recovery of fixed cost components through non-distance related charges (for example, as part of a two-part tariff) can give rise to significant distortions of competition (in particular, through what is referred to in the energy sector as 'pancaking'). Such an approach would be highly undesirable for *en route* charging - other than as part of the development of an approach of the kind set out in Section 9 (ie. origin/destination/distance charging) – as the negative effects on competition are likely to be significant.

- The potential for the development of coherent and uniform congestion charging arrangements should be considered as part of a more general evaluation of potentially desirable longer term developments to air traffic flow management arrangements at the European level.

SECTION 9
ALTERNATIVE OPTIONS FOR ATM CHARGING STRUCTURES

9.1 Introduction

In this section we consider some alternative charging structures that might be promoted via the Single European Sky implementation rules.

At a general level, member state ANSPs are already able, if they so wish, to harmonise their *en-route* charges (i.e. create an ACB), subject always to any constraints imposed by ICAO rules and principles. Thus, for example, two or more adjacent member states could agree to set a common *en-route* rate and to accommodate overall cost recovery by means of adjustment to terminal charges.[141] This would be simplest where the cost structures of the ANSPs are broadly similar, but uncertainties over cost measurement and indeterminacies in the allocation of costs imply that significant rebalancing of charges (between terminal and *en-route*) could be implemented without manifest violation of cost-reflectivity principles.

However, it is necessary to achieve a degree consistency in harmonisation of charging in order to ensure that the European ATM network works well as a whole. Hence, for the purposes of developing Single European Sky implementation rules, it is also necessary to consider how such consistency can be promoted.

The three approaches to charging structures that are considered are:

- Development of current charging arrangements into a more coherent point-of-origin / point-of-destination / distance structure.
- The introduction of differentiation between *en-route* charges in upper and in lower airspace.
- The introduction of differentiation of charges by ACC, whether the airspace controlled by the ACC is at upper or lower levels.

Charging for use of electricity and (to a lesser extent) gas transportation networks is increasingly moving toward entry/exit systems in Europe. The basic idea of entry/exit charging is that system users pay charges at the points at which they are connected to the relevant system (in the Nordic countries, the payments are explicitly called 'point of connection' charges). Appendix 2 provides an assessment recent charging developments in European electricity and gas networks.

DG TREN is heavily involved in exercises in these sectors that are similar in important respects to the Single European Sky initiative. Although the issues raised are by no means identical, there are nevertheless a number of similarities that can potentially be

[141] Cost recovery issues could alternatively or additionally be addressed through revenue sharing arrangements.

exploited in thinking about ATM, and can be used to inform the Single European Sky deliberations to an appreciable extent.

Since the entry/exit terminology may be unfamiliar in the context of ATM, and since entry and exit points can be defined both for airspace as a whole (i.e. at airports) and for airspace blocks (i.e. as when calculating distances travelled within the block), leading to potential ambiguities, from this point on we will refer to the charging structure being considered here as *origin/destination/distance charging*.

9.2 Origin/destination/distance charging

In summary, under this approach, users pay:

- A charge depending upon point of origin (entry)
- A charge depending upon point of destination (exit)
- An *en-route* charge.

Current ATM terminal charges would be subsumed in the origin and destination charges. The origin/destination charges are effectively fixed in nature, although they could be expressed in terms of euros per km if there is a common initial distance (e.g. 20km or 80km) to which they are deemed to apply (the two are economically equivalent). Origin/destination charges can, as now, be levied according to a weight formula, which ideally should be harmonised both amongst member states and with the weight element in the *en-route* charges.

En-route charges could be levied as now, but alternatively could be differentiated along lines discussed in subsequent sub-sections on other options. **In the former case, which will be considered first, the approach can be summarised by the following formula for the charge levied on a particular flight:**

$$\textbf{Charge} = \{PO + PD + UR.[D - E]/100\}.[MTOW/50]^{\mu},$$

where:

PO is the point of origin charge coefficient,
PD is the point of destination charge coefficient,
UR is the en-route unit rate,
D is the distance travelled by the flight, in kms,
E is a distance exempt from en-route charges (e.g. from 40km to 160km),
MTOW is maximum take-off weight, and
μ is a parameter to be determined (currently set at 0.5 for the en-route component).

In effect, the *status quo* is a variant of this approach. The point to be made, however, is not that the *status quo* is entirely satisfactory (although it is an option), but rather that the framework can be developed in ways that would allow it to better contribute to the realisation of Single European Sky objectives. **The important conceptual step to**

make, in contemplating such developments, is to set aside any notion that the origin/destination charges should be linked, on grounds of cost reflectivity, to the costs incurred by specific operational units handling the initial and final phases of a flight.

In short, the latter notion is, in our view, a misreading of the cost-reflectivity principle. What matters is that the charges as a whole reflect the costs of a flight that originates at a particular point, travels a particular route, and terminates at a particular destination. Precisely where those costs fall is a separate issue, and is one that, for our immediate purposes, can be considered to be secondary.[142]

9.2.1 Current terminal arrangements: the PWC Report

Issues concerning the appropriate determination of terminal charges were extensively covered in the earlier PWC Report for the Commission. The main concerns of PWC were centred on the lack of harmonisation and on relative arbitrariness in the way that such charges were set, not on the existence of such charges as such, or of their place in a wider framework for ATM charging. More specifically, PWC were concerned that, in some states, terminal charges were relatively low in relation to relevant costs, leading to cross-subsidisation between *en-route* charges and terminal charges.

The crucial question here, however, is: what are the relevant costs? As already stated, these are not necessarily costs that can be attributed to operational units handling the relevant phase of the flight.

There are, therefore, two issues to be considered:

- What principles should govern how charges levied at points of origin and destination are calculated?

- How can the application of these principles be harmonised on a network-wide basis?

9.2.2 How more explicit tariff disaggregation can help

If origin/destination charges are regarded as components of a wider structure of gate-to-gate ATM charges, rather than as payments for a distinct, well-defined service based upon particular ways of organising ATM activities, then some potential blockages to harmonisation can be removed. Consider, for example, the argument that, since the costs of providing *en-route* ATM services are partly fixed, the tariff structure should reflect this by being split into fixed and variable parts. If this argument were accepted, one way of allocating the relevant fixed costs would be to terminal (origin/destination) charges. **Indeed, origin/destination charges are likely to be the most effective way of (implicitly) introducing two-part tariffs.**

[142] There are immediate analogies with other sectors. For example, entry charges in electricity do not correspond with costs actually incurred at the point of connection. Rather, they are set such that, when combined with the relevant exit charge, they reflect the costs of transmitting power from one location to another.

The last point can be illustrated by considering its implications relative to those of a different option, namely introducing a fixed charge for use of each airspace block. To repeat points made in Section 8 above, whereas every flight has one point of origin and one point of destination, flights of similar distances may pass through a smaller or larger number of airspace blocks, depending upon the route and the organisation of airspace. The larger the number of blocks along a flight of given distance, the higher would be the 'fixed' charge, an effect sometimes referred to as pancaking. Not only is this an arbitrary and non-cost reflective outcome, it is directly contrary to the high level principles of the EC Treaty, since a domestic flight of, say, 500km may be charged significantly less than a cross-border flight of the same distance.

In contrast, origin and destination charges can be used to introduce fixed cost elements without leading to the distortions associated with pancaking. Since each and every flight has one point of origin and one point of departure, any 'fixed' payments are independent of the number of airspace blocks passed through: all that matters is distance, and the artificial effects of boundaries are eliminated.

For the avoidance of doubt, we stress that we are not specifically arguing here for a two-part tariff (i.e. charges with fixed and variable components). Rather, the general point is that the architecture of this type of tariff structure is sufficiently flexible to accommodate sub-options in ways that satisfy both cost-reflectivity and non-discrimination criteria (whereas the imposition of a fixed charge for each ACB flown through would be manifestly discriminatory).

Since this approach is somewhat different from others that have been more extensively explored, some repetition of the principles may be justified. Thus, rather than starting from attempts to determine a tariff structure on the basis of attempted breakdowns of the accounting costs of existing operations (a process that encounters difficulties because of poor information and a certain degree arbitrariness in cost allocations), it starts by defining 'components' of the service as they might be perceived by users. Origin, destination and route are the most basic of these, although they are not the only aspects of significance (for example, timing/delays is an obvious other factor that can be reflected in charges). Given these building blocks, it can then be asked whether there is a division of charges amongst the three, defined components which might satisfy criteria such as cost-reflectivity, transparency, simplicity, non-discrimination, etc.

9.2.3 The split between origin/destination and en-route charges

As discussed in Appendix 2, in seeking to develop charges for use of European electricity transmission grids, regulators in Member States and the Commission have taken the view that there is arbitrariness in the relative proportions of revenues that can be recovered from entry (generation) charges on the one hand and exit (load) charges on the other hand. The response has, therefore, been to seek to agree a harmonised split of revenues, whilst leaving to subsidiarity the way in which Member State regulators determine more local differentiation of charges, provided only that the outcome conforms with the harmonised division of revenues. It is, therefore, natural to ask whether a similar approach could be adopted in relation to the division of either (a) total

revenues as between the origin/destination charges and distance/en-route charges or, almost equivalently, (b) the corresponding, permissible cost allocations.

In our view, this is likely to be a feasible task, although it would require further cost analysis work to verify the position. PWC approached the issue via detailed examination of disaggregated accounting information, but progress was limited by various, underlying information problems. This is very much a supply-side-only approach to charge differentiation, with charges closely related to specific operational activities. For reasons given earlier, however, in our view decisions concerning unbundling and charge differentiation should more appropriately reflect a balance of both supply-side and demand-side factors, not least to avoid adverse effects from the introduction of artificial boundary discontinuities in charges (where a sharp change in charges, dictated by operational boundaries on the supply side, does not reflect any corresponding sharp change in costs, and where inefficient substitution may therefore be artificially encouraged by the inappropriate discontinuity at the boundary). And, on the demand side, it is not necessarily of great interest to ANS users precisely how ANSPs organise themselves or their activities. What matters to users are the gate-to-gate service, the quality of service, the charges levied, and the alternative, substitute services that may be available.

On this basis, an alternative, simpler, more direct way forward is to work with overall unit cost information, as reflected in the current unit *en-route* rates and terminal charges, and examine the extent to which these costs are influenced by a set of drivers that include variables that can be clearly linked to the relevant stages of flight paths (i.e. departure/take-off, arrival/landing, and *en-route*).

Thus, at the most basic level of all, costs could be related to:

- The number of departures + arrivals within the relevant territory (naturally linked to origin/destination charges)
- Kilometres controlled within territory (naturally linked to *en-route* charges)
- GDP/head or wages (as a proxy for input costs, particularly labour costs)

The question is then largely an empirical one: how closely are observed costs across territories linked to these potential drivers? For example, if observed cost differences are little influenced by arrivals/departures numbers, this would tend to suggest that relatively little weight be given to origin/departure charges and that some of current terminal charges be absorbed into *en-route* charges, along the lines of option E in Table 4.2.2.1 of the PWC Report. On the other hand, if departures and arrivals are strongly linked to costs, this might suggest the opposite movement, implying at least some shift in revenue recovery from *en-route* charges to origin/destination (currently terminal) charges.

EUROCONTROL has already undertaken extensive work on factors influencing costs. Although the motivation has been somewhat different – the aim being to arrive at performance benchmarks for ANSPs – this expertise could be developed and used to address the somewhat different, and probably less complex, question raised here: to

what extent can observed cost variations be linked to variables that would define a more formalised origin/destination/en-route charging structure?

To the extent that simple measures of the number of departures/arrivals and kilometres controlled within a given territory do exhibit a relatively robust relationship with observed costs, the resulting relationship can be utilised to calculate bounds on the appropriate split between terminal (origin/destination) and *en-route* charges for ANS within a given territory.

Whether or not such robust relationships exist is, of course, an empirical question. Similarly, the precise way in which results can be used depends upon what any such exercise shows. It is possible, for example, that, in a cross-section analysis, ATM costs will be found to be approximately subject to constant returns to scale, with departures + arrivals and kilometres controlled jointly accounting for most of the variation in costs (when the latter are corrected for input price variations). In that case, it would be realistic to contemplate harmonisation of a specific, formulaic split between revenues from origin/departure charges and from *en-route* charges on a Europe-wide basis. Alternatively, there may be economies of scale, in which case harmonisation might require only that the split falls within certain defined boundaries, which themselves may depend upon the particular mix of traffic in a particular airspace block (e.g. departure + arrivals in relationship to kilometres controlled).

As stated above, a considerable amount of relevant technical work has been done on cost variations. What evaluation of the charging structure issue requires, therefore, is not the development of new techniques and data, but rather a reframing of the questions that are posed. Thus, to repeat, in relation to origin/destination (terminal) charges, the central question is: how significant a cost driver is the number of arrivals + departures in the territory of a Member State?

9.2.4 Example and preliminary results

Consider the relationship:

$$Cost = \alpha + \beta\,(departures + arrivals) + \gamma\,(kilometres\ controlled) + other\ effects$$

If the coefficients β and γ are well determined, they provide a basis for disaggregating charges as between points of origin/departure and *en-route*. Further, the extent to which the two variables 'explain' variations in costs can be interpreted as an indicator of the potential scope for increasing cost reflectivity by further disaggregation.

We have attempted some preliminary statistical work on this matter. Because of time and data constraints, the cost information used is limited to costs allocated to *en-route* charges under the current arrangements. This has two implications:

- Current allocations between terminal and *en-route* charges are not harmonised, and there is therefore 'noise' in the data as a result.

- Since terminal costs are excluded, it is to be expected that the effect of arrivals and departures on costs will be damped. Indeed, if current charges were cost-reflective, the estimated coefficient β could be expected to be zero.

The latter point implies that any observed, significant relationship between arrivals + departures and (*en-route*) costs can be interpreted as an indicator that the current charging structure is unbalanced, at least in relation to the criterion of overall cost reflectivity.

Since input prices (e.g. wage rates) can be expected to have a strong influence on costs, the data sample in the analysis was confined to EU/EEA Member States. Although this does not eliminate the likely effects on costs of input price variations among ANSPs, it can be expected to substantially reduce the importance of the omitted variable compared with a data sample covering all EUROCONTROL countries.

In terms of preliminary results, we find that:

- There is no evidence of significant, *long-run* economies of scale in ATM provision.[143]

- Both arrivals and departures + kilometres controlled have significant and well-defined links to *en-route* costs.

The second of these results confirms the finding of the PWC Report, arrived at by different methods, that the current charging structure may be unbalanced by virtue of terminal charges that are set at rates that are unduly low and *en-route* charges at rates that are unduly high.

The results also suggest that there may be significant scope for harmonising current, national *en-route* unit rates, without fundamentally undermining cost reflectivity. This could occur through the development of ACBs, in which participants in a block agreed to set a common, *en-route* rate within the airspace. Given that origin/destination charges could, and on the evidence should, vary from ANSP to ANSP, relatively modest levels of encouragement might be sufficient to encourage the development of such blocks.

More speculatively, we can see little in the data that would preclude full harmonisation of *en-route* charges over a wide area of Western Europe, encompassing at least: Benelux, Germany, France, Italy, Spain, and the UK. More generally, it may be feasible to reduce *en-route* charges across all EUROCONTROL member states to just three or four rates, linked reasonably closely to well-defined characteristics of the airspace that affect costs.

[143] Unit costs are, of course, negatively related to volume of traffic controlled in the *short-run*, as can be observed from the upward adjustments to unit rates that ANSPs make when they are confronted with dips in traffic volumes. Of course, as noted above, we recognise that this result is preliminary and based on one study and that there are a number of other studies in this area.

9.2.5 Key points

(a) Cost reflectivity

Given potential confusions about what the principle of cost-reflectivity might mean, it is emphasised once again that the origin/destination (entry/exit) charges do not necessarily have to closely reflect those costs associated with a distinct flight phase, mutually exclusive of flight phases covered by *en-route* charges. Thus, arrival/departure numbers might add to total costs because they tend to increase complexity in airspace beyond the normal limits of terminal control, but nevertheless still within the airspace of the relevant member state. Allocating such extra costs to origin/destination charges may be a simple, effective and appropriate way of reflecting complexity. Whether or not this is the case is an empirical issue.

(b) Distributional effects

One of the major difficulties in any rebalancing of charges is that it may have major distributional effects on network users. Thus, a shift towards greater revenue recovery through origin/destination charges relative to *en-route* charges will tend to have adverse effects on airlines flying over shorter routes – since the distance component of such flights contributes relatively less to total ANS charges. Such distributional effects may have implications for important matters such as the strength of competition on some routes and regional development.

In this context, we note that there are mechanisms by which any redistributive effects of a formalised origin + destination + en-route charging structure, characterised by a smaller revenue contribution from *en-route* charges than now, can be mitigated:

- As explained earlier, the weight factor, which can be rationalised in terms of the efficient recovery of costs that are fixed in the short- to medium-term, is subject to some uncertainty. As discussed in Section 8, according to the PWC Report, its value could reasonably be raised on the basis of arguments deriving from willingness to pay. Since short routes are normally serviced by smaller aircraft, this would offset at least some of the distributional effects of rebalancing.

- For reasons given, we do not believe that there is any requirement to link origin + destination charges to a specific set of operations corresponding to a given flight phase. It would be feasible, therefore, to reduce the distances covered by *en-route* charges by, say, increasing the distance exempted from such charges for each flight (e.g. to, say, the first and last 80kms of the flight). This would shorten the distances used in the calculation of *en-route* charges, which would have a beneficial effect for airlines flying short routes.

  We note, however, that any such decision would have unwanted, inter-ANSP effects in cases where an ACB boundary is crossed in the first or last 80km. Consequential compensatory arrangements would therefore likely need to be made.

- By allowing ANSPs to vary, according to defined principles, origin and destination rates among airports in their territory, it is possible that such charges could be lower at regional and at less congested airports. As well as potentially improving the effects of cost-reflectivity – since substitution amongst airports *is* a feasible option for most airlines – this would tend, on average, to benefit the smaller airlines which rely more on these airports than do bigger companies.

9.2.6 Encouragement for restructuring

As has been stressed throughout this Report, what matters in terms of incentives for improvements in efficiency, including efficiency achieved via restructuring, are the relationships between total revenues on the one hand and total costs and other performance measures on the other hand, together with the non-financial pressures that may be applied on ANSPs including, importantly, by system users. The *structure* of charges does not, in and of itself, have very much bearing on the matter.

It is, however, possible to develop arrangements in which choices concerning charging structures are linked to incentives for performance improvements. For example, since the rate/speed at which charges are, over time, adjusted towards costs is of some importance for financial returns, ANSPs that agreed to join together in an airspace charging block, with a common *en-route* rate, might be afforded the opportunity to obtain more favourable (to them) charge-adjustment arrangements.

At the end of Section 6 we set out a general benefit sharing approach that was sufficiently broad that, whilst providing improved incentives and risk sharing arrangements within a common framework, allowed for the accommodation of at least some national regulatory preferences. **One possibility is to permit longer contract/regulatory lags, within this framework at least for some predefined period, in circumstances where ACBs are either formed or extended.** On the other hand, the view might be taken that the encouragement and pressures surrounding the benefit-sharing approach would themselves be sufficient to stimulate efficient consolidation and that no further steps are necessary.

9.2.7 Next steps?

As a result of its similarities with the existing structure of ANS charges, it is unlikely that member states would have objections to a formalisation in terms of an origin + destination + distance (*en-route*) structure. The principal issues therefore occur in relation to:

- Harmonisation of the distance exemption limit for *en-route* charges (the 20km to 80km issue).

- Rebalancing of revenue recovery as between the origin/destination components.

- Consideration of any other adjustments (e.g. to the weight factor) that might be deemed desirable to mitigate adverse distributional effects of rebalancing.

Rebalancing of charging structures is nearly always controversial, and the normal policy response is to implement changes gradually. **The Commission could, for example, set a maximum fraction of revenue that can be recovered by member state ANSPs from en-route charges, with a view to adjusting that fraction over time in the light of evolving cost and performance information.** Since the traffic mix (take-offs and landings, overflights, kilometres controlled, average weight factor) varies from one member state to another, it would be inappropriate to impose a uniform maximum. **It would, however, be feasible to prescribe a maximum that is formulaically linked to traffic characteristics**, particularly since the objective is a gradual, directional shift rather than an instant leap to different charging levels.

9.2.8 Summary

The existing division between terminal and *en-route* charges for ATM is, in fact, a form of origin/destination/en-route charging. The problems to which it gives rise appear to be much more to do with the way in which existing charges are set than with the structure of charges *per se*.

Assessment of the origin/destination/en-route option is, therefore, a natural starting point when assessing potential reforms to ATM charging structures. The question is: can it be developed in ways that can overcome the perceived problems, or is it better abandoned altogether?

The PWC Report went some way in addressing relevant issues. What we have suggested here is that there should be explicit recognition that the *status quo* is a less than satisfactory implementation of a particular structure of charges that is not only potentially coherent but is also close to the preferred approach to charging in other European network industries (i.e. electricity and gas). Given this recognition, it can be asked: how can this framework be developed most effectively? An improved implementation of origin(entry) + destination (exit) + *en-route* can then be compared with other approaches to the ATM charging structure, with a view to determining which option is preferred overall.

Perhaps the most important conceptual breakthrough that can be made is to divorce thinking about the division between origin/departure charges and *en-route* charges from any organisational divisions of operational responsibilities. If arrivals, departures, and kilometres are simply considered as possible bases on which charges can be levied, it can be seen that they might effectively be used to reflect costs, no matter where, in operational terms, those costs actually fall. Whether or not this is the case is then an empirical matter. More specifically, it depends upon whether or not arrivals/departures and kilometres controlled, individually and together, account for substantial proportions of observed cost variations among national territories.

If assessments do indicate that these variables are robustly correlated with costs, then it is possible to develop harmonisation rules that will address many of the perceived problems of existing arrangements. Such harmonisation may be as simple as mandating, at the EC level, a required split in revenue recovery, or mandating bounds on such a split, as between the various charges, and it may feasibly be combined with

permissive approaches to certain further (but not all) forms of charge disaggregation at the member state level, provided that the mandated constraints are not violated.

9.3 Charging differentiation based upon a split between upper and lower airspace

As noted in section 2, this option has been explored by the CRCO task force. In many ways, the underlying approach is similar to that underlying the potential development of the origin/destination/en-route structure described above. In upper airspace, there are fewer vertical movements and the airspace is less complex. Hence, ATM costs per km controlled tends to be lower, and dividing the airspace for charging purposes at, say FL285, arguably provides a way of reflecting such cost differentiation.

The approach can be summarised by the following formula for the charge levied on a particular flight:

$$\textbf{Charge} = \{PO + PD + URL.[DL - E]/100 + URU.DU/100\}.[MTOW/50]^{\mu},$$

where:

PO is the point of origin charge coefficient,
PD is the point of destination charge coefficient,
URL is the en-route unit rate in lower airspace,
URU is the en-route unit rate in upper airspace,
DL is the distance travelled by the flight in lower airspace, in kms,
DU is the distance travelled by the flight in upper airspace, in kms,
E is a distance exempt from en-route charges (e.g. from 40km to 160km),
MTOW is maximum take-off weight, and
μ is a parameter to be determined (currently set at 0.5 for the en-route component).

There are two ways of viewing the upper/lower differentiation option. First, it can be seen as an *alternative* to the kinds of developments outlined in the origin/destination/en-route option. Second, it can be considered as a *supplement* to the first option. These perspectives will be considered in turn.

9.3.1 Upper/lower charging differentiation, complexity and tariff rebalancing

Whereas the origin/destination/en-route option deals with costs arising from complexity by 'pushing them downwards', anchoring them to charges defined by particular airports or groups of airports, the upper/lower approach seeks, in effect, to 'leave them behind' by establishing what are likely to be lower *en-route* charges in the upper level. If the upper/lower option is seen as an *alternative* to origin/destination/en-route option, by implication the existing balance between terminal charges and lower airspace *en-route* charges is left relatively undisturbed, except perhaps for some minor tidying up and harmonisation. On this basis, much of the cost of dealing with complexity (e.g. vertical movements) would be allocated to lower-airspace *en-route* charges.

One result of this is that lower airspace unit *en-route* rates could be expected to rise, which is nothing more than the usual tariff rebalancing effect that occurs when charging structures are reformed. Thus, once the overall revenues of a regulated undertaking are determined – and this is a matter of controlling average charges, whether by linking them to allowable costs, by imposing a price cap, or by some other means – reduced revenues from one component of a tariff structure implies that increased revenues will be recovered from other components of the tariff.

Although it is difficult to be fully confident about the outcome in the absence of detailed modelling, the outcome might reasonably be expected to increase inter-ANSP differences, at least in absolute terms, in the lower airspace (*en-route*) unit rates. This is because, at the moment, there is some averaging of the costs associated with complexity across all *en-route* charges, but if these costs were concentrated in lower airspace *en-route* charges, existing cost differences attributable to complexity – which, on the basis of the analysis of section 9.2.4 are responsible for significant variations in current, *en-route* unit rates – would be spread over a smaller number of service units.

At this point, one argument that has been put forward is that the higher charges in lower airspace would lead to a situation in which 'local' airlines would place greater pressures on 'local' ANSPs to reduce charges by increasing efficiency and reducing costs. As discussed in Section 2 of this Report, the thrust of this line of reasoning is that changing the tariff structure will give rise to a process that will, somehow or other, provide encouragement to greater efficiency.

There are three comments to be made on these points:

- Irrespective of the validity of the argument, similar points could be made in relation to tariff re-rebalancing via an origin/destination/en-route approach. The argument is, therefore, not so much a point in favour of a particular charging option (i.e. upper/lower differentiation) but more a point in favour of all those charging options that have similar distributional effects.

- Second, whatever 'common-sense' intuitions may suggest, there are some obvious difficulties with the economic logic of the argument. For example, if there is increased resistance to higher charges in the lower airspace, the obvious first defence of an ANSP could be to avoid the difficulty by limiting any rebalancing in the first place. In effect, the anticipated resistance is a *disincentive* to rebalancing, and upper-airspace charges may not fall to anything like the extent anticipated. This leads back to the fundamental economic point: in the end, the pressures and incentives for performance improvement will need to come from regulation. **Measures will need to be put in place to promote desired adjustments, which will not happen automatically simply by giving ANSPs opportunities to differentiate charges.**

- Third, to the extent that all lower-airspace unit rates increase by similar amounts, 'local' airlines based in one territory will not be unduly disadvantaged relative to 'local' airlines based in another territory (although European airlines as a whole will

likely be slightly worse off relative to overseas airlines). In such circumstances, the pressures to abate unit rate increases may be limited. If, however, unit rates for lower-airspace increase by more in more complex airspace blocks – as we have suggested might well be the case – any cost reduction pressures that exist will be exerted unevenly across airspace blocks. Specifically, the pressure from airlines to reduce costs may be greater in territories where complexity is greater which appear extremely arbitrary.

9.3.2 Specialised upper-airspace ATM providers

A stronger argument in favour of upper/lower charging differentiation is that it might facilitate the development of specialised upper-airspace ATM providers serving FABs/ACBs above FL285. This is perhaps best seen by considering the way in which such a provider might operate under different charging arrangements.

For simplicity, assume that a FAB is established that covers a large block of a cross border upper airspace, and also some part of lower airspace. A European FIR covers the upper airspace, which allows for charging independently of national boundaries, but national FIRs remain in lower airspace.

Taking account of the Maastricht charging scheme and of the CEATS and NUAC charging options that have been put forward and discussed in their respective business plan documents [144], three 'models' can be identified, which are as follow:

(a) National unit rate cost allocation

In this case, national providers have control over their own national lower airspace. It is assumed that, for part of their upper airspace, they have entered a joint service provision agreement of one sort or another.

The agreement specifies the shared resources required to operate the FAB. The operating costs are centralised and aportioned according to an agreed cost allocation rule (e.g. proportion of dedicated controller positions, of dedicated sectors, etc.) as in the Maastricht case [145]. The apportioned costs are then included into the national ANSP's costs.

For revenues, the volume of provided service units in the upper FAB is calculated by member state according to national FIR boundaries, even if the boundaries of operational sectors in the upper airspace do not fully correspond.

[144] Ibid pp.35-36, and in particular Nordic NUAC, Phase 1 report, Appendix 3 on 'Institutional and Organisational Issues, November 2002, pp22-25
CEATS business plan, 3 February 2003, Cash flows, Payments for infrastructure
[145] ICAO's policies on charges for airports and air navigation services, ICAO, 6th edition 2001.

(b) The average unit rate FAB model

In this case the member states enter into an agreement to restructure their upper airspace into a FAB [146]. An average single unit rate is jointly decided and attached to the FAB according to cost allocations made by the ANSPs. What this means is that the FAB is also an ACB Each member state charges a national unit rate in their lower airspace. Therefore if the upper airspace provider operates part of the lower airspace of a member state, it obtains revenues according to the lower national unit rate.

The charging process might involve initially passing on all costs of restructuring – resectorisation in the upper airspace, closure of assets in the lower (etc.) - to the upper airspace and hence to the average upper rate. Then, over time, the upper average rate could be gradually cut as the restructuring costs are written off.

(c) FAB/ACB unit rate autonomy

This model introduces the possibility of full independence in the management of the upper FAB/ACB rate and the lower national rates. The FAB agreement need no longer take the form of inter-ANSP contractual arrangements: the upper airspace could be controlled and managed by a separate company, of which ANSPs and possibly other investors could be shareholders.

Thus the relevant entity could have control over the financial aspects of operating the FAB. It would pay the cost of using infrastructure to the national ANSPs, decide upon using particular suppliers of equipment and support services (eg: meteorological services), and decide on the way to finance them –buy, lease, outsource.

The entity would obtain European certification. It would set its own unit rate and might offer its services to other member states. It might also offer to extend its activities to the lower airspace

(d) Discussion

The ability to set different unit rates above and below FL285 could be expected to facilitate the implementation of the latter two of the above three 'models'. It is, of course, possible that unit rates in upper airspace could be harmonised (so as to create an ACB) without such an ability to differentiate. However, in the absence of upper/lower differentiation, this would require simultaneous harmonisation of lower unit rates. And, whilst we think that this would in many cases be feasible if the origin/departure/en-route approach were followed, here we are concentrating on the upper/lower approach as an *alternative*, not a supplement. On that basis, it would appear likely that ACB creation in upper airspace would be facilitated if it did not require simultaneous harmonisation in lower airspace.

[146] For each country there may be several FABs and a FAB may cover part of the lower national airspace.

There are, however, some clear weaknesses associated with the upper/lower differentiation approach, and these are of immediate relevance when evaluating it against other options.

9.3.3 Limitations of the upper/lower differentiation approach

(a) The number of ACBs

The effect of introducing a horizontal division at FL285, with different charges above and below that FL, is to double the number of ACBs. That is, from a charging perspective, it would lead initially to greater fragmentation of the airspace architecture.

There are, therefore, crucial questions as to what would happen following this initial impact effect, which were raised and discussed in Section 2 above. It is to be expected that there would be some consolidation in ANS provision for upper airspace, and some consolidation in charging arrangements (i.e. harmonisation of upper *en-route* charges in large ACBs). Nevertheless, in the absence of some consolidation in lower airspace the number of ACBs would, as a matter of arithmetic, increase. **In an obvious sense, whilst there would be consolidation of provision in upper airspace when upper airspace is viewed separately, there would be greater fragmentation in airspace as a whole.**

(b) Non-cost reflectivity

Much more significantly, there is a strong argument that a substantial differential in unit rates above and below FL285 would not be cost reflective. **This follows from the fact that the dividing line is, in terms of cost differences associated with ATC, arbitrary.** We have seen no evidence to suggest that this flight level corresponds with any significant change in costs, and the fact that upper/lower distinctions have been made at FLs covering a quite wide range of altitudes suggests that it does not. Similarly, existing, specialised upper airspace control centres have not, to date, converged on FL285 as the altitude level that marks out their area of control.

What this means is that any significant difference in unit rates above and below FL285 would not, within an altitude range of say FL250 upwards, correlate at all closely with the relevant costs. **Within this range, the charging structure would not be cost reflective.** This raises two immediate problems.

First, to the extent that flight plans can be varied to take advantage of any substantial unit charge differential, the resulting pattern of substitution would be inefficient: **it would detract from, rather than enhance, network performance.** Flight plans would be changed to take advantage of lower unit rates above FL285, but the cost savings would be modest. The savings to airlines would be matched by shortfalls in revenues to ANSPs, and, since the latters' costs would be little affected, unit rates would presumably be raised in subsequent periods.

In response to this problem, it might be responded that, in practice, there would not be any major adjustments in flight plans, since other commercial factors drive airlines to

seek high altitudes; but in that case the best that can be said is that reform of the charging structure will have no material effect on network operations. The case for significant upper/lower differentiation must then rest upon the reasoning that it will somehow trigger a beneficial business dynamic: it can not rest upon any orthodox economic foundation.

Second, if it is accepted that a charging structure is not cost-reflective within a range of altitudes around FL285, such a structure is open to the charge that it is discriminatory. For example, it could be argued that, since (a) there is no evidence to suggest the existence of significant difference in costs in controlling flights between FL195 – FL 285 and flights above FL285, and since (b) shorter-haul flights spend a greater proportion of their time between FL195 and FL285 than do longer-haul flights, the effect is to discriminate against the former, with potentially distorting effects on competition in airline markets.

(c) Reduced encouragement for restructuring at lower levels

There is a further argument to the effect that the ability to differentiate *en-route* charges on either side of FL285 might actually **reduce** pressures to improve ANSP performance at lower airspace levels.

Consider an ANSP responsible for a small volume of airspace that, as a result of high unit charges, is suffering from loss of traffic volumes. The principal reason for such loss of traffic is likely to be that overflights are 'bypassing' the ANSP by means of flight re-routing in upper airspace. The resulting commercial pressures will fall on ANSP operations as a whole.

If, however, the ANSP can respond by reducing unit rates only in the upper airspace, unit rates for lower airspace, for which acceptable alternatives are less likely to be found, can be raised. That is, unit charges will be reduced where demand elasticities are high and increased where demand elasticities are low. The commercial pressures resulting from the loss of traffic are thereby abated.

Such a charging strategy would, as just explained, be impeded to some extent by competition law. An alternative strategy, with similar economic effect, would be to join a joint venture with neighbouring ANSPs to create a large FAB/ACB, with its own unit-rate, in upper airspace. Although this might look to be exactly the kind of development that the Single European Sky seeks to encourage, closer examination reveals that the effects are not necessarily benign. Specifically, the pressures to improve performance in lower airspace are eased. By isolating and addressing the 'competitive threat' in upper airspace, unit rates in lower airspace can be raised and costs can be recovered without the necessity for substantial performance improvements.

Given these limitations, our view is that, on balance, reliance on charging differentiation between upper and lower airspace, as defined by FL285, is not a strong alternative to further development of the existing structure of charges.

9.3.4 Upper/lower differentiation as a supplement to the origin/destination/en-route approach

If costs associated with airspace complexity are recovered largely through origin and destination charges, the case for using altitude as a proxy for complexity weakens. In effect, in terms of costs, airspace becomes much more homogeneous, at all altitudes. To the extent that altitude is used as a differentiating variable, the resulting variations in charges are likely to be relatively limited in magnitude.

Nevertheless, there remains a case for a permissive approach to possible differentiation of charges above and below FL285. This case is based upon the facilitation of new structures of ATM provision in upper airspace, as discussed above. **We think it should be recognised, however, that the contribution that can be expected from such developments is limited.** Charge differentials that can be reconciled with cost-reflectivity and non-discrimination principles will be modest, and, whilst upper airspace restructuring will be assisted by the ability to establish upper airspace ACBs independently of co-operation at lower airspace levels, the impact of such charge differentials on ANSP strategies in relation to lower airspace will tend to be muted, and may even be perverse in some cases.

9.3.5 Summary

Since a substantial differential between unit rates above and below FL285 would not, on the available evidence correspond to any corresponding differential in costs, such a charging outcome would be open to challenge on the basis that it was not cost reflective and was discriminatory. In our view, any charge differential that could be robustly defended against such challenge would need to be modest in magnitude.

For this reason, we are of the view that upper/lower differentiation should not be seen as an alternative to the development of the origin/destination/en-route structure in ways that allocate more of the costs of complexity to origin and destination charges. Rather, the possibility of charge differentiation by altitude would be a useful, if limited, supplement, aimed chiefly at allowing some developments in the re-organisation of upper airspace to proceed more easily.

9.4 ANSP/ACC charging differentiation

Finally in this section, we consider an approach where unit rates would vary from ACC to ACC. In other words, to each ACC would be attached a uniform unit rate. What this means is that OABs/FABs would be fully aligned with ACBs in all parts of airspace, which could lead, for example, to:

- Separate upper airspace charges associated with upper airspace ACCs (e.g.MUAC).
- Multiple charging zones in large member states that have two or more ACCs.

- Cross-border charging zones.

9.4.1 ACC pricing and the competition paradigm

From a charging (but not an operational) perspective, charge differentiation by ACC would, like the upper/lower airspace approach, lead initially to increased fragmentation of airspace. In this case, however, such fragmentation might be seen in a more positive light. The reason for this is that arguments for the ACC approach appear to be closely related to the notion that it is possible to introduce some degree of competition (*'network competition'*) in provision among ANSPs/ACCs, and that such an outcome can lead to the usual benefits that flow from competition, including improved efficiency, better service, and lower charges.

The concept of such *'network competition'*, whereby different providers compete to some extent for customers, is not to be dismissed out of hand. Such competition plays an important role in the communications sector, and it appears on a limited scale in other networks. For example, there is a degree of competition between long-distance gas pipeline companies in the US; and in the earlier stages of railway development there was sometimes quite fierce competition between rival companies.

(a) Substitutability and complementarity

The significance of network competition as a factor depends upon the pattern of substitutability and complementarity among services provided by the various ANSPs or ACCs.

- Services are *substitutes* if measures to increase the demand for one (e.g. a lower charge) *reduce* the demand for the other.

- Services are *complements* if measures to increase the demand for one *increase* the demand for other.

The competition paradigm applies when services are relatively strong substitutes for one another. Customers then have a choice of provider. If price rises, or if quality of service falls, they switch to another provider, and the quantitative significance of the switching is substantial. In the absence of collusion, this ability of customers to switch maintains strong pressure on providers/suppliers to improve their performance in meeting the requirements of customers.

It is immediately obvious how this type of reasoning might be used to support ACC charging. If reliance on a particular ACC becomes unattractive, either because the unit rate is seen as being high in relation to neighbouring ACCs, or because lack of capacity is causing congestion and delays, an airline might seek an alternative flight plan that 'bypasses' the relevant provider. Such substitution/switching reduces the traffic handled by the ACC and reduces revenues. Financial pressure is then placed on the parent ANSP to rectify the position by improving performance (which might possibly mean closing an inefficient ACC).

However, one difficulty that arises is that many ANSPs are not fully commercial, and caution should be applied in proceeding on the assumption that the reactions to loss of business and revenues would necessarily be similar to those of private companies. The result may, for example, simply be increased financial support from the host government.

Equally important, there is an empirical question concerning the strength of the substitution effects.

(b) The position in networks

Networks tend to be characterised by complex patterns of substitutability and complementarity. We give two examples:

- When ACCs are responsible for different levels of airspace, the dominant pattern is likely to be that of complementarity. A typical flight path will include phases in both lower and upper airspace. Thus, measures to promote traffic in lower airspace (e.g. constructing more airport capacity) can be expected, on average, to *increase* traffic in upper airspace. It is true that there is some substitutability at the margin, as was discussed above in relation to upper/lower charging differentiation. In response, say, to lower unit rates in upper airspace, there may be adjustments in flight plans to increase the time spent at higher altitude. The scope for such adjustments is, however, limited, and the dominant pattern is the joint use of both upper and lower airspace.

- For ACCs controlling airspace at similar levels, the pattern is less clear. A reduction in the unit charge by one ACC may, by leading to re-routing, cause decreases in traffic in some neighbouring ACC control areas and increases in traffic in others. Substitutability might be expected to be more important, and more significant quantitatively, where the ACC controls a relatively small part of airspace but, because of its geographic position, typically handles, or could handle, a significant volume of overflights. Where the ACC control area is much larger, the costs of bypass will tend to be substantially greater, and substitutability will be significantly lower.

9.4.2 The likely effects of ACC charge differentiation

On the basis of the above remarks, it can be expected that ACC pricing might put greatest pressures on smaller ANSPs. Larger ANSPs, responsible for multiple ACCs, can not be expected to encourage the individual ACCs to compete with each other. Pricing will be co-ordinated across large volumes of airspace, and the costs of bypassing such blocks are likely to imply that substitution effects are immaterial.

In relation to smaller ANSPs, we have already discussed possible responses to 'competitive pressures' in circumstances where upper/lower charging differentiation is feasible. We think it is likely that ACC pricing would, in practice, provide some encouragement to restructuring. However, at this point, attention should perhaps be explicitly drawn to an obvious point that has so far only been implicit: **restructuring,**

whether in terms of ATM operations or the charging architecture, is a potential means to an end, not an end in itself. The underlying aim is service provision that most effectively serves the requirements of network users and, ultimately, their customers, the travelling public.

Thus, it is possible to envisage restructuring whose chief motivation and effect is to reduce financial pressures on ANSPs by limiting 'competition'. More specifically, consolidation in upper airspace, where 'competition' is potentially stronger, can potentially be used as a means to 'protect' service provision in lower airspace from financial pressures that might otherwise occur.

9.4.3 Summary

As in the case of upper/lower charging differentiation, arguments to the effect that ACC pricing in particular, and reliance on competition among providers more generally, would itself give rise to significant pressures for improved ATM performance are, on the economics, unconvincing. In broad terms, there does not appear to be a high degree of substitutability among the services offered by individual ANSPs/ACCs, even in upper airspace. For example, application of the Commission's tests for a single European ATM market – as used, for example, by DGCOMP – would certainly be failed. 'Competition' among ACCs/ANSPs is therefore, at best, weak, and can be expected to remain so.

Where competitive effects are stronger – such as in relation to the upper airspace of smaller ANSPs – 'defensive' strategies, to mitigate already weak competitive pressures, are available. In particular co-ordination with neighbouring ANSPs/ACCs in upper airspace, whether in terms of charging agreements alone (which, in the limit, might be seen as a form of collusion) or in terms of operational consolidation, might be expected to be viewed as 'easier options' to ATM strategies aimed at improving performance. The greatest danger perhaps is in that limited restructuring, particularly in upper airspace, might be used as a means to weaken pressures to improve performance in the management of lower airspace, where a substantial fraction of overall costs are incurred.

A permissive approach to ACC pricing is, therefore, at best, a limited supplement to other measures that more directly address existing performance weaknesses in European ATM. As for the upper/lower differentiation option, there do not appear to be any convincing grounds for believing that this option, by and of itself, could be expected to lead to material improvements in network performance.

9.5 General conclusions on charging structure options

Consistent with economic analysis of networks in other sectors, both theoretical and applied, it is unlikely that the structure of charging for ANS, by and of itself, will have substantial effects on overall ANS performance. The main pressures for performance improvements are much more likely to come from a combination of more effective regulation of charge levels (rather than charge structures), the separation of Europe-wide co-ordination functions and their allocation to specialist network managers/ system operators, and direct pressure on providers exerted by airlines (emanating ultimately from the pressures of airline customers in competitive airline markets). Well constructed charging structures can be expected to have only a supporting (but nevertheless useful) role in guiding future developments in ATM.

Our general view on charging structures is that there is merit in developing existing arrangements in ways that reflect costs of complexity by means of origin and destination charges (i.e. reformed terminal charges), which should be set according to principles that will ensure greater consistency throughout the European network. In relation to the en-route component of ATM charges, whilst there are some arguments against allowing greater differentiation – so as to better translate what limited 'competitive' pressures there are in upper airspace into pressures to improve ATM performance at lower levels - on balance, we do not think that these are strong enough (because inter-ANSP competition itself is weak) to warrant anything other than a permissive attitude. That is, upper/lower airspace differentiation (defined by FL285) and charging by ACC should be allowed where ANSPs wish to introduce them. The magnitude of any such differentiation should, however, be limited by prohibitions against anti-competitive discrimination, whether those prohibitions derive from general European competition law or from sector-specific regulations.

BIBLIOGRAPHY

Airservices Australia *"Australian ATM Strategic Plan"*, Ed. 1.01, November 2001

Airways Corporation of New Zealand Limited *"Airways' Air Navigation Plan for New Zealand"*, May 1999

An Industry-wide Joint Proposal by the European Association of Aerospace Industries (AECMA), Airports Council International–Europe (ACI-EUROPE), the Civil Air Navigation Services Organisation (CANSO), Association of European Airlines (AEA), European Business Aviation Association (EBAA), European Regions Airline Association (ERA), International Air Carrier Association (IACA) and the International Air Transport Association (IATA) *"An Industry Consultation Body for the Single European Sky"*, April 2003.

Armstrong M., Cowan S. and Vickers J., "*Regulatory Reform: Economic Analysis and British Experience"*, The MIT Press, October 1994

ATC/EUC, *"ATFM improvements"*, Air Traffic Controllers European Unions Coordination

Averch, H., and LR Johnson *"Behavior of the firm under regulatory constraint"* American Economic Review, 52, 1962, 1053-69

CANSO, *"Position Paper on the Communication from the European Commission to the Council and the European Parliament on the Implementation of the Single European Sky"* February 2002

CEATS, *"Operational concept document",* CEATS, 2002

CEATS, *"Presentation background"*, CEATS, 2002

CEATS, *"Business plan"*, CEATS, 3 February 2003

Civil Aviation Authority, *"NATS' Application to Re-open the EUROCONTROL Charge Control"* CAA Decision, March 2003

Commission of the European Communities, *"Communication from the Commission to the Council and the European Parliament: the creation of the single European sky"* COM (1999) 614 final dated 1 December 1999

Commission of the European Communities, *"Amended proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL laying down the framework for the creation of the Single European Sky"* COM (2001) 123 final/2 dated 30 November 2001 - 2001/0060 (COD)

Commission of the European Communities, *"Amended proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the provision of air navigation services in the Single European Sky",* COM (2001)564 final/2 dated 11 December 2001 - 2001/0235 (COD)

Commission of the European Communities, *"Amended proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the organisation and use of the airspace in the Single European Sky"* COM (2001)564 final/2 dated 11 December 2001 - 2001/0236 (COD)

Commission of the European Communities, *"Amended proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the interoperability of the European Air Traffic Management network",* COM (2001)564 final/2 dated 11 December 2001 - 2001/0237 (COD)

Council of European Energy regulators *"Principle on the management and allocation of available transfer capacity of interconnections"*, CEER position paper, 30 September 2002

Council of European Energy regulators, "*Role and Work of the Council of European Energy Regulators: Proposals Presented at the VI Madrid Forum"*, Workshop on the Internal Market for Gas for Candidate Countries, Brussels, November 2002.

Council of the European Union, *"General rules for granting Community financial support"*, Regulation 2236/95 of 18 September 1995(amended by Regulation 1655/99 of 19 July 1999).

Council of the European Union, *"Common position adopted by the Council with a view to the adoption of the Regulation of the European Parliament and of the Council laying down the framework for the creation of the Single European Sky",* 11 March 2003 - 2001/0060 (COD)

Council of the European Union "*Common Position adopted by the Council with a view to the adoption of a Regulation of the European Parliament and of the Council on the provision of air navigation services in the Single European Sky",* 11 March 2003 - 2001/0235 (COD)

Council of the European Union, *"Common Position adopted by the Council with a view to the adoption of a Regulation of the European Parliament and of the Council on the organisation and use of the airspace in the Single European Sky"* 11 March 2003 - 2001/0236 (COD)

Council of the European Union, "*Common position adopted by the Council with a view to the adoption of a Regulation of the European Parliament and of the Council on the interoperability of the European Air Traffic Management network"*, 11 March 2003 - 2001/0237 (COD)

ETSO, *"Evaluation of congestion management techniques for cross-border transmission"* ETSO, November 1999

ETSO, "*Outline proposals for a Co-ordinated Congestion Management Scheme based on the ETSO Vision*", ETSO, September 2002,

ETSO, *"ETSO Position Paper on Locational Signals and European Transmission Charging"* ETSO, Prepared for the Regulators forum In Rome , October 17-18, 2002

European Commission *"A Single European Sky: Broadening horizons for Air Travel"* (2002)

European Commission Directorate General of Energy & Transport *"Single European Sky: Report of the high level group"* November 2000

EUROCONTROL Performance Review Commission, "*An assessment of Air Traffic Management in Europe during the calendar year 2002"* PRR 6 Version II, , May 2003

EUROCONTROL Performance Review Commission *"A comparison of performance in selected US and European en-route centres"*, May 2003

EUROCONTROL/CRCO "*Preliminary study: differentiation between upper and lower airspace –charging area defined independently from UIR/FIR boundaries",* internal memorandum

EUROCONTROL Enlarged Committee for Route Charges, *"Minutes of meetings of the Task Force on Possible Pricing Mechanisms"* minutes of 13 February 2001.

EURCONTROL CRCO *"Guidance on the Rules and Procedures of the Route Charges System"* (July 1999)

EUROCONTROL Performance Review Unit *"Year 2000 European ANS Providers Benchmarking Report*" ACE/2000 Data Analysis Working Group, September 2002

EUROCONTROL Performance Review Unit *"Fourth Performance Review Report",* PRR4, 2000

EUROCONTROL Economics *"Description of the Process for Economic Evaluation of ATM investments",* Issue 1.0/Document 12, 10 February 2003 (unpublished)

EUROCONTROL Economics *"Approach to Assess the Benefits and Costs of ATM investments"*, Issue 1.0/Document 20, 12 June 2003 (unpublished).

EUROCONTROL *"Report of the Economics Workshop -30 September 2002",* EUROCONTROL HQ, Brussels(2002)

EUROCONTROL *"CFMU Annual Summary 2002"*, Version 1.2, February 2003

EUROCONTROL Performance Review Commission, "*Special Performance Review Report on Delays (January – September 1999)"* PPR2, November 1999

EUROCONTROL Experimental Centre "*Analysis of unused ATFM slots 2000"*

EUROCONTROL, *"ECIP status report 2002"*, (2002)

EUROCONTROL "*Off to a successful start: CRCO implementation of Route per State Overflown (RSO)",* Skyway 1998

European Parliament and Council*, "Directive 96/92/EC of the European Parliament and of the Council of 19 December 1996 concerning common rules for the internal market in electricity"*, OJ L 27, 30.01.1997

European Parliament and Council, *"Directive 96/48/EC of 23 July 1996 of the European Parliament and of the Council of 19 March 2001 on the interoperability of the trans-European high-speed rail network"* OJ L235, 17.9.1996
European Parliament and Council, *"Directive 2001/16/EC of the European Parliament and of the Council of 19 March 2001 on the interoperability of the trans-European conventional rail system"* OJ L110/1, 20.4.2001

European Parliament and Council, *"Directive 2001/14/EC of the European Parliament and of the Council of 26 February 2001 on the allocation of railway infrastructure capacity and the levying of charges for the use of railway infrastructure and safety certification"*, OJ L75/39, 15.3.2001

European Parliament and Council *"Proposal for a Regulation of the European Parliament and of the Council establishing a European Railway Agency"* COM (2002) 23 final 23 January 2002

Held, Alvarez and Prat, "*Presentation for the Madrid Regulatory Gas Forum"*, 30/31 October 2002.

IATA, *"ATC in Europe: An IATA Action Plan for the New Millenium"*, International Air Transport Association CO219/9/99

IATA '*Quality of ATM Service including service level agreements'*, Draft Policy, Technical Operations Policy manual – Part B resolutions 13.04.04

ICAO*, "ICAO's policies on charges for airports and air navigation services"*, 6th edition 2001

Jaquard Philippe, *"Independent Study for the Improvement of ATFM"*, (IGACEM), 11 September 2000

Keyworth T, Decker C, Slater D & Yarrow G, *"The Regulation of Radioactive Waste Management in the UK"*, The Regulatory Policy Institute, February 2002

Laffont J.J and Tirole J., "*A Theory of Incentives in Procurement and Regulation*" The MIT Press, March 1993 1st.ed

Maastricht UAC, *"Annual Report",* 2001

NAV CANADA *"NAV CANADA business plan 2002-2005"*, 2002, Available on NAV CANADA website

NAV CANADA *"Backgrounder on rate stabilization account",* Available on NAV CANADA website

NAV CANADA *"Announcement of revised service charges"*NAV CANADA press release of July 21 2003

Nordic NUAC *"Project phase 1 report, from project vision",* NUAC, September 2002

Nordic NUAC, *"Institutional and Organisational Issues"*, Phase 1 report, Appendix 3 on 'November 2002

PriceWaterhouseCoopers "*Study of the Terminal Charges for Air Traffic Control Services*" Final Report – March 2001

Schubert Francis, "*The Financing of Cross Border Air Traffic Services, a legal perspective"* Skyguide, 2003.

Vickers J. and Yarrow G., "*Privatisation: An Economic Analysis"*, The MIT Press, (1989)

Vogel S. K. "*Freer Markets, More Rules: Regulatory Reform in Advanced Industrial Countries*" (Cornell Studies in Political Economy), , Cornell Univ Press, (September 1996)

Wilmer, Cutler & Pickering *"Study for the European Commission on the Regulation of Airspace Management and Design",* 14 May 2001

APPENDIX 1
REVIEW OF THEORY AND PRACTICE OF PRICE REGULATION

The objective of this appendix is to review some of the major general themes in price regulation, as they have developed in theory and practice across various regulated segments of the communications, energy, transport and water sectors. The review is not intended to be comprehensive.[147] Rather, the aim is to focus on material and on experience that is likely to be of greatest use in considering options for future approaches to ATM charging.

A.1.1 Cost of service regulation

Cost-of-service regulation is the name usually given to a regulatory approach that proceeds by:

- determining the relevant, or possibly relevant, products and services to be supplied, and estimating ranges for the likely volumes of supplies;
- estimating the costs of supply at the expected volume levels, including, within the category of capital costs, provision for a reasonable return on capital; and
- most crucially, setting charges/prices that are closely aligned to such costs over time (i.e. charging durations are short).

Since an allowable rate of return is included in the cost assessment, the approach is also known as *rate-of-return* regulation.

The chief advantage of this approach is that, for any given level of costs, those costs are closely reflected in charges/prices. Other things equal, such cost-reflectivity is good for *allocative* efficiency in the economy.

The chief disadvantage of cost of service regulation is that it provides poor incentives for reducing costs and for innovation generally. Any benefits from improvements in efficiency or from innovation will be very quickly passed on via the pricing mechanism, implying very little reward for the supplier.

Speaking broadly, **cost of service regulation has benefits in promoting *static* efficiency in resource allocation, but has major drawbacks as a means of encouraging dynamic efficiency.** In an economic environment in which suppliers are already operating efficiently (or in which efficient levels of costs can be estimated with confidence) and in which there is very little change, cost of service regulation can work well. On the other hand, if *either*

[147] For a more extensive analysis see *Regulatory Reform: Economic Analysis and British Experience* M. Armstrong, S. Cowan and J. Vickers, The MIT Press, October 1994

(a) existing suppliers are not considered to be fully efficient and efficient levels of costs can not be estimated with confidence, *or*

(b) the relevant market or sector is subject to major economic change,

cost of service regulation, in its 'pure' form, is unlikely to perform well.

In this context, it can be noted that:

- condition (a) has been a major factor in leading many countries to reject cost of service regulation when privatising utilities -- the established state-owned enterprises were widely believed to be relatively inefficient, and

- condition (b) can be expected to be a more prominent factor influencing regulatory decisions in the context of the Single European Sky -- a combination of growing demand, technological change, and regulatory/legal change (the Directive and its implementation) imply a dynamic economic environment.

A.1.2 Price-cap regulation

Strictly speaking, price-cap regulation involves some initial setting of maximum prices for designated products and services, such that those maximum prices are thereafter independent of either the conduct or the performance of the regulated undertaking for some relatively long period. In practice it is very frequently not actual prices that are so *predetermined*. Rather the caps tend to be set in terms of:

- price indices of 'baskets' of goods and/or services, or

- average revenues (total revenues divided by some measure of output), or

- total revenue, or

- some combination of the above (e.g. a total revenue cap to cover fixed costs and an average revenue cap to cover variable costs).

These apparently technical issues can be important for outcomes. For example, if total revenue is capped, revenue will be independent of fluctuations in the output of the business. This means that a provider with significant fixed costs is sheltered from risk associated with demand volatility, and that prices tend to rise when demand/output falls, and to fall when demand/output rises – a result that is very similar to the outcome with cost-of-service regulation. On the other hand, if it is average revenue that is capped, prices will largely be invariant to the level of demand/ output, and risks associated with demand volatility will fall much more on the provider – some fraction of fixed costs will not be recovered when demand is low, whilst fixed costs will be over-recovered when demand is high. A mixed cap, containing both total revenue and average revenue constraints, will have effects lying somewhere between the two, with demand volatility

risk shared between customers and providers in proportions reflecting the weights given to the two constraints.

The key aspect of price cap regulation is that average charges/prices are set so as to be independent of *controllable* (by the undertaking) costs for some significant duration. This does not necessarily mean that prices are fixed for the relevant period, since they can be indexed to *non-controllable* variables such as the rate of inflation -- as in RPI - X approaches -- or even to elements of costs that are deemed to be beyond significant influence by the relevant undertaking -- as in RPI - X + Y approaches, where Y is a variable reflecting non-controllable costs. The duration of the controls is a matter for determination, but a typical period is of the order of five years, in which case the arrangements can be interpreted as a five-year 'contract' between suppliers and users considered as a collective (i.e. not between suppliers and individual users), or alternatively as a form of cost-of service regulation with a five-year regulatory lag.

It is to be stressed that price-cap regulation is generally concerned with the setting of average prices (i.e. with the revenue requirement issue). Within this general framework, it is possible to implement variants that allow, or require, the detailed *structure* charges/prices to reflect the structure of costs. It is also possible to proceed more simply, by disaggregating the revenue requirement in a way that reflects broad cost divisions, for example as between fixed and variable costs. Thus, the control formula might be written so as to allow for a certain level of revenue (possibly indexed to inflation) to remunerate fixed costs and a *per-unit* allowance that links part of allowable revenue to the volume of supply.

The divorce between average prices and costs that is implied by price-cap regulation has the obvious advantage, relative to cost-of-service approaches, that it gives strong incentives to improve cost efficiency, since, at least until prices are adjusted, any reductions in costs will feed directly into higher profits.

The sharing of risks is also different as compared with cost-of-service regulation. Customers are more protected from risks associated with cost changes, which are borne to a greater extent by the supplier, although this depends partly upon the precise form of the arrangements.

It is often argued that price-cap regulation has disadvantages in terms of allocative efficiency (since cost changes are not quickly reflected in charge/price changes). Whilst true, this may not be of great quantitative significance in circumstances where the overall elasticity of demand for services is relatively low, particularly if the structure of prices reasonably reflects the structure of costs. For example, if the overall demand for rail travel is low but there is greater sensitivity in relation to time of journey, a price-cap arrangement that distinguishes between peak and off-peak periods may impose little burden in terms of lost allocative efficiency (because the relative price that is most important in influencing decisions on usage *does* reflect the relative costs of provision).

The more significant drawback of the price-cap approach is that it may create incentives to degrade the *quality of service*. Thus, costs can be reduced not only by becoming more efficient, but also by, in effect, offering lower quality of services to

customers. In practice, therefore, implementations of price-cap regulation tend to pay close attention to quality of service issues and, speaking broadly, the approach is only likely to be attractive where quality of service can be clearly defined and monitored.

This last point is one of the reasons why a pure form of price-cap regulation is problematic for ATM. Safety is clearly a highly important, though by no means the only, dimension of quality of service, and it can be difficult to monitor. For example, it might be difficult to detect, *ex ante*, changes in arrangements that led to small, but potentially very harmful, degradations in safety.

Other limitations of the price-cap approach include:

- Perverse incentives in relation to cost-reduction towards the end of the charging period, since reductions in costs might be perceived as likely to lead to reductions in prices at the upcoming review. Given that, once set, the revised charges/prices can be expected to persist for an extended period, the incentive to 'pad' costs in the period immediately preceding a review can be strong.

- **Unless there is a strong and credible commitment to ensuring that, when charges/prices are periodically re-set, the undertaking will be allowed to recover efficiently incurred costs, including a reasonable rate of return on capital, there can be problems of under-investment (see further below).**

A.1.3 Hybrid arrangements

Pure cost-of-service and price-cap regulation represents two ends of a spectrum of possible charging/prices arrangements and, in practice, most practical and effective forms of price/charge setting tend to lie somewhere between the two. For example:

- In the US, cost-of-service regulation has long been modified by (a) the existence of regulatory lag and (b) the practice of disallowing costs which have been assessed as inefficiently incurred. The effect is to restore greater incentives for cost-reduction. Increasingly, these *ad hoc*, adjustments to the framework have also been supplemented by more explicit *'incentive regulation'* provisions that have been introduced into the arrangements.

- Under price-cap regulation, the 'pricing period' tends not to be set for very long durations (e.g. of, say, 15 years or more), and price reviews tend to give considerable weight to cost levels, including a rate of return on the capital base or 'regulatory asset value' (RAV), when charges/prices are reset. This tends to prevent charges/prices from deviating too far, for too long, from underlying costs.

The real question, therefore, is not whether to opt for one or other of the 'pure' approaches to charge/price setting, but how best to resolve the various trade-offs that arise in particular cases among such considerations as allocative efficiency, cost efficiency, dynamic efficiency, the allocation of risk, quality of service, and so

on. Since the trade-offs differ across different types of communications, energy and transport networks, the ways in which those trade-offs are resolved also tend to differ.

In this context, it can be noted that the practical convergence of regulatory approaches to hybrid solutions is very much in accord with the implications of recent economic analysis, and specifically with the implications of *principal-agent* analysis (see further below). This tends to imply that, in the face both of inevitable demand and cost uncertainties, and of asymmetries of information, the best arrangements will tend to have the feature that part of, but not all of, any cost changes should be reflected in changes in charges/prices. However, the most appropriate 'sharing arrangements' will tend to vary according to the circumstances of the particular situation.

There is an important underlying point here. The same objectives and general principles, applied to different circumstances (e.g. different networks, or different parts of the same network that nevertheless exhibit significantly different characteristics) will, rightly, lead to variations in the specifics of regulatory arrangements (e.g. charging mechanisms). In section 6.2 we will set out an approach based upon a common pricing formula which nevertheless allows for adaptations to be made so as to reflect specific circumstances.

For the EU, a Single European Sky policy should mean that there are common objectives and that common principles be applied. This is quite consistent with the existence of variations in implementation at Member State level. Indeed, if there are variations in the relevant economic circumstances, the application of common objectives and principles positively requires some variation in implementation. What is important then is to ensure that actual variations in practice do, appropriately, reflect local circumstances, and that they are, therefore, consistent with the common objectives and principles. That common principles can lead to varying implementations does not provide an endorsement of varying implementations that do not reflect common principles.

On this basis, it is possible to identify three sets of tasks in the development of appropriate 'hybrids':

1. Establishing the common objectives and principles to be applied.

2. Applying those objectives and principles to develop the appropriate 'hybrids'.

3. Monitoring compliance, in the sense of ensuring that any specific arrangements that are developed are consistent with the objectives and principles.

A.1.4 Analytical work: the principal-agent approach

Much recent economic work has been based upon what is known as the principal-agent framework, and, as noted, there is clear linkage and correlation between this work and the actual evolution of regulatory practice. This linkage is two way: sometimes

theoretical work has led to improvements in practice, sometimes best-practice regulation has been ahead of, and stimulated, theoretical work.

The analytic, economic framework of principal-agent analysis has tended to develop in a highly abstract way, focusing very heavily on the technical detail. This can obscure the fact that it deals with, or has relevance to, many practical issues. Moreover, even at its most abstract, the approach can provide a useful guide to practical thinking, by clearly identifying the factors and trade-offs that are relevant to good practice.

Broadly speaking, the approach assumes that there is a *'principal'* who sets an *'incentive structure'* or *'contract'* for an *'agent'*, who is generally a commercial decision maker. Each party has its own (differing) objectives, and the parties may have different attitudes to risk. The aim of the principal is to determine the contract that best contributes to his (the principal's) own objectives, taking account of the likely conduct of the agent in the face of the incentives provided. Specifically in regard to this last point, the arrangements must ensure that:

1. The agent is able to obtain an average return sufficient, taking account of risk, to make the contract worthwhile (the *'participation constraint');* and

2. Proper recognition is taken of the fact that the conduct of the agent will likely be significantly influenced by the terms of the contract (the '*incentive compatibility constraint'*).

The approach has greater generality than this simple, hierarchical specification may suggest. There can, for example, be multiple principals and/or multiple agents. In regulatory contexts, the principal is normally the regulator, whether independent or a government department, and the agent is the regulated provider(s) of the relevant service. However, the contracts can in principle also be set by the purchaser of the relevant service (e.g. in public procurement examples), or by 'contractual' agreements between users and providers, possible overseen by a regulatory authority. Extensive theoretical development of the approach is set out in *Laffont and Tirole*.[148]

In ATM, there has traditionally been little distinction between providers and 'regulators', the two activities having been bundled together in past periods. The increasing separation between provision and regulation is, however, moving the sector towards a structure more closely aligned to best practice in other network industries, and hence the insights of the principal-agent approach are becoming increasingly applicable to the types of issues with which this report is concerned.

In particular, three sets of highly related issues, which need to be addressed simultaneously, are highlighted by the approach. These can be summarised in the following questions:

[148] *A Theory of Incentives in Procurement and Regulation* J.J Laffont and J. Tirole, The MIT Press, March 1993 1st.ed

1. To what extent is it appropriate to incentivise providers?

2. How is the burden of risk best allocated as between providers and their customers?

3. What information is available to the parties when determining the contract?

   For example:

   a. With what precision can outputs, including quality of service, be defined and measured?
   b. How good is the overall quality of the information that is available?
   c. What are the incentives on the parties to produce and reveal information?
   d. How is relevant information distributed as between the parties?

At this point there is some divergence between theory and practice, in that the theoretical literature usually takes a relatively simple approach that focuses on the *asymmetry* of information between regulator and regulated (e.g. the provider is assumed to know a lot more about its own cost structure, and its own potential for cost reduction and for service quality improvement, than does the regulator). In practice, however, this is often not the main issue, at least in the early stages of the development of regulation of monopoly providers. Rather, the more significant practical problem is often one of generally poor information: a*ll* parties are poorly informed about the relevant trade-offs, options and opportunities.

A major reason for this is that unreformed arrangements for service provision -- particularly when they are based on 'cost-plus' approaches to charging -- often place very little pressure on providers to obtain information about their own cost structures or about the possibilities for providing innovative services to their customers. Rather, the pressures are often the other way round: *it is better not to know the detail of costs*. More information on such matters might reveal egregious inefficiencies that would show the provider in a bad light, or might indicate the extent of cross-subsidies that neither government or the provider would want to come to light. It would, of course, be possible to attempt to withhold information once it was obtained, but the risks involved in this strategy can be avoided by not knowing in the first place.

Thus, whilst the economic literature places much emphasis upon the implications of the distribution, among parties, of existing information (the asymmetric information issues), there is a more fundamental question concerning lack of pressures/incentives to *discover* new information. The centrality of *discovery* in the economic process was a theme of the Austrian School of Economics, but the notion tends no longer to be found in mainstream thought other than in application to R&D intensive industries such as pharmaceuticals and micro-processors. In fact, the concept is useful in a wide variety of contexts, and in none more so than in network industries that are undergoing a process of reform.

Given this, one of the early tasks in reform processes is to seek means of improving the quality of available information, at least until such time as higher quality

information is produced, automatically, from improved arrangements. In another context, we have called one aspect of this the *information-formatting* role of regulation. [149]

A.1.5 Incentive regulation

Developments in economic analysis reflect, and are reflected in, the numerous attempts to implement *incentive regulation* across sectors and across countries. *Incentive regulation* is a broad term, which encompasses diverse approaches that range from very modest attempts to introduce marginal incentives to improve the performance of service providers to much more radical approaches that seek to establish strong links between financial returns and performance, similar to those to be found in competitive markets.

In popular usage, the term is often applied only to the latter, 'high-powered' incentive arrangements, and most particularly to arrangements that seek to provide strong incentives for cost cutting. This is both wrong and unduly restrictive. Incentives might, for example, be provided to increase investment or to increase safety, leading to *increases* in operating costs, if, in the circumstances, that was the implication of the underlying policy objectives and principles. Similarly, particularly where quality of service is difficult to measure or the provider has only weak control over outcomes, there are situations in which it is appropriate that incentives be 'low powered'. Throughout this report, whenever we use the terms 'incentives' or 'incentive regulation' we do so in this more general, neutral (in relation to strength and direction) sense.

Two general points can, however, be made about incentive regulation -- one of principle and one of practice:

- In all cases, the aim is to establish some link (whether weak or strong) between the charges made for service provision (and hence the financial revenue of service providers) and the performance of providers in responding to the requirements of their customers.

- Experience suggests that it is sometimes the case that, contrary to the implications of much, but not all, theoretical analysis, even incentive arrangements that are very modest in their financial implications can have substantial effects on performance, particularly when incentives are being introduced for the first time. Explanations for this phenomenon are not well developed, but one possibility is that the exercise of *specifying* and *measuring* performance improvements that are of value to customers can lead to a shift in the business focus of the service provider. Put simply, **the provider becomes more *customer-focused*, and this alone can lead to major changes in business conduct.**

Among the various types of incentive regulation that can be identified are included: cost-of-service regulation with regulatory lag; price-cap regulation with fixed contract duration; 'benefit sharing' arrangements, possibly with maximum and minima on

[149] Keyworth T, Decker C, Slater D & Yarrow G, 'The Regulation of Radioactive Waste Management in the UK', *The Regulatory Policy Institute*, February 2002

financial exposures (see section 6.2); and yardstick regulation. The last of these may be less familiar than the others in discussions concerning ATM, and some explanation at this early stage may be warranted.

(a) Yardstick regulation

Yardstick regulation can be implemented when there is a multiplicity of providers who are comparable, as for example when a number of monopolistic providers offer similar services but within different economic areas (e.g. local or regional water service providers, local or regional electricity or gas companies). The approach seeks to link the charges made by any one provider to the performance of *other* providers. This replicates that property of competitive markets which links the rewards obtained by any one supplier to its success in meeting customer requirements *relative to the success achieved, in the same task, by other providers*. In effect, a form of competition is established -- which again may be either weak or strong, depending upon the precise charging rules adopted -- among service providers, even though each enjoys monopoly power within its own geographic area.

Perhaps the simplest example of such an approach in the ATM context would be an arrangement in which a *single en-route* charge per mile, for a given type of aircraft, was applied across all European airspace. Such a charge could be based upon the average cost of all relevant providers (e.g. sufficient to allow aggregate cost recovery), and could be combined with a pre-set sharing rule that reflected, at least to some extent, any cost variations deemed to beyond the control of the providers. Once established, however, the arrangement would establish incentives to reduce costs since, for example, a 10% reduction in the costs of any one provider relative to others would lead, even on a cost-plus basis, to a much smaller percentage reduction in its charge revenues.

In such an example, it can be noted that it is the set of sharing rules, which might be a revenue pooling and sharing agreement, that determines the strength of the incentives. All options are possible, from simple cost plus to 'high powered' incentives that allocate to a provider virtually all of the benefits of cost reductions that it achieves. The extent to which providers are incentivised can, by appropriate choice of rules/agreement, differ as between individual providers.

Under such arrangements, it can be noted that:

- Incentive arrangements can, in effect, be detached from issues concerned with the detail of the charging structure.

- It is, therefore, possible to have a very simple, common charging structure whilst at the same time leaving substantial discretion as to the extent to which they wish to incentivise the local provider(s).

- The complexity lies in the revenue sharing/pooling arrangements, which require inter-provider agreement, not in the charging structure facing network users.

A.1.6 Investment issues

The starting point of the modern economics literature on economic regulation was the observation that the imposition of controls on prices/charges was liable to lead to a distortion in investment incentives. Specifically, pure cost-of-service regulation in the traditional US-style might be expected to lead to over-investment and 'gold-plating' of service provision, a result that became known as the Averch-Johnson effect.[150]

On the other hand, stimulated by the adoption in the UK of what was, in effect, a reasonably close approximation to 'pure' price cap regulation (the first BT price cap), later work has shown that this latter type of price control tends toward bias in the opposite direction, towards under-investment. Two possible reasons for this are:

- A price cap with, say, a five-year contract duration (i.e. reviews are conducted every five years) is similar in effect on the regulated undertaking to a five-year supply agreement at pre-determined prices. Such a contract says nothing, however, about what prices might be expected at the end of the relevant period. Hence, investments in assets that are expected to be productive beyond the end of the contract/pricing period, are necessarily subject to regulatory uncertainty. Worse, the regulator has substantial influence on what those prices will be and therefore appears equivalent, in the eyes of the supplier, to a monopsonistic buyer.[151] Once the investment is made, the regulator/monopsonist will be in a strong bargaining position, since the relevant capital costs are likely to be non-recoverable/sunk. In the limit, the regulator could even influence prices down to average variable costs, in which case continued operation would make no contribution to the recovery of previously incurred investment costs. Recognising this problem of possible *regulatory opportunism*, the regulated undertaking will have diminished incentives to invest, particularly in long-term projects. This is an example of an investment *hold-up* problem.

- In pursuit of allocative efficiency, a regulator may be tempted, or may be required, to impose a tariff that reflects short-term marginal costs. In such circumstances a firm will under-invest since, by so doing, it can create capacity constraints, higher short term marginal costs, and hence higher prices and profits (see *Vickers and Yarrow* [152]).

Concerns about investment incentives (which are a key issue for ATM) have been a major influence in the development of hybrid approaches, which seek to address both shorter term issues (e.g. use of existing capacity) and longer term issues (e.g. expansion and contraction of capacity) simultaneously. As already stated, the most appropriate form of charging arrangements will depend upon particular circumstances and contexts -- there are no 'magic formulae' with general applicability. However, a

[150] Averch, H., and LR Johnson 'Behavior of the firm under regulatory constraint,' *American Economic Review* 52, 1962, 1053-69

[151] A monopsonistic buyer is the equivalent of a selling-side monopolist, however, in this case the monopsony is the only buyer in a market.

[152] *Privatisation: An Economic Analysis* J.Vickers and G. Yarrow, G , The MIT Press, (1989)

number of general implications concerning the approach to investment can be drawn from the type of work discussed above:

- Allowing providers to secure revenues on a cost-plus basis, including a reasonable return on capital, will, if unsupported by other measures, tend to lead to 'gold plating'. That is, if all project costs can be automatically recovered, there will be a tendency to incur extra expenditures that, if exposed to a rigorous cost/benefit evaluation would not be justified. Although such an approach encourages a high (possibly excessive) level of investment, it does nothing to promote high quality investment. Thus, projects may be developed at excessively high costs, or investment may be directed to inappropriate areas or lines of business activity. Particularly in complex networks, measures to counter-act these tendencies would imply highly intrusive regulatory intervention in the investment process; and the track record of such intervention is not good.

- Price-cap regulation, if unsupported by other measures, can not be expected to address investment issues. There needs to be longer term reassurance that regulators will not behave opportunistically. Policy credibility and confidence in the stability of regulatory decision making are key to investment performance.

- Arrangements that are successful in promoting efficient investment are likely to be characterised by some unbundling of the way in which pricing and investment issues are addressed. In all the above cases, it is *linkage* between investment decisions and prices that causes the problems. Thus:

    - In the Averch-Johnson analysis of rate-of-return regulation, it is because higher investment leads to higher prices that 'gold plating' (excessive investment expenditure) emerges.

    - In the hold-up case, it is the threat that prices will be set artificially low that discourages investment.

    - In the short-run marginal cost pricing case, it is because investment leads to lower prices that investment is discouraged.

A.1.7 Network externality issues

One of the principal features of networks is that they are characterised by the existence of externalities. An externality in this context is defined as an economic effect that falls on providers and users from network transactions to which they are not parties.

This is a complex area, and each network has its own features. Rather than attempt a general discussion, therefore, we will move straight to application of the general principles of regulatory good-practice to the specific case of ATM.

Suppose that an ANSP in one member state decides to add to capacity, so as to increase the number of flights that it can handle. Such a decision may have effects on other blocks of airspace and on the costs of other ANSPs. For example, if there is lack of

congestion elsewhere, both before and after the capacity expansion, the effect may simply be to change the number of flights in those other blocks. The volume effect here could go either way. By relieving a congestion constraint, the investment may increase the number of flights, including flights in other airspace blocks. Alternatively, flights that had been diverted so as to travel greater distances in other airspace might, following the investment, no longer be so diverted. The precise effects on volumes therefore depend upon the specifics of the relevant circumstances. **The bottom line, however, is that the investment of one ANSP can be expected to have effects, at fixed prices/charges, on the revenues of other ANSP.**

Similarly, the investment may have effects on the costs of other ANSPs. One such impact occurs via any change in cost related to the changes in volumes just described – the variable cost effect. Another type of impact occurs when relieving congestion in one part of airspace, by expanding capacity, has the effect of creating greater complexity in other parts of airspace. Since, other things equal (eg: number of movements), complexity tends to raise costs, 'negative cost externalities' will then be imposed on other ANSPs.

Where they exist, externalities have adverse consequences for both the operation and the development of networks. It is important, therefore, that they be reduced whenever this can be achieved without incurring costs that are disproportionate to the resulting benefits, and there are basically three ways of doing this, in descending order of the level of centralisation required:

1. 'Internalise' the effects by consolidation of providers of network services.

2. Establish a system-wide set of prices/charges and set of investment incentives that reflect the effects.

3. Establish, by agreement, initial 'rights' and then permit different providers to enter into contractual negotiations with one another.

These approaches are not mutually exclusive, and it may that some combination is the most appropriate way forward (just as hybrid approaches to price regulation have become the preferred option of best-practice regulators). For example, an initial framework agreement might establish 'rights' but also go some way toward setting out pricing signals and investment incentives in order to deal with some, but not all, of the external effects (or, alternatively, the latter may be set out in another documents, such as a Directive). Subsequent contractual agreements among providers might then lead to some consolidation of service provision. Such a development can be observed occurring in electricity networks as adjacent system/network operators combine their activities – a process that is, in a number of respects, similar to potential joint ventures among ANSPs.

APPENDIX 2
DEVELOPMENTS IN OTHER EUROPEAN NETWORK INDUSTRIES

In this appendix we briefly review the recent developments in the European networks for electricity and rail in regard to network management and charging principles. The first part of this appendix (A.1) presents possible ways of addressing the problems with the current ATFM process in Europe as identified in the main report above, based upon the approaches to system operation adopted in the electricity and rail networks. The second part of this appendix (A.2) discusses the recent developments in the charging principles in the European-wide internal markets for gas and electricity.

A.2.1 System operation in the European electricity network

As with the initiatives relating to more efficient use of existing capacity in ATFM, the development of congestion management approaches and principles to allow for flows across the borders of different nation states has been a major focus of the creation of the internal electricity market in Europe.[153]

In electricity, issues of system operation and congestion management have been addressed both by the Council of European Energy regulators (CEER) and the European Transmission System Operators (ETSO) organisation.[154] A first principle of the CEER in undertaking its work is that:

> *Congestion management method(s) should operate in an economically efficient manner and, when it is possible, provide appropriate economic signals for both efficient economic dispatch of existing plant and efficient investment in additional network infrastructure and/or generation*[155]

Similarly, ETSO propose that congestion management approaches adopted in each member state provide the appropriate economic signals regarding short term congestion management, they note that:

> *It is absolutely essential that the procedures adopted provide the appropriate economic signals to generation, demand and TSOs, both in terms of short-term operational actions and longer-term investment decisions.*[156]

[153] "Directive 96/92/EC of the European Parliament and of the Council of 19 December 1996 concerning common rules for the internal market in electricity", OJ L 27, 30.01.1997

[154] The Council of European Energy regulators (CEER) brings together energy regulators from fifteen European Union (EU) and European Economic Area (EEA) Member States and was formally constituted on March 7th, 2000. The CEER acts as a focal point for contacts between regulators and DG TREN. The European Transmission System Operators (ETSO) represents 32 independent transmission system operators in the European Union.

[155] "Principle on the management and allocation of available transfer capacity of interconnections", CEER position paper, 30 September 2002, page 1

[156] "Evaluation of congestion management techniques for cross-border transmission" ETSO, November 1999, page 14

(a) Defining a 'transmission product'

In ETSOs view an initial step in congestion management involves the development of some form of flow or transmission 'product' which can be allocated and traded among participants. The advantage of the creation of a 'transmission product' for national transmission system operators is that it provides a measure of assurance that system users will operate at a given level, rather than above or below it. For system users the creation of a 'transmission product' introduces a certain level of guarantee or assurance that the available capacity requested in the future will be there when required. This, of course, depends upon how 'firm' or binding the commitments of the transmission product are on system operators.

Therefore, an important recommendation of ETSO is the proposal that the transmission product be firm on system operators. While recognising that this exposes system operators to additional risks, which may lead to conservative and risk-averse behaviour, they nevertheless see 'firmness' on system operators as important for efficiency as it introduces an incentive for system operators to accurately predict and manage transmission capacity.

(b) Potential Congestion Management Methods

In electricity, ETSO have also assessed the merits in terms of efficiency of a number of congestion management methods for co-ordinating system operation across the European electricity network, including:

(i) Capacity Curtailment

Capacity curtailment as a congestion management technique involves restricting the amount of trade between different states to the maximum published net transfer capacity. This is in many ways similar to the current ATFM system of imposing regulations on *en-route* slots relative to the available declared capacity for each ACC sector. ETSO conclude that capacity curtailment has the advantage of not requiring a particular capacity allocation or cost allocation mechanism in the event of system congestion, as those transactions which would cause overloads are simply rejected according to a pre-defined priority rule. However, a major drawback with capacity curtailment identified by ETSO is the lack of economic incentives it conveys to system users and system operators, and that this has implications in terms of the promotion of efficient trading and overall system operation.

(ii) Auctions for capacity

On the basis of its analysis, ETSO see auctions as potentially forming the basis for several congestion management methods. They note that the principal advantage of auctions as a congestion management technique is that they allow market participants to resolve amongst themselves issues of congestion at different parts of the network through the bids that they submit. They note that auctions are also economically efficient, as the bids submitted reflect the real market value of capacity as perceived by participants. An additional advantage of capacity auctions is that they may provide

efficient signals for system operators as to the areas where further development of the network may be required. While recognising the additional system complexity introduced by auctions, ETSO concludes that this complexity is not insurmountable and that methods exist to handle these situations.

(iii) Capacity re-dispatch

An alternative method of congestion management considered by ETSO involves the re-dispatching of capacity within the system to relieve areas of the congestion. In effect this means redispatching capacity internally within a local system until the physical limit of redispatching possibilities is reached, which extends the possibility for cross-border flows beyond the fixed capacity through an internal reorganization of capacity. In some ways this is conceptually similar to the Miles in Trail approach to congestion management adopted in US ATFM, that is where local congestion management measures are adopted – through defining the distance between two consecutive aircraft in a given flow – to alleviate congestion at a number of bottlenecks elsewhere in the network.

(c) Priority rules for capacity curtailment close to real time

Finally, ETSO consider the merits of the various priority rules for capacity curtailment. ETSO note that as a general principle: "..it is generally accepted that 'market-based' methods (essentially auctions, market splitting and re-dispatch) are preferable to rule based methods such as 'first-come-first-served"[157]

The 'first come-first served method', which is the current priority rule adopted for ATFM, is assessed by ETSO as encouraging participants to make longer term forecasts and has the benefit of providing system operators with 'better and sooner' information on the volume of future capacity in advance. However, a major drawback of the 'first come-first served' priority rule in ETSO's view is that it necessarily limits the amount of short term activity that can occur, which it considers an important requirement to ensure overall system efficiency. In order to address this weakness, ETSO outlines three possible solutions:

- the introduction of penalties charged to users who do not use their requested capacity
- the setting aside of some capacity for short-term trading
- the introduction of a 'use-it-or-lose-it' rule to allow for the release of capacity that is unused in the short term

Consistent with the discussion above, the possible applications of these findings for European ATFM might involve the introduction of penalties for those system users who book *en-route* slots but do not use them. Alternatively the introduction of a 'use-it-or-lose-it' rule for booked capacity might ensure that potential unused slots are reallocated within the system close to real time.

[157] "Outline proposals for a Co-ordinated Congestion Management Scheme based on the ETSO Vision, ETSO, September 2002, page 2

A.2.2 System Operation in the European railways internal network

Issues relating to system operation also arise in the context of the European railways network and are addressed in the recent European Directive regarding the allocation of railway infrastructure capacity. [158] Specifically, paragraph 30 of the Railways infrastructure directive highlights the issue in relation to the European railways network:

> *A lack of information about other railway undertakings' requests as well as about the constraints within the system may make it difficult for railway undertakings to seek to optimise their infrastructure capacity requests.*[159]

The position taken to system operation in the Railways directive is similar to that adopted in the electricity directive which sees co-operation between infrastructure managers in different parts of the network as the key mechanism that will facilitate the efficient operation of the network. Article 15 of the Railways directive discusses the principles that should be observed when developing cooperative measures in the allocation of infrastructure capacity on more than one network. The key points of this Article are that:

- Infrastructure managers are responsible for developing procedures to enable the efficient creation and allocation of infrastructure capacity
- The procedures developed to coordinate the allocation of infrastructure capacity shall involve representatives of infrastructure managers whose allocation decisions have an impact on more than one infrastructure manager.
- Decisions or activities as to the allocation of capacity shall only be taken by representatives of infrastructure managers

The railways directive allows for a number of different possibilities for dealing with a congested network in the short term including:

- the possibility of charging to reflect congestion (to reflect the scarcity of capacity) at different points of the network
- the use of a priority criteria to allocate infrastructure capacity which should be set out in the network statement. The precise priority criteria which is to be implemented is left open in the Railways Directive, however, it does note that any criteria must take account of the importance to society of each service relative to that service which is being excluded.

In terms of the specific congestion management techniques that can be adopted by infrastructure managers, the Railways directive allows for the possibility that charges can be levied on those users whose actions disrupt the operation of the network and for compensation for those undertakings that suffer as a result of this disruption.[160] Further the Railways Directive allows infrastructure managers to levy an appropriate charge for

[158] "Directive 2001/14/EC of the European Parliament and of the Council of 26 February 2001 on the allocation of railway infrastructure capacity and the levying of charges for the use of railway infrastructure and safety certification", OJ L75/39, 15.3.2001

[159] Point 30 of the Preamble to Directive 2001/14/EC on railway infrastructure capacity

[160] Article 11(1) of Directive 2001/14/EC on railway infrastructure capacity

capacity that is requested but is not used, and this is seen as providing incentives for the efficient use of capacity.[161]

A.2.3 Charging principles in European energy networks

As noted in Section 9 above, charging for use of electricity and (to a lesser extent) gas transportation networks is increasingly moving toward entry/exit systems in Europe. The basic idea of entry/exit charging is that system users pay charges at the points at which they are connected to the relevant system (in the Nordic countries, the payments are explicitly called 'point of connection' charges). In electricity, for example, the points of connection are the locations at which power enters the transmission grid from generators (entry points, generally labelled G) and at which power leaves the grid destined for customers (exit points, generally labelled D (demand) or L (load).

In consequence of the physical characteristics of power flows, distance factors can be incorporated into the G and L charges automatically, such that the sum of the two charges is higher or lower according to the distance between them, the direction of power flows, and other factors that affect the costs of transportation. However, users also may pay other charges for the services supplied by the 'system operator' (SO) or for other costs that their activities impose. Thus, energy losses may be greater when power is carried over longer distances, and this cost variation may be reflected in charges. It is this latter tri-partite division – entry + exit + distance related charges – that is of particular interest since, in form, it resembles the current structure of ATM charges: terminal charge at point of origin + terminal charge at point of destination + *en-route* charges.

Entry/exit charging structures appear to be emerging as the preferred way forward for the pricing of EU energy networks. Thus, in a recent presentation from the Council of European Energy Regulators (November 2002), which works with the Commission on the relevant issues, Sergio Ascari of the CEER Brussels Office summarised, for the benefit of candidate countries, CEER's proposals for the development of European charging structures at the VI Madrid Forum, together with the rationale for those proposals.[162] These are, as summarised:

- Harmonisation of the tariff system: entry-exit except if proven infeasible.
- Transparency of available capacity: data at interconnection points and borders to be published
- Capacity allocation and congestion management: non discriminatory rules
- Monitoring of interoperability progress after TSOs initiative

[161] Article 12 of Directive 2001/14/EC on railway infrastructure capacity

[162] *Role and Work of the Council of European Energy Regulators: Proposals Presented at the VI Madrid Forum*, Workshop on the Internal Market for Gas for Candidate Countries, Brussels, November 2002. See also Held, Alvarez and Prat, *Presentation for the Madrid Regulatory Gas Forum*, 30/31 October 2002.

In arriving at these proposals, the CEER had assessed alternative tariff systems against a number of criteria, established at the V Madrid Forum:

- *Competition and tradability (including non-discrimination, particularly as between large and smaller shippers).*
- *Adaptability*
- *Cost reflectivity/cost recovery*
- *Simplicity/transparency*

Similarly, within the process that is developing the arrangements for cross-border transmission of electricity, ETSO (the association of European Transmission System Operators in electricity) has stated, in a paper for the CEER Forum in Rome on 17/18 October 2002,[163] that, in the context of seeking to link markets in Member States:

34. What we are looking for is a way of linking a number of separate mechanisms. To do this, it can be thought of that the charges for transmission are made up of:

1. A Local Entry (G) or Exit charge (L)

2. An additional non transaction based charge for transmitting across from one local/regional market to another

35. This means that the emphasis should be on apportioning costs incurred in transmitting between one local/regional system and another. It will also need to be considered if it is appropriate to have a compensation mechanism for transit between the different areas.

36. Charges could take the following forms:

- *Local G and L in each local market for injecting /offtaking determined by subsidiarity*
- *Some form of "transmission fee" between one local market and another*

37. To define the latter charge, the following steps are required

a) Decide what costs are to be included (sunk investment, losses, congestion, operation and system services).

b) Determine what part of these costs should be assigned to cross border trade

c) Determine who causes the costs

[163] ETSO Position Paper on Locational Signals and European Transmission Charging

d) *Determine a mechanism for obtaining Funds and transferring to the local market which incurs the costs.*

Whilst, as already implied, the underlying economic position in ATM differs in a number of important respects, the similarities in the problems being addressed by regulators and DG TREN in the energy sector suggest to us that these other streams of work can, at a minimum, inform the Single European Sky deliberations to an appreciable extent.

It may also be worth noting that, although an entry + exit + *en-route* arrangement leads to charge differentiation that is less closely aligned with operational activity boundaries than some other options, it *is* closely aligned with the structure of services required by, and supplied to, users, namely control from a point of departure, to a point of destination, along an allocated flight path. That is, it fits better with a demand-side approach to service differentiation (unbundling) than with a supply-side approach.